The Resemblance Structure of Natural Kinds

A Formal Model for Resemblance Nominalism

Javier Belastegui Lazcano

Supervisors: Dr. Thomas Mormann
Dr. Jon Pérez Laraudogoitia

A thesis presented for the degree of
Doctor of Philosophy

University of the Basque Country
November 2020

Declaration of Authorship

I hereby declare that this dissertation, entitled "The Resemblance Structure of Natural Kinds: A Formal Model for Resemblance Nominalism" and written under the supervision of Thomas Mormann and Jon Pérez Laraudogoitia, is the result of my own original research. To the best of my knowledge and belief, it contains no material previously published or written by another person, except where it is explicitly acknowledged in the text. No part of this thesis has been submitted for a degree at any other university or institution.

Javier Belastegui Lazcano

Para Ane,
que se sabe todas mis canciones.

Abstract

The aim of this thesis is to better understand the ways natural kinds are related to each other by species-genus relations and the ways in which the members of the kind are related to each other by resemblance relations, by making use of formal models of kinds. This is done by first analysing a *Minimal Conception of Natural Kinds* and then reconstructing it from the ontological assumptions of *Resemblance Nominalism*. The questions addressed are:

1. What is the *external structure* of kinds? In what ways are kinds related to each other by species-genus relations?

2. What is the *internal structure* of kinds? In what sense are the instances of a kind similar enough to each other?

According to the Minimal Conception of Natural Kinds, kinds have two components, a set of members of the kind (the extension) and a set of natural attributes common to these objects (the intension). Several interesting features of this conception are discussed by making use of the mathematical theory of concept lattices. First, such structures provide a model for contemporary formulations of *syllogistic logic*. Second, kinds are ordered forming a complete lattice that follows *Kant's Law* of the duality between extension and intension, according to which the extension of a kind is inversely related to its intension. Finally, kinds are shown to have *Aristotelian definitions* in terms of genera and specific differences. Overall this results in a description of the specificity relations of kinds as an algebraic calculus.

According to Resemblance Nominalism, attributes or properties are classes of similar objects. Such an approach faces Goodman's *companionship* and *imperfect community* problems. In order to deal with these a specific nominalism, namely *Aristocratic Resemblance Nominalism*, is chosen. According to it, attributes are classes of objects resembling a given paradigm. A model for it is introduced by making use of the mathematical theory of similarity structures and of some results on the topic of *quasianalysis*. Two other models (the polar model and an order-theoretic model) are considered and shown to be equivalent to the previous one.

The main result is that the class of lattices of kinds that a nominalist can recover uniquely by starting from these assumptions is that of complete coatomistic lattices. Several other related results are obtained, including a generalization of the similarity model that allows for paradigms with several properties and properties with several paradigms. The conclusion is that, under nominalist assumptions, the internal structure of kinds is fixed by paradigmatic objects and the external structure of kinds is that of a coatomistic lattice that satisfies the Minimal Conception of Kinds.

Agradecimientos

My greatest intellectual debt is to my supervisor Thomas Mormann. The originality, rigour and elegance of his proposals have been a constant source of inspiration for my work. So much so that many parts of this thesis can be seen as attempts to extend some of his ideas and defy others. He has not only made invaluable comments to this thesis, but has also transmited his fascination with mathematics to me. With exemplary tolerance, trust and wisdom, he has given me freedom to explore different possibilities while preventing me from making important mistakes. He has supported me even after retiring, at more than ten thousand kilometres away from this university. For all these reasons, I am greatly indebted to him.

Quiero agradecer también a Jon Pérez, que aceptó amablemente mi petición de incorporarse como co-director para poder finalizar este proyecto.

La presente investigación ha sido financiada con dinero público. Por ello, agradezco al Gobierno Vasco por haberme concedido la ayuda para mi contrato predoctoral con la Universidad del País Vasco, y a ésta última por contratarme.

During the Winter of 2019, I made a brief stay at the Munich Center for Mathematical Philosophy (MCMP), with a grant provided by the institutions just mentioned. I would like to express my deepest gratitude to Hannes Leitgeb, whose work strongly influenced crucial parts of this thesis, for inviting me to join the MCMP. When I arrived at the MCMP I was warmly welcomed by many wonderful philosophers, including Matteo, Michal, Albert, Fausto, Ivano and Neil. I want to thank them for their help and for all I learnt from their discussions. Special thanks go to Bruno, whose criticism to several parts of this thesis was invaluable.

I am extremely grateful to Mattias and Sophie, for reading and commenting on many parts of this thesis and for constantly transmitting to me their philosophical enthusiasm. I would also like to extend my gratitude to those who have made questions and objections in the talks I have given including the ones in the Jornadas de Avances, thanks to Adrian, Guglielmo and José Vicente.

Mi más sincero agradecimiento a todos los miembros del núcleo y periferia del grupo Praxis. Mila esker, en primer lugar, a su I.P. Andoni, por su constante ayuda en el ámbito administrativo, y en segundo lugar, a sus estudiantes y seniors, entre otros a Caterina y Nasim, Maite Arraiza, Ekai, Iñaki y Hannot. En particular, gracias a Sergio y a Ainhoa, con quien he compartido despacho durante estos cuatro años, tanto por sus sabios consejos como por soportar mis divagaciones frecuentemente. Gracias a Amaia y a Maite, por su apoyo incondicional y por las interminables conversaciones y paseos.

La realización de esta tesis no habría sido posible sin el apoyo de mis amigos, tanto de los que pertenecen al ámbito académico como de los que no. Entre otros, gracias a Gorka, Lourdes, Alba, Davi y Charly. Tampoco habría sido posible sin el de mi familia, el aita, la ama, mi hermana y Simone, que han cuidado de mí y me han proporcionado una educación formidable. Gracias en especial a Julen, cuya ambición, creatividad y curiosidad ilimitadas han sido una fuente inagotable de motivación para mí.

Por último, gracias Ane. Gracias por cenar conmigo, por cantar conmigo y por hablar conmigo. Si no me he roto durante estos años ha sido gracias a ti. Gracias por estar ahí.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

SOCRATES: The first consists in seeing together things that are scattered about everywhere and collecting them into one kind, so that by defining each thing we can make clear the subject of any instruction we wish to give.

Phaedrus
PLATO

1.1 The Problem of Natural Kinds

Natural kinds are *objective kinds or sorts of things.* They are ways things are supposed to be arranged, organized or sorted 'out there in the world'. These ways things are sorted into impose a minimally stable structure in the world. The initial metaphor is that of a giant dovecote or a big wardrobe where every object in the world has its place. When we classify objects, some of these ways of sorting them seem to be more appropriate than others. Examples of kinds could be pinguins, tulips or water. In contrast, garbage, grue things, all those liquids that do not taste bitter to a human being who is 35 years old or the things that were on my desk this morning, are not natural kinds. There is something odd about these latter groupings. Although they might have been chosen according to some pragmatic criteria and even if they may support successful inferences in some restricted contexts, they are arbitrary collections of individuals that strongly depend on our interests and conventions. Or at least, this is how the realist sees the problem[1].

A famous fragment of Borges' tale "El Idioma Analítico de John Wilkins" helps to illustrate this point[2]:

> "Esas ambigüedades, redundancias y deficiencias recuerdan las que el doctor Franz Kuhn atribuye a cierta enciclopedia china que se titula *Emporio celestial de conocimientos benévolos.* En sus remotas páginas está escrito que los animales se dividen en a) pertenecientes al Emperador b) embalsamados c) amaestrados d) lechones e) sirenas f) fabulosos g) perros sueltos h) incluidos en esta clasificación i) que se agitan como locos j) innumerables k)dibujados con un pincel finísimo de pelo de camello l) etcétera m) que acaban de romper el jarrón n) que de lejos parecen moscas."

There is something clearly wrong in this classification. It has several structural deficiencies. For instance, some animals do not belong to any of the classes

[1] Kinds have also been called 'types', 'sorts', 'secondary substances', 'species and genera', and so on.

[2] A rough translation would be: "These ambiguities, redundancies and deficiencies remind us to the ones that doctor Franz Kuhn attributes to some Chinese encyclopedia entitled *Emporio celestial de conocimientos benévolos.* In its remote pages it is written that animals are divided into a) owned by the Emperor b) embalmed c) trained d) piglets e) mermaids f) mythical g) loose dogs h) included in this classification i) which move crazily j) countless k) drawn with a very thin paint brush made of camel hair l) etcetera m) which have just broken the vase n) which look like flies when seen from far away."

mentioned (unless they go under "etcétera" (i)). Classes (j) and (l) have been given vague boundaries on purpose. It is not clear how literal the description of class (j) is. Class (e) appears to be a proper species of (f), but the latter one has no other proper species apart from (e). This is strange, are there no other fantastic animals apart from mermaids? Some classes, such as (m), appear to be empty, whereas others such as (h) and (n) are coextensional. In fact, some classes seem to have nothing to do with the fact that their members are animals, such as (b), (k) or (i). Some of these cases might be acceptable, nevertheless their combination suggests that this classification is purely arbitrary: there does not seem to be any rational criterion that has been used to divide animals into classes and the classification seems to serve no clear purposes. This example shows that not every classification of a given domain of objects is equally acceptable. Moreover, one could construct many other pseudo-classifications of the same domain such as this one. The question is, why is this one unacceptable?

It seems that, given what we know about the nature of the entities to be studied, some ways of classifying those objects are plausible and some of them are not. It is as if the world itself put some constrains regarding which ways of classifying those objects are adequate. Again, this is how the realist sees the problem. The pressing philosophical question is whether there is some criterion that allows us to distinguish between natural and non-natural kinds, or between more or less natural kinds. It should come as no surprise that philosophers strongly disagree about what natural kinds are, just as they disagree about the nature of causation, identity, modality or existence. These conceptions can be thought of as different proposals for an adequate description of the structure of natural kinds. Kinds could be essentialist [34], causal [12], promiscuous [32], conventional or conceptualistic [52], interactive (as Hacking has suggested for social kinds), qualitative, functional [37], historical [36], and so on.

There are some assumptions regarding what kinds look like that seem to be common to almost all theories. For instance, given some kinds, it is usually some other entities, let us call them 'objects', the ones that are classified into kinds. One says that this or that object is of this or that kind or sort of thing. Moreover, objects are clasified into kinds according to some criteria. Usually, these criteria are based on how similar the members of a given kind are to each other, on the features that these objects seem to share, on the relations in which they stand to each other or on the roles they fulfil. It is also acknowledged that some of these kinds are more specific than others. These relations between kinds lead to 'universal claims' about all the members of a given kind. For instance, one can say that the kind of hammers is a species of the kind of tools, the converse being false (given that other tools, such as screwdrivers, are not hammers). A fortiori, whenever we consider a specific hammer a, we can safely say that a is also a tool. There are other more theoretical roles that kinds are supposed to fulfil, as will be seen in the next Chapter.

However, once we move beyond these platitudes, most theories of kinds disagree. As suggested, the main division is that between *natural kind realists* and *natural kind anti-realists* [11]. This classical debate is linked to the more specific problem of scientific realism. It is not easy to give a clear formulation of these views. For example, both [10] and [134] distinguish the following position from other stronger commitments:

Natural Kind Realism There are natural kinds that reflect mind-independent natural divisions in reality.

The strongest of these realist views is *essentialism*. Most essentialists hold that objects belong to a kind in virtue of sharing some properties that are essential to them and that are therefore exemplified by all the members of the kind. A fortiori, each kind has a general essence consisting in those properties whose exemplification by an object is a necessary and sufficient condition for its membership to the kind. This view is quite strong because it has many consequences regarding the modal features of the members of a kind. For instance, it is thought to imply that an object could not have not belong to the kinds it actually belongs to. Some essentialists even argue that these essences make natural laws metaphysically necessary. In contrast, according to *causal theories*, kinds are bundles of properties that co-occur in virtue of some underlying causal mechanism (geological processes, natural selection processes, and so on). These theorists think that not all the members need to share exactly the same properties, and thus they would qualify the previous analysis as "most of the members of a given kind tend to have such and such properties in common". According to these theorists, there may not be a specific set of properties which are jointly sufficient and individually necessary to belong to the natural kind.

Why do realists think that the world is already divided into kinds? Plausibly, one of the reasons is a variant of the popular *No-Miracles Argument* for scientific realism. Suppose that you wake up in an exotic jungle. It is full of bizarre beings that you have never seen, but that look roughly like the plants and animals that you are familiar with. After some time wandering, you have started to conceptually categorize the 'living beings' in the jungle. The red insect-like plants make a soft noise before dropping what seem to be edible seeds. The fox-like creatures exhibit aggressive behaviour when you get close enough to them. By trial and error and by carefully observing your environment, you learn which plant-like things you can eat and which animal-like creatures are dangerous. Some days later, you finally get out of the jungle. The simplest explanation for your success through the jungle is that you have been roughly tracking kinds of things that you did not know about. You are probably wrong regarding many things. After all, you are developing a non-theoretical classification. You may have wrongly thought that the red insect-like plants and the orange insect-like plants were closely related to each other, that they are species of a common kind. Or you may have wrongly thought that all the exemplars of the fox-like creatures you found belong to the same kind. Maybe you were only tracking similar behaviours or some similarities in superficial observable features and only some of the exemplars you met happened to be members of a common biological kind. But now suppose that there was no objective fact regarding similarities between the entities in the jungle. There are no real differences or commonalities between them, not even in their behaviours. Whatever similarities you think you are finding are just imposed by your cognitive system over the inhabitants of the jungle. How likely is it that you would survive the trip? Given that the plants you ate or the animals you avoided had nothing in common, all the patterns that you followed to get out of the jungle had no objective basis whatsoever. How likely is it that all the seeds you ate were edible and had similar effects on your organism if the plants they originate from had nothing in common? Your successful trip would be no more than a very lucky coincidence. Realism is a plausible explanation to how we can manage to make roughly successful predictions regarding the behaviour of the entities that surround us (see [133] for a similar argument based on predictive power).

Nevertheless, some philosophers cast doubts on the very notions of naturalness and objectivity that realists appeal to (e.g. see [11] for some constructivist positions). The term 'natural' prompts many misunderstandings. If 'natural' were understood as opposed to psychological, social, cultural (and also abstract) or human made, then this would exclude by definition all the kinds of entities studied by social sciences. But as said, philosophers understand by 'natural' something closer to an *objective difference*. In this restricted sense, it may be argued that even *artefact kinds*, like kinds of hammers or kinds of buildings, are also natural kinds of some sort (e.g. archaeologists need to classify artifacts, but they will only accept some of these classifications). Classifying individual artefacts by their function may be more natural than classifying them by the number of scratches they have. Being natural does not imply being actual either[3]. I will not purse these matters at length, but it is useful to remind ourselves of this to prevent misunderstandings[4].

In any case, the notion of naturalness is usually explained in terms of that of 'objectivity'. But what does it even *mean* to say that "there are objective differences or divisions in the world"? Traditionally objectivity has been understood as *existential independence* (in the sense of [77]). The basic idea is that kinds would exist independently of the existence of minds or rational agents capable of thinking about them. There are possible worlds where our familiar kinds exist but where no minds exist. But this criterion faces well known difficulties [11]. For instance, it seems to exclude the possibility that there are natural kinds of psychological states or processes, kinds of psychiatric disorders or kinds of social processes. In other words, the kinds of objects studied by economics, psychology, sociology and other social sciences would be considered as non-natural. Although these kinds of entities are non-natural in the sense of belonging to the 'social world', there are reasons to consider the classifications made in the social sciences to be more or less objective (e.g. see [23] for the case of psychiatric disorders). Another notion (again in [77]), that of *ontological independence with respect to identity*, is also problematic. For example, it sounds plausible to say that identity conditions for kinds (or even tokens) of mental states and psychiatric disorders refer to the subjects in which they occur. In that case psychological entities would depend for their identity on minds. But surely differences between kinds of psychological processes like episodic memory, deductive reasonings or attention are objective. One could argue analogously for

[3]For instance, one can easily distinguish between mythological or fictitious objects, like Polyphemus or Elrond, and mythological or fictitious kinds, like the kinds *Cyclops* or *Elf*. Of course, there are no such things in our world. Mythologists will probably recognize only some of them as mythological kinds.

[4]For example, [119] has argued that species of minerals recognized by the *International Mineralogical Association* (IMA) are not traditional natural kinds. Some of the reasons he gives are based on an understanding of 'natural' as opposed to 'artificial' or 'human made', which is the notion of 'natural' that the IMA uses. A crystal formed by SiO_2 'found in nature' and one formed by technicians in Bell Labs would belong to the same kind *quartz*. That is allegedly wrong since only the former one is natural. But the same could be said concerning any instance of a new chemical element that was synthesized in a laboratory. This is not the notion of naturalness that will be assumed in this thesis. If the two particular entities have the same chemical composition and crystalline structure, then they plausibly belong to the same natural kind, regardless of whether the instances have been created in a laboratory.

the case of social entities or facts. For instance, E. Durkheim proposed suicide as an example of a social fact, the paradigmatic kind of event that sociologists should study. The kinds of structures of kinship and of exchange of goods studied by anthropologists may also be considered to be objective. It seems simply wrong to say that there is no objective difference between a nuclear family and an extended one, or between a religious ceremony and a penal institution. One could think that what is objective, in the sense of being independent of the existence of minds, are the *membership conditions* of the natural kind. Often the criterion according to which some objects belong to a kind instead of another one is the result of an explanation that makes use of theoretical principles. But once more, whether something is a mental token of a given kind surely depends either for its existence or for its identity on some minds.

Some philosophers also question the idea that nature is 'carved at its joints'. According to them, there are no joints at all[5]. It is us who put them in the first place. The world (if it makes sense to talk about such a thing) is a non-structured collection of things. The similarities and differences that we allegedly observe (or theorise about) among objects depend much more on our attention, interests, prejudices, conventions or cognitive capacities than on the structure of the world itself. The belief that this is not so is just the result of habits and a psychological tendency to expect that objects come already hierarchically organized into kinds. The classificatory structure of the world is at most a psychological projection[6]. A beautiful metaphor for conventionalism can be found in [138]:

> "(...) we are facing questions of realism at the metaphysical level, in the sense I have just explained: realism about the structure of the world, not about its content. And in this connection the opposition is not between Quine and Meinong—between the desert and the jungle. It is between Quine and Aristotle, between the desert and the garden, so to speak—and I mean a natural garden, like the Garden of Eden, with its tidily organized varieties of flora and fauna neatly governed by natural laws that reflect the essence of things and the way they can be, or the way they must be. To the extent that you believe that the world is like a garden in this sense—that it comes structured into entities of various kinds and at various levels and that it is the task of philosophy, if not of science generally, to "bring to light" that structure—to that extent you are a realist. But if you think that the Edenic tree of the knowledge of good and evil is a fiction and that a great deal of the structure we are used to attribute to the world out there lies, on closer inspection, in our head, in our organizing practices, in the complex system of concepts

[5]Other philosophers, like Dupré [32], think that there are *too many* joints in the world.

[6]The strongest anti-realist thesis would hold that the world is a *giant blob* in which there are no real *distinctions* among objects. That is to say, it is ontologically indeterminate whether "two" given objects are identical or not. This may be due to the fact that whether something is a particular or not is also indeterminate. In this extreme case, even the distinctions among objects would be the products of our cognitive system, linguistic conventions or cultural upbringing. This thesis seems to require some sort of indeterminacy of identity, but there are several strong arguments against it, e.g. [38]

> and categories that underlie our representation of experience and our need to represent it that way—then, to that extent you are not a realist." [138]

In that paper Varzi[7] makes use of the concept of boundary to argue by analogy that there are no real divisions in nature. From a plane, the coast of Great Britain can seem to be a line that clearly separates what is inside the island from what is outside of it. But as the plane approaches the coast this boundary gets fuzzier and fuzzier, until it disappears leaving behind a confusing mass of rocks, animals and plants. It is the limitations of our cognitive system what make us believe that there is a determinate boundary that clearly separates the two sides. Something analogous would happen in the case of natural kinds.

However, a world without natural kinds is odd, more than it may seem at first sight. Suppose that all the resemblances we observe between objects are the result of choices made by our cognitive system. Consider the following disturbing scenario:

> "When Smith wakes up in the morning he sees many familiar objects in his kitchen: his two mugs of coffee, his cactus, his cat Pipo, and so on. A couple of hours later, Smith is lucky to perceive the world as it really is. It is difficult to convey what he discovers. He notices something odd in his kitchen. Everything has stopped being familiar. The objects are so detailed that Smith wonders how he has possibly considered them to be *of the same kind*. He compares his two mugs of coffee. Obviously the mugs are completely different objects, as shown by the hue and saturation of their colour, the enormous difference in millimetres of their height, the differences in the shapes of the printed images, the way coffee is poured and splashes from the coffee pot in each of them, and so on. In fact, it is unclear why someone would group them under the same name. Smith looks fascinated at his kitchen: each of the tiles, and each of the parts of a tile, is radically different from the rest. "One would have to arbitrarily ignore important properties of objects in order for him to come up with the strange idea that they *have something in common*" he thinks. Smith calls his cat "Pipo", twice. But he doubts that he has actually used the same word. He, like Funes the Memorious, remembers having made two clearly different sounds. In his mind the two words have clearly different spellings. "Why are the first and the second sounds the same word, and how are they different from "house"?" he asks. Smith, as a good reader of Locke, realizes that he should have named each object with a different name *to avoid misunderstandings*."

In Smith's world, the world as it really is, objects do have properties. Although there are differences, given that every object is radically different from

[7] Varzi does think that the world is minimally structured, namely, mereologically.

the rest, objects have no common properties and the properties they have are too different from each other. To put it differently, there are only trivial kinds. There is a unique kind for each object. Before having this experience and due to his cognitive limitations, Smith could not discover what the world is like. For instance, it may be that his limited visual perception forced him to ignore or identify real differences. However, during the experience Smith has cognitive capacities which are vastly superior to the ones that any known measurement instrument could afford him. He just needs to carefully look at any region in the space to see with unlimited precision everything that happens in each of the parts of that region. A dual but more monotonous scenario could be devised where all the objects in the world are equally similar to each other. Imagine a world where *all* objects are perfect duplicates. Notice that it will not do to imagine a world full of people that are all clones. After all, each of their parts should also be a duplicate of the rest. The closest picture one can get at is an abstract one of atomic entities that are all copies of each other, or of a world of points that are spatially indistinguishable. Objects are numerically distinct from each other, but that is all. All objects are of the same kind, one cannot find any differences among them. In that world there is at most one kind of object. These two extreme scenarios suggest that if it is our cognitive systems what force us to group objects together and if the resemblances noted are not really grounded in the world, the resulting pictures are disturbingly odd.

Of course, not all sorts of anti-realism make the same claims made by Varzi. A different sort of conventionalism is found in Hacking [52], which is a direct sceptical attack to the notion of natural kind (based on a previous historical analysis of different conceptions of natural kinds in [51]). Hacking reviews most of the contemporary theories of kinds, and from this plural mess of positions and notions he extracts pessimistic conclusions regarding our contemporary state of "scholastic twilight". Hacking's main argument seems to be some sort of pessimistic induction: there are several different notions of natural kind (Whewell's, Mill's, Venn's, Russell's, Putnam's, and so on) and theories of natural kinds which are incompatible with each other (scientific essentialism, causal theories, and so on). Each of these relates the notion of a natural kind with some philosophical concepts instead of others (causation, induction, laws, and so on) and none of them seems to have given us any real understanding of how science works. Hacking's conclusion is that it is highly probable that there are no such things as natural kinds. At least there are relevant kinds or sorts of objects. At most, kinds are just, as Whewell put it, "classes denoted by a common name about which there is the possibility of general, intelligible and consistent, and probably true assertions", in other words, extensions of general terms. According to Hacking, the notion of natural kind should be exorcised:

> Take any discussion that helps advance our understanding of nature or any sicence. Delete every mention of natural kinds. I conjecture that as a result the work will be simplified, clarified, and be a greater contribution to understanding or knowledge. Try it. [52]

But Hacking's argument is a non-sequitur: from the plurality of conceptions and theories of natural kinds it does not follow that the concept is sterile. The same could be said of many other fundamental philosophical concepts, such as causation or knowledge. For the time being, there is no much that we can say to counter Hacking's challenge. Nevertheless, this thesis is written with the aim of showing that kinds are not simply "classes denoted by a common name about which there is the possibility of general, intelligible and consistent, and probably true assertions". I hope that something more informative can be said about them.

1.2 Aim, Methodology and Results

1.2.1 Aim of the Thesis

There are already many theories of natural kinds. There are also many discussions concerning whether a given sort or kind of object is natural or whether the sorts or kinds described by classifications in science (such as the biological species) are indeed natural kinds. In contrast, the aim of this thesis is not to defend a specific theory of kinds nor to debate whether a specific sort or kind is natural. The aim is rather to assume a Minimal Conception of Kinds in order to explore two structural features of kinds by making use of mathematical models, in order to further our understanding of several specific principles about kinds[8]. These are:

i The *external structure* of natural kinds, that is to say, the different ways in which kinds are ordered by species-genus specificity relations.

ii The *internal structure* of natural kinds, that is to say, the fact that the members of a given kind have several common properties or are sufficiently similar to each other.

This Minimal Conception of Kinds will be introduced in Chapter III. It can be formulated as follows:

Minimal Conception of Kinds Every kind has as members some objects (the extension) sharing certain sparse attributes (the intension). More strongly, all the objects share all these attributes, and these attributes are all those sparse attributes shared by these objects.

Any model involves some ontological assumptions regarding the primitive or fundamental entities to be described. The formal model used to explore the Minimal Conception in Chapter III is based on the theory of concept lattices and appeals both to objects and attributes in order to explain what kinds are. Since these attributes behave like universal entities, the question arises whether one could give a nominalist model of kinds that satisfies the Minimal Conception and does not posit these attributes as primitive entities. In order to explore this issue, I choose an ontological position known as "resemblance nominalism". The nominalist only assumes that objects are in categorical resemblance relations to each other. Properties are constructed as derived entities, they are usually considered to be classes of similar objects. Despite the fact that I favour

[8]Since I will be only concerned with the structure of kinds I will make use of both scientific and vernacular examples as needed.

resemblance nominalism, I want to highlight that *the aim of this thesis is not to defend resemblance nominalism* as the best account of the nature of properties or something like that. The aim is philosophically more modest, since for the most part (except when it comes to the *formal* objections to it) I will assume that such an account is at least defensible. Instead, the more specific aim of this thesis is:

Aim of the Thesis To develop a formal model for kinds that satisfies the Minimal Conception and is based on the ontological assumptions of resemblance nominalism.

Resemblance nominalists do not posit universal attributes and they offer as surrogates classes of similar objects. The nominalistic kinds are defined in terms of these properties. A fortiori, the nominalist reconstruction is done in two steps:

1. First, the natural attributes are reconstructed as collections of similar objects.

2. Second, the natural kinds are reconstructed as pairs consisting of a set of objects and a set of attributes (where these attributes are the ones obtained in the first step).

The first step will lead us from a model based on similarity relations to a structure representing objects and properties in exemplification relations (what is called a 'formal context'). The second step leads us from this structure to the lattice of nominalistic kinds. More specifically, the whole process can be divided into three stages, as follows:

i I select a class of mathematical structures that provide a model for the Minimal Conception of Kinds, with the aim of giving an explanation of the external structure of kinds. This model is very rich: it provides a semantics for syllogistic logic, it implies the Kantian Law of the duality between extension and intension, it has as special cases the hierarchical tree-like models of kinds and it allows for the definitions of kinds in terms of specific differences. We can call such a model a *realist model*. This is done in Chapter III.

ii I select a class of mathematical structures that provide a model for a specific conception of resemblance nominalism, namely aristocratic resemblance nominalism, with the aim of explaining the internal structure of kinds. This model respects the basic formal properties of categorical similarity and will be shown to have several equivalent formulations. Moreover, this model sheds some light on the controversial notion of a paradigmatic object. We can call such a model a *nominalist model*. This is done in Chapter IV.

iii I show which subclasses of realist models can be reconstructed from the nominalist ones. The properties of the corresponding lattice of kinds are studied in detail. In other words, the models used by the realist are 'translated' to models used by the nominalist. This is done in Chapter V.

At this stage, one could object that such a 'reconstruction' is an idle formal exercise. At best, it would be a convoluted and esoteric way of rephrasing something that we already know. At worst, it would give an obscure reformulation of the problem that misses its crucial relevant features, or a replacement of the original problem by something completely different (as Strawson argued time ago). I think that such a view is misleading. By making formally explicit those assumptions that connect some concepts to others we are lead to general principles that, despite being evident or trivial once formulated, could have gone unnoticed. As an example, in order to give the nominalistic reconstruction of kinds a Minimal Conception of Kinds and the corresponding realist model for kinds were chosen in Chapter III. A closer look at this model lead me to focus on a principle about the external structure of kinds that is no longer discussed in the literature on kinds, namely Kant's Law of the duality between extension and intension. Furthermore, in the chapter it is also shown that definitions of kinds in terms of specific differences can be given in the model. However, although Kant's Law and the form of Aristotelian definitions have been known for centuries, one could have hardly guessed just by looking at their informal formulations that they followed directly from the algebraic structure induced by the order relations among kinds (as defined by the Minimal Conception).

Moreover, the 'translation' given is made systematically rather than just giving simple paraphrasing templates (as are usually found in the metaphysics literature, at least on the topics related to this thesis). A fortiori, this results in a detailed explicit analysis of the conditions needed for such a reconstruction to succeed. This is not to say, of course, that the proposal in this thesis is free of problems. In fact, the results are limited, as it will be argued later on. But again, these limitations could have hardly been found just by looking at the informal formulations of the problem. Nevertheless, and independently of whether these results turn out to be of any philosophical relevance, the point I want to make is that such an attempt of reconstruction can lead to more insight into the topic just by forcing us to make explicit all the assumptions involved. In any case, since a formal approach to kinds is uncommon, it is convenient to consider in more detail the methodology that is used in this thesis.

1.2.2 Methodological Remarks

Regarding the methodology to be used, as was said, in Chapters III-V I will approach the problem of natural kinds using formal models. This is not to say that all there is to this thesis are the mathematical models. Philosophy proceeds by arguments, and there are many kinds of non-deductive inferential relations. One argues by analogy and by considering the explanatoriness of whole philosophical theories, introduces informal thought experiments and counterexamples, compares the obtained results to those found in other academic disciplines, and so on. Moreover, the plausibility of the assumptions introduced in the models has to be argued somehow and one cannot be expected to do so deductively.

That being said, this is a PhD thesis on mathematical philosophy. I am firmly convinced of the usefulness of mathematical models to make us better understand philosophically interesting phenomena and I am sure that the problem of natural kinds will not be an exception in this regard. Given that a formal approach to natural kinds is uncommon, it will be useful to consider some general features of this methodology. Formal models have already been fruitfully applied to philosophically interesting phenomena such as truth, modality, vagueness, confirmation and induction, and so on. Although formal methods are more popular in philosophy of science (see for instance [71] for an overview), semantics and epistemology, there are now standard formal treatments of metaphysically fundamental topics like modality, mereology or causation. Application of these formal methods has allowed us to shed light on specific principles and inferential relations between these principles that may have passed unnoticed to the unaided thinking. For instance, think about de re/de dicto and scope distinctions, the many applications of possible world semantics or the Bayesian approaches to the paradoxes of confirmation (and other epistemological problems). By analogy, one can expect that formal models will be useful for the problem of natural kinds too. As for the benefits of model-building, [145] has a more elaborate discussion on the role of models in philosophy. As the author contends, models can be used to keep the philosophical theory consistent, make qualitative predictions (e.g. some logical principles may not hold in the domain at issue), build counterexamples, make crucial distinctions, allow for computational simplicity, and so on.

Not everyone is optimistic regarding the development of an informative and strong formal framework for natural properties and natural kinds though. Take as an example the following quote by Williamson himself [144]:

> ”Metaphysically universal generalizations of logic are the structural core of metaphysics. We need the best logic we can get. Logic restricted to natural properties and relations is pathetically weak. Imagine such a restricted version of Leibniz's Law for identity: it says that identicals have all their natural properties in common. Someone claims to be Nicolas Sarkozy. We point out that Sarkozy speaks French and he does not. He agrees, but objects that since we have not shown speaking French to be a natural property, we have not refuted the identity. Metaphysics based on weak logic wastes its time taking crank theories seriously. It needs a strong logic with laws of unrestricted generality. For example, a better version of Leibniz's Law says that identicals have all their properties (however unnatural) in common. If all natural properties and relations satisfied strong structural generalizations, expressible as laws of higher-order modal logic not satisfied by all unnatural properties and relations, that might be some reason to formulate higher-modal logic with correspondingly restricted quantifiers. But the extensive literature on natural properties and relations has produced no such strong characteristic logic of natural universals. Rather, some of the most informative principles of higher-order modal logic depend on the absence of any naturalness restriction. The most obvious example of a logical principle of higher-order logic that depends on unnatural properties and relations is the comprehension schema (...)” [144]

I think that Williamson's pessimistic attitude regarding the possibility of a successful formal framework for natural kinds is premature. Hopefully the proposals made in this thesis will have shed some light on this vexed topic.

Regarding the methodology of modelling, some general comments are in order. First, every model has some limits to its representational power. As philosophers of science have repeatedly pointed out, models usually introduce representational artefacts. They include features that do not correspond to anything in the world. From the point of view of metaphysics this may be particularly dangerous, after all, we want to have a clear view on which components in the representation correspond to something in the world and how, and which ones do not. But sacrificing a bit of adequacy may be worthy enough if it fosters simplicity and fruitfulness. Besides, the lost adequacy may be recovered in the future when we have better models at our disposal. Sometimes models are not fine-grained enough and do not allow to make some distinctions we can pretheoretically identify. Moreover, models systematically ignore features of the domain that are selected by the researcher as irrelevant for the purposes of inquiry. Although we should always try to look for models which are closer to the features of the phenomenon we are studying, we have to keep in mind that a complete match will never be obtained. Nevertheless, it is this same feature of models what allows them to be fruitful. We only need to remind ourselves of Borges' tale "On Exactitude in Science" about the perfect map of the world. If one wants a perfect model of the world, he should take the world itself as a model. But the world is clearly useless if it is to be used as a tool to understand the world. That is the whole point of using models. The choice of a good model involves balancing all these aspects, especially the tension between adequacy, simplicity and fruitfulness. So it is, as Williamson himself puts it, more an art than a science.

Second, whenever a formal model for some philosophically interesting phenomenon is introduced (part-whole, possibility, causation ...), philosophical disagreement about some of the basic principles embodied by the model tends to lead to the development of weaker or alternative models of that same phenomenon. A plurality of alternative models quickly emerges. New questions arise regarding the differences and similarities between the conceptions embodied in these models, about their fruitfulness for tackling the original problem or about the possibility of finding translations or correspondences between them. Although this plurality can lead to some skepticism regarding how unified the original phenomenon was, it can show us how naive and coarse our original conceptions were and can also span richer discussions concerning the specific conditions and principles that have to hold for some of the interesting philosophical theses about the phenomenon to be true. This plurality of models allows us to carve more carefully the different species of the original phenomena or to consider what happens when additional parameters are introduced. Such a plurality is usually a sign of theoretical maturity and is to be welcomed.

Lastly, when I refer to 'formal' models I do not intend to limit myself to different logical systems or to theories falling under the subject of 'mathematical logic'. For instance, formal epistemology uses probability theory and related mathematical theories like decision theory or game theory besides epistemic logic. Once we have accepted the use of other mathematical theories than logic to model philosophically interesting phenomena (like degrees of confirmation or utilities), it does not seem reasonable to restrict our formal repertoire in advance. Of course, one has to argue for the appropriateness of the mathematical tool that is being proposed by showing that it is philosophically fruitful regarding the problem at hand. In particular, this thesis will use concepts from order theory and the theory of similarity structures to develop models for natural kinds. The reader will judge whether these tools were adequate for these purposes.

1.2.3 Main Results

The main results of this thesis have already been mentioned. Broadly speaking, they are:

i An appropriate model for a Minimal Conception of Kinds based on the theory of concept lattices is selected and is used to analyse the external structure of kinds. (Chapter III)

ii An appropriate model for Aristocratic Resemblance Nominalism based on the theory of similarity structures is given and is used to analyse the internal structure of kinds. (Chapter IV)

iii Several specific results concerning the nominalist reconstruction of a certain class of realist models are given. (Chapter V)

Let us break down each of these results into further pieces. Regarding the first result in Chapter III, it contains the following:

1. Kant's Law of the duality between extension and intension, according to which the extension of a kind is inversely related to its intension, is shown to hold in the concept-lattice model of kinds. In contrast, the hierarchy condition is shown to be a special case of the model. The relations between this model and others present in the literature, such as Thomason's algebraic model of kinds and Corcoran and Martin's syllogistic logic are highlighted.

2. Two new operations of specific difference are introduced into the model and shown to be defined in any such concept lattice. The properties of these operations are studied. This allows for definitions of kinds in terms of genera and specific difference. Moreover, each specific difference is shown to induce a non-classical internal negation of kinds, whose properties are also studied. A comparison between the picture given by the model and the classical Aristotelian one is made.

3. A brief sketch of a modal expansion of the model is given, in order to allow for changes in the extension of a given kind across different possible worlds.

These results give more insight into the species-genus specificity relations according to which kinds are ordered and provide a realist model of kinds to be reconstructed by nominalist means in the following chapters.

About the second one in Chapter IV, it is further subdivided as follows:

1. The basic formal properties of similarity structures are introduced in different but mathematically equivalent ways, including two axiom systems for similarity in terms of already known similarity operators. Fundamental results (by Brockhaus, Mormann and so on) concerning similarities are reviewed (including Brockhaus Theorem).

2. The basic properties of categorical similarity are defended against some objections. More specifically, Tversky's criticism of symmetry is considered at length.

3. A new subclass of similarity structures, namely the pure similarities of order 1, is introduced and suggested as a plausible model for aristocratic resemblance nominalism. Several properties of this class of structures are proven. More specifically, it is shown to be a subclass of the class of similarity structures to which Brockhaus Theorem of quasianalysis applies. Two alternative models, the topological polar model by Mormann and Rumffitt, and a new order-theoretic model, are introduced.

Concerning the third one in Chapter V, it can be further subdivided as follows:

1. It is proven that the three resemblance nominalist models introduced in Chapter IV, namely the polar model, the order-theoretic model and the similarity model, are mathematically equivalent. The corresponding set-theoretic realist structures to which they are equivalent (the polar contexts) are also introduced. Under the assumption that each of them is independently plausible as a model of the nominalist ontology, the fact that they are equivalent suggests that the choice was appropriate. This result clearly delimits the class of realist structures that can be reconstructed (in a unique way), and therefore the class of realist models that the nominalist can reconstruct.

2. It is proven that the concept lattice that can be obtained starting from a pure similarity of order 1 is (co)atomistic. This means that every natural kind can be obtained by combining its infimae species (or its maximal general). Moreover, it is also proven that every such complete lattice can be obtained uniquely in that way starting from a similarity, for it is shown that each such similarity (that satisfies a stronger indiscernibility axiom) uniquely corresponds to a specific kind of '(co)atomistic' order. This gives us a description of the class of specificity orders between kinds that the nominalist can mimick and allows for finding more specific kinds of orders that can be reconstructed just by adding the corresponding axioms to the similarity structures. Furthermore, the quasianalysis is shown to give several equivalent representations of objects: the same object can be represented as the set of its paradigms or as the set of its attributes (which are classes of similar objects). A fortiori, the nominalistic kinds also have different equivalent representations.

3. Several concepts and results concerning similarities of order 1 are generalized to similarity structures (called 'simple' similarities) where there is an arbitrary number of paradigms for each property and each paradigm can have an arbitrary number of properties. Although this results (in a sense) in a loss of uniqueness regarding some corresponding realist contexts, it leads to a more plausible model (for aristocratic resemblance nominalism) than the one given by pure similarities of order 1.

The combination of these three results gives a picture of what the natural kinds look like for the nominalist. The first result shows that different intuitions on how to model the nominalist assumptions about properties (understood as classes of similar objects) lead to mathematically equivalent models. The nominalist reconstructs each realist property as a maximal set of similar objects that is determined by a special object, called a 'paradigm'. The corresponding realist structures are described by making use of the correspondence with the polar model.

The second result gives more details regarding the external structure of nominalistic kinds. It turns out that positing paradigms is some sort of qualitative atomism. This means that each kind can be obtained as the overlapping of its maximal genera (or as the sum of its infimae species). This includes as particular cases, atomic complete boolean algebras and atomic trees, but also other kinds of structures. The quasianalysis provides several equivalent ways to present a given object as a bundle of paradigms or as a bundle of properties. This representation extends to the kinds themselves.

The third result allows for weakening the original assumptions about the nature of paradigmatic objects. Paradigms as posited in Chapter IV, are strange creatures. They are objects that are in one-one correspondences with properties. This means that each such paradigm has a unique property, and each property has a unique paradigmatic instance. It is instructive to think of paradigms as duals to haecceities. Whereas an haecceity is a property had by a unique object, a paradigm is an object that has a unique property. In any case, these qualitatively point-like or atomic entities look suspicious. Ordinary objects have usually several properties, and if properties have paradigmatic instances then it is highly likely that they have several of them. The third result shows that there are more general models in which each paradigmatic object has several properties and each property has several paradigmatic instances. This allows for the formulation of a more plausible aristocratic nominalism.

These results allow for a limited answer to the questions that motivated this thesis in the first place, namely a search for a better understanding of both the external and internal structure of kinds: under nominalist ontological assumptions regarding the nature of objects and resemblance relations, the internal structure of kinds is determined by the similarities among these objects and their paradigms, whereas the external structure of kinds is that of a complete (co)atomistic lattice that satisfies the requirements of the Minimal Conception of Kinds. The limits of the approach taken in this thesis and the problems that remain open will be briefly discussed later on.

1.3 Summary of the Chapters

The structure of the thesis is the following one:

- i Chapter I: Introduction.
- ii Chapter II: Theories of Natural Kinds.
- iii Chapter III: The Hierarchical Structure of Natural Kinds.
- iv Chapter IV: The Resemblance Structure of Natural Attributes.
- v Chapter V: The Resemblance Structure of Natural Kinds.

vi Chapter VI: Conclusion.

vii References.

The aim of Chapter II is to introduce the problem of natural kinds. First, I present a state of the art to show why the topic of kinds is philosophically relevant. Kinds have been posited in order to deal with several philosophical problems, such as the epistemological *problem of induction*, the metaphysical *problem of universals*, the semantic *problem of the reference of natural kind terms* and the methodological problems related to the *role of classifications*. Second, I expand on this state of the art by reviewing the main theories of kinds, namely *Scientific Essentialism* and *Cluster Theories*. I argue that, despite the fact that these approaches make substantive assumptions about the nature of kinds, they leave several questions regarding their structure unanswered. For instance, they are silent on the specific principles that the specificity relations between kinds satisfy, between these and the fact that the members of a kind are supposed to share some natural properties, and so on. I also consider *Conceptualism* briefly. Despite the fact that conceptualist approaches to kinds (which take kinds to be concepts) are not very popular today, theories like that of conceptual spaces do suggest specific principles about the structure of kinds and formal models for them that provide a starting point for such a discussion. In contrast, I suggest to follow an alternative path of studying the external and internal structure of natural kinds by exploring a minimal conception through a formal model. The external structure of kinds concerns the ways kinds are ordered by species-genus specificity relations, whereas the internal structure of kinds concerns the ways the members of a kind are related to each other by similarity relations (or by sharing natural attributes). This *Minimal Conception of Natural Kinds* is introduced in the next chapter.

The aim of Chapter III is to discuss the possible *external structure of natural kinds* by making use of a formal model. Traditionally, kinds have been thought to be ordered according to species-genus relations. The chapter starts by analysing a condition (which goes back to Porphyry) found in the literature that requires kinds to be *hierarchically ordered* in a tree-like fashion. This condition will turn out to be too strong, but it will pave the way for a more nuanced analysis of the possible order relations that may hold between kinds. In this chapter the first serious use of modelling techniques is made. First, the lattice-theoretic model of kinds developed by Thomason is considered. As an argument for its plausibility, it is shown that it provides a semantics for the syllogistic logic of Corcoran and Martin. Under the assumption that kinds are at least related to each other as described by the classical Square of Opposition, this provides a reason for taking the lattice-theoretic model to be a plausible model for kinds. Second, a *Minimal Conception of Kinds* based on some features traditionally assigned to kinds is described and assumed throughout the whole chapter. According to this conception, kinds are bi-dimensional entities constituted by an *extension* (the objects that are members of the kind) and an *intension* (a set of natural properties exemplified by the members of the extension). Furthermore, a detailed formal model for it (by making use of *formal concept lattices*) is introduced. This model is shown to be a special case of Thomason's and therefore provides a semantics for syllogistic logic. Moreover, in contrast with the other two, this one does give some information regarding the ways kinds are related to each other. In particular, the traditional *Kantian Law* of the duality between extension and intension is shown to hold in the model. Thirdly, several novel aspects of the model are discussed, including *Aristotelian definitions* in terms of specific differences and *internal negations*. Finally, a comparison between the model and the Aristotelian conception of kinds is drawn and a sketch of a modal approach is also given. Roughly put, this results in a model of an 'Aristotelian world', whose fundamental ontology consists of particular objects related to universal attributes by exemplification relations.

Chapters IV and V work in tandem. Their aim is to explore the *internal structure of natural kinds*. According to the model introduced in the previous chapter, what makes several objects members of a kind is the fact that they share some natural attributes, which are plausibly understood to be universals. Therefore, the question arises whether a nominalist could give an alternative account of kinds that still satisfies the basic constraints of the Minimal Conception. First, I start by discussing some features of resemblance relations. I briefly introduce the theory of *similarity structures* and I reply to some standard objections to the notion of *categorical similarity*. The most important objections come from the empirical work by A. Tversky in psychology, that seems to show that similarity is neither a reflexive nor a symmetric relation. Then the basic forms of resemblance nominalism are introduced, namely egalitarian, aristocratic and collectivist resemblance nominalisms. After a discussion of Pereyra's egalitarian approach, *aristocratic resemblance nominalism* is chosen. According to this conception, properties are represented as maximal classes of similar objects. However, not all the objects in the class fulfill the same roles, some of them are more typical instances of the property and therefore play a more fundamental role in structuring the class. These special objects are the *paradigms*. This version of nominalism is the metaphysical analog of the prototype view of concepts. However, since this version of nominalism is usually considered to be underdeveloped and technically defective (some objections by Pereyra are reviewed), a discussion on the nature of these paradigmatic objects is given. Finally, three different models for aristocratic resemblance nominalism are introduced. The first one is a topological model called the *polar model* taken from contemporary research by Rumfitt and Mormann on vagueness and conceptual spaces. According to this model, objects are represented as points in a space and each object is thought of as being arbitrarily close to its paradigms. The second one is based on Mormann's suggestion that the specialization order of weakly-scattered spaces is a good model for the prototypicality order among objects. I extended this model by (putting it upside-down and) adding an axiom to it and I made a proposal regarding what the attributes would be (fixed ultrafilters). The third one is based on a special class of similarity structures, the *pure similarities of order 1*, which is introduced in this thesis for the first time (although it is based on previous work on quasianalysis by Brockhaus and Mormann). According to this model, paradigms are objects that have a special role as building blocks of the resemblance relations that hold among all objects. I briefly argue for the material adequacy of these models with respect to aristocratic resemblance nominalism. Although the models make use of apparently different structures, in the next chapter it will be shown that they are actually equivalent.

In Chapter V I reformulate the nominalistic challenge to universals as the quest for constructing a *surrogate for a realist structure of universals from a purely nominalistic ontology*, by devising a structural mapping from the former one into the latter one. In other words, the aim of the nominalist is to find a 'translation' from the model of the realist (that appeals to universal entities) to a structure constructed from its own model (that only appeals to resemblances between objects). The idea is that if such a mapping can be constructed, then the model of the nominalist will have at least as much explanatory power as that of the realist. From the point of view of the resemblance nominalist, this requires us to deal with *Goodman's objections* to Carnap's *quasianalysis* by putting to work the model introduced in Chapter IV. For these purposes, a reductive result is proven for a limited class of realist structures that shows how the nominalist can reconstruct the structures the realist starts from just by assuming the similarity model (including surrogates for the latter's entities, formal relations and basic ontological principles). This result shows that the previous three models for aristocratic nominalism are all of them equivalent to each other. Once this surrogate structure is in place, it is shown how the corresponding *nominalistic lattice of natural kinds* is built (by following Chapter III) and which properties this lattice has. In other words, the results in this section give a description of the class of external structures of kinds that the aristocratic nominalist can recover. More specifically, it is shown that the class of lattices that can be so reconstructed is that of *(co)atomistic lattices*. Moreover, several equivalent representations of nominalistic kinds can be given by making use of quasianalysis. This lattice of kinds will be the surrogate for the small Aristotelian world introduced in Chapter III. In this way we get a nominalistic model that satisfies the requirements of the Minimal Conception of natural kinds. Moreover, in the meantime several possible external structures for kinds will be explored, which will be used to show that some features of the traditional picture of kinds still have some serious work to do. The similarity model given in the previous chapter makes some strong assumptions about the paradigmatic entities. In particular, it requires that paradigms are entities having just one property. This is implausible. Accordingly, a generalization of some of the previous notions and results is given for a broader class of similarity structures (the *simple similarities*). This class of structures gives a generalized aristocratic model where each paradigmatic entity can have several properties and each property can be exemplified by several paradigms.

Chapter 2

Theories of Natural Kinds

The species in which the things primarily called substances are, are called *secondary substances*, as also are the genera of these species.

Categories, 2a13-2a18
ARISTOTLE

The aim of Chapter II is that of giving a state of the art for the problem of natural kinds by reviewing the main competing theories, namely *Scientific Essentialism* and *Cluster Theories* (in particular, causal theories). I will also consider the *Conceptualist* approach, for it will be relevant in the following chapters. The essentialist programme is presented and some reasons for rejecting it are advanced. The cluster theories of kinds are briefly considered and some objections reviewed. The upshot of this chapter is that, despite the fact that they make very strong metaphysical assumptions about kinds, the discussions generated by these theories are still too general. They have not lead to consider more specific principles about kinds. For instance, these theories are silent on the ways kinds are related to each other by specificity relations, on how these order relations are related to the fact that the members of a kind share some attributes, or on how having some common attributes makes the members of a kind similar to each other.

I think that taking into account such structural principles would improve our understanding of kinds. For this purpose, I will propose to study the structural features of kinds through a *Minimal Conception of Kinds* which makes very few assumptions about the external and internal structure of kinds. The main features of this conception will be explored in the remaining chapters of the thesis.

2.1 Arguments for Natural Kind Realism

This section considers the problem of natural kinds from four different sides: the epistemological, metaphysical or semantic one. It serves two purposes. On the one hand, the problems that helped to introduce the notion of natural kind in the contemporary discussion allow us to formulate some arguments in favour of natural kind realism, or at least in favour of adopting the notion of a natural kind (whatever these kinds turn out to be). On the other hand, they can be used to highlight the philosophical relevance of the problem of natural kinds by revealing its connections to other philosophically fundamental problems such as the nature of induction, causation, natural laws, universals, essences, and so on[1].

[1]Of course, the roots of the problem of natural kinds can be traced back to Aristotle's treatment of species and genera or even to Plato's Theory of Forms. It reappears under different guises through the history of philosophy say, in the medieval discussions about universals or in the empiricist epistemological theories of general ideas.

2.1.1 Epistemological Argument: The New Riddle of Induction

Since the devastating criticism by Hume, the problem of induction was considered to be concerned with a non-circular and non-mysterious justification of our inductive inferential practices. Goodman [47] replaced this problem by that of explaining what makes possible for some predicates (the projectable predicates) to be useful in making successful inductive inferences. Goodman's argument is known as the *grue paradox*. Let us suppose that we are examining emeralds and that we want to induce a proposition concerning the colour of all those emeralds based on the observations made before a given instant of time t. Let us suppose that all the emeralds examined before t are green. Therefore, we can induce that "all emeralds are green". Now let us define a new predicate 'grue' as follows: "x is grue iff (x has been examined before t and x is green) or (x has not been examined before t and x is blue)". The problem is that each of the observations made before the instant t of the form "a is a green emerald" goes along with an observation of the form "a is a grue emerald". As a consequence, the hypothesis $h =$ "all emeralds are green" has in time t the same degree of confirmation as the hypothesis $h' =$ "all emeralds are grue". But this is clearly problematic. Let t' be a time instant after t and let b be an emerald we find in t'. From h we deduce that b will be green. From h' we deduce that b will be blue, for b has not been examined before t. These two predictions contradict each other (assuming every emerald is coloured and no emerald can be bicoloured), but both are equally justified, because they have the same degree of confirmation [22]. Furthermore, if we observe a green emerald in t' this will confirm the hypothesis h and refute the hypothesis h', whereas observing a blue emerald in t' will confirm h' and refute h. If we distinguish kinds from properties, we can still easily devise predicates that designated, respectively, natural and nonatural kinds instead of properties (e.g. 'tulipose' as "x is a tulipose iff (x has been examined before t and x is a tulip) or (x has not been examined before t and x is a rose)". The point is the same.

Intuitively, the hypothesis we thought we were confirming was h and not h', the prediction we wanted to make was that the next emerald would be green and not blue and the reason why the predicate 'grue' is unacceptable is that there are no grue objects in the world. What seems to be happening is that 'grue' is not a natural predicate, it does not correspond to an objective kind of thing but is a term whose extension has been artificially selected with the sole purpose of making up the philosophical sceptical challenge. But Hempel's syntactic confirmation theory puts no reasonable constriction to which predicates we should accept or reject. Besides, mentioning a specific time instant t or the fact that the object is being observed by someone are irrelevant to make the point. The thing is that there are no grue objects in the world. The question this paradox leads us to is this: what distinguishes the predicate 'grue' from the predicate 'green'? Why does the latter one allow us to make successful inductive inferences, in contrast to the former one? What do we want to say when we say that whereas the former one is an artificial predicate the latter one does correspond to a real or natural property or kind of thing in the world? Moreover, if we suppose that the propositions express natural laws of the form "all P-s are Q-s" and that these are confirmed by their instances of the form $Pa \wedge Qa$, the problem of distinguishing projectable and non-projectable predicates has consequences for the problem of the confirmation of laws. Hempel's classical raven paradox is tightly linked to Goodman's paradox, both having examples of predicates (e.g. negations of predicates, such as 'non-raven') which do not allow for the propositions expressed by the sentences in which they occur to be confirmed.

The problem seems to be related to the absence of cohesion among the members of the extension of 'grue', which is reflected in the suspicious disjunctive definition of the term. That is why Quine [107] proposed that projectable predicates would be those whose extension is a natural kind or class, in such a way that the extension of the term would have to include sufficiently similar objects, thus excluding the adoption of 'grue'. He appealed to some ideas from Carnap [17] according to which a natural kind is a set of individuals which are all of them more similar to each other than to any non-member of the set. However, he rejected this solution as defective. He concluded that the distinct scientific disciplines would substitute the pre-theoretical notion of similarity by other theoretical notions, and as a consequence the natural types or kinds would be revised too. Thus, the vernacular or 'manifest' natural kinds would be more and more different from the corresponding scientific natural kinds.

However, alongside his pessimistic conclusions, Quine also suggested that we have innate standards of similarity, which would explain why we tend to classify objects according to their similarities. The corresponding 'quality spaces' would be studied by psychologists. Although Quine's remarks were very sketchy, cognitive psychologists have indeed developed models of conceptual categorization based on similarity relations among the items. These models represent this categorization function of the mind mathematically as a high-dimensional space in which the items are represented as points, their dissimilarities get represented as distances, the dimensions of the space correspond to respects of comparison and the concepts are represented as certain regions of the space having specific geometric features. As a specific example, philosophers like Peter Gärdenfors [42], [43] have followed this conceptualist route and developed a whole framework according to which natural properties are in fact regions in a *conceptual space*. According to Gärdenfors, natural properties would correspond to *convex regions* in a given domain. Despite the fact that there are also psychological models of categorization that take similarity relations to be derived, such as the attribute model by Tversky [137], the conceptual spaces approach is developing fast as an alternative research programme (see [27] for a comparison between the two approaches).

The epistemological questions related to kinds are: what makes natural kinds (or their predicates) projectable? What is the relation between natural kinds and natural laws? Our first argument is therefore epistemological. We can reconstruct it as an argument for the best explanation:

Epistemological Argument from Induction and Confirmation

1. We can make successful inductive inferences about the properties and behaviour of the members of the extension of some general terms. Some of the observations we make about these objects confirm or refute generalizations about their general terms. [Pr]

2. We cannot perform successful inductive inferences about the properties and behaviour of the members of the extension of artificially made gen-

eral terms (e.g. 'grue', 'non-raven'), nor confirm or refute successfully generalizations that involve them. [Grue and Raven Paradoxes]

3. The best explanation for (2) is that some of our general terms correspond to natural kinds. That is to say, they correspond to kinds of objects which have similar properties and behaviour independently of our conventional choices. [C]

What Goodman's and Hempel's paradoxes seem to show is that only some of our general terms are fruitful for the formulation of laws and the prediction of phenomena, but syntactic considerations are not enough to determine which terms these are. The most plausible explanation of this fact is that the world constraints significantly the extension of these terms. The world seems to be minimally structured into kinds of things.

2.1.2 Metaphysical Argument: Universals and Sparse Properties

The *Problem of Universals* is plausibly the most ancient metaphysical puzzle. Nevertheless, an adequate description of the problem itself is still in dispute. Some have introduced it as the quest for an answer to the "One over Many Problem" [2], that is, to the question "how is it possible for several particulars to have the same property?". We seem to have some pre-theoretical intuitions regarding facts that involve objects "having the same nature" or "sharing the same property". We are used to say that two objects are of the same colour or have the same size. Furthermore, we think that some of these claims are objectively true, but what in the world makes them true? The main task is to answer the following question: how is it possible for several objects to have the same property?

In [2] Armstrong criticized nominalist and Platonist answers to the problem of universals in order to develop his own Aristotelian theory. According to him, universals are repeatable entities that are wholly (not partially) present in the particular objects that instantiate them. As a consequence, since several distinct particulars may instantiate the same universal, universals may have different spatial locations in the same time instant. So, in contrast with particulars, universals can be (and usually are) multiply located. In contrast with nominalistic theories, universals are sui generis entities and neither collections of primitively similar particulars, nor mereological sums, nor predicates or concepts. In contrast with Platonist theories, Armstrong's universals are necessarily instantiated (*Instantiation Principle*). Besides, necessarily, every particular instantiates at least one universal (*No Bare Particulars*). Armstrong's theory was intended to be compatible with scientific naturalism, that is why it is taken to be an *a posteriori scientific realism of universals*. That is to say, universals are discovered empirically by scientific research and not by an a priori analysis of the meaning of predicates or general terms. The similarity among particulars would consist simply in their instantiating a common universal (well, in their instantiating similar universals). Armstrong made use of universals to give philosophical explanations of other phenomena, such as natural laws or possible worlds.

The most general division that is drawn here is:

Particularism/Nominalism (necessarily) there are no universals.

Universalism/Realism (necessarily) there are universals.

Traditionally, particularists have been called *nominalists* and universalists *realists*. We will make use of these terms indistinctly.

At first sight it may seem that natural kinds would be universals for Armstrong. This is not the case. Armstrong's universals are the determinate magnitudes and relations of fundamental physics (whatever these turn out to be). In [3] he expressed some doubts regarding kinds as substantial universals, he thought that they at most supervene over particulars and universal attributes:

> But it is not clear that we require an independent and irreducible category of universal to accommodate the kinds. (...) Given all the states of affairs, where these are conceived as involving nothing but particulars ('thin' particulars, mere individuators), properties and relations, then, it may be hypothesized, all the kinds of things that there are, supervene. And if they supervene, they are not an ontological addition to their base". [3], pp. 65-68

However, not everyone agrees with the way Armstrong formulated the problem of universals. In particular, David Lewis did not like it much. But he did think that the entities posited by Armstrong might be useful to deal with many traditional philosophical problems, such as similarity, causality, natural laws, induction and so on. So what he did in [73] was to characterize properties functionally. He distinguished two different conceptions of properties associated with very different theoretical roles. Under the *sparse* or *natural conception*, there are few properties, just enough for them to fulfil the following tasks:

- Natural properties carve nature at its joints. That is to say, their predicates adequately correspond to or reflect the *objective* structure of reality.
- Natural properties ground or explain the *similarities* among objects.
- Natural properties provide their bearers with *causal powers*.
- Natural properties occur in *natural laws*.
- Natural properties are expressed (or referred to) by projectable predicates, and therefore can be used to make *sucessful inductive inferences* about their bearers.
- Natural properties form a *minimal ontological base* over which the rest of the properties supervene.

In contrast, under the *abundant conception* there is an enormous amount of properties. These properties can be distinguished from each other in such a fine-grained way that they may be used as semantic values to give a compositional account of the meaning of linguistic expressions. In fact, given the Comprehension Schema there are as many abundant properties as would be needed for such semantic purposes. Examples of abundant properties which are clearly non-natural [73] are *not golden*, *golden or wooden*, *self-identical*, *owned by Fred*, *grue*, and so on. Lewis himself proposed sets of possibilia as abundant properties and suggested several candidates for natural properties (primitively natural sets, sets of similar objects, and so on), including Armstrong's universals.

The point that Lewis wanted to make is that while these two conceptions require different entities, Armstrong's universals seem to satisfy only the sparse conception. To put it briefly, for Lewis positing universals is one way to ground the distinction between natural or sparse properties and non-natural or abundant properties. Universals would be one of the candidates for sparse properties, but they are not the only ones. Instead of universals, tropes (particular properties) or classes of similar objects could be used as candidates for natural properties. Although Lewis himself says that he will take 'property' to mean a class or set of possibilia, the use that he makes of the concept of a property looks that of a *functional* concept, characterized by the theoretical roles it fulfils. Under this conception, every position regarding the problem of universals accepts the existence of properties in some (albeit possibly weak) sense. A sparse property would be whatever fulfils the theoretical roles he lists. The problem of universals becomes one of arguing for the entities to be considered as best candidates to fulfill that role.

Other philosophers, such as Pereyra [109], have chosen the inverse problem, the "Many over One Problem", as the main concern for the dispute. The question for him is then "how is it possible for the same particular to have different properties?". That he reformulates the problem of universals in that way is no surprise, given that he defends a position known as *resemblance nominalism.* According to resemblance nominalism, there are no universals. The role of natural or sparse properties is best fulfilled by classes of similar objects. For an object to have a property is for it to belong to a certain resemblance class, no further entities are needed. Therefore, the emphasis is put not on how it is possible that the same property is had by different objects, given that this is explained by the fact that properties are just collections of objects, but on how it is possible that the same object has several properties, given that this requires appealing to resemblance facts involving objects.[2]

In contrast, Lowe [77] seems to think that the problem of universals is just a part of the more general puzzle concerning the categorial structure of the world. In order to explain several metaphysical phenomena, such as qualitative change and the persistence conditions of entities, their identity conditions, their natures or essences, their modal features, the status of natural laws, and so on, Lowe argues for a theory that requires the existence of both particular and universal entities. He favours an Aristotelian 4-category ontology obtained by crossing the two dichotomies particular-universal and substantial-non substantial, defined as follows:

- x *E-depends ontologically on* y iff necessarily (if x exists then y exists).

[2] In an attempt to get a clear formulation of the problem, Oliver [97] considers three possible treatments of it consisting in giving some explanation to the sentences of the forms "a and b are of the same type/have a property in common", "a and b are both F", "a has a property" and so on. We may be trying to find out the ontological commitment of these sentences, or to give a conceptual analysis of them or to find their truthmakers (the entities that ground their being true). Oliver argues that the problem of universals is this latter one. I will not consider the topic of truthmakers in this thesis though.

- x *I-depends ontologically on* y iff identity conditions for x involve y.
- x is a *particular* iff x can instantiate some y and x cannot be instantiated by any z.
- x is a *universal* iff x can instantiate some y and x can be instantiated by some z.
- x is a *substance* iff there is no y such that x I-depends ontologically on y.
- x is *non-substantial* iff x is not a substance.

The four categories are related by the two formal ontological relations of *instantiation* and *characterization.* Substantial particulars are *objects*, which on the one hand instantiate substantial universals or *kinds*, and on the other hand are characterized by non-substantial particulars or *modes.* Non-substantial universals or *attributes* are instantiated by modes and characterize kinds. An object *exemplifies* an attribute iff it is characterized by a mode that instantiates that attribute. As an example, $Pelusa$ is an object which instantiates the kind Cat and which exemplifies the attribute of $Fluffy$ by being characterized by the trope $Fluffy_{Pelusa}$. In sum, for Lowe the problem of universals becomes one of choosing the best theory of ontological categories. In particular, Lowe thinks that kinds (natural or not) are *substantial universals*, or as Aristotle called them, 'secondary substances'.

In this thesis I will adopt the basic distinctions from Lowe's 4-category ontology, the reason being that is the most clear and comprehensive system of ontological categories that I know of. Note that what I will accept is merely Lowe's conceptual distinctions, since they allow me to chart finely the different positions. That does not imply, of course, that I will accept his strong realism. But it does imply that I will assume his distinction between attributes and kinds. More reasons for making this distinction will be given in Chapter III.

Whatever the most adequate formulation of the problem really is, we have our second argument, which is metaphysical:

Metaphysical Argument from Universals and Sparse Properties

1. There are some metaphysically relevant facts that require explanation: that some objects are more similar to each other than to other objects, that similar objects produce similar effects in other objects, that truthmakers for certain sentences seem to require the existence of entities that are not objects, that there are entities of different ontological categories, and so on. [Pr]

2. To explain the facts mentioned in (1), we appeal to some entities shared by these objects, namely, the properties. But only some of these fulfill the role of *sparse* properties, others seem to be too gerrymandered for these purposes. [Theoretical Role of Sparse Properties]

3. The best explanation for (1)-(2) is that only some of the properties shared by objects can explain some of the facts mentioned in (1). These properties are the natural properties. [C]

Of course, the main metaphysical discussion concerns the nature of these natural properties, whether they are particulars or universals. This is one of the ways to introduce the problem of universals, to be considered in the next chapters. Contemporary essentialists usually accept some sort of realism of universals for natural properties, and some of them also for natural kinds. Thus, the metaphysical side of the debate goes further than the issue of whether there are real or genuine natural 'divisions' among the objects in the world. The additional question is [10]: what kind of entities are these kinds or sorts of objects?

2.1.3 Semantic Argument: Reference of Natural Kind Terms

In the 70-s Saul Kripke and Hilary Putnam formulated some powerful objections to the descriptivist theories of reference and meaning of proper names and natural kind terms. Descriptivist theories assumed that the reference of proper names and natural kind terms was mediated by some descriptions (or disjunctions of descriptions, senses or modes of presentation) that the speaker knew a priori, that expressed that the referents had some (bundles of) properties necessarily and that constituted the meaning of the term. Kripke [67] pointed out several counterexamples and problems for each of these theses. He proposed an alternative picture according to which those terms refer directly by being introduced first by an initial baptism. Then their use is socially extended from this baptism by historical-causal chains of speakers that transmit the term to each other. In this proposal, the reference of the term completely exhausts its meaning.

However, the semantic theses defended by Kripke [66], [67] seemed to have major epistemological and metaphysical consequences. One of Kripke's arguments against descriptivist theories starts by showing that the notions of a priority and necessity (respectively, a posteriority and contingency) are not co-extensional. The paradigmatic examples of a posteriori necessarily true propositions would be some empirically discovered identities expressed by statements of theoretical identity, such as *heat = kinetic energy of molecules* or *gold = the chemical element of atomic number 79*. Before [66], it was largely assumed that some propositions could express contingent identities among entities. For instance, *the First Postmaster General of the United States = the inventor of bifocals* would express a contingent identity, since the First Postmaster General of the United States, i.e. Benjamin Franklin, could have not invented the bifocals. This statement was in conflict with the principle of the necessity of identity $x = y \to \Box(x = y)$, that could be proven in a modal logic assuming as a premiss the necessity of self-identity $\Box(x = x)$ (the rest of the steps in that deduction not being very questionable). What Kripke argued was that strictly speaking every identity holds necessarily. The key was to refer to the objects in a specific way. The theoretical notion that does the trick is that of a *rigid designator*, a term that refers to the same entity in all the possible worlds (in which that entity exists). Proper names would be clear examples of rigid designators. In that way, if the descriptions 'the First Postmaster General of the United States' and 'the inventor of bifocals' were replaced by proper names (or were rigidified) like 'Benjamin Franklin' and 'B. F.', we would have a necessary identity. Being rigid, these names would refer to the same individual object (that person we call 'Benjamin Franklin') in every possible world (in which he exists). Kripke's arguments are grounded on some metaphysical intuitions of the speaker: *Benjamin Franklin* = *B. F.* would simply imply that that object (the person) is identical to itself, and there is no possible situation in which we would say that *that object* would not have been itself. In some sense there is a conceivable scenario, which is epistemologically indiscernible from the actual one, in which we discover that the object that we had been referring to as 'Benjamin Franklin' and the one that we had been referring to as 'B. F.' were not the same object. A speaker may not know that they are names for the same person. But this situation would not be a world in which, *supposing that* 'Benjamin Franklin' and 'B. F.' refer to some individual object a, this object is distinct from itself. That situation seems to be unintelligible. The notion of rigid designator fulfilled the role of explaining what is going on when we directly refer to an object, thus explaining the deictic character of proper names.

Alongside the principle of the necessity of identity, Kripke discussed the following essentialist theses, where the box is the operator for metaphysical modality:

1. If x originates in y, then $\Box(x$ originates in $y)$. [Origin Essentialism]
2. If x is constituted by y, then $\Box(x$ is constituted by $y)$. [Constitution Essentialism]
3. If x is a K, then $\Box$ (x is a K). [Natural Kind Essentialism]

For instance, (1) if Jaime is son of Elvira and Jacinto, it is not possible that Jaime would not have been son of Elvira and Jacinto. (2) If x is constituted by some carbon molecules $y_1, \ldots, y_n$, then x could not have been constituted by other molecules. (3) If Jaime is a *homo sapiens*, then Jaime could not have belonged to a different biological species. Each of these conditionals would be a priori justified. Since the antecedents can only be justified a posteriori, their truth could only be established by empirical research, then that would make the conclusion a necessarily true a posteriori proposition (since at least one of the premisses requires a posteriori justification). The consequents express the attribution of some property or relation to some objects necessarily, that is to say, they are cases of de re modality. For instance, if x is a member of natural kind K, then in every possible world in which x exists, x is a member of natural kind K. Apparently, the assignment of some necessary (de re) properties to individual objects would be equivalent to attributing essential properties to them. From that it follows that (1)-(3) would express that the origin, constitution or membership to a natural kind by an individual object a would be essential properties of that object a. Although 'being identical to itself' may not seem a very interesting essential property, theses (1)-(3) would point at a stronger essentialism.

Kripke used two different kinds of arguments to defend his theses. First, these theses would be justified a priori, on the basis of our metaphysical intuitions. For instance, supposing that this table is made of wood, this same table could not have been made of ice. We could discover that we were wrong to think that this table is made of wood (e.g. it is made of ice but a clever guy painted it in such a way that it tricked us). But supposing that it is in fact made of wood, it seems inconceivable that it would have been made of ice. If this table (that I am pointing at right now) is made of wood, to suppose that it may have been made of ice is just to imagine a different table. Second, essentialist theses could be deduced in some way from the semantic theory of direct reference. We will deal with these arguments later on.

Although these theses concern individual objects (they attribute *individual essences* to objects), (3) establishes a strong link between an individual object and the natural kinds to which it belongs. (3) says that belonging to a natural kind would be an essential property of every member of that kind. But (3) says nothing about whether the natural kind itself is (really) defined by some properties that are essential to its members. It could be that the natural kind itself has an essence (a *general essence*) consisting in some properties that are such that their being exemplified by some object is a necessary and sufficient condition for it to belong to the natural kind. That would explain why the object essentially belongs to the natural kind. But this goes further than (3). However, Kripke extended his semantic theses to the case of the natural kind terms. As a consequence, the essentialist theses also applied to natural kinds. Kripke did not say much about natural kinds, but he held that natural kind terms would be rigid designators, just as proper names are. Therefore, Kripke's arguments could be applied to cases of theoretical identity between natural kinds such as $Water = H_2O$.

On the other hand, Putnam [103] made several objections to the descriptivist theories of the meaning of general terms. His aim was to introduce a new theory of the meaning of general terms (specifically, of natural kind terms) which did not involve 'mentalistic' theses according to which the intension or meaning of a general term is determined by the mental states of the speakers. Putnam's arguments are the widely known *Twin-Earth thought experiments*. These start by assuming an internalist semantic position according to which the meaning of a general term is determined by the mental states of the speaker. Supposing that the intension of the term determines its extension, Putnam presents a counterexample (the Twin-Earth scenario) in which a speaker that is in the same psychological states as the actual speaker uses the same term ('water') to refer to a liquid substance which is macrostructurally indiscernible (e.g. same taste, same colour, and so on) but microstructurally distinct (e.g. $XYZ \neq H_2O$) from our actual water. That is to say, the term has a different extension in that world (XYZ molecules). So two speakers in the same psychological states may use the same term with different extensions. Under the intuitive assumption that the substance in Twin-Earth would not be water, Putnam argues (among other theses) that the intension of the term is neither determined by the psychological states of the speakers nor mediated by the descriptions that these speakers associate with the term. To the contrary, the intension is fixed by the environment and some social mechanisms (e.g. sample baptisms and the division of linguistic labour) guarantee that the correct use of the term is transmitted and that the kind of thing is directly referred to. The thought experiment is as follows.

In the actual world, we have a substance that we call 'Water' (at least whenever we find it in its liquid state). Chemists have found that this substance is constituted by H_2O molecules, or that it has a molecular microstructure that we call '$\mathrm{H_2O}$'. Let us imagine a possible world, we call it 'Twin-Earth', in which there is a liquid which is superficially indistinguishable in its observable properties and behaviour from our liquid. The inhabitants in Twin-Earth call it 'Water' too. Let us imagine also an individual called 'Oscar' in our world and an exact duplicate of Oscar in Twin Earth named 'Twin Oscar'. In particular, Twin Oscar is in the same psychological states as Oscar is. Before the discovery of the molecular structure of water the psychological states of Oscar and Twin Oscar related to water are equal. Now, let us suppose that the extension of 'Water' in Twin Earth is the set of all the XYZ molecules or the set of all the macroscopic samples of a liquid which has molecular structure XYZ (where 'XYZ' is a molecular formula different from '$\mathrm{H_2O}$'). Since the term 'Water' has different extensions in Earth and in Twin Earth, and assuming that the extension of a term is determined by its intension or meaning, we conclude that the intension of the term cannot be determined by the psychological states of the speakers (because these are in the same psychological states).

An important part of the argument consists in the intuition that a substance which is macroscopically indiscernible from our water but which has a different molecular structure (say XYZ) would *not* be water. According to Putnam, the reason is that natural kind terms have an indexical component, they are rigid terms just as proper names are. We introduce a term like 'water' with an ostensive definition: we take a sample of the liquid and we say "this liquid is water". When we talk about water, we either talk about this sample or about any other liquid sample which is sufficiently similar (in some relevant respects) to this liquid. So when we ask whether the liquid XYZ in Twin-Earth is water, what we are asking is whether that would be the same kind of liquid as this liquid we have in our world (x is water iff $x =_L a$, where a is this liquid in our world). What necessary and sufficient conditions that liquid should satisfy to be the same kind of liquid as this one is something discovered by that part of our linguistic community devoted to make research about its properties (the chemists). That is why the ordinary speaker does not need to know these conditions. He makes a *deferential* use of the term and he trusts that the corresponding specialists will know the theoretically relevant properties that something needs to have to be of the same kind. A fundamental part of Putnam's proposal is that the (transworld) relation *is the same kind of liquid as* is a theoretical relation. That is to say, the choice of the relevant physical properties is made according to what the actual scientific theory about that natural kind says. In sum, what Putnam's argument seems to show is that:

$$\forall x \; \Box (x \,\text{is water} \rightarrow x \,\text{is}\, H_2O)$$

In every possible world we consider, if something is a macroscopic sample of water or a molecule of water, then it also has molecular structure H_2O or is a molecule of H_2O. Plausibly, the converse is necessarily true too, but here there are some problems since 'water' seems to exclusively refer to the liquid state of that substance, and not to that substance in a solid or gas state [11]. In any case, Putnam's argument rests on an intuition: we would not say that that liquid in Twin Earth is the same kind of thing as this liquid here, since that would require them to have the same microstructure. We believe that every possible world in which we found water would be a possible world in which there is a liquid with the same microstructure as this liquid. In other words, we believe that having a specific microstructure is a necessary condition for being this kind of liquid, and therefore that is an essential property of this liquid.

Kripke's and Putnam's arguments hint at a strong connection between the meaning and reference of natural kind terms and natural kind essentialism.

Nevertheless, and ignoring essentialism for now, the semantic theses put forward give us another argument in favour of natural kind realism:

Semantic Argument from Natural Kind Terms

1. The reference of our natural kind terms is not mediated by descriptions determined by the psychological states of the speakers. [Antidescriptivism, Twin Earth Scenarios]

2. The best alternative explanation to (1) is that our natural kind terms directly refer to kinds or types of things in the world, that (at least partially) fix their meanings. The use of these terms is grounded on social linguistic mechanisms such as the baptism of samples, the division of linguistic labour and the social transmission of the uses of some terms from some speakers to others. [Direct Reference Theory]

3. There are natural kinds, they are the references of our natural kind terms. [C]

Kripke's and Putnam's theses and arguments have been used by contemporary essentialist theories of natural kinds, in particular by scientific essentialism.

2.1.4 Naturalist Argument: Scientific Classifications

It is usually acknowledged (at least among scientific realists) that it is the task of scientific disciplines to find out what kinds of things there are and to discover that in virtue of which several objects belong to the same kind. The way scientists describe kinds is by developing *classifications* for the entities in the domains they study. These classifications are supposed to modify some of our pre-theoretical beliefs about these kinds of things (e.g. many general names for plants do not refer to species but to entire genera). Prima facie, one could distinguish between *vernacular natural kinds* like *water* and *tulips* and *scientific* or *theoretical natural kinds* like the molecular compound H_2O and the genus $Tulipa$. Examples of scientific kinds are described by the most successful scientific classifications, such as cladistic biological taxonomy or the periodic system of chemical elements. There seems to be no one-one correspondence between each vernacular kind and each scientific kind[3]. Scientific classifications often show that vernacular classifications are coarser than theoretical ones. For instance, the extension of a vernacular kind may in fact include members of several overlapping theoretical kinds. Or they may show that a vernacular kind approximately corresponds to a scientific kind which happens to have many different species (e.g. $Tulipa$). It can also happen that there is no scientific kind corresponding to a given vernacular kind (e.g. $jade$)[4].

Moreover, fruitful classifications are made using sorting criteria (see [36]) that are usually based on more fundamental theoretical principles. For instance,

[3]Philosophers often start discussing vernacular natural kinds and try to obtain knowledge about scientific kinds, or they accept that both vernacular and theoretical kinds are in some sense genuine natural kinds. But if Putnam's semantic hypotheses are correct, then to each vernacular kind term there will correspond a scientific kind term which has the same meaning, and therefore there will be no real distinction between vernacular kinds and scientific kinds.

[4]These interactions between folk and theoretical classifications are further discussed by Dupré [32].

historical classifications sort entities in virtue of the historical relations among them, e.g. phylogenetic classifications of species and historical classifications of languages by descent. *Qualitative classifications* sort entities by their shared qualitative features, e.g. the periodic system of elements or the phenetic classifications. *Functional classifications* sort entities in functional kinds, that is to say, they group together entities that fulfil a similar role, e.g. a classification of artefacts like weapons or tools or an ecological classification.

But there are features common to every classification. When we classify a domain of entities we represent it with a smaller structure, consisting of the classes (or the general terms or nouns) into which we map each of the elements of the domain. The qualitative parsimony achieved by a classification makes salient the most basic structure of the domain, that of similarities and differences among entities. Furthermore, this grouping seems to be an indispensable step towards understanding. By locating each entity into its corresponding class, we obtain a simpler picture of the domain and the relations among the former ones are more easily tracked. The quantitative parsimony obtained can be quite remarkable, given that we can reduce a domain of an infinite amount of entities to a domain of finite cardinality. If several classifications of the same domain are combined, they give us even more clues regarding the ways in which the distinct classes of entities are interrelated. For instance, some classes are more specific than others, some classes are the overlapping of several others, one classification as a whole may make finer distinctions than another one, and so on.

When one classifies a domain of entities one maps each object into a class. This map preserves the similarities among the elements as distinctions between the classes, allowing us to move 'one layer of abstraction up'. In the trivial case, we start from a domain of entities whose only structure consists in the identities and distinctions among some entities and we map them to the finest classification possible. This is the procedure we follow when we assign proper names, labels, postal codes, telephone numbers, passwords, and so on. However, more fruitful classifications make use of coarser differences. This involves a trade-off. The more coarse-grained the differences are, the more general our classification becomes and the smaller the resulting number of classes is. In exchange, we will lose information by ignoring certain differences. Conversely, the more fine-grained the differences are, the more specific and informative our classification becomes. In exchange, the number of classes will be higher and the resulting classification will be more unmanageable.

Assuming a minimal realism, one expects these distinctions and order relations among the classes in the classifications to correspond to distinctions and order relations among kinds in nature. This being so, it is reasonable to suppose that we should be able to make some reliable (although admittedly simple and mostly categorical) surrogative inferences [129] about the domain of entities from our classifications, which may further imply even more specific inferences. As a consequence, these propositions render scientific classifications empirically testable. For example, the simplest propositions are syllogistic, they have the forms "All K-s are K'-s" or "some K-s are not K'-s. One may think that such propositions are easily testable. This is far from being true. For instance, cladists have developed a whole arsenal of techniques to infer the phylogenetic relationships among species from the homologue traits shared by their members [140]. If classifications can be testable, then they should be revisable in the light of new empirical evidence. That is indeed the case. We may discover that some of our hypotheses were wrong. Biological taxonomies are in constant revision. For example, when a fossil taxon is discovered, some of the phylogenetic relationships between already known taxa as depicted in the trees have to be changed.

Thus, a classification has at least two main theoretical aims, as several authors such as Hempel [59], Ereshefsky [36] and Mayr [83] have highlighted. One is to organize all the entities in a systematic way that shows the relationships between the kinds to which they belong. As a consequence of this systematic organization, we obtain a quantitatively more parsimonious domain of entities and our understanding of the relationships among them is improved. The second one is to allow us to make surrogative inferences in the form of empirically testable (categorical) propositions. This allows us to formulate empirical hypotheses concerning the domain of entities. As a consequence of the empirical testability of these propositions, classifications may be revised in the light of further evidence.

It was once assumed that classifying is just a step towards proper theorising, which would start with measurement and continue up to the formulation of scientific laws in the form of (differential) equations. After all, familiar arithmetical operations with real numbers cannot be meaningfully applied to classifications. Of course one can use numerals to represent kinds, in fact, many nomenclature codes use numerals for their classifications. But these are just like ZIP codes or proper names. One cannot meaningfully perform arithmetically basic operations over them[5]. It is true that not many traditional statistical concepts of measures of central tendency apply to 'categorical properties', and therefore that the kind of inductive inferences that one can make from such information is rather limited. Stevens [127] famously argued that the only adequate statistical measure of central tendency for categorical variables was the mode. Arguably, counting the number of times that a specific item appears in a sample does give some information concerning the population studied, but this is far away from the whole apparatus of statistical techniques that scientists generally want to use to infer interesting (causal) correlations between properties. Being so close to a task made in ordinary life, and devoid of the alleged rigour, stability and precision given by the numerical structures traditionally associated with measurement, disciplines like biology that invest too much time in systematic "stamp collecting" (to use Rutherford's unfortunate phrase) have been at times considered not to be proper sciences.

Fortunately, this is not the case now. Philosophers of science are putting more and more attention to how classifications are made. This has lead to the development of several realist theories of kinds. In fact, contemporary scientific realist positions regarding natural kinds take these classifications to be tracking natural kinds. They can be divided into three big camps[6]. The first group is formed by essentialist theories, such as the *Scientific Essentialism* by Ellis [34]. The second group is formed by causal theories that stem from Boyd's *Homeostatic Property Clusters Kind Theory* (HPC) [12]. The third group is more varied but it contains *Promiscuous Theories* like the one by Dupré [32] and

[5]For instance, it makes no sense to add a class with itself to obtain bigger classes, there is no natural 'zero' class which behaves as the zero does for real number addition and multiplication , it is unclear what difference could be found between 'positive' and 'negative' classes, classes need not be necessarily linearly ordered, and so on.

[6]See [1] for a similar classification.

pragmatist theories of different sorts, that are closer to conventionalism in some respects. What these three theories have in common is a minimal naturalistic concern. That is to say, according to these authors, we should look for natural kinds in the classifications of our most successful scientific theories, such as the standard model of particles in physics, the periodic system of chemical elements, the mineralogical classification and the biological classifications made by the cladistic school of taxonomy. Our fourth argument starts from the premiss of scientific realism to conclude that plausibly there are natural kinds:

Argument from Scientific Classifications

1. The most successful scientific theories are closer than the previous ones to a true description of the structure of the world. [Scientific Realism, Ontological Thesis]

2. Some of these theories ground (more or less accepted) classifications of particular objects into general kinds or types. [Pr.]

3. There are natural kinds, they are those entities represented by the classes that occur in the classifications grounded by the most successful scientific theories. [C]

These are only some of the arguments that several authors have used to appeal to natural kinds. They are not the only ones. Moreover, this does not mean that there are no alternative explanations (like conventionalist or conceptualist explanations) for each of these problems. But what they do show is that these other explanations will have to give a coherent answer to each of these problems and that is a minimal constraint to be satisfied by any adequate account of natural kinds. In what follows, I will assume that a (minimal) Natural Kind Realism constitutes a good answer to the four challenges mentioned so far. The problem now is to get some more information regarding what these kinds are supposed to be. In order to get a clearer view on what kinds are supposed to be, let us consider now some of the main theories of kinds, namely essentialism, causal theories and conceptualism[7].

2.2 Scientific Essentialism

The purpose of this section is to introduce a popular version of natural kind essentialism, namely, *scientific essentialism*. Natural kind essentialism started with the semantic arguments put forward by Kripke [67] and Putnam [103]. Contemporary authors who have defended such theories of kinds are Ellis [34], Bird [10] and Tahko [134]. There are other sorts of essentialism, such as the *sortal essentialism* that comes from Wiggins [139][8], a version of which is defended by Lowe [77]. Generally speaking, essentialist theorists tend to emphasize the semantic and metaphysical sides of the problem of natural kinds.

[7] It should be clear that a brief chapter such as this one cannot make justice to the full variety and richness of these accounts. A sufficiently informed analysis of these theories would require at least one chapter for each and preferably a whole PhD thesis. This is a task I cannot accomplish here without sacrificing what I consider to be a fruitful research proposal that will be developed in the next chapters. Nevertheless, the chapter is written with the aim of conveying the reader why I decided to leave these theories out of the rest of the thesis.

[8] Wiggin's essentialism is in fact conceptualistic.

2.2.1 Theoretical Roles of the Concept of Essence

The distinction between essential and non-essential or accidental properties was introduced by Aristotle. In Aristotelian metaphysics, this distinction fulfils several theoretical roles, such as explaining the nature of objects, the possibility of change and the possibility of our knowledge of necessary truths. The problem of change consists in that the same object x can have a property P in instant t, and another property P' incompatible with P in t', where $t < t'$. By the *Indiscernibility of Identicals* it would follow that $x\,in\,t \neq x\,in\,t'$. But what we want to say is that there is an object x which exists both in t and in t', even though it has suffered qualitative change. If this were not so, change would be impossible. The Aristotelian answer to the problem consists in making the distinction between essential and non-essential properties, and then restricting the *Indiscernibility of Identicals* to essential properties. That is to say, if $x = y$, then x and y have the same essential properties. The same object can exist in different time instants with different properties if these are accidental to it. The problem of the possibility of knowledge consists, for Aristotle, in explaining how it is that we have knowledge of necessarily true universal propositions. These propositions involve definitions of those objects of which they give us knowledge. Since giving a definition of a kind of object is giving the necessary and sufficient conditions under which the object is the kind of object that it is, by giving the definition we are giving the essence of the object. Knowledge is therefore knowledge of essences of objects. The third problem refers to the nature of objects. Aristotle's hylemorphic explanation of what substantial particulars (primary substances) are appeals to the matter-form distinction. A substance is composed of matter and form. Matter is pure potentiality, whereas form is that which determines matter as certain thing instead of another one. The form of an object is its essence, it is what makes objects distinct from each other. Thus, the form is supposed to fix the identity conditions for the object.

The theoretical roles of the Aristotelian distinction are still present in the contemporary literature in some way or another. The problem of change has to do with the general persistence conditions of objects [77]. The problem of our knowledge of universal necessarily true propositions has to do with the knowledge of de re modal propositions, and more specifically, with our knowledge of natural laws [34]. Contemporary essentialists, like [39], also hold that giving a real definition of an entity consists in giving its essence. The problem of the nature of objects is thus linked to the identity conditions for objects [77]. The distinction between essential and non-essential properties is so central that it is linked with the rest of metaphysically fundamental problems, like those of reality, identity, existence, modality, change and the distinction between particulars and universals. So it is reasonable that the following ones should be the theoretical roles most frequently attributed to the concept of essence:

i. The essence determines the conditions under which the entity is that entity and no other thing. The essence distinguishes one entity from another one. In other words, the essence determines the *identity conditions* of the entity.

ii. The essence determines the conditions under which the entity persists or continues to be the same thing, without getting extinct or being transformed into a distinct entity. It determines what is a substantial change and what is a mere accidental change. In other words, the essence determines the *persistence conditions* of the entity and it is invariant under accidental changes.

iii. The essence determines the properties that the entity could or could not have had. It determines how the entity could have been like. In other words, the essence determines the *de re necessary properties* of the entity.

iv. The essence determines the conditions under which the entity exists or could have (not) existed. In other words, the essence determines the *existence conditions* of the entity.

v. The essence determines on which other entities it depends ontologically. In other words, the essence determines the *relations of ontological dependence* of the entity with respect to other entities.

vi. The essence is the *real* or *fundamental nature* of the entity. The essence is what the entity really is.

Through the history of philosophy empiricist and nominalist authors have shown a sceptical attitude towards the distinction between essential and non-essential properties. Usually these authors have proposed to use classes, general concepts or general terms as surrogates for kinds or types and definitions as surrogates for essences. According to them, one defines a term or a concept, not a kind of thing. For instance, Locke introduced the distinction between *nominal essence* and *real essence* to argue that what philosophers propose as definitions of the essence of an object are only necessary and sufficient conditions that objects should satisfy to be appropriately represented by the corresponding general ideas. Not any definition will do of course, since there are psychological constraints on how these general ideas are formed. But in any case, our definitions are of those general ideas. Since these are obtained as the result of a process of abstraction from qualities or observable properties of objects, they cannot adequately correspond to the fundamental structure or properties that make the objects have the observable properties that indeed have. The real essences of objects, if there are such things, are unknowable. This sort of objections among others (say, about the explanatory uselessness and ad-hocness of appealing to essences) directed the attention of philosophers towards the definitions of concepts and general terms and favoured a conceptualist approach to essence. Thus, conventionalism and conceptualism became the standard positions.

Some authors seem to still identify the general essence of a kind with the necessary and sufficient conditions that objects have to satisfy to belong to it. In other words, they identify the essence with the *membership conditions* of the kind. But this latter view, if it is not further developed, has some well known problems. One can give necessary and sufficient conditions for something to belong to a class of objects without these conditions having to express the essence or common nature of its members, simply because these conditions could be stipulated or conventionally chosen. For instance, the *Comprehension Axiom Schema* guarantees that given a set S and a predicate P, there is a subset A of S consisting in exactly those elements x in S which satisfy P. Therefore, for the members of S, satisfying P is a necessary and sufficient condition for belonging to A. But P does not have to express any essence or general nature common to the members of A. Moreover, P could be just the result of a stipulation. For example, suppose that S is the set of all blue or green things and P is the general term 'grue'. Then P does not express any essence common to the members of the extension of 'grue' and the term does not give objective membership conditions, but still the definition of 'grue' states necessary and sufficient conditions for membership to the extension of the term. Therefore, assuming that there are necessary and sufficient conditions for belonging to a kind does not commit us to essences. In particular, a definition does not need to express the general essence of objects.

However, the discussions on the feasibility of modal predicate logic lead to the re-emergence of a modal approach to essentialism, which is known as the *modal conception of essence*. According to this conception, a property is essential to an object iff the object has that property necessarily, in other words, in every possible world (in which the object exists). This is the conception of essence that comes from Kripke [67]: essences are exactly the *de re* necessary properties of objects. Thus we had that[9]:

$$a \text{ is essentially } P \Leftrightarrow \Box Pa$$
$$a \text{ is accidentally } P \Leftrightarrow \neg\Box Pa$$

Alongside the development of first-order modal logics and Kripke's arguments for necessary a posteriori true propositions, this conception of essence combined with naturalism and scientific realism, resulting in a theory of kinds known as *scientific essentialism*. The idea is that it is not enough for the necessary and sufficient conditions to hold contingently, they must hold metaphysically necessarily. An object belongs to the kind iff there are some properties which are essentially had by that object. Equivalently, the object belongs to the kind iff there are some properties which are such that the object has them in every possible world (in which the object exists).

However, Kit Fine rejected in [39] the equivalence between essential properties and de re necessarily exemplified properties. He famously objected to this identification by arguing that it is essence what grounds de re necessity and not vice versa. He attacked the right-left direction of the previous equivalences. His arguments start by choosing some properties that objects exemplify necessarily (in every world in which they exist) and then showing that intuitively they have nothing to do with the nature of the objects chosen. For instance, in every possible world in which Socrates exists, the singleton $\{Socrates\}$ also exists and Socrates is a member of that singleton. This is because if Socrates exists, then the singleton $\{Socrates\}$ necessarily exists and he is necessarily a member of this set. But it is not essential in any sense to Socrates that he is a member of $\{Socrates\}$. Nothing in the nature of Socrates makes him be a member of some specific set. In contrast, it is in the nature of the singleton $\{Socrates\}$ that Socrates is its only member. The problem is that the modal conception of essence is indifferent towards the difference in the sources of these two necessities. According to another example, and accepting the necessity of distinctness, if Socrates is distinct from the Eiffel Tower (and we think so) then he is so necessarily. But it is not essential to Socrates that he is distinct from the Eiffel Tower. Therefore, not every property necessarily exemplified by an object is an essential property of the object.

[9]As it is well-known, to allow for contingently existing objects we have to weaken the formulation to the following one: a is essentially $P \Leftrightarrow \Box(\exists x\, x = a \to Pa)$.

Based on these counterexamples Fine argued for taking the notion of essence as fundamental and modality as derived from it. According to him, we should favour the *definitional conception of essence* over the modal conception. The essence of an entity, its nature, is given by its real definition and not by those properties that it happens to exemplify necessarily. According to this conception, we should take the last description from the previous list as being the basis for the rest: the concept of essence is primitive. That being so, metaphysics can be thought of as the science of essence[10]. However, in order to discuss scientific essentialism, I will mostly stick to the modal conception according to which the general essence of a kind is exhausted by the modally de re necessarily exemplified properties that fix the membership conditions of the kind. The reason is simply that this is the usual way in which the discussions are framed in the literature on kinds.

2.2.2 The Scientific Essentialism Programme

The minimal thesis that all essentialists seem to be committed to is the following, proposed by Tahko [134]:

Natural Kind Essentialism There are genuine and mind-independent natural kinds which are defined by their essential properties.

Of course, essentialism does not imply that we already know all the essential properties of a natural kind. This is a 'minimal' thesis because it does not hold that every genuine natural kind is defined by some essential properties shared by all and only its members. The reason is that essentialist theses seem to be at odds with the current conception about what biological species are. That being so, essentialists usually restrict their theses to chemical kinds, such as chemical elements or compounds. Some authors like Ellis [34] bite the bullet and hold a restricted version of essentialism. For Ellis the real natural kinds are just the fixed kinds (elementary particles or chemical isotopes), whose members share only some intrinsic essential properties. This excludes biological species and possibly also minerals. This minimal claim will be further developed.

Essentialist theses are usually divided into two kinds, depending on whether they are concerned with *individual essences* or with *general essences* of objects. The individual essence of an object is its individual nature, it is what makes that object be the specific object that it is instead of another. As explained, the individual essence of an object is traditionally understood to be the collection of all those properties, the essential properties, that make that individual be that individual instead of another one. As a consequence, these properties are necessarily (de re) exemplified by the object in every (metaphysically) possible world in which the object exists[11]. Typical candidates for individual essential

[10] Note that according to Lowe [77], since every entity has an essence, essences are not entities, on pain of infinite regress. Instead, according to Lowe, essences are the conditions under which the objects 'defined' by them exist or are distinct from one another.

[11] The notion of metaphysical necessity (or "logical necessity in its broad sense") I am referring to is the one that appears in the literature. It is a necessity between the logico-mathematical necessity and the nomological necessity. These are necessarily true propositions that are neither tautologies nor analytic propositions, but which are weaker than nomologically necessary propositions. Nevertheless, some scientific essentialists collapse the distinction between nomological and metaphysical necessity.

properties of an object would be its origin, its constitution and its membership to a certain natural kind (the ones suggested by Kripke [67])[12].

In contrast, the general essence of an object is a nature that this object shares with others. It is what makes that object be the sort or kind of object that it actually is. Traditionally, it is the collection of all those properties that make the object be the kind of object that it is. Therefore, these properties are shared by all the instances of that kind of object and having them is a necessary and sufficient condition for belonging to the kind in question. In fact, it is by exemplifying those properties essentially that the object necessarily belongs to the kind that indeed belongs to. Keeping this distinction in mind, scientific essentialism holds *at least* the following theses, possibly restricted to specific domains of entities:

Individual Kind Essentialism If x is a K, then x is essentially a K.

General Kind Essentialism For every natural kind K, there are some natural properties $P_1,\ldots,P_m$ which are such that x is a K iff x essentially exemplifies $P_1,\ldots,P_m$.

General Kind Essentialism says that kinds have general essences. The general essence of a kind are those properties that determine its membership conditions and are essentially exemplified exactly by the members of the extension of the kind. All these essential properties are part of the individual essence of each of these objects. In contrast, *Individual Kind Essentialism* says that being a member of the kinds to which the object belongs is also an essential property of the object, and therefore part of its individual essence. In particular, *Individual Kind Essentialism* forces the extension of the kind to be invariant across possible worlds (up to restrictions of the extension to the corresponding domain of the world)[13].

In a different sense, the thesis is a minimal essentialism because it is not committed to any specific relation between natural kinds, natural laws, the (non) intrinsicness of properties, the categorical or dispositional nature of properties, and so on. In contrast, consider Ellis's scientific essentialism [34]. Ellis's theory is a very strict Aristotelian metaphysical picture of reality. First, every universal is a natural kind. Some of them are substantial natural universals (e.g. *electron*), others are natural properties (e.g. *electric charge z*) and still others are natural processes (e.g. *precipitation reaction*). Each of the universals is instantiated by a particular of certain kind: instances of substantial universals are individual objects (e.g. this electron a), instances of natural properties are tropes (e.g. the electric charge c of this electron a) and instances of natural processes are particular events and processes (e.g. this particular chemical reaction f). Individual objects instantiate some natural substance iff they essentially exemplify some causally intrinsic properties. As a consequence, they necessarily

[12]By the necessity of identity, every object is necessarily identical to itself. It can be discussed whether *being identical to itself* is an essential property or not. In any case, if it is essential to the object then it is generally considered to be so trivially.

[13]There is some ambiguity regarding the talk about general essential properties. We are talking about whether some individual objects exemplify some properties essentially, not about whether the kinds themselves exemplify these properties. But if one holds that kinds are entities that can have attributes, e.g. as Lowe does [77], then this distinction becomes important.

instantiate the substance. Second, many of these intrinsic properties are dispositional (e.g. inertial mass is the resistance to being accelerated) and therefore manifest themselves only under specific circumstances. These conditions are the causal processes in which the objects get involved. Third, natural laws are simply the expression of those causal processes (e.g. a force gets applied to a mass). Since natural laws are just those causal processes and objects are involved in those processes by virtue of their dispositional essential properties, natural laws are metaphysically necessary. Laws could not be otherwise, since they are fixed by the essential nature of the objects they involve. This conveys a very different conception of nature. The world of the dispositional essentialist is *dynamic*. Objects behave necessarily as they do due to their natures, and not because their behaviour is determined by some laws that extrinsically impose some contingent properties over them. As a consequence, a world where negatively charged particles attract each other is metaphysically impossible, for essentially negatively charged objects necessarily repel each other. Science discovers these essentially exemplified intrinsic properties by discovering first the natural kinds to which the corresponding objects belong. These objects show their nature in the causal processes in which they take part and scientists express these processes in the form of laws.

But it would be wrong to assume that every scientific essentialist agrees with all the theses that Ellis defends, although some seem to share most of them (e.g. Bird seems to agree on many [10]). Nevertheless, it is convenient to consider in more detail condition (3). It contains the basic core properties of kinds that most essentialists would accept [34]:

i. Natural kinds are objective. They do not depend on our interests, language, practices, perceptual apparatus, choices, and so on. Whether a particular object belongs or does not belong to a natural kind is neither chosen arbitrarily nor by convention. [Objectivity]

ii. Natural kinds are categorically distinct from each other. There cannot be a gradual merge or fusion of one kind into another in such a way that it is indeterminate to which of these kinds a particular object belongs. [Categoricity]

iii. The identity of natural kinds ontologically depends on the intrinsic properties of its members, not on the extrinsic relations among them. [Intrinsicness]

iv. If two particulars x, y which are members of a kind K have different intrinsic properties and these properties cannot be acquired nor lost by members of kind K, then they belong to some proper species K', K'' of K. [Speciation]

v. If a particular belongs to several natural kinds K, K', then these two are proper species of a common genus K''. [Hierarchy II]

vi. Natural kinds have real essences and essential properties. [Essentialism]

The *Objectivity* and *Categoricity* conditions have been briefly discussed and we cannot go deeper into that. *Hierarchy* (and in some sense *Speciation* too) will be treated in Chapter III. The remaining conditions are *Intrinsicness* and

Essentiality. Although Ellis does not make the following reasoning explicit (as far as I have been able to check), these two conditions give the essentialist identity conditions for natural kinds. We will denote by K the natural kinds and by P the properties. Ellis holds that:

> If $K \neq K'$, then there is a property P such that it does not hold that (P is an essential property of K iff P is an essential property of K').

Where it says "it is an essential property of K" we will interpret it as "is an essential property of each and every particular instance x of K". Since for Ellis all essential properties are intrinsic and assuming contraposition, we have:

> If for every intrinsic property P it holds that (P is an essential property of every member of K iff P is an essential property of every member of K'), then $K = K'$.

If two natural kinds are identical then they are coextensional. And if they are coextensional then their instances have exactly the same essential intrinsic properties. Therefore the converse also holds. The identity conditions for kinds are:

Essentialist Identity Conditions for Natural Kinds For every (intrinsic) property P it holds that (P is an essential property of every member of K iff P is an essential property of every member of K') iff $K = K'$.

So two natural kinds whose members have exactly the same essential intrinsic properties are identical. What makes kinds distinct are the essentially exemplified properties of their members, whatever the latter ones are (i.e. even if the extension differs from some world to another, the properties are invariant). The general essence of a natural kind is formed by those properties that the members of the kind essentially exemplify in virtue of their belonging to the kind. Not every essentialist holds that the essential properties are intrinsic to their members though. For instance, some philosophers (see [11]) have tried to make essentialism compatible with current biological taxonomy by defending that the ancestry of a species is essential to it. Nevertheless, they still share something like the conditions just mentioned. After all, if the essence of a kind is its real definition and this is the set of necessary and sufficient properties that objects have to exemplify to belong to the kind, then what distinguishes one kind from another is precisely this general essence.

Which existence conditions essentialists will give for natural kinds will depend on their attitude towards the problem of universals. Ante rem realists will hold that the kind exists even if it is not instantiated by any actual particular object, whereas in rebus realists of kinds will ask for the kind to be necessarily instantiated. But since most (all?) scientific essentialists are naturalists, they will probably hold the latter. In sum, this allows us to give a neat summary of the *essentialist programme*:

Core of the Essentialist Programme

1. There are natural kinds that ground or are grounded on objective differences in reality. [NKR]

2. For every natural kind K, there are some natural properties $P_1,\ldots,P_m$ which are such that x is a K iff x esentially exemplifies $P_1,\ldots,P_m$. [Natural Kind General Essentialism]

3. Two natural kinds K, K' are identical iff for every property P it holds that (every member of K essentially exemplifies P iff every member of K' essentially exemplifies P). [Identity Conditions]

4. If an object essentially exemplifies a property, then it necessarily exemplifies that property. [Essentiality implies De Re Necessity]

5. If an object instantiates a natural kind, then it necessarily/essentially instantiates that kind. [Natural Kind Individual Essentialism]

Some common additional commitments are that kinds are universals, that necessarily every kind is instantiated by some particular, that necessarily every particular instantiates a kind, that there are a posteriori necessarily true propositions, that laws are metaphysically necessary and so on. But the previous ones are enough to tackle the arguments for and against scientific essentialism, given that they form the minimal core of scientific essentialism. In the following sections I will briefly cover some of these arguments.

2.2.3 Cost-Benefit Arguments

According to Mumford [93], the best argument for scientific essentialism would be a costs-benefits argument: essentialism is to be adopted because among the best explanations for several phenomena it is the one with less ontological costs[14]. I take it as an argument for the best explanation.

The main problem with this argument is that natural kinds from special sciences, in particular those described by biology, do not seem to be essentialist. Any theory of natural kinds that excludes such paradigmatic classifications is so revisionist that it makes the ontological costs too high. Among scientific classifications those made by biological systematics are the most fruitful ones, so much that biological classification has been taken as a paradigm by the rest of scientific classifications. Most philosophers of biology think that essentialism is simply incompatible with evolutionary theory (see [36]). Some authors tried to avoid this problem by appealing to some sort of origin essentialism. In these accounts, what would be essential to biological species are their phylogenetic descendant relations to other species. At first sight, this seems to suit the most popular taxonomic approach to species, the cladistic one. Nevertheless, it is still a controversial thesis. A similar point is usually made regarding the classifications made by other special sciences (psychological states and processes, diseases, and so on). Although there is still a debate about whether they track natural kinds at all, an essentialist position risks excluding them from start. Essentialists can solve these issues by taking relations or extrinsic properties as fixing the membership conditions of kinds. To be in certain relations to other entities would be part of the nature of the instances of the kind. However, this manoeuvre requires more work and will definitely make the ontological costs of essentialism higher. In contrast, a minimal scientific essentialism (as the one introduced by Tahko [134]) does not require taking *all* natural kinds to be

[14]To be clear, Mumford [93] does not accept essentialism.

essentialist. So long as the existence of kinds in other disciplines is accepted, such an approach will have to be combined with non-essentialist theories of kinds. This pluralism may be more plausible but it makes the costs-benefits argument weak from the other point of view, by keeping the costs that come with essentialism while reducing the benefits that a general theory of kinds is supposed to provide.

Ellis also argues that, given that scientific essentialism implies scientific realism, this is a good reason to accept the former one. But there are other weaker theses that also imply scientific realism. Natural kind realism already implies scientific realism without any commitment to essences. Furthermore, other non-essentialist theories of kinds, such as causal theories, also imply some sort of restricted scientific realism. Therefore, implying a restricted realism is not in any form a virtue exclusive to essentialism. One could argue that the strong necessitarian conception of natural laws, according to which natural laws are metaphysically necessary, is the best theory of natural laws. One may assume essentialism and argue that the natural laws are grounded on the essences of the objects, as Ellis [34] does. But one could also appeal to strong necessitarianism as the best account of natural laws, and argue that as a consequence essentialism must be true. This strategy does not seem very appealing to many, since in general philosophers find the claim that natural laws are metaphysically necessary to be too strong. The idea that there could not have been other natural laws than the ones that actually hold sounds implausible. Since Hume's criticism of the notion of natural necessity, natural laws have been considered to be contingent. Even necessitarians like Armstrong who think that natural laws involve some sort of genuine necessity (among universal attributes) still defend that natural laws themselves are contingent. Such an argument relies on the plausibility of necessitarian theses as independent from essentialism, but there are currently strong alternatives to necessitarian views (for instance [69]).

In any case, the costs-benefits argument does not seem to be a very powerful move for essentialism. So let us consider now other more interesting arguments[15].

2.2.4 Semantic Arguments

The next package of arguments comes from the Kripke-Putnam discussions on the semantics of natural kind terms. Kripke's first argument in [66] consists in arguing that we know the following conditional a priori, where K is a natural kind:

$$\text{if } x \text{ is a } K, \text{ then } \Box(x \text{ is a } K)$$

This conditional is simply the thesis of Kripkean Natural Kind Individual Essentialism. One may argue that it is not just an intuition what grounds these principles. At least in the case of the *Necessity of Identity*, Kripke gave a modal proof of it from uncontroversial premisses[16]. However, the necessity of identity is acceptable even to anti-essentialists, given that it is considered to be at most

[15] Moreover, one could wonder whether we should not include in this costs-benefits analysis the *epistemological* costs of essentialism. After all, there is still no satisfactory story regarding how we come to know the natures of things.

[16] Some philosophers like [76] have objected to the argument.

a trivial essentialist principle. But Kripke did give an argument for the individual essentialism. It is a 'proof' that the author suggests in [67] for origin essentialism. This argument would only appeal to the theory of direct reference (names are rigid designators), the Necessity of Distinction and the Indiscernibility of Identicals (and some additional uncontroversial premisses). Moreover, the argument could be adapted for other kinds of essentialism, such as natural kind essentialism. It can be schematically put as follows:

Kripke's Argument for Origin Essentialism

1. Let a be a table and let m be some piece of wood (directly referred to). [Pr.]
2. In world w, a is made of m. [Pr.]
3. Suppose that there is a world w' in which a is made of a different piece of wood m'. [Pr.]
4. Suppose that m and m' do not depend on each other, e.g. they do not overlap. [Pr.]
5. There is a world $w'' \neq w$ in which a is made of m, and a table b is made of m'. [Pr.]
6. $\forall x, y\,(x \neq y) \rightarrow \Box(x \neq y)$. [Necessity of Distinctness]
7. In world w'', $m \neq m'$. [3,6]
8. $\forall x, y\,(x = y) \rightarrow (Px \rightarrow Py)$. [Leibniz Law]
9. In world w'', $a \neq b$. [5,8]
10. In world w', $a \neq b$. [9,6]
11. If x is made of wood m' in w', then $x \neq a$.
12. In world w', a is not made of wood m', this contradicts (3).

This argument was criticized by Salmon [117]. Thus formulated[17], it is clear that (11) does not follow from the rest of the premisses. The argument shows at most (10), that tables a and b are necessarily distinct. But this is not enough to guarantee that if something is made of m' then it must be distinct from a. The crucial issue is, as Salmon says, that an additional essentialist premise is required, such as "if it is possible that a table x is made of a piece of wood y, then necessarily every table made of y is the table x and no other" or "if it is possible that a table x be made of wood according to a plan P, then necessarily every table that is made of y according to plan P is table x and no other" (see examples P2 and P2' in [117]). But clearly, to infer origin essentialism from these premisses is to beg the question.

Another argument consists in applying the theses of rigid designators and the necessity of identity to the case of natural kind terms. A reconstruction can

[17]Premisses 4 and 5 introduce some additional difficulties, but they can be ignored [117].

be found in Soames [126], who introduces the kind *water* by lambda abstraction as:

$$[\lambda x.\forall y(\text{instances of } y \text{ are composed by two hydrogen atoms and a oxygen atom}) \leftrightarrow y = x](water)$$

What the previous identity statement establishes is that certain natural substance that we refer to by 'water' is identical to an entity that we rigidly refer to by the definite description 'the entity whose molecules are composed by two hydrogen atoms and one oxygen atom'. By the *Necessity of Identity* we deduce that this identity is necessary. We are then allegedly lead to conclude that in every possible world, water or every molecule or sample of water has the same molecular structure.

The problem is that it is not so clear what this identity statement means. It does not seem to be such a simple case as that of $Hespherus = Phosphorus$ or $Tully = Cicero$. Now we are establishing an identity between natural kinds themselves, and we do not know what these are. If the reference of 'water' is the set of all macroscopic samples of certain liquid and H_2O is the set of all sums or combinations of particular molecular compounds consisting in two hydrogen atoms bonded in a certain way with an oxygen atom, then the statement seems to establish and identity between these two sets. But if this identity holds necessarily by rigidity, then these two sets are identical in every possible world. This seems to be false since there could have been less or more water molecules or samples than the ones that actually exist [11]. Of course, the statement cannot be establishing an identity between entities of different ontological categories (e.g. 'water' refers to a particular but H_2O is a universal), that would be a category mistake.

Soames says that 'water' is a rigid designator that refers to an abstract entity. But if the entity is abstract, then it is implausible to hold with naturalism (as scientific essentialists do) that it is discovered by empirical investigation. After all, in what sense could one discover *a posteriori* that two abstract entities are identical? It seems more plausible to interpret what Soames wants to convey as the suggestion that the reference is an instantiated concrete universal (like Armstrong's), or at most reducible to a complex of concrete universals. Therefore, according to direct reference theory, the reference of water is the concrete universal *water*. But now the proposition expressed by the statement is $water = H_2O$, and therefore what it says is that some concrete universal is identical to itself, which is an innocuous tautology (this point is made by Lowe [78]). One needs something stronger to get an essentialist conclusion. It seems more plausible to say that the statement should be interpreted as establishing the reduction of the universal *water* to some complex of universals (e.g. *being composed of hydrogen*, *being composed of oxygen*, ...). Some philosophers, like Hawley and Bird [57], defend that natural kinds are reducible to such complex universals. But this still does not get us to essentialism unless we add the premiss that the instances of these universals (e.g. samples of water, molecules) essentially exemplify these latter universals or are essentially composed or constituted by exemplars of these universals. However, that requires *Natural Kind Individual Essentialism* and/or *Constitution Essentialism*, and therefore it is hardly an advance.

The moral of the story is that under the direct reference theory an identity statement is just that: the trivial statement that says that certain entity is (necessarily) identical to itself. One cannot obtain stronger conclusions from this. If $a = b$ holds, then we cannot conclude that a is essentially P unless we already know that b (i.e. a) is essentially P. To sum up, there is no proof from allegedly uncontroversial premises such as the *Necessity of Identity* and the *Indiscernibility of Identicals* for the condition and it does not follow from direct reference theory by taking natural kind terms as rigid designators either. Similarly, one cannot obtain *General Essentialism* from these arguments either.

Let us consider now Putnam's arguments. The Twin Earth thought experiment by Putnam [103] is considered to be one of the most ingenuous and influential arguments from the XX-th century. Although the aim of the argument was to refute the internalist descriptivist theory of the meaning of natural kind terms, it also seemed to show that the reference of those terms were determined by some microstructural properties which would be essential to the members of the corresponding kinds. What the argument seems to conclude is that in every possible world we consider, if something is a macroscopic sample of water, then it has molecular structure H_2O. Plausibly the converse is true too[18].

Let us assume an S5 modal logic with constant domain. Let $=_L$ denote a binary equivalence relation defined over the domain of possible individuals $=_L \subseteq D \times D$. It represents the theoretical relation of *being the same kind of liquid as*. Let 'W', 'H_2O' be predicates denoting the properties of *being water* and *having a H_2O molecular composition*. Let the individual constant a denote the actual sample of water referred to by ostension in a given world w. Then the following is the simplest version of the argument as presented by Salmon [118]:

Putnam's Argument (Salmon's Version)

1. $\Box\forall x\,(Wx \leftrightarrow x =_L a)$. [Transworld Baptism of Sample a]
2. H_2Oa. [Empirical Discovery]
3. $\forall x, y\,(x =_L y \rightarrow (H_2Ox \leftrightarrow H_2Oy))$.
4. $\Box\forall x(Wx \rightarrow H_2Ox)$. [Essentialist Conclusion]

Salmon argues that, as it stands, the third premiss is too weak to deliver the essentialist conclusion. It has to be modally augmented somehow and de dicto necessities will not suffice. For instance, it is not enough to require $\Box\forall x, y(x =_L y \rightarrow (H_2Ox \leftrightarrow H_2Oy))$, since the premisses are compatible with a counterexample consisting of the world w where H_2Oa, H_2Ob, Wa, Wb, $a =_L b$ and world w' where $\neg H_2Oa$, $\neg H_2Ob$, Wa, Wb, $a =_L b$, and wRw'. In other words, a and b could be the same kind of liquid in both worlds, namely *water*, while having both a XYZ chemical composition. Appealing to the Kripkean strategy of making $=_L$ transworldly ($\forall x, y\ x =_L y \rightarrow \Box x =_L y$) is not enough either. Even if its extension is invariant across possible worlds it may happen that both a and b are XYZ in a world w' so long as they are both water in w'. To get the conclusion one needs something stronger that will appeal to de

[18] There are some additional problems here. It seems that 'water' refers exclusively to the liquid state of the substance [11]. If so, then 'water' is a phase sortal.

re necessities sooner or later[19]. However, such a move will be tantamount to making assumptions that are not free of essentialist content. For instance, if one replaces (3) by something like the following:

$$\forall x, y\,(x =_L y \to (H_2Ox \to \Box H_2Oy))(3*)$$

then the inference can be shown to be valid. For suppose that $\Box\forall x(Wx \to H_2Ox)$ is false in a world w, where H_2Oa holds. Then there is a world w', wRw' such that $\forall x(Wx \to H_2Ox)$ is false in w'. That means that $Wb \wedge \neg H_2Ob$ is true in w'. By the first premiss, $\forall x(Wx \leftrightarrow x =_L a)$ is true in w' and therefore $b =_L a$ holds in w' too. By symmetry and (3*), $(H_2Oa \to \Box H_2Ob)$ is true in w'. But $\Box H_2Ob$ is false in w' given that $w'Rw'$, therefore H_2Oa is false in w' too. Since $=_L$ is reflexive and H_2Oa holds in w, by (3*) $\Box H_2Oa$ is true in w and therefore H_2Oa holds in w', which implies a contradiction. But a premiss like (3*) clearly begs the question concerning essentialism. It says that if we have to samples of the same liquid and one of them has a actually a specific molecular composition then the other one will have that same molecular composition in every possible world. Since the theoretical relation of *being the same liquid as* is reflexive, this is basically individual essentialism in disguise.

Note that, if one wants to get an essentialist conclusion, *is the same*$_L$ *as* must be a transworld relation (thus relating entities from different possible worlds) and the worlds must be all of the metaphysically possible worlds. According to Putnam *is the same*$_L$ *as* is a theoretical relation. Putnam later [104] tried to explain that he had only considered nomological possibility and that he disagreed on extending the conclusion to a more general metaphysical possibility. But if so, unless one collapses metaphysical possibility to nomological possibility, one cannot get the desired essentialist conclusion. Scientific essentialists identify nomological and metaphysical possibility, but this is far from being accepted.

Notice also that there is an apparently simpler version of the argument that makes use of Leibniz Law as follows. In world w, liquid b has at least one property that a does not have, namely having microstructure XYZ. By Leibniz Law, $b \neq a$ and by the *Necessity of Distinctness* $\Box(b \neq a)$. Therefore, the liquid in w is necessarily distinct from this one. But since w is an arbitrary world, any liquid with a different microstructure from this one is a different liquid. Therefore, this liquid has the same microstructure in every possible world. Putnam sometimes writes as if this was the intended argument, for instance when he says that "supposing that water in the actual world is H_2O, a substance in a world which was not H_2O would not be water". But there are obvious problems with this argument. It implies that every property of a is necessarily exemplified by a. For if b had a property that a lacked, then by Leibniz Law b would be distinct from a. But obviously, the essentialist is looking for a distinction between essential and non-essential properties, he does not want to take every property as essential. He could restrict Leibniz Law to essential properties, but if he did so, then he could not use the principle in the argument, because that would require him first to determine that having that chemical composition is essential to water (which is what he was supposed to infer). By contrast, relativizing Leibniz Law to worlds as $x = y \to (x\,is\,P\,in\,w \to y\,is\,P\,in\,w)$ blocks the original argument precisely because it loses its transworld

[19] I may be providing a reconstruction which is somewhat different from Salmon's original analysis, but in any case I think that it adequately captures his main point.

force. Suppose that $XYZb$ and $\neg XYZa$ in w. Then it follows that $a \neq b$ is true in w and by the *Necessity of Distinctness* $a \neq b$ in every possible world. But we can still have $\neg XYZb$ and $XYZa$ in w', for some possible world w'.

Of course, the argument could be cast in different forms by adopting additional premises. The crucial point is that it will make use sooner or later of a disguised version of the intuitive premiss that two liquids with different composition in (possibly) distinct possible worlds are ipso facto liquids of different kinds. The premiss says that the persistence conditions for a involve its microstructure. If these are thought to be grounded on the essence of things, then this suggests that the hidden intuitive premiss is essentialist. This is not a proof by cases, there are plenty of other versions left. Nevertheless, we can see a pattern in the previous arguments. It seems that all these arguments rest on essentialist intuitions. Therefore, one could argue, as Bird [10] does, that we should forget about Kripke and Putnam's arguments and consider the intuition itself instead.

2.2.5 Metaphysical Arguments

Intuitions about De Re Impossibilities

Thus, the argument is that the intuitions are sufficiently compelling for us to take essentialism to be true. The essentialist will insist on some well chosen intuitions. Socrates could not have been a stomp, a sense data, the Spanish civil war, the musical note C, the redness of your T-shirt, the number π, the surface of my desk, the Nazi ideology, Kubrick's *A Clockwork Orange*, and so on. That is to say, even though Russell could have been a poached egg, it may be that Socrates could not have been an action, an event, an abstract entity, and so on[20].

Certainly, essentialism implies these theses, but that is because it holds the stronger thesis that implies that an object could not have not belonged to whatever kinds it belongs to. However, this has a suspicious consequence: the essentialist has to hold also that an object could not have belonged to some natural kinds that are very similar or close to the ones that it belongs to. For instance, Socrates could not have been a *Homo Neanderthal.* At least in my case, my intuitions regarding this scenario are far less clear. Or consider origin essentialism. Suppose that a person S was born from gametes a and b. By origin essentialism, this is so necessarily. But let us suppose that we could have made a clone b' of b in vitro. Could not S have been born from a and b'? How inconceivable is this situation? Or let us suppose that in the possible world w', the legend of the stone golem from Prague is true. A rabbi in Prague has created a humanoid creature from a piece of clay by a spell in order to defend his city from the outsiders. Given that he used a spell, could he not have made the same golem from a piece of copper or play dough instead of using clay?

Essentialism has often been accused of being too strong for this reason. A universe where an object has a minimally different property from a property it essentially exemplifies seems to be conceivable. But according to standard modal essentialism, called *rigid* essentialism by Roca-Royes [108], this is metaphysically

[20]Surely some of these intuitions are stronger than others, because they concern *ontological categories.* One can reject natural kind essentialism and accept a more general essentialism restricted to ontological categories, see Mackie [79].

impossible. The reason is that an object essentially exemplifies a property P iff it exemplifies P in every possible world. One way to put it is to say that the feature of being an essential property is *unstable*: a slight difference in the property makes it non-essential. Some essentialists have weakened their theses to allow for some stability, these are called *flexible* essentialism by [108]. According to these positions, an object essentially exemplifies a property P iff it exemplifies in each possible world a property P' which is sufficiently similar to P. But these weaker positions fall prey to *Chisholm's paradox*. This argument starts by assuming the intuitive premiss that a given object could have had a slightly different origin or be constituted or composed by exactly the same constituents except for one of them (all its constituents except one of them are the same). The argument applies this intuition once and again to finally arrive at a possible world where that very same object lacks all the properties that were supposed to be essential to it or to one where it has properties that are radically different from the ones it started with. By the transitivity of accessibility that last world will be accessible to the original one, thus accepting the possibility that an object may have lacked all those properties that are allegedly essential to it or the possibility that the object may have had radically different properties to those that it actually has (in fact, this sort of arguments were considered by Salmon to refute the claim that modal necessity is not even S4).

One can formulate instances of Chisholm's paradox for kind essentialism based on intuitions that are in conflict with essentialist ones. Consider the following one. An object may have had slightly different determinates of the same determinable, even when this property is supposed to be essential to the object:

i. For every natural kind K, there are some natural properties $P_1,\ldots,P_m$ which are such that x is a K iff x essentially exemplifies $P_1,\ldots,P_m$. [Natural Kind General Essentialism]

ii. It is possible that for some natural kind K at least one of the properties $P_1,\ldots,P_m$ is a determinate property P_j of a determinable property P. [Pr.]

iii. Let x be a member of K. Then by (i) and (ii), x has the property P_j essentially.

iv. But x may have had a determinate property P_k of determinable P and which is approximately similar to P_j, while still being an instance of K.

v. Therefore, property P_j of x is not essential to x, (i) is false.

For instance, let x be a member of the kind *electron*. The essentialist must say that it is essential to x to have certain electric charge $e(x) = -1.602176565 * 10^{-19}$. The *electric charge* $-1.602176565 * 10^{-19}$ is a determinate property of the determinable property *electric charge*. But x may have had an electric charge that was slightly different from that, say $e'(x) = -1,602176566 * 10^{-19}$. Therefore, the property $e(x)$ is not essential to x. We can obtain a big difference between the determinates P_k and P_j by introducing intermediate steps that involve determinate properties that are more similar to each other, as in the Chisholm cases.

Premiss (i) is general essentialism, it is not in dispute. (iii) and (v) just follow from previous premisses. The essentialist must attack (ii) and/or (iv). Rejecting (ii) is too revisionary for an essentialist. It would mean that no essential property can be a determinate property. But many theoretically interesting properties posited by scientific theories are magnitudes and therefore determinable properties. So we are left to analyse in more detail condition (iv). The essentialist cannot say that it is essential to members of K that they exemplify both a determinate P_i and another different determinate P_k of the same determinable property P, since these exclude each other [21]. The essentialist could answer that the essential property is *having some determinate property of the determinable property P*. If P were a ratio magnitude, we would be saying that it is essential to x that it has a property with some value in the real line $\mathbf{R}$. There are at least two ways to interpret this. On the one hand, the essential property is that of having at least one determinate of P, but it is indeterminate which one it is. On the other hand, the essential property is having a determinate of P, but which one it happens to exemplify may change form one world to another. Regarding the former, if the property is essential to the object, it seems clear that it has to be a specific property P. How could be essential for an object to have some determinate property while this one being of an indeterminate value? Regarding the latter, the problem is that one has to restrict the possible values that the magnitude could have. If not, any value from the real line $\mathbf{R}$, as big or as small as we wished, could be exemplified by the object. But this is surely incompatible with natural laws. Say, the value of P_i could be $10^{10^{\cdots^{10^{10^{\cdots}}}}}$, thus plausibly contradicting the necessity of some natural law. So a more plausible answer for the essentialist would be to say that the essential property is that of having a determinate property with value in some specific interval $[a, b]$ of the real line $\mathbf{R}$. Now we refine the previous argument:

> Let x be a member of K^* and let P_i be the determinate property exemplified by x. Then P_i has a value in $[a, b] \subseteq \mathbf{R}$, for some $a, b \in \mathbf{R}$. Let $0 < \varepsilon \in \mathbf{R}$ be as small as we wish. Then x could have had a property P_k approximately similar to P_i, i.e. with value in $(a - \varepsilon, b + \varepsilon)$–$[a, b]$. Even so, x would still belong to K^*.

So we could ask whether an object with a determinate property with a value slightly outside the inverval, i.e. $a-\varepsilon$ or $b+\varepsilon$, for a quantity $\varepsilon > 0$ as small as we wished, would still belong to the natural kind. If the difference ε was sufficiently small, we would say that the object would still belong to the kind even if its property P_i had no value inside the interval $[a, b]$. To reply, the essentialist should have to enlarge the interval of possible values to $[a - \delta, b + \delta]$, for some big enough $\delta > 0$. But how do we know which value of δ we have to choose as a limit for the values of all the determinate properties that the member of the kind K^* could exemplify? Without essentialist constraints the value δ could simply be an approximation and we could say that members of the kind just tend to have properties with values in $(a - \delta, b + \delta)$. But we are looking for essential properties. If δ is not big enough, we can reformulate the argument. But δ could also be too big, so big that it forced some non-members of K^* to be members of it. To sum up, there are versions of Chisholm's paradox that

[21] That no object can have two determinate properties of the same determinable in a given instant is usually accepted.

are problems for a flexible version of natural kind scientific essentialism, since the intuitions on which these scenarios are based are in conflict with essentialist intuitions.

There is another problem in appealing to modal intuitions, as Bealer [5] clearly saw. This problem is specific to scientific essentialism. Scientific essentialists accept Kripke's main conclusion, that conceivability does not imply possibility. Given the *Necessity of Identity*, there are a posteriori true propositions whose negation is conceivable but metaphysically impossible. For instance, take any proposition expressed by an identity statement which is discovered a posteriori. The main moral of the story is that we should not put too much weight on our ability to conceive certain scenarios. Essentialists usually argue that although it may appear conceivable that natural laws could have been different or that objects could have had different properties, this does not imply that these facts are metaphysically possible. If we want to discover what is or is not possible, we should make empirical research. Now consider the converse, whether inconceivability implies impossibility. Assuming contraposition, this is equivalent to asking whether possibility implies conceivability. Kripke does not deny this, but why would we be justified in inferring from the fact that something is metaphysically possible that we must be able to conceive it? One could ask, if our intuitions regarding what is metaphysically possible are unreliable (given that metaphysical possibility is far more restricted than what we can conceive as possible), why should our inability to conceive an scenario be a reliable source of knowledge of metaphysical impossibilities? How could one justify such an asimmetry between our ability to know better metaphysical impossibilities than possibilities? If this is so, we cannot reliably appeal to our modal intuitions to argue for *Individual Essentialism*, since we could be equally suspicious about arguing from unconceivability to metaphysical impossibility, which is what is required for establishing that thesis. But the conditional is justified by appealing to modal intuitions that support the inconceivability of certain scenarios (say Twin Earth), so there seems to be a problem in the scientific essentialist programme. It takes as evidence the very same kind of intuitions whose epistemic value undermines [5][22].

Substantial Change

As Bird [10] says, even if we assumed Kripke-Putnam essentialist theses, one would still face a further problem. Some essentialists accept the criticism by Fine [39] regarding the equivalence between essential properties and de re necessarily exemplified properties. Recall that, if one accepts Fine's criticism, then Kripke's and Putnam's essentialist theses do not assign essential properties to objects, but only necessarily de re exemplified properties. In particular, they would not show that a certain object essentially belongs to a natural kind or that a macroscopic sample of some liquid has some microstructure essentially. One needs an additional step to secure that these properties are so exemplified in virtue of the nature of their bearers[23].

[22] One could add that there are anti-essentialist views, such as Lewis's 'contextualism', which try to explain away the essentialist intuitions as being the effect of context. According to these theories, these possibilities are excluded in a given context because they are 'too remote' or 'irrelevant'. Thus, there are at least alternative explanations for the essentialist intuitions.

[23] After noticing this Bird argues that according to Fine what we need is that the essentialist claim be true in virtue of the identity of the objects in question. But according to Bird, one

There seems to be a close connection between general and individual essentialism of natural kinds. Suppose that there is a general essence or nature shared by all the instances of a given kind, in other words, that general essentialism is true. This essence fixes or determines those properties that are modally de re exemplified by all the instances of the kind. These latter properties seem to determine the membership conditions for the latter one. So if these properties are necessarily exemplified by any given instance of the kind, it would seem that any such an instance could not cease to belong to the kind without going extinct. It seems then that general essentialism of kinds implies Kripkean individual essentialism of kinds. But if general essentialism of kinds implies individual essentialism, then if the latter one is false the former is in serious trouble. A classical objection to individual essentialism consists in arguing that it is too strong. It seems conceivable that objects could transform in such a way as to change the kind of things they are (see Lowe [**?**]). Let us call this the *Metamorphosis Objection*:

Metamorphosis Objection

1. If General Essentialism of Natural Kinds is true, then Individual (Kripkean) Essentialism of Natural Kinds is true.

2. Objects could stop instantiating those kinds they are actually instances of. In other words, Individual Essentialism is false.

3. Therefore, General Essentialism of Natural Kinds is false.

Second premiss says that the very same object x can belong to a kind K in time t and then cease to belong to K in time t', where $t < t'$. The object survives the change of kind. Literature, fairy tales and myths are full of stories where objects transmute or get transformed in such a way that they change their kinds: frogs that become human princes (or the other way around, unlucky people like Gregory Samsa that become big ugly insects), alchemists that transmute substances into gold, beasts that become statues during the day, wizards that transform themselves into animals at will, and so on. Of course, given what we know about the world, some of these events seem to be nomologically impossible. But without additional argument, nomological possibility is considerably restricted compared to metaphysical possibility. In these tales we are considering worlds were most of the known natural laws fail. Thus, according to individual essentialism it is metaphysically impossible that during the odd days of January my cat gets continuously transformed into a geranium through the day and reverts this transformation by the next morning. But this is clearly conceivable, as examples from literature suggest.

Now, the previous objection considered conceivable scenarios in which objects changed their properties and moved from one kind to another. But it seems that there are actual cases in which this occurs, as Bird argues in [10]. Some

---

can argue that this is indeed the case for the examples of the sort 'necessarily water is H_2O', given that these are 'identity statements'. However, this looks like an equivocy fallacy. By 'identity' Fine refers to the nature or essential properties of the object. By 'identity' Bird refers to the relation of identity that necessarily holds between an object and itself. This is a problem since even the most anti-essentialist can accept the necessity of identity while rejecting every substantive essentialist thesis.

objects seem to change or transmute from one kind to another while still being the same objects. But if an object essentially belongs to a natural kind, then it does so necessarily and therefore these cases are metaphysically impossible. It is of no help to relativize the kind membership to a time instant t (as it is done with PII), because if this property is essential to the object then the object will have it in every instant in which it exists. Let us consider one of his examples:

> "In this example a heavy atomic nucleus changes its kind when it undergoes alpha or beta decay. A nucleus contains 92 protons and 146 neutrons; it emits an alpha particle, and as a result the nucleus now contains 90 protons and 144 neutrons. Or a nucleus with 55 protons and 82 neutrons emits a beta particle (an electron) and becomes a nucleus with 56 protons and 81 neutrons. In such cases the kind of the nucleus is governed by the number of its protons (although the isotopes, governed also by the number of neutrons, are also kinds, subkinds of the elemental kinds). In both cases the decay processes described lead to new kinds (a uranium nucleus yields a thorium nucleus, and a caesium nucleus yields a barium nucleus). And it is natural in each case to regard the nucleus as having retained its identity in the process (.....)" [10]

In this example, an atomic nucleus changes its natural kind, say from being an instance of $Uranium$ to being an instance of $Thorium$ by ceasing to have some properties (the number of its protons changes). Let us grant that such is indeed the case. Nevertheless, Bird says, one could explain these cases away by rejecting natural kind individual essentialism:

If x is a K, then x is essentially a K.

While at the same time keeping natural kind general essentialism:

> For every natural kind K, there are some natural properties $P_1, \ldots, P_m$ which are such that x is a K iff x essentially exemplifies $P_1, \ldots, P_m$.

It may be a contingent property of these objects that they belong to K while it being essential to them that they have these properties once they belong to K. In other words, one could accept that kinds have general essences while at the same time allowing for the possibility that some objects change their kinds. Apparently, individual and general essentialism do not imply each other. It is convenient to highlight that it is not enough to reject *Individual Essentialism* to avoid cases of transmutation, one has to reject its modal version. If x can change its kind, then there is a possible world in which x does not belong to the kind. This is what the previous examples seem to imply.

However, I think that there is a conflict between holding *General Essentialism* and rejecting the principle that the objects should necessarily belong to the natural kind. The most obvious way of formalizing these two theses in S5 leads to inconsistency. To keep matters simple, let us formulate the claim using first-order modal logic. Let $K, P_1 \ldots P_n$ be monadic predicate symbols.

We assume that the logic is first-order modal logic S5 with constant domain. We will formulate *Individual Essentialism* as:

$$\forall x\,(Kx \to \Box Kx)[IE]$$

For *General Essentialism*, some attempts could be:

$$\forall x\,[Kx \leftrightarrow (P_1x \wedge \cdots \wedge P_nx)]$$
$$\forall x\,\Box[Kx \leftrightarrow (P_1x \wedge \cdots \wedge P_nx)]$$
$$\forall x\,[Kx \leftrightarrow \Box P_1x \wedge \cdots \wedge \Box P_nx]$$
$$\forall x\,\Box[Kx \leftrightarrow \Box P_1x \wedge \cdots \wedge \Box P_nx][GE]$$

The first one is a non-starter. It is not enough for the objects to contingently exemplify those properties in some world. The second attempt is the de dicto reading, but is clearly not enough. It just says that in each possible world, the K-s are exactly the ones that exemplify $P_1 \ldots P_n$. So an object x is a K iff it exemplifies all these properties in the actual world, and whenever it exemplifies these properties it will be a K. But x could fail to exemplify some of these properties in some worlds. Now, if it is part of the nature or essence of x to be a $P_1 \ldots P_n$, then if x lacked some of these properties in some worlds it would not be that very same x, for what it is to be such an x involves having those properties. As it is traditionally put, these properties would be just accidental to the object. So the *de dicto* reading does not seem to be appropriate to capture the essentialist commitment. If those properties are part of the essence of the kind, then the object should exemplify them in every possible world in which it exists. Therefore, something like the third formula is required, which gives a de re reading of the thesis. This may still not be close enough, since the biconditional should hold necessarily. Why? Because if having such and such properties is having the essence of the kind, then this correspondence between a kind and its essence should hold in every possible world. It cannot be the case that the essence of the kind changes from one world to another, so we must have something like the last formula. In logical terms, the predicate K is defined in terms of other predicates $P_1 \ldots P_n$. For something to be a K it is necessary and sufficient that it is also necessarily a P_1, and necessarily a P_2 and necessarily ... and necessarily a P_n. These latter predicates denote the properties that are determined by the essence of the kind, and therefore fix its membership conditions.

It can be proven that (GE) implies (IE) in first-order modal predicate logic S4 with constant domain[24]. Let (W, D, I) be a S4 model and g a variable assignment. We informally abbreviate as usual $I(\phi, w) = 1$ as "ϕ is true in w" and $I(\phi, w) = 0$ as "ϕ is false in w". Suppose that $\forall x, \Box[Kx \leftrightarrow \Box P_1x \wedge \cdots \wedge \Box P_nx]$ is true in world w but $\forall x\,(Kx \to \Box Kx)$ is false in w. Then $Kx \to \Box Kx$ is false in w under a variable assignment g^x_a identical to g except that it replaces x by a, for some individual constant a. In other words, Ka is true in w, whereas $\Box Ka$ is false in w. But then there is a possible world $v \in W$ which is accesible

[24]The model will assume the *modal conception* of essence. Fine has convincingly argued against it, but since the properties we will be considering can be taken to be those necessarily exemplified by the instances of natural kinds in virtue of the kind's essences, we can still use it without any problem. We will use a constant domain semantics, an extension to a variable domain semantics should not be too difficult by making the appropriate auxiliary assumptions.

to w, i.e. wRv, and in which Ka is false. But $\Box[Ka \leftrightarrow \Box P_1 a \wedge \cdots \wedge \Box P_n a]$ is true in world w by instantiation. By axiom T, w is accesible to itself, i.e. wRw, therefore $Ka \leftrightarrow \Box P_1 a \wedge \cdots \wedge \Box P_n a$ is true in w too. Therefore $\Box P_1 a \wedge \cdots \wedge \Box P_n a$ is true in w. So $\Box P_i a$ is true in w, for each $i \in 1, \ldots, n$. But since v is accessible to w, $Ka \leftrightarrow \Box P_1 a \wedge \cdots \wedge \Box P_n a$ is true in v too. Therefore, $\Box P_1 a \wedge \cdots \wedge \Box P_n a$ is false in v. And so $\Box P_k a$ is false in v, for at least one $k \in 1, \ldots, n$. If so, then there is a possible world $u_k \in W$ such that u_k is accesible to v and $P_k a$ is false in u_k. By S4 it follows that wRu_k and therefore $P_i a$ is true in u_k, for each i. So in particular $P_k a$ is true in u_k, which is a contradiction. Thus (GE) S4-implies (IE). As were formulated here, rejecting *Individual Essentialism* while keeping *General Essentialism* leads to inconsistency. Therefore, if there are actual cases of transubstantiation, *General Essentialism* is false.

Nevertheless, this may be thought to be unfair to Bird, for he may propose his thesis as a revision of *General Essentialism.* He says that the Aristotelian need only be committed to the weaker thesis that objects will have the essence of a kind just during that time in which they belong to the kind. When an object changes from one kind to another, it will change its essence accordingly. Somewhat before that he makes the following remark:

> "More importantly, the claim that all members of a kind have some property, the essence of that kind, is consistent with the anti-Aristotelian claim that some entities can change their kind. An anti-Aristotelian kind essentialism requires only that when they change kind they lose or acquire the relevant kind essence." [10]

I am not completely sure that I understand how Bird's thesis would go. The idea seems to be that necessarily, for each natural kind K there are some properties $P_1, \ldots, P_n$ which are such that an object x instantiates K during T iff x is P_1 during T and ... and x is P_n during T, where T is some time interval. We exchange the modal operators that force the object to have those properties in every possible world by temporal operators that say that the object has these properties during some given time interval. So if an object ceases to have some of the P_k properties during time interval T, then it will cease to be a K.

However, a version of (GE) according to which objects can change their essence by changing from one kind to another is at odds with the theoretical roles that essences are supposed to play. The whole point of appealing to general essences is to explain the possibility of qualitative change and fix the modal facts that involve those objects, alongside their identity and persistence conditions [77]. Consider identity and persistence conditions. If objects can change their kind then they will change their identity and persistence conditions and, a fortiori, they will not be the same objects. For if x is K in t, and then x is K' in t', the conditions under which x in t is identical to x in t' are different from those under which x in t is identical to itself in t. How then are they supposed to be the very same object? Or consider the classical problem of qualitative change. An object x may be P in t and Q in t', while P and Q being incompatible properties. Therefore, by the *Indiscernibility of Identicals*, x in t is distinct from x in t'. Aristotle's solution to the puzzle was precisely to invoke the distinction between essential and accidental properties in order to restrict the *Indiscernibility of Identicals*. The object x in t has some accidental properties, like being P, that get lost in t'. Nevertheless, x in t is identical to x in t' because it has some properties that are invariant under the qualitative change, namely its essential properties. The phenomenon of qualitative change is explained by assuming that only some of the properties of the object retain its identity through time, while others may change without affecting it. These essential properties are the ones that determine the persistence conditions of the object. So they are the ones the object has in virtue of belonging to a given kind. But if x in t has different essential properties that x in t', x in t is not the same object as x in t'. In such a case not even this version of the *Indiscernibility of Identicals* is preserved and the phenomenon of qualitative change is left unexplained. In sum, one cannot appeal to the essence-accident distinction to solve the puzzles related to qualitative change and identity and persistence conditions, precisely because which are the essential properties of objects changes through time. If these objections to Bird's proposal are sound, then Bird's original metamorphosis or transubstantiation counterexamples are still a threat to *General Essentialism of Natural Kinds*. An essentialist that finds these counterexamples convincing enough has to take a different route. The essentialist can reinterpret these counterexamples as cases in which an object gets extinct and a new one appears. The new problem is how to determine whether the entity is the same one or not. I am unsure whether some general arguments can be given here. Unfortunately, we cannot deal with this topic too.

In any case, we considered several arguments against scientific essentialism. The cost-benefits arguments face the fact that essentialism is at odds with the kinds of special sciences. The semantic arguments seem to assume only the semantic theses of direct reference theory, but all of them smuggle in essentialist assumptions. The metaphysical arguments appeal to intuitions that certain scenarios are impossible. However, they conflict with either instances of Chisholm's paradox or either possible or actual cases of substantial change. Now it is time to take a look at other theories of kinds.

2.3 Cluster Theories

Essentialism is not the only game in town. A popular alternative today is given by the various *Cluster Theories* of natural kinds. Many of these theories appeal to causality or counterfactuals as the ground for the co-occurrence of the properties shared by the members of the kind. Some current examples are Boyd's *Homeostatic Property Clusters Theory of Natural Kinds* [12], the *Simple Causal Theory of Kinds* by Khalidi [64], [65] and the *Theory of Stable Clusters* by Slater [123]. Weaker proposals are Dupré's *Promiscuous Realism* [32]. Generally speaking, cluster theorists tend to emphasize both the epistemological side of the problem of kinds and the naturalistic approach to kinds that pays a close attention to how scientific classifications are made.

2.3.1 Homeostatic Property Clusters

Cluster theories of kinds are weaker than essentialism. Kinds are still considered to be related to a cluster of properties, a set of properties shared by the members of the kind. However, it is not generally assumed by cluster theories that all of the members of the kind share all the properties in the cluster. Plausibly, the first one to propose such an approach to kinds was Bertrand Russell in [116]:

> "If a "natural kind" is defined by means of a number of properties $A_1, A_2, \ldots, A_n$ (not known to be interdependent), we may, for some purposes, consider that an individual which has all these qualities except one is still to be considered a member of the kind -for example, Manx cats are cats in spite of having no tail. Moreover a great many distinctive characters are capable of continuous modification, so that there are borderline cases where we cannot say definitely whether a given character is present or absent. A natural kind is like what in topology is called a neighbourhood, but an intensional, not an extensional, neighbourhood. Cats, for example, are like a star cluster: they are not all in one intensional place, but most of them are crowded together close to an intensional centre. Assuming evolution, there must have been outlying members so aberrant that we should hardly know whether to regard them as part of the cluster or not. This view of natural kinds has the advantage that it needs no modification before incorporation in advanced science." [116]

Russell proposes to consider kinds as clusters of properties that not all the members of the kind need to share. This allows for 'outlying aberrant members' that belong to the boundary of the kind, so to speak. It is interesting to note that Russell makes use here of a spatial metaphor that presents kinds as certain *regions* in a space whose points are properties. This spatial conception of kinds will reappear again in the following section on Conceptualism.

More recently, in several works such as [12], Richard Boyd has proposed a *Homeostatic Property Clusters Theory of Natural Kinds* (HPC). The idea is that the members of a natural kind share some (similar) properties because there are some causal mechanisms that make those properties cluster together (the properties co-occur). These property clusters are maintained by some sort of 'homeostatic equilibrium'. Thus Boyd's theory is a *causal* cluster theory: it explains the co-occurrence of some properties by appealing to causal relations.

Boyd's theory has several interesting consequences. First, natural kinds have a stronger dependence on our inferential practices than it is usually acknowledged. This thesis is called *accommodation* by Boyd, and allows him to introduce certain conventional elements in the HPC theory of kinds without eliminating the objective basis of kinds, which rests on the causal mechanisms. The basic idea is that we in some sense select the application conditions of natural kind terms so that they fit the causal structure of the world. In other words, in order to give some explanations about the objects of a certain domain, we appeal to causal relations among them. These causal relations connect certain properties of objects together. Natural kind terms are chosen to refer to those properties and therefore to allow for making inferences that are grounded on those causal relations. This allows for a partial explanation of the projectability of natural kind terms: the inferences involving these terms are projectable because they are supported by causal relations between the properties clustered. The projectable terms are the ones that track the causal structure of the world, so that the latter ones guarantee the fruitfulness of the inferences made by using the former.

Second, the coherence or unity of the kind is grounded on the causal mechanisms that make the members of the kind similar to each other. In other words, similarity rests on causality, not vice versa. Moreover, these causal mechanisms fulfill an explanatory role. Instead of appealing to the general essence shared by the members of the kind, HPC theorists appeal to the causal mechanisms that cluster the properties together. This is supposed to make HPC closer to naturalism, given that whereas scientists appeal to causal relations to explain why a given object belongs to a certain kind, they do not seem to appeal to essences for such purposes.

Third, the members of a natural kind do not need to share the same or exactly similar properties. The cluster of properties is kept 'in equilibrium' or in 'homeostasis' by the causal mechanisms. This means that the causal mechanisms *tend* to make the properties in the cluster occur. However, it usually happens that not all of the properties in the cluster occur. A fortiori, natural kinds have 'vague boundaries'. The reason is that, if there are no necessary and sufficient conditions that objects have to satisfy in order to belong to the kind, there will be some objects for which it will be indeterminate whether they belong or not to the kind in question, because they will have many (though not all) of the properties in the cluster. This is in contrast to the clear-cut membership conditions assumed by essentialists.

Finally, the causal mechanisms associated with a kind can change over time. Thus, the properties common to the members of the kind can change over time too. This last feature makes the causal approach very different from essentialism, that would not allow a change in the general essence of the kind.

2.3.2 Objections to the Role of Causality

Cluster theories are a very recent approach to natural kinds, thus their advantages and disadvantages are still to be considered. A prima facie advantage of Cluster Theories over Scientific Essentialism consists in their being able to handle the kinds of special sciences. Given that they make weaker commitments than essentialist accounts, causal theories have been invoked as a first approximation to molecular compounds, biological species, mental states, diseases, psychiatric disorders, and so on. Moreover, at first sight it seems that the special sciences ground their explanations on some causal mechanisms (or just causal relations) posited as explanations of several phenomena. If we consider that the choice of the relevant properties strongly depends on the theoretical principles that ground the classification, we get a very tight connection between classifications, the theoretical principles chosen as criteria for making them, the theories in which the classifications occur and the causal mechanisms posited to explain why the principles chosen are the most adequate ones.

Generally speaking, cluster theories attempt to avoid the strong conclusions of essentialism by generalizing the notion of kind and weakening their commitments. This allows them to account for several features found in scientific classifications, such as the vagueness in the membership conditions for kinds or the fact that sometimes there is no set of properties shared by all the members of the kind. However, some of these theories such as (HPC) do so by appealing to causality, which makes them the target of the first sort of objections. It is not clear that causality can be used to explain all cases of classification it purports to. In some scientific domains, demanding the existence of causal mechanisms is too much to ask for. For instance, in the case of physics, it does not seem to be any sort of causal mechanism that makes all the electrons have the same properties [21]. In the context of biology, it has also been argued that successful biological classifications do not necessarily appeal to causal mechanisms. For instance, Ereshefsky and Reydon [37] have argued that the following kinds are not explained by (HPC):

1. Non-Causal Kinds: the members of some kinds belong to them in virtue of being similar without there being any causal mechanisms behind (e.g. kinds of microbes and kinds of stars).

2. Functional Kinds: the membership conditions of some kinds are given in terms of the functions fulfilled by their members (e.g. the kinds *Predator* and *Gene*).

3. Heterostatic Kinds: some kinds have estable differences among their members, which are appealed to in order to classify those members (e.g. the differences between males and females of a common species).

It is not clear that even cladistic taxonomies, which are the most successful ones in biology, fit the HPC picture at all. According to these authors, Boyd's homeostatic theory is still too close to traditional theories of natural kinds in giving some role to similarity relations that do not hold for historical classifications in biology. This feature of contemporary theories of kinds is called 'similarity fetishism' by Magnus [80]. In other cases, such as psychiatric kinds [23], homeostatic causal mechanisms seem to be absent too. According to Cooper, natural kinds are just clusters of co-occurrent properties.

Some authors, like Magnus [80], answer these objections by making further distinctions. For instance, he distinguishes between HPC-types and HPC-tokens. Members of a HPC-token kind share the same causal mechanism, whereas members of a HPC-type kind only have similar causal mechanisms. For example, members of a biological species would form a token kind, since they share the same lineage. But a Martian tiger and an earthly tiger would belong at most to a type kind, since the causal mechanisms that made them appear are at most similar. According to Magnus, chemical compounds are HPC-type kinds. But such a strategy introduces kinds of causal mechanisms based on similarity relations. However, these kinds of mechanisms should themselves be individuated by appealing to causal relations. This leads to a different problem for causal theories put forward by Craver [25]. If kinds are individuated in causal terms, then how are kinds of causal mechanisms themselves individuated? We can say that two causal mechanisms are approximately similar or even duplicates from each other. But is there a further 'higher order' causal mechanism that clusters together these causal mechanisms into one kind? The account seems to face a dangerous infinite regress[25].

[25] There may be other pressing objections. For instance, if there are abstract entities, then there are kinds of abstract entities. If mathematical entities, meanings, possibilia, values and so on turn out to be abstract (or if mental entities turned out to be causally unrelated to physical entities), then the causal accounts will not have a story to tell about how it is that these entities belong to their corresponding kinds. Abstract entities are non-causal by definition. With respect to these entities, similarity scores better.

The second sort of objections come from the lack of details of the proposals. There are many aspects of these theories that remain unexplained. For instance, what does the vagueness of kinds consist in? What is the notion of causality at work? What are the causal mechanisms and how do they specifically relate to the co-occurrence of the properties in the cluster? How is the notion of equilibrium or homeostasis to be understood? Due to some of these problems, several descendants to Boyd's original theory have appeared. Some recent examples of very different theories are the *simple causal theory of kinds* by Khalidi [65] or the *theory of stable clusters* by Slater [123]. These theories differ on whether they put more emphasis on causality than on properties or on whether they do without causality completely.

For instance, consider vagueness. Membership conditions of kinds are supposed to be vague. It is enough for some objects to have some properties for them to belong to the kind, but there may not be a fact regarding *which* of these properties is enough to have. A fortiori, whether some object belongs or not to a given kind may be indeterminate. But what is this vagueness supposed to be? If the vagueness is purely semantic or epistemic, then this only concerns the representational limits of our terms or concepts, such a thesis is still compatible with kinds having clear-cut membership conditions. In contrast, Cluster Theories may be thought to be committed to some sort of ontological indeterminacy in the world. However, the thesis that the world is itself vague is a very controversial position that is still hard to make sense of, as the discussions on the indeterminacy of identity show. Causal theorists have not said much in this regard.

Moreover, it is not always clear what notion of causality these theories are making use of. What theory of causality are based on? What basic principles about causality are accepted? Khalidi [65] makes a more specific suggestion and chooses a conception of causality based on the model of acyclic directed graphs, as the one proposed by the Structural Equations Modelling approach. Kinds are collections of properties (core or derivative properties) that form a network of causal relations. However, he gives no detailed account about how the approach is supposed to work, or about how the usual features of kinds such as their specificity relations or the entities they sort are represented in such a model. Although this is an interesting and plausible proposal for the causal approach, it is still very sketchy.

Moreover, it is not clear how appealing to causality *implies* that kinds will have vague membership conditions. One can conjecture that this may have something to do with the fact that if the occurrence of some set of properties causes the occurrence of another set of properties, then the former one will increase the probability of the latter. Is the vagueness then represented by this probability? Then the previous doubts re-emerge. What conception of probability is at work? Does the probability represent frequencies or propensities based on these causal relations or does it represent our epistemic uncertainties about these causal relations? In both cases, how is this to be understood as a case of ontological indeterminacy? Furthermore, note that, if kinds are clusters of properties, it is not easy to avoid making use of sets in order to represent these clusters. But sets have their identity conditions fixed by Extensionality (two sets are identical iff they have exactly the same members). How can then the membership conditions of such clusters be vague if the identity conditions of the representing sets are clear-cut? One may propose other fancy structures like fuzzy sets to avoid this problem, but then how are these fuzzy sets to be made compatible with the formal framework for causality? In short, several crucial features of these approaches are still very sketchy.

2.3.3 Promiscuous Realism and Practice-Oriented Theories

There are weaker theories than the ones just considered. The most famous proposal is *Promiscuous Realism* by Dupré [32]. According to Dupré, there are many objective similarity relations among objects, and these induce overlapping classifications. In fact, there are so many similarity relations that the choice of the relevant ones is always strongly dependent on our interests. The main objection that is usually made to Dupré's thesis is that he imposes no theoretical constraint on the choice of the relevant similarities. As a consequence, folk or pre-theoretic classifications are on a par with scientific ones (so long as they are fruitful for the purposes for which they were devised). But scientific realism is committed to the thesis that scientific theories are successfully revising our pre-theoretic beliefs about (and in particular our pretheoretic classifications of) the world. If that is the case, then some of our folk classifications do not adequately represent the structure of the world.

Criticisms of the previous approaches to the problem of kinds have lead many philosophers to adopt a more 'practice-oriented' view. These philosophers insist on paying a closer attention to how classifications are made in science. Such a case-by-case analysis reveals important differences between classifications made in distinct disciplines and for different purposes. This leads to what we could call a *natural kind pluralism*. Thus, some of these theories should not be considered 'Cluster Theories' any more. An important example is given by Ereshefsky and Reydon's approach to kinds based on the notion of a *Taxonomic Research Programme* [37], which is based on previous work by Ereshefsky on biological taxonomies [36]. The core idea is that instead of discussing what natural kinds are, we should consider whole classificatory programmes. Every classificatory or taxonomic programme has at least three components:

1. Sorting Principles: these are rules that determine how to sort things into kinds. For example, sort according to similarities induced by common causal mechanisms, sort according to most recent common ancestor, and so on.

2. Motivating Principles: these describe the aims of the programme. For instance, making inferences, giving clear-cut identity conditions for kinds, producing stable kinds, and so on.

3. Classifications: these are the result of applying the sorting principles to a specific domain. For example, a cladistic classification, a morphological classification, and so on.

Besides these descriptive features, they propose the following normative criteria for comparing the different research programmes:

1. Internal Coherence: the sorting principles should promote the motivating principles of the programme.

2. Empirical Testability: the motivating and sorting principles should be empirically testable.

3. Progressiveness: compared to rival programmes, the programme should allow us to make new empirically successful classifications and extend the ones we already had.

These normative constraints can be illustrated, as the authors do, by appealing to the disputes between taxonomic schools in biology[26]. The three main taxonomic schools are *evolutionism*, *pheneticism* and *cladism*. Every taxonomic school accepts the modern synthetic theory of evolution. What distinguishes them is the choice of the species concept that they take as sorting principle and the kind of properties or similarities they consider to be relevant or good evidence for the claim that some entities belong to a common taxon. We will focus only on pheneticism and cladism.

Prima facie, higher taxa are collections of species that are nested in successively more general taxa. *Monophyletic taxa*, also called 'clades', contain an ancestor species and all its descendants. Thus a clade contains all those species that have a species K as their most recent common ancestor (most recent in comparison with the rest of species). As an example, *Reptilia* is not a clade for despite it contains the most recent common ancestor of birds and crocodiles (compared to lizards), it does not contain birds. In contrast, *polyphyletic taxa* contain only some species independently of whether they have a most recent common ancestor or not. From an epistemological point of view, there are several kinds of features, properties or similarities that can be taken as evidence for the corresponding taxa. *Homoplasies* are common features that appeared independently due to natural selection instead of being inherited from a common ancestor. Therefore, if one considers homoplasies as evidence for classification, the resulting taxa will be polyphyletic. In contrast, *homologies* are shared features that are inherited from a common ancestor species. Homologies result from *cladogenetic* speciation processes, in other words, speciation processes due to geographical isolation where an ancestor species is splitted into (at least) two distinct species.

The *cladistic school of taxonomy* only accepts homologies as evidence for monophyletic taxa, those taxa being produced only by cladogenetic processes. The only criterion they accept is that based on the phylogenetic relations among the species. Accordingly, cladists give very clear necessary and sufficient conditions for belonging to a common monophyletic taxon or clade. Two species belong to the same clade iff they share a most recent common ancestor. This definition allows the same species to belong to several nested clades, which can then be represented in the form of a tree called a 'cladogram'.

In contrast, *pheneticists* are suspicious towards the appeals other schools make to theoretical principles in taxonomy. They think that the choice of the relevant similaries or properties used to classify biological entities should not be guided by theory, and therefore that taxonomy should be theoretically neutral for several reasons. For instance, making taxonomic principles and methods independent of theories guarantees that the resulting classifications will be more stable and useful for distinct subdisciplines in biology. Moreover, in practice taxonomists do not have access to the genealogical history of the entities they

[26]I follow the terminology in [36] and [140].

classify and therefore cannot know which of the features they use as evidence are homologies and which ones are not. According to them, evolutionists and cladists commit some sort of circular reasoning. Although they lack knowledge about the phylogenetic relations among the entities, they take some features and reject others as evidence of common ancestry, and then they appeal to common ancestry to explain and justify why they took such features as evidence. This being so, pheneticists put no constraints over the choice of relevant features or over the taxa to be considered as valid. As a result, some of the phenetic taxa are polyphyletic and some of the chosen features are homoplasies. The crucial contribution of the phenetic school was the introduction of a rigorous methodology that consisted in selecting some exemplars, recording the chosen relevant features and plotting them in a multi-dimensional space where the points represented the biological entities (organisms, taxa, and so on) to be classified, the distances represented the degrees of dissimilarity and the dimensiones corresponded to the features[27]. Classifications were the result of applying different statistical techniques over the space that consisted in partitioning the space by clustering the biological entities that were sufficiently similar.

The three schools appeal to some similarities or shared features as evidence for the adequate classification of organisms and species. But this is just the *epistemological* side of the classification. Ontologically speaking, they make very different claims. Pheneticists are ontologically agnostic, they assume that overall similarity should be enough for most taxonomic purposes regardless of the theoretical principles that explain why those features are there in the first place. Cladists explain the sharing of features as a consequence of the phylogenetic relations among the species. As a consequence, the identity conditions for species (and higher taxa) are given only in terms of phylogenetic relations.

Ereshefsky and Reydon consider the Phenetic Research Programme to be a non-progressive programme in biology. Whereas its motivating principles involve developing stable and theoretically neutral classifications that may be used by the distinct biological disciplines, its sorting principles require sorting organisms and species according to overall (morphological, ecological, behavioural, genetic, ...) similarity. According to the authors, this programme is non-progressive because it is not internally coherent. Whereas the programme purports to develop theoretically neutral classifications, the choice of the relevant properties used to sort organisms and species must in the end follow theoretical considerations (namely, those of evolutionary theory). In contrast, the Cladistic Research Programme is considered to be a progressive research programme, because it satisfies all the previous desiderata. Its motivating principles are to develop classifications reflecting the phylogenetic relations between the species and the processes of cladogenesis that caused their splitting (i.e. to reconstruct the tree of life). Its sorting principles sort organisms according to the *cladistic species concept*[28] and species according to their sharing a most recent common ancestor. Since only the use homologous traits (and parsimony principles) as evidence for inferring the phylogenetic relations is allowed, the motivating

[27]In fact, these spatial models were later on transferred to psychology, we will take a look at them in the following section on Conceptualism.

[28]The cladistic concept was introduced by Hennig. Roughly speaking, two organisms belong to the same phylogenetic species iff they belong to a minimal clade that has been formed and maintained by some causal processes (e.g. interbreeding, ecological or developmental forces, ...).

principles and the sorting principles are consistent with each other.

The proposal of Taxonomic Programmes is not committed to a specific theory of natural kinds and it is closer to how science actually works than other approaches. Moreover, it is compatible with a modest taxonomic pluralism, since several taxonomic programmes may satisfy all the previous criteria. While still being compatible to some sort of realism, it imposes enough normative constraints over the adequacy of classifications to rule out too weak approaches, such as promiscuous realism. Nevertheless, I think that the approach is limited in at least two ways. On the one hand, by replacing kinds for whole taxonomic programmes, it sweeps the problem of what natural kinds are under the rug. The only information that we get about kinds is that their membership conditions are described by some programme-specific sorting principles. On the other hand, by focusing on a very coarse conception of taxonomy, it ignores the *structure of classifications*. But this structure is plausibly the part of the programme that represents the specificity relations that hold among the kinds being described, so in the end we get no information concerning these either. Consider the analogy with scientific theories. There is a difference between discussing the theoretical criteria needed to choose between two theories and discussing the structure of the scientific theories themselves. The former may inform us about how theoretical virtues such as simplicity and fruitfulness motivate the choice of a certain theory, but it leaves many components of the theory unexplained. For example, it does not explain whether a theory is to be conceived of as an axiomatic system or as a class of intended models, or how the terms and sentences of such a system relate to the entities and laws being represented by them. Furthermore, it does not explain what the structure of the magnitudes represented by these terms is, or what structure causal relations have.

To sum up, some cluster theories appeal to causal mechanisms or relations to ground the fact that members of a kind have some properties in common. But there are kinds whose members do not fit this pattern. Furthermore, these accounts leave many crucial aspects of the proposal (the notions of vagueness and causality at work) unexplained. One can retreat from causal to weaker theories that only require as a necessary condition that the members of the kind exemplify some similar properties that tend to co-occur. Moreover, it may be that any other additional structure that kinds may have varies according to the 'kind of natural kind' considered. More recent practice-oriented approaches seem to point at some sort of natural kind pluralism by focusing on whole taxonomic programmes, but in exchange, they also leave several features of kinds unaddressed.

2.4 Conceptualism

There is a long tradition in philosophy (going back to the medieval nominalists and to Locke) that identifies kinds, or species and genera, with *concepts*. For convenience, I will subsume any such position under the label *Conceptualism*. Conceptualism is compatible with some of the views just reviewed. For instance, some conceptualists such as Wiggins [139], are also essentialists. Moreover, there are big differences between what concepts are supposed to be according to these authors. For some (e.g. Frege), concepts are abstract entities. For others (say Gärdenfors [42]), concepts are mental entities. This has implications concerning which discipline is the one studying concepts (say logic or psychology). Generally speaking, conceptualists of the latter sort tend to emphasize the epistemological side of the problem of natural kinds.

I will focus mainly on the latter views. The reason is that the psychological theories of concepts appeal to similarity relations between objects to explain concept formation, and the approach I will propose later on appeals to similarity relations as a fundamental component of the structure of kinds. Moreover, such theories make use of formal models of conceptual categorization. This will give us a very different picture of kinds from the ones just discussed and will be an introduction both to the properties of similarity and to the formal models of kinds, which will prove useful in the next chapters.

2.4.1 Conceptual Categorisation

In this section I want to briefly consider the role that similarity plays in conceptual categorization according to psychological theories. There are several reasons for this. First, it counter's Quine's famous view that overall similarity is scientifically useless. Second, it gives a useful background to considering Tversky's objections to the main properties of categorical similarity in Chapter IV. Third, some popular spatial models of attributes, namely Gärdenfors' *conceptual spaces approach* and Carnap's *attribute spaces*, are inspired by the spatial models of conceptual categorization given in psychology. These models have a lot in common with the one I will propose later on in Chapters IV and V and can be interpreted as conceptualist formal models of kinds. Finally, one of the approaches to be examined in this thesis formally, namely aristocratic resemblance nominalism, can be interpreted as being a philosophical analogue to the psychological accounts of categorization based on exemplars or prototypes. Therefore, looking at the latter can give us some clues about the former one.

Quine famously objected to the concept of similarity, which he considered to be equivalent to that of kind, by suggesting that it is scientifically useless:

> "We cannot easily imagine a more familiar or fundamental notion than this, or a notion more ubiquitous in its applications. On this score it is like the notions of logic: like identity, negation, alternation and the rest. And yet, strangely, there is something logically repugnant about it. For we are baffled when we try to relate the general notion of similarity significantly to logical terms. (...) It is a mark of maturity of a branch of science that the notion of similarity or kind finally dissolves, so far as it is relevant to that branch of science." [107]

Despite Quine's remarks, similarity is still a fundamental theoretical notion in cognitive psychology. In the case of psychology, it has to be emphasized that similarity has not been replaced by some 'theoretically more acceptable' notion. The reason is that similarity plays a major empirical role in those theories. First, similarity is an indispensable part of the process of concept formation and categorisation that happens while learning [54]. According to most theories of concepts, concept formation works by first detecting similarities among items, then developing concepts (usually around some exemplars or prototypical instances) and finally categorizing new instances by comparing them in terms of their similarity to either previously seen exemplars or idealized prototypical instances of the concept. Once we have developed the appropriate concepts, the relations that hold among them allow us to make inferences about their instances.

Second, similarity relations ground the analogical inferences by which we transfer our knowledge about the properties and behaviour of some objects to others. The simplest analogical inferences have, generally speaking, the following form:

1. a is P.
2. a is similar to b.
3. Therefore, b is P.

These inferences allow us, under suitable circumstances, to transfer what we know about certain objects to others. Analogies can be found in ordinary, literary and scientific discourse. A famous example is the analogical reasoning from premises such as "the fact that the sun attracts the planets causes the planets to revolve around the sun" and "the nucleus of a hydrogen atom is similar to the sun of the solar system", to the conclusion "the fact that the nucleus attracts the electron causes the electron to revolve around the nucleus". As the example shows, analogical inferences can be based on similarities between relational structures, which allows for inferences involving complex entities. The most known account is Gentner's *structure mapping theory* [44], which was developed for the purposes of explaining how these structural similarities allow for analogical inferences.

Third, psychologists have developed rich formal models of similarity, such as spatial models [96], feature-based models [137], structural models [44], transformational models [55], and so on[29]. These models have been successfully applied to empirical data gathered in experiments to explain many different psychological phenomena[30]. For example, apart from being used to explain conceptual categorisation, these models have been adjusted to deal successfully with some tricky features of similarity, such as its context-dependency.

There are several competing theories of concepts in psychology (see [54]). According to the *classical theory of concepts*, also called the *definitional theory of concepts*, concepts are particular mental entities with definitional structure. Concepts have (or can be identified with) some necessary and sufficient conditions that objects have to satisfy to fall under them. These conditions are stated by an adequate definition of the concept. The extension of the concept is the set of all the objects that fall under the concept, in other words, the set of all the objects that satisfy the definition. Plausibly, these necessary and sufficient conditions can be understood to be intensionally simpler concepts of which the

[29]See [54] for a comparison between these models from the point of view of psychology, and [27] for a comparison between Tversky's and Gärdenfors' approaches, from the point of view of philosophy. We will pay some attention to Tversky's model in Chapter IV.

[30]In fact, based on these spatial models, the psychologist Roger Shepard even proposed that there is a fundamental law, the *Universal Law of Generalization* for learning, that structures the process of conceptual categorization [122].

other ones are composed. In other words, most concepts are *complex* concepts, composed by simpler ones. This latter fact helps to explain the compositionality of the meaning of concepts and the analytic inferences that can be made with them. For instance, the concept PET can be thought of as being composed of the concepts ANIMAL and DOMESTIC. From this it follows that every pet is also an animal. Conceptual categorisation consists in checking whether a given new object satisfies the necessary and sufficient conditions for it to fall under the concept. Such a conception makes relatively easy to check whether a concept has been misapplied, namely when the object thus categorised does not satisfy the definition of the concept, and therefore helps explaining the normative dimension of our use of concepts. The classical theory of concepts is basically the one inherited from the traditional philosophical conceptions of kinds and concepts and is also the account of concepts that has proven most useful in analytic philosophy[31]. This should not be surprising, after all, this theory is based on the traditional philosophical conception of a concept, and as suggested, many philosophers have identified concepts with kinds. For example, when Kant talks in his Logic Lectures about Aristotelian species and genera, he treats them as concepts (whether Kantian concepts are indeed psychological entities or abstract meanings is a different story that we do not need to deal with).

The limitations of the classical theory of concepts are well-known [54]: there are not many sucessful definitions of concepts, inferences among concepts are supposed to be analytic (and the analytic-synthetic distinction was questioned by Quine), some concepts seem to have vague conditions for application, and so on. In fact, as empirical evidence on conceptual categorisation started to accumulate against the classical theory, psychologists suggested different theories for the structure of our concepts. Particularly important are the experiments by E. Rosch [113], which point at certain 'typicality effects'. It seems that we have a tendency to think about some objects of the extension of a concept as being more typical instances of the concept than others. Thus, a particular chair is a more typical instance of the concept FURNITURE than a stove is and a particular robin is a more typical instance of the concept BIRD than a penguin is. These typicality effects are correlated with other psychological phenomena. For instance, subjects can recall more quickly more typical instances of a concept than less typical ones. These discoveries prompted the emergence of two new theories of concepts. According to *prototype theories of concepts*, we store in our memory a list of features (the prototype) typically found among some similar objects. Prototypes are some sort of idealized mental summaries of the properties typically had by the objects falling under the concept (prototypes are usually explained in terms of probability). According to *exemplar theories of concepts*, what we store are copies of some particular items (exemplars) that we have previously encountered during our concept learning process. In fact, the two main psychological models of similarity, namely the attribute model by Tversky [137] and the spatial model by Nosofsky [96], were developed as an attempt to give an explanation of these typicality effects.

According to these theories, conceptual categorisation consists in mapping each newly encountered item to the corresponding category, by comparing how similar the item is either to the prototype or to the stored 'exemplars' of the

[31] In fact, one can think of the concept lattice of a context to be introduced in Chapter III as giving a model of the classical view of concepts. This seems to be what Wille was thinking when he introduced them.

kind in question. That is to say, concepts are structured around a prototypical description or a collection of typical exemplars, and new items fall under the concept depending on the degree to which they satisfy the description or their degree of similarity with respect to the exemplars. Despite being empirically more plausible, these theories have their problems too. For instance, some concepts (such as mathematical or logical concepts) which clearly have definitional structure appear to have typicality effects too. Moreover, these theories are allegedly unable to explain how it is that some of our concepts have definitional structure and they cannot explain compositionality. Paradigms, exemplars or prototypes are supposed to be ill-suited for this tasks. This is because the paradigm of a given concept does not seem to be a function of the paradigms of the concepts from which it is composed[32].

Despite being historically influential as an approach to the problem of natural kinds, the conceptualist position is not very popular nowadays. Nevertheless, the approach has its advantages. If properly developed, it could be used to deal with the epistemological side of the problem of natural kinds, namely the problem of induction. Despite being more popular in psychology than in philosophy, spatial formal models of concepts such as the ones just mentioned have been recently applied to semantics and formal epistemology. An important example is the theory of conceptual spaces by Gärdenfors [42] and [43], which was originally proposed as a way to distinguish natural from non-natural concepts in order to deal with Goodman's infamous grue puzzle.

2.4.2 Conceptual Spaces

As was said, the conceptualist approach is rich in formal models and some of them can be found in the philosophical literature. One is Carnap's *attribute spaces*, introduced in his last system of inductive logic [19]. Another one is Gärdenfors more popular account of *conceptual spaces* [42], [43]. The conceptual spaces approach is starting to attract more attention among philosophers. Recently, they have been applied to several philosophical problems, such as giving a foundation for cognitive semantics [43], vagueness or the structure of scientific theories (see [146]). A detailed comparison between these two approaches can be found in Sznajder's work [131] and [132]. We follow her in order to briefly compare the two.

Carnap's attribute spaces were spatial models introduced to give a semantics for the language of his last system of inductive logic. Given a classical language consisting of several families of monadic predicates, these were to be interpreted by regions of a space whose points would be maximally specific attributes. Each family of predicates would partition a previously given domain of individuals, in such a way that each individual would be assigned a point in the space each of whose coordinates would be a specific attribute falling under the predicate of one family. One can think about the points of the dimension of a space (e.g. the colour space) as (universal) determinate attributes (e.g. a specific hue of red), about the regions as determinable attributes (e.g. redness), about the dimensions of space as more general determinable attributes that are independent from each other (e.g. colour or size) and about the distances between points

[32]There are other theories of concepts like the theory-theory, weaker versions of the classical theory or hybrid views, but we need not consider them here (see [54] for an overview).

as degrees of dissimilarity between the specific attributes. Moreover, Carnap also considered the sizes of the regions as an additional parameter. Although he was clearly inspired by the corresponding psychological models, the aim of introducing the spaces was normative, it was part of the project of explicating the notion of confirmation in his logic of induction.

By contrast, Gärdenfors introduced in [42] his conceptual spaces with the descriptive aim of providing a representation of a level of cognition between the symbolic and the connectionist ones. The spaces introduced by Gärdenfors are structurally similar to Carnap's. They are called 'domains'. Domains can be perceptual, as for example the domains of colour, size, shape, taste, ..., but they may also be theoretical, as the domains of length, mass, temperature, and so on. The most studied example of such a domain is the domain of colour. It can be represented as a 3-dimensional solid (a spindle) whose dimensions represent the hue, saturation and luminance (or brightness) of colours, and whose distance represents the degrees of dissimilarity among colours. Whereas luminance increases with the vertical axis of the space (the white colour being at the top and the black colour being at the bottom), saturation decreases the closer the points of the space are to the vertical axis. Hues can be roughly identified by looking at the colour circle formed by the points in the surface of the spindle that are at exactly the same distance from the middle.

Due to the descriptive aim of the project, the choice of the dimensions for these domains are informed by empirical psychological research (e.g. hue or saturation for the colour domain). Generally speaking, a property (e.g. red) is represented as a region in (one of the dimensions of) the space. A concept (e.g. apple) can be represented as a collection of regions from possibly different spaces. The distinction between properties and concepts here roughly corresponds to our distinction between attributes and kinds. Thus, one may take Gärdenfors' conceptual spaces approach as a *conceptualist model of kinds*.

One of the first aims of conceptual spaces was to suggest as an empirical hypothesis that natural properties were convex regions in a conceptual space. This was the P criterion, which was coupled with a different hypothesis regarding more general natural concepts:

P Criterion A *natural property* is a convex region in a conceptual space.

C Criterion A *natural concept* is represented as a set of regions in a number of domains alongside an assignment of salience weights to the domains, and information regarding how the regions from different domains are correlated to each other.

A convex region is a region that contains every point that is between any two points already in the region. In the case of Euclidean space, convex regions have the nice shapes we usually associated with the notion of convexity (say round-like shapes). Thus, a natural attribute such as *Green* would be a convex region, whereas a natural concept (i.e. a natural kind) such as *Apple* would be a collection of convex regions from different domains (e.g. *Green*, *Sweet*, *Round*, and so on).

Gärdenfors proposed the convexity requirement as an empirical hypothesis to be tested. Although the empirical evidence suggests that the requirement is necessary, it is not sufficient[30], since there are still many convex regions that do not correspond to natural properties in a given space (e.g. in the colour space).

However, the choice of such regions is made easier by the fact that Gärdenfors combines the spatial model with the Roschian approach to prototypes. Each natural property is structured around a prototypical instance, following the prototype and exemplar models of categorization. To explain conceptual categorization, first several points of the space are chosen as representing paradigmatic objects. Then the space is divided into regions induced by the distances from the points of the space to each of the paradigms (what is known as a 'Voronoi tessellation'). The resulting division represents a classification induced by the prototypical objects and can be taken to be a first approximation to the natural attributes, since with a suitably chosen metric (say the Euclidean one) the resulting Voronoi cells are convex.

Gärdenfors [42] does not seem to select a specific mathematical framework for explaining the notion of a conceptual space. The reason is that the appropriate similarity relation may have different formal properties from one domain to another, depending of the respects of comparison (i.e. the dimensions) involved. Some parts of his work suggest that it must be at least a space induced by a basic 'betweenness' relation, so that convexity can be formulated for it. When quantitative domains are considered, it seems that there is also a metric distance at play. Other more recent works like [43] seem to appeal to either vector or normed spaces. Moreover, the author makes heavy use of different coordinate systems (like the polar one) for different purposes (e.g. explaining prepositions). The best guess would be that he is proposing a generalization of the Euclidean vector space $\mathbf{R}^n$ (possibly an L^p space). Formally the model is similar to those of other spatial approaches in psychology, so for most purposes, a conceptual space must satisfy at least the following requirements of a metric space:

Definition 1. *Let S be a set and $d\colon S^2 \to \boldsymbol{R}$ a real valued function. Then (S, d) is a* metric space *iff $\forall x, y, z \in S$:*

- *i* $d(x, y) \geq 0$. *[Positiveness]*
- *ii* $d(x, y) = 0 \Leftrightarrow x = y$. *[Indiscernibility]*
- *iii* $d(x, y) = d(y, x)$. *[Symmetry]*
- *iv* $d(x, z) \leq d(x, y) + d(y, z)$. *[Triangle Inequality]*

A metric allows for the definition of comparative similarities, such as:

$$T(x, y, z) := d(y, z) \leq d(x, z)$$

By interpreting the distance as dissimilarity, the condition says that y is more similar to z than x is iff y and z are closer to each other than x and z. Thus the distance function induces a comparative triadic similarity relation between points. Furthermore, the distance function provides a richer notion of similarity, since the similarities themselves can be added to each other. Since the similarity or dissimilarity between two objects x and y is represented by the distance function $d(x, y)$, we can understand the distance between two objects as their *degree of dissimilarity*. Thus, (ii) and (iii) are requirements analogous to reflexivity and symmetry. For instance, (ii) says that two objects are maximally similar iff they are identical, which is again a version of the Identity of Indiscernibles. In contrast, (iii) says that the dissimilarity of x to y equals the dissimilarity of y to x. Since the metric distance is a function and by (i)-(ii), any two different objects will be dissimilar to each other *up to some degree of dissimilarity*. In other words, any two objects are comparable by dissimilarity if one chooses a coarse enough degree of dissimilarity[33].

By taking products of spaces one can get spaces with several 'dimensions', which represent the respects of comparison. A point of each space can be thought to represent a specific determinate attribute and an object is represented as an n-tuple of points (each from one coordinate), in other words, as a bundle of attributes. That is most clear when the dimensions are 'unrelated' to each other, as when a dimension corresponds to size while the other corresponds to colour. An object a corresponds to a pair (x, y) where x is its specific size and y is its specific colour.

In order to introduce the notion of convexity, one needs to assume a betweenness relation among points. If one already has a metric, a betweenness relation can be defined as follows:

$$B(x, y, z) := d(x, z) = d(x, y) + d(y, z)$$

In words, point y is between points x and z iff the distance from x to z is the sum of the distances between x and y and y and z. In terms of similarity, an object y is between objects x and z iff the degree of (di)similarity between x and z can be obtained by adding the degree of (dis)similarity between x and y to the degree of (di)similarity between y and z.

The previous notion of betweenness is standardly known as the *geodesic betweenness*. This is not the only one, there are others. For example, in the euclidean space we also have the *affine betweenness*, which is the one usually considered:

$$B(x, y, z) := \exists t \in [0, 1]\ y = (1 - t)x + tz$$

In any case, from a betweenness relation one can define the notion of convexity standardly as follows:

Definition 2. *Let (S, d, B) be a metric space with a betweenness relation $B(x, y, z)$ and $A \subseteq S$. Then A is* convex *iff for all $x, y \in A$, for all $z \in S$, if $B(x, z, y)$ then $z \in A$.*

It is easy to see that $\varnothing$ and S are convex, that arbitrary intersections of convex sets are convex, and that the union of a chain of (upwards nested) convex sets is convex. However, the union of two (non comparable) convex sets need not be convex, and the complement of a convex sets need not be convex either. For counterexamples, one can think about the usual convexity in the Euclidean plane (for the union, take two disjoint convex sets). Usually the singletons $\{x\}$, which represent the points, are also convex. By excluding certain collections as non-properties by making use of the betweenness relation based on the (dis)similarity, the geometric structure has a way to distinguish natural from non-natural properties.

Moreover, the spatial representation allows for interpreting several operations among concepts as spatial relations. For instance, two concepts can be

[33]We will consider some of these properties in more detail in Chapter IV, where we deal with the properties in similarity.

said to be coinstantiated by a common object iff the corresponding regions overlap in the point representing that object. One concept is said to imply another concept iff the region corresponding to the former one is included into the region representing the latter. This allows for representing basic syllogistic inferences. In principle, similar inferences can also work for complex concepts. If concept K is a subconcept of concept K', it is plausible to say that whatever concept K' implies is also implied by concept K. Moreover, the more geometric structure the space has, the richer the relations between concepts can be.

Conceptual spaces are smoothly combined with the prototype or exemplar models of categorization. In order to do this, one selects a certain subset of the space, to be thought of as the set of all exemplars or prototypical instances. Then one defines the Voronoi tessellation induced by this set as follows:

Definition 3. *Let (S, d) be a metric space and $A \subseteq S$. Then the* Voronoi tessellation *induced by A is defined as the family of the sets of the form $V_p := \{x \in S \mid \forall q \in A - p\ d(x, p) \leq d(x, q)\}$, for each $p \in A$.*

The Voronoi tessellation is a covering of the space by regions obtained by measuring the distance from the different points in the space to the prototypical ones. Each class is fixed by a prototypical item and contains all those items that are more (or equally) similar to it than to the other prototypes. Since an item can be at equal distance to several prototypes, it can belong to several such classes. The set of all the points that are at equal distance from several prototypes is the *boundary* of the tessellation. In other words, the tessellation represents the categorization process: each object x is compared by similarity to each exemplar p, if it is sufficiently similar to p it is included under the corresponding concept. Under additional constraints, these classes turn out to be convex for some of the distance functions.

Apart from the convexity requirement, Douven and Gärdenfors [30] have argued for several 'design principles' that a system of natural concepts should satisfy. Such principles are chosen by analogy with an optimal conceptual scheme that would be developed to allow for a system to make correct, sufficiently fine-grained and successful classifications under limited constraints (e.g. limited memory or perceptual capacities), we quote:

1. Parsimony: The conceptual structure should not overload the system's memory.
2. Informativeness: The concepts should be informative, meaning that they should jointly offer good and roughly equal coverage of the domain of classification cases.
3. Representation: The conceptual structure should be such that it allows the system to choose for each concept a prototype that is a good representative of all items falling under the concept.
4. Contrast: The conceptual structure should be such that prototypes of different concepts can be so chosen that they are easy to tell apart.
5. Learnability: The conceptual structure should be learnable, ideally from a small number of instances.

Parsimony requires that the overall system of concepts does not have too many concepts, in order not to overload the memory. However, it needs to have enough concepts for it to allow for informative inferences involving them. Moreover, prototypical instances are chosen in such a way that objects falling under a given concept (and thus similar to the corresponding prototypes) are maximally similar to each other, whereas prototypes from different concepts are maximally dissimilar to each other. By linking concepts to prototypes, the concept learning process is easier, since one can develop the concepts from those prototypical instances and then recall the concepts again by just remembering the prototypical instances. In other words, an optimal system of concepts involves a small amount of informative concepts that correspond to a few prototypical instances for each, in such a way that similarities among the instances of a common concept and disimilarities among prototypes of different concepts are maximized. Such a system makes much easier the tasks of remembering concepts (through their prototypes), making inferences, categorising each new instance under the corresponding concept, choosing the appropriate prototype and learning the corresponding concepts. The aforementioned conditions have empirical support too (see Douven [31] for some of these empirical results concerning the representativeness and contrastness of prototypes in the colour space).

These models hint at how a conceptualist theory of natural kinds could make use of them to explain what kinds are. According to such a theory, a kind is a concept. Thus, a kind is composed of other simpler concepts, which are represented as regions in different conceptual spaces. Inferences among concepts can be given in terms of inferences among the simpler concepts from which they are composed, which in turn can be given in terms of spatial relations among them (inclusion, overlapping and so on). Moreover, concepts are developed from prototypical instances, and categorisation works by comparing new instances to the prototypical instances stored in memory. Such an approach can be used to deal with the epistemological dimension of kinds, by appealing to concept formation and categorisation.

2.5 Conclusion of Chapter II

It is time to make a diagnosis of the state of the art. There seem to be some assumptions about kinds that are common to all the previously listed theories, namely:

i Kinds are related to some entities, the *objects*, which are members of the kind.

ii Kinds are related to some entities, the *properties* or *attributes* (possibly relations), which are shared by the members of the kind.

iii Properties shared by the objects make the members of a given kind *similar* enough to each other.

iv Kinds have (possibly vague) *membership conditions.*

v *Membership conditions are related to the properties shared* by the members of the kind, although they may be fixed by causal relations, essences or other processes.

vi Kinds are related by *specificity relations*. In other words, some kinds are more specific than others.

The theories disagree on the rest of the assumptions regarding what kinds are supposed to be. Whereas essentialists appeal to general essences as the ground of the membership conditions for kinds, causal theories appeal to causal mechanisms for the same purpose. In contrast, conceptualists appeal to epistemic subjects that select and store in memory certain objects in order to develop the corresponding concepts.

The suspicion I have is that the discussions that involve essentialism and causal theories are too general. What I find missing in the literature are specific principles concerning kinds. It is true that each theory proposes some principles that hold about kinds. In this sense, the richest account is that of essentialism. However, natural kind essentialists have not gone very far in extracting the specific consequences of their proposals. So far, the most detailed view I have found is that by Brian Ellis, which was mentioned in the section of essentialism. The most specific principle that Ellis discusses is that of the *hierarchy condition*. I think that this a very good example of a principle about kinds that may prove useful, given that it is informative with respect to the way kinds are ordered by specificity relations. Consider by analogy a principle that, according to many philosophers, holds for determinate and determinable properties: no object can have, at the same time, two distinct determinates of the same determinable property. Whether it really holds or not, it does give us some more information about the determinate-determinable distinction. The debate about determinate-determinable properties can be structured around such specific principles by testing them against specific counterexamples. This criticism extends to cluster theories too. Cluster theories attempt to avoid the strong conclusions of essentialism by generalizing the notion of kind and weakening their commitments. This allows them to account, for instance, for the alleged vagueness to be found in the membership conditions for kinds. In exchange, however, some of these theories such as (HPC) lack suggestions concerning specific principles about kinds. Most of them leave the conception of causality at work unspecified and they do not say much about what this vagueness of kinds is supposed to be.

In contrast, conceptualist approaches do discuss several specific structural principles involving concepts, such as the convexity constraint, the more general assumptions made by the spatial models or the design principles we just mentioned. These approaches are very rich in formal models that can be interpreted as models of kinds. Unfortunately, the connections between the conceptualist picture of kinds and the topics discussed in the mainstream literature on natural kinds have not been explored. This may be related to the fact that many authors dealing with the topic of natural kinds identify themselves as realists and therefore would not want to take kinds to be 'simply' concepts, understood as mental entities. Taking kinds to be concepts seems to put at risk the most fundamental idea about natural kinds: that kinds are *natural*, in the sense of being *objective*. If kinds are concepts, then they are dependent both for their identity and for their existence on minds. If that is so then this suggests that nature is not already carved at its joints. This does not follow though. For instance, if conceptual universalists are right, then our most fundamental natural concepts are shared by different cultures. If that is the case then, the hypothesis that these concepts have been abstracted from natural kinds in the world looks promising. If that is the case, then kinds and concepts are distinct, and the former allow for (or even cause) the development of the latter. But if so, why would we identify kinds with concepts in the first place? Is not it better to just focus on kinds instead of looking at the mental entities that we use to represent them? In any case, the question arises whether one could take inspiration from the conceptualist approach to kinds without committing oneself to the stronger thesis that kinds are concepts (mental entities).

As I was suggesting, the first two theories reviewed leave some basic questions regarding the *structure* of natural kinds, such as the following, unaddressed:

i How are kinds related to each other by specificity relations? What specific principles do the species-genus relations between kinds satisfy? Do kinds overlap with each other?

ii Can kinds combine or compose with others resulting into new kinds?

iii How are similarities among the members of the kind related to the properties shared by them? If similarities are grounded on sharing properties, how does this work? If it is the other way around, how are properties obtained from similarities?

iv How are the specificity relations between kinds related to the properties shared by their members?

v Do any of the principles of the traditional Aristotelian conception of species and genera, such as the syllogistic relations between species, the definitions in terms of genera and specific difference, the existence of a summum genus, and so on hold for kinds?

Many of these basic questions remain unanswered even after choosing one of the theories discussed (the exception are some of the conceptualist models, which do put emphasis on the formal principles). In other words, what I think is missing in the current literature is *an account of the formal structure of kinds*. It is useful to compare the situation to that of modality. Philosophers discuss about the nature of possible worlds. Some philosophers defend modal realism, others different sorts of ersatzism and still others combinatorialism. However, these proposals have to be coupled with some sort of theory about the 'structure of metaphysical modality', so to speak. The sort of principles these theories have to preserve are the ones given by predicate modal logics, which are formal models that describe the consequences that the different properties of accessibility relations have for our valid reasoning about modality. Unless one couples these metaphysical theories with the formal principles described by these formal models of modality, talk about possible worlds is severely impoverished. For instance, think about principles such as "if it is necessary that p, then it is the case that p", the de dicto-de re distinction or the more controversial Barcan formulae. The conceptual machinery of possible worlds has proven to be extremely useful for metaphysical purposes in part because it was based on a previous formal study of the mentioned principles. Some of the metaphysical theories might seem to be more compatible with a literal interpretation of these principles than others. The other theories may have to describe additional construction steps (say, construct possible worlds as maximal consistent sets of propositions) in order to satisfy these principles. In any case, our understanding of modality is enriched by considering these basic structural principles. Something similar seems to me to be happening in the case of kinds.

Some philosophers, mostly philosophers of science (say [64], [80], [37]) seem to share this diagnosis: the discussion on kinds is too general. However, as we saw, they draw a very different conclusion. They think that it is best to examine in detail first the classifications found in different scientific disciplines in order to arrive at a more adequate conception of kinds. In fact, this concern has lead to the development of most cluster theories and more generally to the practice-oriented approaches. Such a project is in principle compatible with the formal one I want to pursue in this thesis, and I am sure that it will be very fruitful for our conceptions of kinds. Nevertheless, the two projects can be developed independently. Let me reason again by analogy. Although it is certainly fruitful to consider what principles concerning composition, causality or modality assume the specific scientific theories, formal models of composition (e.g. mereologies), causality (e.g. structural equations models) or modality (e.g. modal logics) can be developed (and have been developed) independently of these theories. Analogously, although it is it is fruitful to consider what principles concerning kinds assume the specific scientific classifications, formal models of classification or of kinds can be developed (and have been suggested) independently of these.

There is another feature of essentialist theories and causal theories that I find to be somewhat suspicious. Both theories appeal to philosophically loaded concepts, such as essence or causality, to explain classification. However, at least from a cognitive point of view, classifying objects is simpler than detecting causal relations among events or discovering the real nature of objects. In this sense, I think that they both put the cart before the horse. It is true that the fact that classifying is epistemologically prior to discovering essences or causal relations does not imply anything about the ontological priority of kinds over essences or causal relations. Moreover, essentialists will complain that a good classification should necessarily reflect the real natures of the objects to be classified. But still, there are some features of classifications which are independent of any talk about essences and that could give us some clues regarding the structure of kinds. It may be that a substantive and full realist analysis of the concept of natural kind ultimately requires appealing to essences, causal relations, counterfactual stability, natural laws, or other loaded concepts. But it may also be the case that we can get closer to a fruitful explanation of what a natural kind is that does not require us to answer first all these philosophical problems. I propose to see how far we can get on our understanding of what kinds are without appealing to essences or causal relations.

As we saw, according to the minimal conception of natural kinds by Whewell and Hacking, natural kinds are at most ”classes denoted by a common name about which there is the possibility of general, intelligible and consistent, and probably true assertions”. According to Hacking, this is the most informative and uncontroversial explanation we can give of kinds. I think that this minimal analysis can be greatly improved. The task of the following chapters is to explore a *Minimal Conception of Natural Kinds*, which consists of some formal and structural constraints that have been attributed to kinds, namely the ones listed above. In this regard, I think that both the quinean approach and the more traditional conception of natural kinds have been dismissed too quickly.

They contain many interesting features that are still unexplored. The traditional conception of kinds highlights two different aspects of them. On the one hand, kinds are externally structured by being ordered in specificity relations to each other (species-genus relations). On the other hand, kinds are internally structured, their members (extension) being sufficiently similar to each other by having some natural properties in common (intension). This minimal core seems to be accepted by different theories of kinds, although they may differ regarding the facts that ground the intensions of the kinds (essences or causal mechanisms), the nature of the elements in the intension (namely, monadic properties or relations), what the specificity relations are, and so on. This suggests that a proper minimal analysis of kinds should start by considering these two aspects. By focusing on these basic features that are shared by different conceptions of kinds, we may be able to give an informative answer to the question regarding what natural kinds are. My proposal will not be metaphysically neutral though. Since I will make use of natural attributes to explain what natural kinds are, I have to take a stance too on what these natural attributes are. This inevitably leads us to the problem of universals. Quine's approach to natural properties suggests contemporary resemblance nominalism, to which I will commit myself in Chapter IV. But to properly discuss these issues it is expedient to consider first the external structure of kinds. Let us turn then to this topic in Chapter III.

Chapter 3

Hierarchical Structure of Natural Kinds

Intension and extension of a concept have an inverse relation to each other. The more a concept contains under it, the less it contains in it.

Lectures on Logic
KANT

As we saw in the previous chapter, in the current literature one can find many different theories that purport to explain what natural kinds are. Some philosophers like Ian Hacking [52] argue that this embarrassment of riches supports a sceptical stance towards the very notion of a natural kind. In this chapter a different strategy is suggested. Instead of defending a particular theory of kinds, the main purpose of this chapter is to use a formal model based on the mathematical theory known as *lattice theory* to explore the features of kinds that result from the order structure of the specificity relations that hold among them. More specifically, it will be argued that the external structure of kinds is plausibly that of a complete lattice and that such structure can be seen to be induced directly by some minimal assumptions concerning the relations that hold among the members of the kinds and the properties these share. As an application of this model, it is shown that it is enough to induce an Aristotelian conception of definitions of kinds in terms of genera and specific differences. According to this minimal conception, kinds are two-dimensional entities ordered by specificity relations. They are constituted by a collection of objects and a collection of attributes related to each other according to the duality between extension and intension (a principle also known as Kant's Law). This results in a realist model of kinds.

The structure of the chapter is as follows. In section 3.2, it is considered whether any non-trivial principles about the specificity relations between kinds hold. The *Hierarchy* thesis, which says that kinds are arranged forming a tree, is reviewed alongside R. Thomason's model of kinds [135]. Thomason's model makes use of complete lattices, but no further reasons are given for this choice. In order to find out whether there is any reason in favour of using complete lattices as plausible models for kinds, the contemporary approaches to Aristotelian Syllogistic Logic are considered. It turns out that Martin in [82], extending previous work by Corcoran [24] and Smiley [125], gave a complete semantics for syllogistic logic by using the class of bottomed meet-semilattices. Under the assumption that kinds behave as denotations of terms in syllogistic logic and are in the relations described by the Square of Opposition, the model using complete lattices turns out to be a plausible model for kinds. These two approaches suggest that lattices provide an adequate model for kinds, but they do not say much more concerning the relation between kinds, their members and the attributes shared by these. In section 3.3, a formal model for the Minimal Conception of kinds is provided by making use of the *Theory of Concept Lattices* of Rudolf Wille [40] (the use of that theory for modelling kind-like entities was already proposed by Mormann [87]). Concept lattices are complete lattices, and therefore satisfy the requirements put both by Thomason's model and Corcoran's Syllogistic Logic. The rest of the section is devoted to arguing for the material adequacy of the model regarding the Minimal Conception. In other words, its ontological assumptions will be shown to be those (and only those) of the Minimal Conception. In section 3.4 the main application of the chapter is offered. It is shown that the Minimal Conception commits us to a renewed version of Aristotelian definitions of kinds in terms of genera and specific differences. For this purpose, two new operations of subtraction are introduced, which correspond to the specific difference of one kind with respect to another. The operations differ on whether the part being subtracted is the extension or the intension of the kind. The two principles that relate the specificity relations between kinds with the aforementioned operations are discussed in detail. Furthermore, each of these operations induces an internal or term negation that behaves non-classically. The interaction between the negations and the specificity relations are illustrated by making use of the Hexagon of Inner and Outer Negations and a comparison is made between the approach in this chapter and the traditional Aristotelian picture. Finally, in section 3.5 a sketch is given of a modal extension of the model. The chapter ends with some general remarks.

3.1 A Minimal Conception of Natural Kinds

3.1.1 A Minimal Conception of Kinds and Kant's Law

Aristotle's metaphysics of kinds or, as he called them, 'secondary substances', was closely mirrored by his syllogistic logic and his theory of definitions. Kinds are universal substances related to each other as species and genera. Whenever one kind is more specific than another, we say that the former is a *species* of the latter, and that the latter is a *genus* of the former. The different relations between these substances can be tracked by looking at the logical relations between the corresponding terms, which behave according to the principles of syllogistic logic (summarized in the Square of Opposition). This conception of kinds as universal substances was linked to essentialism. For Aristotle, knowledge of kinds is knowledge of their general essences. These essences are found by giving real definitions of the kinds, which state their membership conditions (some necessary and sufficient conditions to be satisfied in order to belong to the kind) in terms of the essential attributes shared by the objects of the kind. The essence of a kind can be defined by giving one of its genus and the specific difference with respect to it. In other words, in order to define a kind, one gives first a more general kind of which it is a species and those essential attributes or properties that distinguish the members of that kind from others. Therefore, already in the Aristotelian picture we find three ontological ingredients. We have the objects or primary substances, which are the entities to be classified into kinds. We have the essential attributes or properties shared by these objects, which are the ones to be described by the definitions of the corresponding kinds. Finally, we have the kinds themselves, the species and genera, into which objects are classified according to their properties. The Aristotelian picture gives a taxonomic methodology: in order to sort a domain of objects into their corresponding kinds, one has to find first the corresponding definitions of kinds, which requires finding the closest genera and the corresponding specific differences.

The metaphysical components of the picture were inherited by medieval thinkers and were subjected to heavy criticism by modern philosophers, who rejected most of them. Nevertheless, its formal features remained quite stable as part of syllogistic logic, as it can be seen by looking at Kant's logical works. Despite the fact that Kant talks in terms of *concepts* and not of kinds, he treats the former simply as the Aristotelian species and genera. In fact, for Kant every concept has two aspects, an *extension* or *sphere*, and an *intension* or *content*. Whereas the extension contains all the entities that fall under the concept (those entities to which the concept adequately applies), the latter contains those characteristics or features associated with the entities of the extension. We have here again a correspondence between the objects, their properties and the kinds (now, the concepts) to which the former belong in virtue of sharing the latter. For Kant, as for Aristotle, the logical relations that hold among concepts are described by syllogistic logic.

Among the principles of logic that Kant discusses in his Jäsche Lectures on Logic [63], there is one that deserves special mention, because it concerns a fundamental relationship that holds between the extension and intension of a concept:

> Every concept, as a *partial concept*, is contained *in* the presentation of things; as a *ground of cognition*, i.e.*as a characteristic*, it has these things contained *under it*. In the former regard, every concept has an *intension* [content]; in the latter, it has an *extension*. Intension and extension of a concept have an inverse relation to each other. The more a concept contains under it, the less it contains in it. [63]

Accordingly, it is usually stated as follows[1]:

Kant's Law of Extension and Intension The extension of a kind is inversely related to its intension.

The principle says that the more instances a kind has, the smaller the number of shared attributes will be. Dually, the more attributes these objects share, the smaller the number of these objects will be. In Kantian terms, it says that the broader the sphere or range of application of a concept is, the poorer its content will be, and vice versa. This applies in particular to two kinds K and K' such that one is a proper species of the other. Then it is clear that the species will have a richer intension than the genus. This difference in intension is precisely the specific difference of the species with respect to its genus. Here is another quote by John Stuart Mill's *A System of Logic* that summarizes this fact:

> From the fact that the genus includes the species, in other words denotes more than the species, or is predicable of a greater number of individuals, it follows that the species must connote more than the genus. It must connote all the attributes which the genus connotes, or there would be nothing to prevent it from denoting individuals not included in the genus. And it must connote something besides, otherwise it would include the whole genus. (...) This surplus of

[1] This law was explicitly formulated by Leibniz (see [130].

> connotation - this which the species connotes over and above the connotation of the genus - is the Differentia, or specific difference; or, to state the same proposition in other words, the Differentia is that which must be added to the connotation of the genus, to complete the connotation of the species.[86]

Kant considered the duality between the extension and intension to be a fundamental principle of logic. Independently of whether this is true or not, what it is clear is that this principle does give some information regarding the external structure of kinds as related to the internal structure.

The aim of this chapter is to explore a Minimal Conception of Kinds that does not appeal to other loaded notions such as essence or causality, in contrast to the theories mentioned in the previous chapter. This Minimal Conception is based on some of the traditional features of kinds just mentioned and will be shown to establish a connection, through Kant's law, between the external structure of kinds and their internal structure. Accordingly, the assumptions to be made throughout the whole chapter are:

Minimal Conception of Kinds Every kind has as members some objects (the extension) sharing certain sparse attributes (the intension). More strongly, all the objects share all these attributes, and these attributes are all those sparse attributes shared by these objects.

That kinds are ordered by relations of specificity will follow from these assumptions. As an example of a theory satisfying the Minimal Conception we have Essentialism of Natural Kinds (say [34]) that was discussed in the previous chapter, the sparse attributes being the properties that form the essence of the kind. However, the converse does not hold, since the Minimal Conception is silent regarding the modal features of the sparse attributes.

Apart from its being weaker than essentialism, there is another reason for considering this Minimal Conception. This conception will turn out to be very informative about the external structure of kinds, because it implies certain constraints on the ways kinds are ordered by specificity or species-genus relations. In the literature on kinds one can find several such constraints. For instance, the *Hierarchy condition*, which states that kinds are ordered forming a tree-like pattern, has been discussed. This condition can be traced back to the Porphyrian tree of categories. However, several philosophers (see [52], [136]) have argued that this assumption is too strong by presenting counterexamples to it. The model can be used to explain why this condition is based on a very narrow view on how kinds can be related to each other.

More interestingly, although the hierarchy condition does not usually hold, the Minimal Conception will be shown to imply Kant's Law. We will see how this duality between extension and intension will force the order structure of kinds to have a specific 'shape', namely that of a *complete lattice*. This will allow us to compare the approach given here to other formal models of kinds, such as Thomason's lattice-theoretic model of kinds and Corcoran's and Martin's syllogistic logic, which also propose such structure to be that of a complete lattice. Moreover, Kant's Law is strongly related to the main novel contribution made in this chapter. Assuming just the Minimal Conception, kinds will be shown to be defined in terms of their genera and specific differences, as in the classical Aristotelian theory of definitions. More interestingly, the discussion will show that there are two different ways to *subtract* one kind from another, by subtracting entities from the extension or from the intension of a kind. These correspond to two operations, whereas one of them behaves like a conditional, the other behaves like a difference. The ways these two subtractions relate to Kant's Law can be made more clear by considering the following two laws, which will be shown to hold in the model:

Law of Specific Conditional the specific conditional or intensional difference between two kinds increases (decreases) as the consequent increases (decreases) and the antecedent decreases (increases).

Law of Specific Difference the specific difference or extensional difference between two kinds increases (decreases) as the kind being subtracted from increases (decreases) and the kind subtracted decreases (increases).

Moreover, the introduction of these two subtraction operations results in two different non-classical *negations* of kinds. Overall, this suggests that the Minimal Conception is already committed to a broadly Aristotelian picture of definitions, without any need of essentialist assumptions.

3.1.2 Attributes and Kinds

Before we consider the external structure of kinds in more detail, I would like to say more about the distinction between the concept of a *kind* and the concept of an *attribute*, or more generally, of a *property*. The idea is that kinds are types or sorts of things, whereas attributes are ways or modes things are or have. This distinction is present in [77] and [34]. For instance, the following are examples of (natural) properties/attributes: *having a mass of value x*, *being old*, *having a horn*, *being spherical*, and so on. In contrast, the following are examples of (natural) kinds: *protons*, *homo sapiens*, *unicorns*, *stones*, and so on. Until now, I have assumed that these notions are different without giving any argument for it. However, there are several features (semantic, epistemological, formal and metaphysical) that seem to ground the distinction between these two concepts.

In the first place, members of a kind usually have several properties in common. In contrast, members of the extension of a property usually share only that property. As a consequence, the properties shared by the members of a kind have a major epistemological role since they can be used to classify the objects and also to guarantee successful inductive inferences concerning new future (or past) specimens and (natural) properties. According to Mill, to know the kind to which some object belongs allows us to discover a never-ending amount of properties had by that object. Natural properties do not seem to satisfy these epistemological roles, it seems that most natural properties are such that the only thing common to all their instances is simply their having that property in common [11].

Second, from a semantic point of view, natural kind terms are mostly either substance sortal or mass terms, terms that express the 'kinds' or 'sorts' of objects. In contrast, predicates for natural properties are not sortals, they are mainly adjectival or characterizing terms. Whereas adjectival terms only have application conditions (necessary and sufficient conditions for their adequate application), sortal terms come with identity conditions too for the objects that satisfy them. Sortal terms are supposed to provide identity and persistence conditions for those objects [139]. Their main theoretical roles are to allow us to count, individuate and reidentify through time those objects that satisfy them. The contrast is sometimes put as follows (Lowe,¿?). Whereas 'tree' is a sortal term, 'green' is merely an adjectival term. If while looking at a forest one is asked how many trees there are, the question can be given a determinate answer (ignoring from the time being the vagueness associated with such ordinary terms like 'tree'). But if one is asked how many green things there are, the question seems to be impossible to answer. Each leaf of a tree is green, but so is each part of the leaf. So how many green things are there? Unless one is given a further clue regarding what kind or sort of green thing one is looking after, the question seems to be unanswerable. Whereas adjectival terms are used to say *how* things look like, sortal terms are used to say *what* things are (and a fortiori how many things there are).

Third, formally and metaphysically, natural kinds are orderly related by species-genus relations or in general, order relations of specificity. Natural properties are related mainly by determinable-determinate relations. These two relations seem to be distinct. To be a determinate of some determinable property is closer to being a member or instance of a kind than to being a more specific kind of another kind. That is to say, the relation between a determinate property and its determinable is more similar to the membership relation than to the inclusion relation. For instance, consider transitivity. If kind K is a species (i.e. is more specific than) of kind K' and kind K' is a species of kind K'', then kind K is a species of kind K''. But if P is a determinate of the determinable P', and P' is a determinate of determinable P'', then it does not follow that P is a determinate of P''. Or consider this other case. An object x may belong to several species of the same kind, at least this does not seem to be impossible. But it has been usually accepted (though not everyone agrees) that an object cannot have two determinate properties of the same determinable at the same time (e.g. x is entirely red crimson and entirely red wine at t). This seems to be related to the distinction between objects and tropes. The instance of a property is a trope. But the instance of a kind is not a kind-trope (e.g. there is no distinction between my 'homo sapiens-eity' and your 'homo sapiens-eity'), it is an object (e.g. me). This has been argued by [77] and [34]. Moreover, some natural properties such as ratio magnitudes (e.g. mass, length) seem to have a richer structure than kinds (see [59]). The *Representational Theory of Measurement* by [128] assumes, for instance, that ratio magnitudes are linearly ordered and that they can be combined by some sort of concatenation operation which is structurally analogous to the sum of the real numbers that we use to represent them. In contrast, kinds do not seem to be equally structured. Another difference seems to be that kinds cannot be relations. There are kinds of relations, there are relations between kinds and some kinds may be reducible to collections of related entities. But there are no relational kinds, in contrast with properties (taken in the most general sense). Tropes and properties may be dispositional or occasional, qualitative, comparative or magnitudes. But this does not hold for kinds[2].

[2]In fact, prima facie it seems more plausible to defend a class nominalism for kinds (kinds

Summarizing:

i. Members of a kind have usually several properties in common, members of the extension of a property usually share only that property. In the former case these properties are used as a criterion to classify the objects and as a basis for sucessful inductive inferences.

ii. Kinds are the reference of natural kind terms, which are sortal or mass terms. Thus, kind terms seem to be used for counting, identifying and reidentifying objects. In contrast, properties (either tropes or universals) are the references of non-sortal or adjectival terms, used to characterize or say how objects look like.

iii. Kinds are related by specificity relations (species-genus), which are different from the determinate-determinable relations in which properties usually stand. These relations have distinct formal properties. For instance, it is more plausible to say that an object may belong to two species of the same kind than to say that it could have two determinates of the same determinable.

iv. Kinds cannot be relational nor dispositional. Instances of kinds are objects (not 'particular kinds'), instances of (universal) properties are tropes. Some properties (e.g. magnitudes) have a richer internal order and algebraic structure than kinds have.

For some of these reasons, authors such as Lowe [77] or Ellis [34] think that kinds and properties form distinct ontological categories. According to Lowe, kinds are substantial universals instantiated by particular objects, whereas properties are attribute universals instantiated by particular modes (tropes). However, all that the previous discussion implies is that the corresponding concepts of kind and property are different. There may be metaphysical reasons for reducing kinds to (bundles of) properties, or for considering kinds as a particular (kind?) case of properties (e.g. monadic properties or 'classificatory concepts or properties'), or for considering kinds as certain combinations of extensions of properties. Nevertheless, it is important to consider them apart conceptually for the reasons just stated.

3.2 The Hierarchical Structure of Kinds

3.2.1 The Failure of the Hierarchy Condition

According to the Minimal Conception, kinds are ordered by *specificity* relations. Let us call this order simply the *species-genus relation.* As Hacking [52] notes, the species-genus relation is a formal binary relation between two kinds. This only implies that one kind is more specific than another iff the former is a species of the latter iff the latter is a genus of the former. This is the sense in which the terms 'species' and 'genus' will be used in this thesis[3]. We will write "kind K is a species of kind K'" as $K \leq K'$, or as $K < K'$ when K and K' are distinct.

are collections of objects) than for properties (which are more plausibly collections of tropes).

[3]Note that the use of the term 'species' in biology is quite different. Whether something is a biological species or not does not depend on its relation to other more general kinds.

It is not uncommon for an individual object to belong to various natural kinds. If the object x is a K and K is a species of K', then x is also a K'. When we state that x is a K we are selecting from the properties of x that make it a K' those that make it specifically a K. So one may wonder whether there is any specific condition that the order structure of kinds satisfies. A popular answer is that this order is *hierarchical* [52], that is to say, that natural kinds are arranged forming a tree-like pattern.[4] This condition goes back to the Porphyrian tree of categories, where the root is the most general kind (the *summum genus*) and the leaves are the most specific kinds in the domain (the *infimae species*)[5]. The hierarchy condition implies that kinds are nested, which results in the shape of a tree. The condition says that if two kinds overlap, then one of them must be a species of the other. By 'overlapping' we can think of the extensions of the kinds having a common member. We can also think about the kinds as having a closest common kind as a species. The condition can be stated as follows:

Hierarchy If K, K' are overlapping natural kinds, then $K < K'$ or $K' < K$ or $K = K'$.

One may wonder whether this idea is not just a forgotten relic of the past. After all, scientific classifications, and in particular biological ones, seem to have moved far away from these Porphyrian trees. Why would someone think that the world is hierarchically structured? Consider the tree of life, where biological species are ordered according to phylogenetic descent. Two species overlap iff there is a species that is the closest phylogenetic descendant of both of them. But if this last species was different from the other two, this would be a case of hybridization, and the tree of a life would not be a tree at all[6]. So given two different overlapping species, one of them is a phylogenetic descendant of the other. Thus it seems that the biological world is hierarchically structured. Furthermore, this hierarchical structure is inherited by the clades to which the species belong. This results again in a tree-like pattern of nested clades. To give a second example, consider the periodic table of chemical elements as a classification of particular atoms. The classification of atoms by chemical elements is refined by the classification by isotopes. There are also classifications, like that by groups, which are coarser than the classification by elements. Take now all the kinds of isotopes, elements and groups, and order them by the inclusions of their extensions. For instance, *Lithium-6* $\leq$ *Lithium*, since every atom of *Lithium-6* will be an atom of *Lithium*. Suppose that two different chemical kinds K, K' overlap. In other words, there is a kind K'' whose extension is included in that of K and that of K'. Since extensions of isotopes/elements/groups are disjoint, that does not happen between kinds of the same classification. Suppose that K is an isotope and K' is an element (or viceversa), and that they overlap. Then at least one K-isotope is also a K'-element, therefore all K-isotopes are

[4]The term 'hierarchy' is misleading. Informally, it only suggests a vague picture according to which a set of objects is ordered in such a way that some of them are clearly above others. But this condition is stronger than that. Nevertheless, I will continue using the term for it is the one that appears in the literature.

[5]This is not historically accurate. Tree diagrams for representing logical relations appear for the first time several centuries later, see [53].

[6]Actually, there are cases of hybridization, as philosophers of biology are well aware of. This is one of the challenges to the idea that life developed in a tree-like fashion.

K'-elements and we have that $K \leq K'$. Analogous arguments can be given for the other two cases. Once again, the chemical world seems to be hierarchically structured.

The hierarchy condition was reintroduced in the contemporary scene as part of the algebraic model of kinds proposed by Thomason [135] (although it was also considered by (Kuhn, 2000)[7]), which will be considered in the following section. However, it has recently come under attack by philosophers of science like Hacking [52], Tobin [136][8], Ruphy [115], Khalidi [64] and Hendry [61], who have presented specific counterexamples from biochemistry, nuclear physics and astrophysics. For instance, Hacking states that:

> "It has been repeatedly argued that natural kinds must, as a matter of logic, be arranged in a tree-like hierarchy. Not so. Bosons, isotopes, and elements are commonly regarded as natural kinds. But since rubidium-47 is a species both of boson and of rubidium, but rubidium is not a species of boson or vice-versa, you cannot put these on a branching tree." [52]

The previous authors argue that if we take a closer look at some examples, the condition seems to be empirically false. For instance, [136] gathers some counterexamples from different sciences. One of them, also considered by [61], comes from chemistry. *Tin* has two allotropes (atoms of *Tin* can be found forming different crystalline structures), *White Tin* is metallic and *Grey Tin* is non-metallic. So the kind *Tin* overlaps the kinds *Metal* and *Non-Metal*, and therefore is not included in any of them. Another example comes from biochemistry. *Albumin* and *Renin* are proteins, *Renin* and *Hairpin Ribozyme* are enzymes, but *Albumin* is not an enzyme and *Hairpin Ribozyme* is not a protein. [64] gives some counterexamples from nuclear physics, astrophysics and fluid mechanics. Let us consider only the first two. Isotopes *Lithium-6*, *Lithium-7* and *Lithium-8* can be classified as species of the kind *Lithium* according to their atomic number, but *Lithium-8* and *Hellium-8* are also species of the kind *Beta-minus Decay Nuclides*. The second example is taken from [115]. Stars can be classified according to independent properties such as temperature, density and mass loss. Classifying by temperature is different from classifying by luminosity. For instance, the star Canopus belongs to the *type F* due to its temperature but also to the class of *luminosity Ib*, Procyon A to *type F* but to the *class IV*, and Antares to *type M* but to the *class Ib*.

Tobin's explanation for these cases rests on a distinction between *intrataxonomic crossings* and *intertaxonomic crossings*. Concerning the former ones, some individual belongs to different natural kinds in the same taxonomy. With respect to the latter ones, some individual belongs to natural kinds of different taxonomies over the same domain of objects. For instance, the same organism can belong to different biological species depending on the species concept being used ('biological', phylogenetic, evolutionist, phenetic, ecological, and so on). Khalidi's examples show similar patterns. We have various classifications of the same objects that obey different criteria (e.g. chemical v.s. nuclear),

[7]For Kuhn, intertaxonomic crossings among concepts from different "lexicons" produce cases of incommensurability; see (Kuhn, 2000).

[8]Tobin discusses other weaker conditions, such as: If K, K' overlap, then both are species of another kind K''. This condition holds necessarily in Thomason's model.

for the classifications are focusing on different properties of the objects. As a consequence, classifications crosscut each other and the same object belongs to various incomparable natural kinds. The choice of the relevant properties for the classification depends on the chosen theoretical criteria. If there are several theoretical criteria in conflict (e.g. the different species concepts) and there is no good reason to choose one over the others, then the objects will be classified in different ways according to the properties considered to be relevant in each case. In short, taxonomic pluralism leads to the proliferation of intertaxonomic crossings.

The previous examples belong to scientific classifications. But one could think that *vernacular* kinds are hierarchically ordered. After all, in our daily life we do seem to have a strong preference for tree-like orderings. Presumably there are psychological and anthropological reasons why we tend to classify things in this way, or that explain why it is such a familiar way for us to do so. Ordinary tasks of classification consist in discriminating objects physically putting them apart into isolated containers such as boxes, folders, and so on. When we nest these containers, we follow a hierarchical pattern. Given two such containers, we can put them both into a single bigger container. But we cannot just place in both of them a container that contains only some of the things that are in one alongside some of the things that are in the other one. Where would we put such an intermediate container? There are many other examples, such as the table of contents of a book or the classification by months, days and hours in a calendar. Shortly put, our daily classifications seem to be hierarchical. However, even this observation is misleading. If we take any collection of objects and classify them we will hardly obtain a tree-like structure. A full example is given below in section three, where a classification of foods and drinks from a feast is given.

In fact, one can give a simple formal explanation for how the crossings identified by Tobin arise. The mathematical explication par excellence of the notion of classification rests on the concepts of an *equivalence relation* and the corresponding *partition* of a domain (e.g. it is the one that [59] tacitly uses). Traditionally, the notion of equivalence captures the relations of similarity between objects[9]. The basic notions are well known:

Definition 4. *Let S be a set and $\approx \subseteq S \times S$ a binary relation over S. Then $(S, \approx)$ is an* equivalence relation *iff $\forall x, y, z \in S$:*

i $x \approx x$. *[Reflexivity]*

ii $x \approx y \Rightarrow y \approx x$. *[Symmetry]*

iii $x \approx y \;\&\; y \approx z \Rightarrow x \approx z$. *[Transitivity]*

Definition 5. *Let S be a set. Then $Q \subseteq \wp(S) - \varnothing$ is a* partition *of S iff :*

i $S \subseteq \bigcup Q$. *[Exhaustiveness]*

ii $\forall K, K' \in Q\; K \neq K' \Rightarrow K \cap K' = \varnothing$. *[Exclusiveness]*

By definition empty kinds are excluded. Whereas *Exhaustiveness* demands that every object belongs to at least one kind, *Exclusiveness* says that no object

[9] I will contest this assumption in Chapters IV and V

belongs to several kinds. It is well known that there is a correspondence between equivalences and partitions:

Proposition 1. *Let $(S, \approx)$ be an equivalence structure. Then the set $Q := \{[x] \subseteq S \mid x \in S\}$ of equivalence classes $[x] := \{y \in S \mid y \approx x\}$ is a partition over S. Conversely, if (S, Q) is a partition over S, then the relation $x \approx' y := \exists K \in Q\ x, y \in K$ is an equivalence relation over S.*

Moreover, this correspondence is unique. If we start from an equivalence relation, get the partition and define a new equivalence relation we will obtain the original one. Conversely, if we start from a partition, get the equivalence relation and define a new partition we will obtain the original one. Since we may have various classifications of the same domain, we may have several equivalence relations on the same domain. An equivalence that is finer than another distinguishes more accurately among the objects:

Definition 6. *Let Q and Q' be partitions of S. Then Q is* finer than *Q', $Q \leq Q' := \forall K \in Q\ \exists K' \in Q'\ K \subseteq K'$.*

It is easy to see that $\approx_Q \subseteq \approx'_Q \Leftrightarrow Q \leq Q'$. So Q is finer than Q' iff every natural kind K in Q is a species of some kind K' in Q'. The refinement order captures the previously mentioned tension between the informativeness of a classification and its unification power, so to speak. The finer an equivalence is (the more distinctions it makes), the more it says and the greater the number of natural kinds is. The coarser an equivalence is (the less distinctions it makes), the more it unifies and the greater the amount of members in each kind is. Accordingly, there are two extreme partitions. Whereas the *universal partition* $1 = \{S\}$ is the coarsest one, for it does not make any distinctions among its members, the *identity partition* $0 = \{\{x\} \mid x \in S\}$ is the finest one, since it distinguishes any element from any other. Between the two, we can find the optimal partitions that balance the informativeness and unifying power. Moreover, given two partitions Q and Q' we can obtain a third partition as their superposition:

$$Q \wedge Q' := \{K \subseteq S \mid \exists K' \in Q\, \exists K'' \in Q'\, K = K' \cap K'' \neq \varnothing\}$$

The superposition of two partitions is another partition that takes into account the distinctions made by both of them. For example, a chart or table of information is the superposition of the partitions COLUMNS and ROWS. In particular, in the Periodic Table we have $ELEMENTS = GROUPS \wedge PERIODS$.

Let us consider now *chains* of partitions, in other words, families of partitions where any two partitions are comparable to each other. These families can be depicted as points in a line. Chains of partitions are very common. To give one example, the Linnaean Hierarchy (that classifies organisms) can be seen as a chain where each partition corresponds to a different rank. It has the form $SPECIES \leq GENUS \leq FAMILY \leq ORDER \leq \ldots$. For a different example, consider some of the classifications of atoms seen in the Periodic Table, like those of partitions $ISOTOPES \leq ELEMENTS \leq GROUPS$. Some authors (e.g. [92]) call these chains *taxonomic hierarchies.* If Q, Q' are partitions (not simply coverings) then the condition of being a taxonomic hierarchy can also be expressed as [92]:

Proposition 2. *Let* $Q, Q' \in F \subseteq L(S)$*. Then* $Q \leq Q'$ *iff* $\forall K \in Q\ \forall K' \in Q'\ K \subseteq K'$ *or* $K \cap K' = \varnothing$*.*

A family of classifications F is hierarchical iff the corresponding kinds are hierarchically arranged. Hierarchies are chains of classifications whose equivalence classes are the natural kinds [10]. With these notions we can explain the difference from [136] between intrataxonomic and intertaxonomic crossings:

Intrataxonomic crossing Let Q be a classification over S. Then Q contains an *intrataxonomic crossing* iff there are two distinct overlapping natural kinds in Q.

Intertaxonomic crossing Let C be a collection of classifications $Q, Q', \ldots$ over S. Then C contains an *intertaxonomic crossing* iff there are two distinct classifications Q, Q' in C such that their overlapping $Q \wedge Q'$ is distinct from Q and from Q'.

Let C be a chain of partitions $Q, Q', \ldots$ over S. Then the intrataxonomic crossings in Q are impossible, since the equivalence classes in Q are disjoint by definition. Intertaxonomic crossings cannot happen either, for any pair of partitions in C are comparable by refinement and so the superposition $Q \wedge Q'$ of two classifications Q, Q' must be one of them. So cases of intrataxonomic crossings violate the transitivity of the equivalences, whereas cases of intertaxonomic crossings violate the condition that the family of partitions should be a chain. Classical classifications preclude both kinds of crossings. This explains why in both cases the hierarchy condition fails[11].

To sum up, both in the case of scientific and vernacular classifications the hierarchy condition fails. This is not to say that it never holds. It may hold for some specific domains but not for others. Many of the previous counterexamples are the result of what Tobin calls *intertaxonomic crossings*, that is to say, overlappings of different classifications over the same domain of entities. Both kinds of overlap can be explained in familiar terms by making use of equivalences and partitions. Since taxonomic pluralism seems to be the rule rather than the exception, this results in systematic violations of the hierarchy thesis. The upshot of these criticisms is that the hierarchy thesis is too strong an assumption. The order structure of kinds, even if restricted to those of a given specific domain, seems to be more complex than that of a tree.[12]

[10] Note that, from a metaphysical point of view, the picture is quite different from that provided by tree models. Natural kinds are now extensionally depicted as certain sets of objects and therefore their identity conditions directly refer to their instances. These sets of objects are maximal collections of pairwise resembling objects. Moreover, the order relations of specificity are not primitive, they are the inclusion relations that hold between the equivalence classes of different partitions. This is quite close to Quine' proposal for an explication of natural kinds in [107]. This suggests that partition models portray kinds as a resemblance nominalist would. We will generalize this approach in the next chapter.

[11] Nevertheless, the distinction between intrataxonomic and intertaxonomic crossings is not so neat as these authors seem to assume. On the one hand, it is difficult to say what kind of crossings we are facing in each counterexample. This is something to be established by checking case by case. On the other hand, crossings of one kind may be suitably transformed into cases of the other kind. I will not explore these matters now.

[12] I take this to be the cautious conclusion to draw since a realistic attitude towards kinds is in principle compatible with taxonomic pluralism. The listed authors discuss this issue at length.

3.2.2 Thomason's Lattice Model of Kinds

Nevertheless, Thomason's insights do not depend on the hierarchy condition. In fact, his model would be better off without it. His idea was that the system of natural kinds $(L, \leq)$, where L is the set of natural kinds and $\leq$ is the partial order *is a species of*, forms a complete lattice that in addition might satisfy the hierarchy condition[13]. But trees can be seen to be a very special class of complete lattices. In contrast, the class of complete lattices is huge. For reasons to be explained in the next section, Thomason's approach is general enough to subsume many interesting classifications as special cases. To understand his proposal we must introduce first some standard concepts from lattice theory.[14]

Definition 7. *Let L be a set, $\leq$ a binary relation over L and let $x, y, z \in L$. Then $(L, \leq)$ is a* partially ordered set *or* poset *iff it satisfies:*

1. *$x \leq x$. [Reflexivity]*
2. *If $x \leq y$ and $y \leq x$ then $x = y$. [Anti-Simmetry]*
3. *If $x \leq y$ and $y \leq z$ then $x \leq z$. [Transitivity]*

A *chain* is an order where any two elements are comparable. Every order $\leq$ induces a dual or converse order defined as $x \leq^{dual} y$ iff $y \leq x$. Some orders have special elements. For instance, an element 1 is the *top* iff it is the biggest element. An element 0 is the *bottom* iff it is the smallest element. A poset is *bounded* iff it has both a top and a bottom.

Definition 8. *Let $(L, \leq)$ be a partially ordered set, $z \in L$ and $A \subseteq L$. Then z is an* upper bound *of A iff $\forall x \in A\ x \leq z$. Dually, z is a* lower bound *of A iff $\forall x \in A\ z \leq x$. An element z is the* join, supremum *or* least upper bound *of A iff z is the smallest upper bound of A. Dually, z is the* meet, infimum *or* greatest lower bound *of A iff z is the biggest lower bound of A.*

We denote the join and meet of A by $\bigvee A$ and $\bigwedge A$, respectively. In particular, we denote $\vee\{x, y\}$ as $x \vee y$ and $\wedge\{x, y\}$ as $x \wedge y$.

Definition 9. *Let $(L, \leq)$ be a partially ordered set. Then L is a* lattice *iff for all $x, y \in L$ the elements $x \vee y$ and $x \wedge y$ exist.*

In other words, a lattice is a poset closed under the binary operations of join and meet, which satisfy for all x, y, z in L:

$$z \leq x \wedge y \iff z \leq x \text{ and } z \leq y$$
$$x \vee y \leq z \iff x \leq z \text{ and } y \leq z$$

If a poset is only closed under meets we call it a *meet-semilattice*, whereas if it is closed under joins we call it a *join-semilattice*. Meets and joins have very nice algebraic properties, which generalize the familiar logical operations of conjunction and disjunction:

Proposition 3. *Let $(L, \leq)$ be a bounded lattice. Then:*

[13]Throughout the chapter, the order $\leq$ will stand for the *is a species of* relation between kinds.

[14]These notions can be found in any standard introduction to lattice theory, such as [26].

1. $x \vee y = y \vee x$ *and* $x \wedge y = y \wedge x$. *[Commutativity]*

2. $x \vee (y \vee z) = (x \vee y) \vee z$ *and* $x \wedge (y \wedge z) = (x \wedge y) \wedge z$. *[Associativity]*

3. $x \vee x = x$ *and* $x \wedge x = x$. *[Idempotence]*

4. $x \vee (x \wedge y) = x$ *and* $x \wedge (x \vee y) = x$. *[Absorption]*

5. $x \vee 1 = 1$ *and* $x \wedge 1 = x$. *[Top]*

6. $x \vee 0 = x$ *and* $x \wedge 0 = 0$. *[Bottom]*

7. $x \leq y \iff x \wedge y = x \iff x \vee y = y$. *[Connecting Lemma]*

The crucial point is that lattices can be introduced both either as ordered sets or as algebraic structures. It is well-known that these two descriptions are mathematically equivalent. This means that whether a given set of entities, say natural kinds, is closed under operations of meet and join is equivalent to these entities being ordered in a certain fashion. In the case of kinds, for them to be closed under these conjunction-like and disjunction-like operations is for any two kinds to have a closest genus and a closest species.

We need some notions to establish when two posets or lattices are structurally similar. As usual, we do this by selecting the mappings that naturally 'copy' or 'preserve' the structure from one poset or lattice to another:

Definition 10. *Let* $(L, \leq_L)$ *and* $(M, \leq_M)$ *be lattices and* $f \colon L \to M$ *a function. Then:*

1. f *is* monotone *iff* $x \leq_L y \Rightarrow f(x) \leq_M f(y)$.

2. f *is an* order isomorphism *iff* f *is bijective and* $x \leq_L y \Leftrightarrow f(x) \leq_M f(y)$.

3. f *is a* lattice isomorphism *iff* f *is bijective,* $f(x \vee_L y) = f(x) \vee_M f(y)$ *and* $f(x \wedge_L y) = f(x) \wedge_M f(y)$.

In the case of lattices, the order isomorphisms are exactly the lattice isomorphisms. Sometimes meets and joins are defined for arbitrary subsets of elements too:

Definition 11. *Let* $(L, \leq)$ *be a lattice. Then* L *is a* complete lattice *iff for all* $A \subseteq L$, $\bigvee A$ *and* $\bigwedge A$ *exist.*

Let us consider some examples of complete and of non-complete lattices:

i. Every finite lattice is complete.

ii. The natural and real numbers with their usual orderings are non-complete lattices.

iii. The family of all subsets of a set forms a complete lattice.

iv. Every family of sets closed under arbitrary intersections and with a top element forms a complete lattice.

v. The set of all the partitions over a domain forms a complete lattice.

These examples hint at how huge the class of complete lattices is and how complex the corresponding ordering relations can be. What Thomason proposes is that natural kinds form a complete lattice.[15] This allows him to introduce the hierarchy condition as:

Definition 12. *Let* $(L, \leq)$ *be a complete lattice. Then* L *is* hierarchical *iff for all* $x, y \in L$*,* $x \leq y$ *or* $y \leq x$ *or* $x \wedge y = 0$*.*

Two kinds are either disjoint or one is a species of the other. In other words, if two distinct kinds overlap then one of them must be a proper species of the other. Trees with bottoms are complete lattices where for each element $x \neq 0$ the set $\uparrow x = \{ y \in L \mid x \leq y \}$ is a chain. It is easy to see that hierarchical lattices are trees with bottom. The most famous example is the Porphyrian tree (with a bottom attached). However, the order structures we just considered are much more general than trees and can account for cases where the kinds of a given domain overlap.

To sum up, Thomason's proposal gives an answer to the question regarding the order structure of kinds, namely, that kinds form a complete lattice. Complete lattices need not be trees, and so the previous worries concerning the hierarchy condition do not apply. However, there does not seem to be any compelling reason for why we should assume kinds to be ordered as a complete lattice. So before we introduce the new model it is worth considering a reason for taking lattices to be a good model of kinds: the class of lattices gives a semantics for accepted contemporary formulations of syllogistic logic.

3.2.3 Contemporary Syllogistic Logic

Consider again the external structure of kinds. The following relations can hold between two non-null kinds $K, K' \neq 0$:

i All K-s are K'-s: $K \leq K'$. [A]

ii No K-s are K'-s: $K \wedge K' = 0$. [E]

iii Some K-s are K'-s: $K \wedge K' \neq 0$. [I]

iv Some K-s are not K'-s: $K \nleq K'$. [O]

Some kinds are species of others and some are not, some kinds overlap and others are disjoint. These relations give rise to the quantificational Aristotelian Square of Opposition. This suggests that the logic describing the external structure of kinds must at least contain syllogistic logic.

Although syllogistic logic is usually studied as a fragment of first order classical logic (namely, the monadic fragment), there are contemporary accounts of syllogistic logic that give independent systems for it. In fact, in such systems many features of monadic first order logic are not even formulated, because the language is not expressive enough. A fortiori, the class of structures that work as a complete semantics for the logic is bigger and includes non-boolean lattices.

[15] This is not strictly speaking true, Thomason did not assume completeness. He also argued that the lattice is not distributive. Indeed, distributivity does seem to be too strong an assumption. For instance, bottomed trees are not distributive. But some kinds may be hierarchically ordered.

We will now see that such class includes all the complete lattices and is not far from being just the class of complete lattices.

Two standard references on this topic are Corcoran's [24] and Smiley's [125], who independently gave a reconstruction of Aristotle's Syllogistic Logic. The authors provided a complete set-theoretic semantics for the logic by interpreting the terms extensionally as sets of objects. In [82], Martin gave a more general semantics by making use of bottomed meet-semilattices. Here I will follow Martin's presentation.

The syntax of syllogistic logic is very simple. The language consists of a denumerable set of common nouns, which form the set of terms $Terms = \{t_1, \dots, t_n, \dots\}$ and a set of four logical symbols $\{A, E, I, O\}$ which are used instead of logical connectives. We let individual variables $x, y, \dots$ range over terms. The set of sentences Sen is obtained by taking each pair of terms and putting a logical symbol in front of them. Therefore, there are at most four kinds of sentences:

i Axy, to be interpreted as "All x-s are y-s". [A]

ii Exy, to be interpreted as "No x-s are y-s". [E]

iii Ixy, to be interpreted as "Some x-s are y-s". [I]

iv Oxy, to be interpeted as "Some x-s are not y-s". [O]

The negation of each such sentence is defined case by case: $\neg Axy := Oxy$, $\neg Ixy := Exy$, $\neg Exy := Ixy$, $\neg Oxy := Axy$. A *syllogistic syntax* is any pair $Syn =< Terms, Sen >$. Next we introduce the semantics.

Definition 13. *Let Syn be a syllogistic syntax. An* order-theoretic model *for Syn is a meet-semilattice with bottom* $(L, \leq, \wedge, 0)$*. An* order-theoretic interpretation *for Syn relative to a model L is a function* $R\colon Terms \cup Sen \to L \cup \{T, F\}$ *such that:*

i If $x \in Terms$*, then* $0 \neq R(x) \in L$*.*

ii If $A \in Sen$ *then:*

- *i If A is some* Axy*, then* $R(A) = T$ *iff* $R(x) \leq R(y)$*.*
- *ii If A is some* Exy*, then* $R(A) = T$ *iff* $R(x) \wedge R(y) = 0$*.*
- *iii If A is some* Ixy*, then* $R(A) = T$ *iff* $R(x) \wedge R(y) \neq 0$*.*
- *iv If A is some* Oxy*, then* $R(A) = T$ *iff it does not hold that* $R(x) \leq R(y)$*.*

In this way, the basic syllogistic propositions are interpreted as order relations between kinds. For example, [A] propositions, of the form "All K-s are K'-s" are interpreted as the relation $K \leq K'$.

An order-theoretic model is *set-theoretical* iff L is a meet-semilattice of sets, where the order is inclusion, the meet is the intersection and the bottom is the empty set. Let a *syllogistic language* be a pair $Lan :=< Syn, \mathbf{R} >$, where $\mathbf{R}$ is the set of all order-theoretic interpretations for Syn. We now define the usual semantic notions:

Definition 14. *Let* $Lan =< Syn, \boldsymbol{R} >$ *be a syllogistic language and let* $R \in \boldsymbol{R}$*,* $X \subseteq Sen$*. Then:*

i R satisfies X *in Lan iff for all* $A \in X$*,* $R(A) = T$*.*

ii X *is* satisfiable *in Lan iff for some interpretation* $R \in \boldsymbol{R}$*,* R *satisfies* X*.*

iii X *is* unassailable *iff for any interpretation* $R \in \boldsymbol{R}$*, there is some* $A \in X$ *such that* $R(A) = T$*.*

iv $X \models_{Lan} A$ *iff for every interpretation* $R \in \boldsymbol{R}$*, if* R *satisfies* X *then* $R(A) = T$*.*

v $\models_{Lan} A$ *iff* $\varnothing \models_{Lan} A$*.*

The basic relations of the square of opposition can be check to hold semantically:

Theorem 1. *Let Lan be a syllogistic language. Then the following hold:*

i $Axy \models_{Lan} Ixy$ *and* $Exy \models_{Lan} Oxy$*.*

ii $\{Axy, Exy\}$ *is not satisfiable and* $\{Ixy, Oxy\}$ *is unassailable.*

iii $Axy \models_{Lan} \neg Oxy$*,* $Oxy \models_{Lan} \neg Axy$*,* $Exy \models_{Lan} \neg Ixy$ *and* $Ixy \models_{Lan} \neg Exy$*.*

Martin compares several natural deduction calculus for syllogistic logic in order to prove the completeness result. We will consider just the formalized version of Corcoran's system, which takes the natural deduction rules to be the usual syllogistic rules of conversion, subalternation and reductio alongside the basic figures of Barbara and Celarent:

Definition 15. *Let* $SYL_C =< BD_S, \vdash C1, C2, Thinning, RAI, PS1, PS2 >$ *be the following set of natural deduction rules for* $\vdash$*, called* Corcoran's Syllogistic Calculus*:*

i From $X \vdash Exy$ *infer* $X \vdash Eyx$*. [C1/Conversion]*

ii From $X \vdash Axy$ *infer* $X \vdash Ixy$*. [C2/Subalternation]*

iii From $X \vdash A$ *infer* $X, Y \vdash A$*. [Thinning]*

iv From $X, \neg B \vdash A$ *and* $Y, \neg B \vdash \neg A$*, infer* $X, Y \vdash B$*. [RAI/Reductio ad Impossible]*

v From $X \vdash Azy$ *and* $Y \vdash Axz$ *infer* $X, Y \vdash Axy$*. [PS1/Barbara]*

vi From $X \vdash Ezy$ *and* $Y \vdash Axz$ *infer* $X, Y \vdash Exy$*. [PS2/Celarent]*

Here $BD_S = \{< X, A > | < X, A >$ is a deduction and$A \notin X\}$, so is the set of deductions in which the conclusion is not already included in the set of premisses. The rest of classical figures such as Darii and Ferio can be defined as rules in terms of these.

Martin proves the adequacy and completeness of the class of bottomed meet-semilattices with respect to Corcoran's logic using standard techniques:

Theorem 2. *$X \vdash A$ iff $X \models A$.*

To sum up, the class of bottomed meet-semilattice provides a complete semantics for Aristotle's syllogistic logic. In fact, it is not far-fetched to take the syllogistic terms to be denoting kinds.

However, complete lattices form a proper species of bottomed meet-semilattices, since the latter ones need not have a top. The two classes of structures are very close though, for it is a standard result that:

Proposition 4. *Let L be a poset closed under arbitrary non-empty meets and with a top element. Then L is a complete lattice.*

Therefore, there are just two steps from the semantics of syllogistic logic to Thomason's model. If the semi-lattice is closed under all *all* non-empty meets (instead of just the finite ones) and a top element is added, then we have a complete lattice. Is there any justification for such a move?

On the one hand, completeness under arbitrary meets is trivially satisfied in the finite case. So objections are only relevant when we are dealing with an infinite amount of kinds. Note that, in the finite case, closure under binary meets implies (by induction) closure under finite meets. Once we accept that the meet of a large but finite amount of kinds is a kind, it seems plausible to generalize this to the infinite case. On the other hand, the addition of a top element corresponds to the existence of a summum genus. It is plausible to assume that, for every domain of entities being classified, there is a kind to which all of them belong. In fact, the usual way in which a domain of entities is selected is by giving something like a summum genus. For example, think about the classifications of *chemical elements*, *diseases*, *minerals*, and so on. Classifications themselves acquired their name from this summum genus.

To sum up, we have here a further argument for modelling the external structure of kinds as a general (complete) lattice, namely, such structures give a semantics for syllogistic logic, which plausibly describes the species-genus relations among kinds. Nevertheless, neither Thomason's approach nor Corcoran-Martin's logic mentions objects or attributes, and apart from the algebraic structure imposed they do not give further information about the order structure of kinds. The rest of the chapter will make use of a model for natural kinds based on the Minimal Conception that gives more insight into the order structure of kinds while still preserves Thomason's basic ideas and is still a model for syllogistic logic.

3.3 The Concept Lattice Model of Kinds

3.3.1 A Model for Natural Kinds

In this section a model for the Minimal Conception of kinds is introduced, based on Wille's Theory of Formal Concept Lattices [40].[16] This model is a more informative special case of the one by Thomason just discussed, as it will be seen soon.[17] One can find some references in the literature making similar proposals.

[16] Concept lattices, semilattices and trees are heavily used models of classifications, see [100].

[17] In fact, the relation between concept lattices and complete lattices is far more interesting than that, since *every complete lattice* is isomorphic to a concept lattice [40].

	Z = 3	Z = 8	Z = 79	N = 8	N = 9
Atom x	X				
Atom y	X				
Atom z		X		X	
Atom p		X			X
Atom q			X		

Table 3.1: Example of Context: Atoms

For instance, Mormann [88] already proposes Wille's theory as a good candidate for explicating Armstrong's realism. There are also slightly more general proposals available, as the ones concerning syllogistic logic just discussed.[18] In what follows, the names of some concepts from the theory of Concept Lattices will be replaced by others more suitable to our current purposes.[19]

Definition 16. *Let S and Q be sets and $I \subseteq S \times Q$. Then (S, Q, I) is a* formal context*. S is the set of* objects*, Q is the set of* attributes *and I is the* exemplification relation*.*

Each context can be represented by a table. For example, let us take a collection of atoms as objects, with their atomic number (Z) and the number of neutrons (N) as attributes. A possible context can be found in Table 3.1. The most basic operation one can perform on contexts is to permute the role of objects and attributes just by putting the exemplification relation 'upside down':

Definition 17. *Let (S, Q, I) be a context and let $I^{-1} = \{(R, x) \mid (x, R) \in I\}$ be the converse of I. Then (Q, S, I^{-1}) is its* dual context*.*

A fundamental reason for choosing contexts is that they provide a definition of the notions of extension and intension:

Definition 18. *Let (S, Q, I) be a formal context, $A \subseteq S$, $B \subseteq Q$. Then $i \colon \wp(S) \to \wp(Q)$ defined as $i(A) := \{P \in Q \mid xIP \text{ for all } x \in A\}$ is the* intension function*, and $e \colon \wp(Q) \to \wp(S)$ defined as $e(B) := \{x \in S \mid xIP \text{ for all } P \in B\}$ is the* extension function*.*

The extension is a function that gives, for each set of attributes B, the set of *all* the objects having *all* the attributes in B. Dually, the intension gives, for each set of objects A, the set of all the attributes had by all the objects in A. Clearly, A is an extension for some B iff $A = ei(A)$ and B is the intension for some A iff $B = ie(B)$.

Definition 19. *Let (S, Q, I) and $A \subseteq S$, $B \subseteq Q$. Then (A, B) is a* natural kind *iff $A = e(B)$ and $i(A) = B$. Here A is the* extension *of B and B the* intension *of A.*

[18] The basic theses in this chapter were strongly inspired by the claims made by [88]. There are other examples of similar proposals, see Swoyer's approach to Leibnizian calculus in [130].

[19] Although the focus here is on monadic properties, one could have n-ary relations in Q by putting n-tuples in S.

Wille calls such a pair (A, B) a *formal concept*, but we call it a *natural kind*. For instance, we have the natural kinds $Lithium = (\{x, y\}, \{Z = 3\})$, $Oxygen = (\{z, p\}, \{Z = 8\})$, $Gold = (\{q\}, \{Z = 79\})$, and the isotopes $Oxygen - 16 = (\{z\}, \{Z = 8, N = 8\})$ and $Oxygen - 17 = (\{p\}, \{Z = 8, N = 9\})$. Although formal contexts may look too simple to obtain any interesting results, this is far from being true:

Proposition 5. *Let (S, Q, I) be a context and $A, A' \subseteq S$, $B, B' \subseteq Q$. Then:*

1. $A \subseteq e(B) \iff B \subseteq i(A)$. *[Galois Connection]*
2. $A \subseteq A' \Rightarrow i(A') \subseteq i(A)$. *[Antitonicity i]*
3. $B \subseteq B' \Rightarrow e(B') \subseteq e(B)$. *[Antitonicity e]*
4. $A \subseteq ei(A)$ *and* $B \subseteq ie(B)$. *[Extensiveness ei, ie]*
5. $A \subseteq A' \Rightarrow ei(A) \subseteq ei(A')$. *[Monotonicity ei]*
6. $B \subseteq B' \Rightarrow ie(B) \subseteq ie(B')$. *[Monotonicity ie]*
7. $eiei(A) = ei(A)$ *and* $ieie(B) = ie(B)$. *[Idempotence ei, ie]*
8. $iei(A) = i(A)$ *and* $eie(B) = e(B)$.
9. $i(\bigcup_{i \in I} A_i) = \bigcap_{i \in I} i(A_i)$ *and* $e(\bigcup_{i \in I} B_i) = \bigcap_{i \in I} e(B_i)$.

From these properties half of the main theorem of the theory follows:

Theorem 3 (Fundamental Theorem of Concept Lattices)**.** *Let L^* be the set of natural kinds of the context (S, Q, I). If (A, B) and (A', B') are natural kinds, define $(A, B) \leq (A', B') := A \subseteq A'$. Then $(L^*, \leq)$ is a complete lattice, where:*

i $(A, B) \leq (A', B') \iff A \subseteq A' \iff B' \subseteq B$.

ii $\bigwedge_{i \in I}(A_i, B_i) = (\bigcap A_i, ie(\bigcup B_i))$.

iii $\bigvee_{i \in I}(A_i, B_i) = (ei(\bigcup A_i), \bigcap B_i)$.

iv $1 = (S, i(S))$ *and* $0 = (e(Q), Q)$.

v $xIP \iff x \in e(P) \iff P \in i(x)$.

Lattices L^* are usually called *concept lattices*. We will call them *lattices of natural kinds*. Note that the whole lattice structure, and a fortiori the whole order structure, follows from the definition of kinds. The extensions and the intensions also form complete lattices. In fact, the lattice of kinds can be seen as the combination of these two lattices. We will make heavy use of these lattices in the last chapter:

Proposition 6. *Let (S, Q, I) be a context. Then $(\boldsymbol{B}_S, \subseteq)$ where $\boldsymbol{B}_S = \{A \subseteq S \mid A = ei(A)\}$ is the* lattice of extensions. *For $\{A_i\} \subseteq \boldsymbol{B}_S$ we have that:*

i $\bigwedge A_i = \bigcap A_i$.

ii $\bigvee A_i = ei(\bigcup A_i)$.

iii $1_S = S$.

iv $0_S = ei(\varnothing) = e(Q)$.

Proposition 7. *Let* (S, Q, I) *be a context. Then* $(\boldsymbol{B}_Q, \subseteq)$ *where* $\boldsymbol{B}_Q = \{B \subseteq Q \mid B = ie(B)\}$ *is the* lattice of intensions. *For* $\{B_i\} \subseteq \boldsymbol{B}_Q$ *we have that:*

i $\bigwedge B_i = \bigcap B_i$.

ii $\bigvee B_i = ie(\bigcup B_i)$.

iii $1_Q = Q$.

iv $0_Q = ie(\varnothing) = i(S)$.

Moreover, the lattice of kinds is isomorphic to the lattice of all the extensions $\mathbf{B}_S$ and to the dual of the lattice of intensions $\mathbf{B}_Q$. This is again an expression of the fact that extension and intension are dual to each other, which turns out to be the explanation for why Kant's Law holds:

Proposition 8. *Let* $\boldsymbol{B}(S, Q, I)$ *be the lattice of natural kinds of* (S, Q, I). *Then:*

i $(\boldsymbol{B}_Q, \subseteq)^{dual} = (\boldsymbol{B}_Q, \bigvee, \bigcap)$ *is the dual of the lattice of intensions.*

ii The extension function $e\colon (\boldsymbol{B}_Q, \subseteq)^{dual} \to (\boldsymbol{B}_S, \subseteq)$ *is an order isomorphism.*

iii The intension function $i\colon (\boldsymbol{B}_S, \subseteq) \to (\boldsymbol{B}_Q, \subseteq)^{dual}$ *is an order isomorphism.*

iv The projection function $\pi_1\colon \boldsymbol{B}(S, Q, I) \to (\boldsymbol{B}_S, \subseteq)$ *is an order isomorphism.*

As usual, we need a concept to determine whether two contexts are structurally similar:

Definition 20. *Let* (S, Q, I), (S', Q', I') *be contexts and* $(f, g)\colon (S, Q, I) \to (S', Q', I')$ *a pair of functions* $f\colon S \to S'$, $g\colon Q \to Q'$. *Then:*

i (f, g) *is a* homomorphism *iff* $(xIR \Rightarrow f(x)I'g(R))$.

ii (f, g) *is an* isomorphism *iff* f *and* g *are bijective and* $(xIR \Leftrightarrow f(x)I'g(R))$.

In some cases, the fact that the lattices of kinds are isomorphic can be directly inferred from the fact that their contexts are isomorphic. Two useful lemmas are:

Lemma 1. *The lattice* $\boldsymbol{B}(S, Q, I)$ *is isomorphic under the mapping* $f((A, B)) = (B, A)$ *to the dual of the lattice* $\boldsymbol{B}(S, Q, I^{-1})$.

Lemma 2. *If* (S, Q, I), (S', Q', I') *are isomorphic contexts under the mapping* (f, g), *then their lattices* $\boldsymbol{B}(S, Q, I)$, $\boldsymbol{B}(S', Q', I')$ *are isomorphic under the mapping* $h((A, B)) = (f(A), g(B))$.

There are some special kinds in the lattice. The extensionally biggest kind is $1 = (S, i(S))$, every object belongs to its extension and it only contains in its intension attributes shared by all objects of the domain. Usually, it has empty intension. To honour tradition, we will call it the *summum genus*. The extensionally smallest kind is $0 = (e(Q), Q)$, every attribute belongs to its intension and it only contains in its extension objects having all the attributes in the domain. Usually, it has empty extension. It will be called the *null kind*.

Each object x induces an extensionally minimal kind $K_x = (ei(x), i(x))$. Its intension is the intension of the object, whereas its extension contains all the objects that have all the attributes that the object has (and possible more). We will call the kinds of the form K_x the *infimae species*. It is debatable whether this name is appropriate, after all, usually these infimae species only contain in their extension an object, whereas traditional infimae species contain several objects. Socrates is an object that determines a kind whose extension contains Socrates and every other object in the domain that has all the attributes that Socrates has. However, the infima species to which Socrates belongs is traditionally something like the kind *human* or *person*, in other words, a kind with a considerably large extension. So I would not put too much weight on this naming. Traditionally, the infimae species are supposed to be those kinds just extensionally above individuals and which are such that no other kind is a proper species of them. However, if the lattice is infinite then such kinds may not exist (because some objects may not be covered by any kind), whereas what we are here calling 'infimae species' *will always exist*. Moreover, the traditional picture does not allow for some of these infimae species to form a chain, since it pictures infimae species as *minimal elements*. But this usually happens in lattices. The appropriate lattice-theoretical notion is instead that of a *join-irreducible element*: an element x in a lattice is join-irreducible iff if $x = y \vee z$ then $x = y$ or $x = z$. Join-irreducible elements can be non-trivially ordered, in contrast with minimal (or atomic) elements. Mathematically this is a very important fact.

Dually, each attribute P induces an intensionally minimal kind of the form $K_P = (e(P), ie(P))$. Its extension is the extension of the attribute, whereas its intension contains all the attributes shared by all the objects that exemplify that attribute. In other words, all the attributes 'implied' by this one will belong to the intension. We will call the kinds of the form K_P the *maximal genera*. Remarks concerning maximal genera are dual to the ones given for the infimae species. Anyways, these special kinds are fundamental in the following sense: every kind is the join of its infimae species and the meet of its maximal genera. In lattice-theoretic terms, the infimae species form a join-dense subset, whereas the maximal general form a meet-dense subset.

Let us give a full example of a lattice of kinds. Let the following context be given in Figure 3.2. The objects are particular foods or drinks found in a feast, such as a specific pear A, a specific cookie B, and so on. The attributes are properties of these foods, such as *being sweet* or *having animal origin*. The resulting lattice can be found in Figure 3.1[20]. Most of the kinds are completely determined either by an object or by an attribute (and we have given them names accordingly). In other words, they are infimae species (noted with a triangle down-side) or maximal genera (noted with a triangle up-side). For instance, the kind $Honey = \{\{HoneyI\}, \{Sweet, Animal-Origin, Highly-Caloric, Artisan, Liquid\}\}$ is an infima species. Other kinds are more interesting. For example, $K = \{\{HoneyI, CheeseC\}, \{Animal-Origin, Highly-caloric, Artisan\}\}$ is something like the kind of "(non-vegan) desserts", whereas

[20]Every Figure in this thesis was made by the author by making use of the free software GeoGebra.

	Sweet	Animal Origin	Highly caloric	Spice	Artisan	Smelly	Liquid	Baked
Pear A	X							
Cookies B	X		X		X	X		X
Cheese C		X	X		X	X		
Oyster D		X				X		
Chicken E		X				X		X
Strawberry F	X					X		
Beer G			X		X		X	
Baked Potato H			X			X		X
Honey I	X	X	X		X		X	
Rosemary J				X		X		
Water K							X	
Sugar L	X		X					

Table 3.2: Example of Context: Food

$K'' = \{\{HoneyI, CookieB\}, \{Artisan, Sweet, Highly - caloric\}\}$ which could be interpreted as the kind of "sweet desserts". Honey belongs to both kinds. This example hints at how complex the relations between kinds could be.

Some relations between kinds hold. For instance, the kind of artisan foods is a species of that of highly caloric ones. However, some highly caloric foods, such as baked potatoes, are not artisan. We have $Artisan \leq Highly - caloric$ where:

$$\begin{aligned} Artisan =& (\{BeerG, HoneyI, CheeseC, CookiesB\}, \\ & \{Highly - caloric, Artisan\}) \\ Highly - caloric =& (\{BeerG, HoneyI, CheeseC, CookiesB, \\ & Baked - potatoH, SugarL\}, \{Highly - caloric\}) \end{aligned}$$

which obeys Kant's Law as expected. This example also shows that even classifications made in ordinary life can be non-hierarchical (for many other examples the reader can consult [40]).

3.3.2 Arguments for the Adequacy of the Model

In this section the main features of the model are explained and its material adequacy with respect to the Minimal Conception is argued for. In other words, the main aim of the section is to show that the ontological commitments of the model are those and only those of the Minimal Conception.

First, the primitive entities are a set of objects S, a set of attributes Q and an exemplification relation I that holds from objects to attributes and not vice versa. A fortiori, entities in S cannot be exemplified, but can exemplify several attributes. Therefore, they are plausibly interpreted to be particular objects. Entities in Q can be exemplified by several objects. Thus, they are plausibly interpreted to be properties. Of course, the model does not prevent us from choosing elements in Q to be sets, mereological sums, predicates, and so on. It does not prevent us from choosing the exemplification relation I to

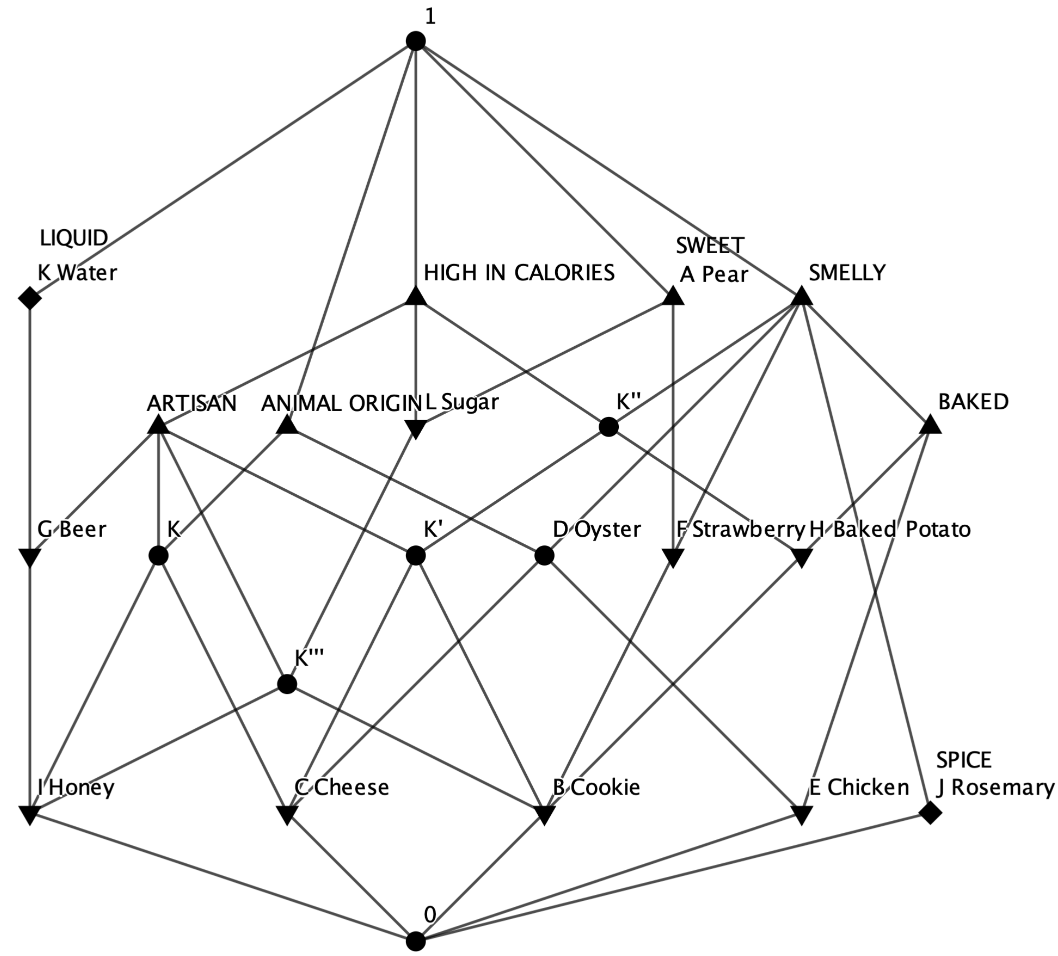

Figure 3.1: Example of Concept Lattice: Food

be the membership relation, the parthood relation, the satisfaction relation, and so on. Although the most straightforward reading of the elements in Q is as universal attributes, nominalists can use their favourite surrogates instead. Many philosophers will require additional principles to be satisfied though. For instance, even though they are not required in the definition of a context, the following two dual principles are often assumed[21]:

$$\forall x \in S \; i(x) \neq \varnothing \qquad\qquad \forall P \in Q \; e(P) \neq \varnothing$$

The one on the left is the *Principle of No Bare Particulars*, it says that every object has at least one attribute. The one on the right is the *Principle of Exemplification*, it says that every attribute is exemplified. Since we are assuming attributes to be sparse, we will require the latter principle. The model is also compatible with the following cases:

- Two attributes may be exemplified by exactly the same objects: $e(P) = e(R)$.
- Two objects may exemplify exactly the same attributes: $i(x) = i(y)$.
- An attribute may be exemplified by every object in the domain: $e(P) = S$.
- An object may exemplify every attribute in the domain: $i(x) = Q$.
- All objects exemplifying a certain attribute may exemplify another given attribute: $e(P) \subseteq e(R)$.
- All attributes exemplified by an object may be exemplified by another given object: $i(x) \subseteq i(y)$.

In other words, there could be distinct coextensional attributes, distinct cointensional objects, and so on. Therefore the following principles may not hold:

- $\forall P, R \in Q \; e(P) = e(R) \Rightarrow P = R$. [Coextensionality]
- $\forall x, y \in S \; i(x) = i(y) \Rightarrow x = y$. [Identity of Indiscernibles]
- $i(S) = \varnothing$. [No General Attributes]
- $e(Q) = \varnothing$. [No General Objects]
- $\forall P, R \in Q \; e(P) \subseteq e(R) \Rightarrow P = R$. [Attribute Independence]
- $\forall x, y \in S \; i(x) \subseteq i(y) \Rightarrow x = y$. [Object Independence]

For instance, the context in Table 3.3 can be used as a counterexample to all of them. We have that $i(y) \subset i(x)$, $i(w) = i(y)$, $e(P) \subset e(R)$, $e(R) = e(M)$, $i(S) \neq \varnothing$, $e(Q) \neq \varnothing$. Of course, whether this counterexample is metaphysically possible depends on further commitments.

Second, following the distinction introduced by David Lewis [73], the attributes are plausibly *natural* or *sparse*, not *abundant* properties. Several features characterize naturalness. Among them, we have similarity, intrinsicness

[21] Of course, these two principles have different names in the Theory of Concept Lattices.

	P	R	W	M
x	X	X	X	X
y	X	X		X
w	X	X		X
z		X		X

Table 3.3: Context of Counterexamples

and simplicity. Sparse attributes are said to ground or be grounded on resemblances or similarities between objects. Lewis argues that intrinsicness can be defined in terms of duplication or exact similarity, the latter one being defined as sharing the same natural properties (so that natural properties turn out to be intrinsic). Under the assumption that one starts with natural attributes, the following accounts of similarity can be given (as Armstrong himself proposed many times)[22]:

$$x \text{ and } y \text{ are } \textit{duplicates} \text{ or } \textit{exactly similar} \iff i(x) = i(y).$$
$$x \text{ and } y \text{ are } \textit{approximately similar} \iff i(x) \cap i(y) \neq \varnothing.$$

The last criterion for sparseness concerns the composition conditions of properties, and thus whether they are closed under formal operations. Natural properties, or at least maximally or perfectly natural properties, are usually taken to be *simple*. Complex properties obtained by combinations from (perfectly) natural properties are less natural or even not natural at all. In the model, attributes need not be closed under any operation over their extensions, although we could force them to be so. In particular, the negation of an attribute (complement of extension of an attribute), the conjunction of attributes (intersection of extensions of attributes) or the disjunction of attributes (union of extensions of attributes) are not necessarily attributes. We cannot manufacture attributes at will. By this I mean that if *Red* is an attribute with extension $e(R)$, there may not be an attribute *non-Red* in the context with extension $e(\textit{non-Red}) = S - e(R)$. In contrast, an abundant conception of properties requires at least some sort of boolean structure over the set of attributes. Nevertheless, natural kinds are closed under joins and meets. This is a crucial difference. These internal operations flow directly from the very definition of a natural kind and their order structure without assuming any closure constraints over the natural attributes. In other words, once kinds are defined as we did, one is forced to accept the existence of meets and joins of kinds, since these satisfy the definitions too. The crucial point though, is that the meet and join operations need not be interpreted simply as classical intersections (or conjunctions) and unions (or disjunctions). Our operations do not have all the formal properties of intersections and unions. For instance, joins and meets do not necessarily distribute over one another. But even if they did, an important difference would remain. To make it more clear, let us take the case of 'disjunctive' kinds. Recall that the intension of a disjunctive kind will be the intersection of the intensions of its disjuncts and that its extension will be the smallest extension including

[22]Intrinsicness and similarity should be considered with care. It is not clear that there is any compelling reason why we should exclude relations from being natural or some similarities from being abundant.

the union of the extensions of its disjuncts:

$$(A, B) \vee (A', B') = (ei(A \cup A'), B \cap B')$$

Let us make use of a Goodmanian example based on folk taxonomy. Let 'Tulipose' be a predicate having in its extension any object that is either a tulip or a rose. If 'Tulipose' does not denote a kind (not the null kind, but no kind at all), then it has no place in our classification. This will happen if, for instance, there are no attributes common to all the members of the set. If so, tuliposes will not form the extension of a kind. So this excludes some Goodmanian predicates from the beginning. Now let us suppose that it denotes a kind. There is indeed a smallest kind $Tulip \vee Rose$ whose extension includes all the roses and all the tulips. It is the closest genus of $Tulip$ and $Rose$. This seems to be a plausible candidate for the denotation of our term 'Tulipose'. But the extension of such a kind is usually bigger than the union of these two extensions. More to the point, if the attributes common to all tulips and all roses are simply those attributes common to every object in the domain, or if there is no such common attribute, then $Tulip \vee Rose$ may happen to be simply the summum genus, say $Plant$. And if the kind denoted is not the summum genus, then it will simply be a bona fide genus of both $Tulip$ and $Rose$, but presented under a rather unusual Goodmanian name. This is a crucial point since it blocks some Goodmanian grue-like objections for kinds. Not every arbitrary collection of objects is the extension of a kind, and in particular, the unions of extensions of kinds are not extensions of kinds.[23]

Third, if required, we could also introduce the hierarchy condition by imposing the constraint that the extensions form a *set-theoretic hierarchy*[24]:

Definition 21. *Let (S, Q, I) be a context. Then it is a* hierarchical context *iff* $\forall P, R \in Q \quad e(P) \subseteq e(R)$ *or* $e(R) \subseteq e(P)$ *or* $e(P) \cap e(R) = e(Q)$.

Proposition 9. *The context (S, Q, I) is hierarchical $\iff$ L^* is hierarchical.*

Therefore, the tree model is a very special case of the more general lattice-theoretic model. Hierarchy requires the extensions of the attributes to be nested. The Minimal Conception can avoid the counterexamples related to the hierarchy condition, while acknowledging that some specific domains may indeed be hierarchically structured.

Although lattices of kinds are not generally hierarchical, there is a principle regarding the specificity relations between kinds that holds in every lattice of kinds, namely Kant's Law of Extension and Intension. It follows immediately $K \leq K' \iff A \subseteq A' \iff B' \subseteq B$[25]:

Kant's Law of Extension and Intension The extension of a kind is inversely related to its intension.

Notice that Kant's Law follows directly from the very definition of natural kinds. The more general the kind is (the greater the amount of its instances is),

[23] Of course, this strategy only blocks the introduction of Goodmanian kinds, not of Goodmanian attributes. For the latter one has to assume that the attributes are natural.

[24] Such families of sets are also used as models of classification, see [100].

[25] The model identifies any two kinds that have empty extension. Thus $Unicorn = Centaur = Fairy$. To avoid this problem the model should be expanded to a modal one.

the less the number of attributes common to all its members will be. Dually, the more specific the kind is, the greater the number of attributes shared by its instances will be. The limiting cases nicely illustrate this duality. The summum genus has the largest extension, that contains every object in the domain, and the smallest intension, which includes the attributes common to every object (usually, this set is empty). The null kind has the smallest extension, which includes only those objects that have all the attributes (usually, this set is empty) and the biggest intension, that contains all the attributes in the domain. The less trivial cases are those of the maximal genera and the infimae species. The intension of a maximal genus is usually fixed by just one attribute, whereas the extension of an infima species is usually fixed by just one object. Roughly speaking, attributes correspond to extensionally rich kinds, whereas objects correspond to intensionally rich kinds.

From a conceptualist point of view, what this law states is that the larger the range of application of a concept is, the poorer its content will be, and vice versa, the richer the content of a concept is, the smaller is range of application will be. Thus, Kant's Law is a duality that reflects the trade-off present in every conception of what a classification is supposed to be: the trade-off between parsimony and informativeness (for more on this, see the next chapter). Consider a family of kinds (from the lattice) that can be used to cover the domain of objects, in other words, a family of kinds such that every object in the domain belongs to at least one of them. The larger the extensions of these kinds are, the fewer kinds will be needed in order to cover the whole domain. The inferences in which those kinds appear will involve many objects and therefore will be very general. However, in exchange, these kinds will have poorer intensions and therefore will support fewer informative inferences about their members. Dually, the richer the intensions of these kinds are, the more inferences about the attributes of their members we will be able to make. However, in exchange, many kinds will be needed in order to cover the whole domain, which will make inferences very specific. Consider again the limiting examples. One can classify all the objects by making use of the infimae species or by making use of the maximal genera. Whereas the former classification is too specific, the latter one is too general. The classification by infimae species is extensionally extremely poor, whatever inferences it will allow us to make will involve very few objects. In contrast, the classification by maximal genera is intensionally too poor, whatever inferences it will allow us to make will involve very few attributes. A good classification balances these two aspects and therefore appeals to kinds that are neither too general nor too specific, in other words, to kinds that do not correspond to attributes nor to objects.

Finally, natural kinds are bi-dimensional entities $K = (A, B)$, where $A = e(B)$ is a set of instances, the set of all the objects that exemplify the attributes in B, and $B = i(A)$ is a set of attributes, which are exactly those exemplified by the objects in A. We can split up the definition of a natural kind that has been given into several theses:

i. A kind K corresponds to a set of objects A, its extension.

ii. A kind K corresponds to a set of attributes B, its intension.

iii. If an object instantiates K, then it exemplifies all the attributes in B.

iv. If an object exemplifies all the attributes in B, then it instantiates K.

v. If an attribute P is in B, then all the objects of K exemplify P.

vi. If all the objects of K exemplify P, then P is in B.

From these considerations, it follows that $A = e(B)$ and $B = i(A)$. The objects in the extension are exactly those that exemplify the attributes in the intension and the attributes in the intension are exactly those that are exemplified by the objects in the extension. Let us see what these theses involve.

Theses (i) and (ii) are assumptions made by most theories of kinds. Here A is the set of objects that instantiate or belong to the kind and B is a set of attributes shared by (some) members of the kind.[26] In principle, both the set of objects and the set of attributes corresponding to the kind could be empty.

Theses (iii)-(iv) are equivalent to $A = e(B)$, that the set of objects of the kind forms its extension. Equivalently put, they require that the set of attributes B fixes the *membership conditions* of the kind. That is to say, for an object x to belong to K it is necessary and sufficient that it exemplifies all the attributes in B. Theses (v)-(vi) are equivalent to $B = i(A)$, that the set of attributes of the kind forms its intension. Thesis (v) says that all the objects belonging to the kind exemplify every attribute in B. Once again, if B fixes the membership conditions of the kind this is to be expected. In contrast, thesis (vi) is a more contentious one. It may happen that all the instances of the kind exemplify P, although P is not a property that is characteristic of the kind. For example, P may be accidentally exemplified by all the instances of the kind, or it may be a necessary consequence of their membership to the kind without being a property essentially exemplified by all of them, or it may simply be theoretically irrelevant. We have assumed that the attributes in the intension are sparse, however that assumption by itself may not be enough.

Another problem is this. It is easy to check that the identity conditions for natural kinds are completely determined both by their intension and by their extension. Two kinds are identical iff they are cointensional iff they are coextensional. The explanation is simply that the model does not take into account the modal features of kinds. These difficulties point at the limitations of both the Minimal Conception and its model. Under stronger commitments for this conception or expansions of the model, some of these difficulties disappear. For instance, essentialists will assume that the attributes in Q are essentially exemplified by the objects, or that I is the relation $\ldots essentially\ exemplifies \ldots$. If the attributes in B form the general essence of the kind, then the exemplification relation I holds essentially of the objects and the premisses are satisfied. The model only describes what happens in the actual world. To allow for distinct but coextensional kinds one has to take into consideration what other extensions that kind could have had. Although we will give a sketch of a modal expansion for this model, issues regarding modality are beyond the scope of this chapter.

Nevertheless, not every theory of kinds will accept the Minimal Conception. It is a common assumption among kind theorists that the membership conditions are fixed by the sparse properties shared by the objects. Some think that these sparse properties are essential to their bearers, others hold that they simply

[26]Causal theorists may object that the collection of members of the kind should not be represented by a set, given that kinds do not have sharp membership conditions.

tend to co-occur as a matter of natural necessity. However, not all of them will accept that the objects share *all* the sparse properties in the intension. For instance causal theorists (say Boyd [12]) will reject some of the assumptions of the Minimal Conception. They will prefer to give identity conditions for kinds in terms of the causal processes that bundle the attributes together. Moreover, they will point out that there is no collection of attributes shared by all the members of the kind and that the collection of objects has 'fuzzy boundaries'. More specifically, they will object to (iii) and (v) by arguing that an object can belong to the kind while not exemplifying all the corresponding attributes. There may not be a set of attributes which is such that all of them are had by all the members of the kind. This results, in particular, in some objects being boundary cases of the kind.

In the following section the main application of the model will be given. The Minimal Conception will be shown to be committed to Aristotelian definitions of kinds in terms of genera and specific differences.

3.4 Aristotelian Definitions of Kinds

3.4.1 Two Operations of Specific Difference

The main application of the model to be given in this chapter concerns the classical *Aristotelian theory of definitions*, which is a rich source of operations on kinds. The aim of this section is to show that kinds can be given Aristotelian definitions in terms of genera and specific differences. More precisely, there are two different ways to subtract one kind from another. One can subtract intensionally a genus from a species by considering the attributes in the species that do not belong to the intension of the genus. In contrast, one can subtract extensionally a species from a genus by considering the objects in the genus that do not belong to the extension of the species. Correspondingly, there are two different ways to define a kind using specific differences. On the one hand, one can get a genus extensionally by taking one of its species and introducing enough objects in its extension. On the other hand, one can get a species intensionally by taking one of its genera and introducing enough attributes in the intension. These operations have a very close relation to Kant's Law. Furthermore, each operation induces a negation that behaves non-classically and whose behavior can be explained by appealing to the Hexagon of Opposition. At the end of this section some differences between the current model and the traditional Aristotelian approach will be highlighted.

According to the traditional theory, one can abstract from a species by deleting part of its intension and correspondingly enlarging the extension. Dually, one can determine or specify a genus by adding attributes to its intension and correspondingly restricting its extension. To define a species one gives a genus and a differentia or specific difference. Starting from a species and a specific difference, the genus can be recovered as the one having in its intension those attributes of the species that are not in the specific difference. To take the standard example, the species *Human* is defined from the genus *Animal* and the specific difference *Rational*.

Generally speaking, one can interpret the meets and joins of kinds as operations of *logical determination* and *logical abstraction*, respectively. For instance, suppose that we take a kind $K = (A, B)$ and that we want to divide it into a species by enriching the intension B. We take a set of attributes B' which forms an intension, we add it to B and then we obtain the corresponding closest natural kind, which has intension $ie(B \cup B')$. This is just the meet $K \wedge K'$, where $K' = (e(B'), B')$. Dually, logical abstraction works by starting from a kind K, selecting only some attributes in B which form an intension by overlapping B with another intension B', and then taking the corresponding natural kind. Again, this is just the join $K \vee K'$. So the operations of meet and join of kinds can be understood as logical determination and abstraction operations that form new kinds from already given ones. Given a kind K we can *abstract* or *generalize* by K', $Abs_{K'}(K) := K \vee K'$, and *determine* or *specify* by K', $Spec_{K'}(K) := K \wedge K'$. As an example, by abstracting the isotope $Oxygen - 16$ from the isotope $Oxygen - 17$ we get the more general kind $Oxygen$, since $Abs_{Oxygen-16}(Oxygen - 17) = Oxygen$.

However, to consider specific differences we have to introduce several new operations that are well-defined in the lattice. If K is a species of K', then we have $B' \subseteq B$ and therefore the remainder $B - B'$ includes those attributes that make the K-s specifically K-s among the K'-s. Given a kind K' as a genus and a specific difference $B - B'$, we can give a definition of a species. But one may object that the process may not be completely satisfactory, given that we are not really subtracting one kind from another. What does it mean to combine $B - B'$ and K' to get the species K? This operation is neither the join nor the meet, since $B - B'$ is not necessarily the intension of a kind. So, is there a natural way to introduce an internal operation corresponding to the specific difference?

Actually, there are two natural options. We saw that both the closure of any set of attributes and the closure of any set of objects induce kinds. Suppose that K is a species of K'. We can define the specific difference as the kind $K' \to K := (e(B-B'), ie(B-B'))$ or as the kind $K' \backslash K := (ei(A'-A), i(A'-A))$. In other words, we can either take the attributes of the species that are not in the genus and then obtain the corresponding kind, or we can take the objects in the genus that are not in the species and then obtain the corresponding kind. The former is the kind induced by the intensional closure over a difference, whereas the latter is the kind induced by the extensional closure over a difference. We can think about the former as the intensional way to subtract a species from a genus and about the latter as the extensional way to subtract the species from the genus. To study the properties of these operations it is convenient to follow a more general strategy. We will first give an abstract characterization and then we will show that they can be found in any concept lattice:

Definition 22. *Let L be a complete lattice. Then $\to : L \times L \to L$ is a* specific conditional *iff it satisfies (1)-(4). Dually, $\backslash : L \times L \to L$ is a* specific difference *iff it satisfies (5)-(8):*

1. $x \wedge y = x \wedge (x \to y)$. *[Modus Ponens]*
2. $x \leq y \iff x \to y = 1$. *[Tautology]*
3. $(x \to y) \wedge (y \to z) \leq x \to z$. *[Transitivity]*
4. $(x_1 \vee x_2 \vee \dots) \to y = (x_1 \to y) \wedge (x_2 \to y) \wedge \dots$. *[De Morgan I]*

5. $x \vee y = x \vee (y \setminus x)$. *[Dual Modus Ponens]*

6. $y \leq x \iff y \setminus x = 0$. *[Contradiction]*

7. $(z \setminus x) \leq (z \setminus y) \vee (y \setminus x)$. *[Triangular Inequality]*

8. $y \setminus (x_1 \wedge x_2 \wedge \dots) = (y \setminus x_1) \vee (y \setminus x_2) \vee \dots$. *[De Morgan II]*

Plausible alternative names for 'specific conditional' and 'specific difference' are *intensional difference* and *extensional difference*, respectively. From now on, we will only focus on the properties of the conditional, since those of the difference can be obtained by duality. Notice that (2) and (6) can be replaced by the equations $x \to x = 1$ and $x \setminus x = 0$, respectively:

Corollary 1. *Let L be a complete lattice where $\to: L \times L \to L$ satisfies Modus Ponens, Transitivity and De Morgan. Then $\to$ satisfies Tautology iff it satisfies $x \to x = 1$ [Identity].*

Proof. Let the conditional satisfy Identity and $x \leq y$, then $x \to y = 1 \wedge (x \to y) = (y \to y) \wedge (x \to y) = (x \vee y) \to y = y \to y = 1$, by De Morgan and Identity. Conversely, let $x \to y = 1$. Then $y = y \wedge (x \vee y) = (x \vee y) \wedge (x \vee y \to y)$ and $x = x \wedge (x \vee y) = (x \vee y) \wedge (x \vee y \to x)$ by Modus Ponens, from which $(x \vee y) \to x = ((x \vee y) \to x) \wedge 1 = [(x \vee y) \to x] \wedge (x \to y) \leq (x \vee y) \to y$ follows by Transitivity. Therefore, $x = (x \vee y) \wedge (x \vee y \to x) \leq (x \vee y) \wedge (x \vee y \to y) = y$. The other direction is trivial. □

Proposition 10. *Let $(L, \to)$ be a complete lattice with specific conditional. Then:*

1. $x \leq y \Rightarrow z \to x \leq z \to y$ *and* $y \to z \leq x \to z$. *[$Monotonicity_I$]*

2. $x \leq y,\ z \leq w \Rightarrow w \to x \leq z \to y$ *and* $y \to z \leq x \to w$. *[$Monotonicity_{II}$]*

3. $x = 1 \to x$ *and* $x \to x = 1$. *[Identity]*

4. $x \leq y \to x$ *and* $y = y \wedge (x \to y)$. *[Weakening]*

5. $x \leq y \iff x = y \wedge (y \to x)$. *[Order]*

6. $(x \to y) \to z \leq (x \to y) \to (x \to z)$. *[$Auto-distributivity_I$]*

7. $x \to (y \to z) \leq (x \to y) \to (x \to z)$. *[$Auto-distributivity_{II}$]*

8. $x \to (y \to z) = y \to (x \to z)$. *[Permutation]*

9. $x \leq y \to z \Rightarrow x \wedge y \leq z$. *[Half-Galois]*

Proof. (1) Let $x \leq y$, by De Morgan $(y \to z) \wedge (x \to z) = (x \vee y) \to z = y \to z$. By Transitivity and Tautology $z \to x = (z \to x) \wedge 1 = (z \to x) \wedge (x \to y) \leq (z \to y)$. (2) Let $x \leq y$, $z \leq w$. By (1), $w \to x \leq w \to y$ and $z \to x \leq z \to y$, so by De Morgan $w \to x = (w \vee z) \to x = (w \to x) \wedge (z \to x) \leq (w \to y) \wedge (z \to y) \leq z \to y$. We apply this again to obtain the other one. (4) By Identity and Monotonicity. (5) By Modus Ponens. (6) By Weakening and Monotonicity. (7) $x \to (y \to z) = (1 \to x) \to (y \to z) \leq [(1 \to x) \to y] \to [(1 \to x) \to z] = (x \to$

$y) \to (x \to z)$ by the previous one. (8) $x \to (y \to z) \leq (x \to y) \to (x \to z) \leq y \to (x \to z)$ by the previous ones, Weakening and Monotonicity. Applying again the inequality we get $x \to (y \to z) = y \to (x \to z)$. (9) If $x \leq y \to z$ then $y \wedge x \leq y \wedge (y \to z) = y \wedge z \leq z$. □

The conditional does not distribute over meets and does not satisfy the other half of the last property either[27]. An explicit counterexample is this. Let the context be $S = \{x, y, z\}$, $Q = \{P, R, T\}$ where $i(P) = \{x, z\}$, $i(R) = \{y\}$, $i(T) = \{x\}$. Then its concept lattice is the pentagon $0 = (\varnothing, Q)$, $1 = (\{x, y, z\}, \varnothing)$, $K_R = (\{y\}, \{R\})$ and $(\{x\}, \{T\}) = K_T \leq K_P = (\{x, z\}, \{P\})$. Then although $0 = K_R \wedge K_P \leq K_T$, we have $K_R \not\leq K_P \to K_T = K_T$.

The last step is to check that the operations in the concept lattices are indeed specific conditionals and differences. Although this fact could be proven directly, we are going to give a slightly more general explanation. It turns out that the two operations can be defined in every complete lattice relative to some chosen join-dense and a meet-dense subsets:

Definition 23. *Let L be a complete lattice, $x \in L$ and $X \subseteq L$. Then x is* join-irreducible *iff if $x = y \vee z$ then $x = y$ or $x = z$. X is* join-dense *in L iff for each $x \in L$ there is a $A \subseteq X$ such that $x = \bigvee A$.*

We obtain by duality the notions of *meet-irreducible element* and *meet-dense* subset. A join (meet)-dense subset allows us to get any element in the lattice as a join (meet)-combination of elements in the set. Let us recall some basic properties of dense subsets:

Lemma 3. *Let L be a complete lattice, $X \subseteq L$ join-dense and $Y \subseteq L$ meet-dense. Let $x \in L$ and $j(x) :=\downarrow x \cap X$ and $m(x) :=\uparrow x \cap Y$. Then:*

i $x \leq y \Leftrightarrow j(x) \subseteq j(y) \Leftrightarrow m(y) \subseteq m(x)$.

ii $x = 0 \Leftrightarrow j(x) = \varnothing \Leftrightarrow m(x) = Y$.

iii $x = 1 \Leftrightarrow j(x) = X \Leftrightarrow m(x) = \varnothing$.

iv $j(x_1 \wedge x_2 \wedge \dots) = j(x_1) \cap j(x_2) \cap \dots$

v $m(x_1 \vee x_2 \vee \dots) = m(x_1) \cap m(x_2) \cap \dots$

We now show that each join-dense (meet-dense) subset induces its own specific conditional (difference):

Proposition 11. *Let L be a complete lattice, $X \subseteq L$ join-dense and $Y \subseteq L$ meet-dense. Let $x \in L$ and $j(x) :=\downarrow x \cap X$ and $m(x) :=\uparrow x \cap Y$. Then $x \to y := \bigwedge(m(y) - m(x))$ is a specific conditional and $x \setminus y := \bigvee(j(x) - j(y))$ is a specific difference.*

Proof. Let $x \to y := \bigwedge(m(y) - m(x))$. First, $x \wedge y = \bigwedge m(x) \wedge \bigwedge m(y) = \bigwedge[m(x) \cup m(y)] = \bigwedge[m(x) \cup (m(y) - m(x))] = \bigwedge m(x) \wedge \bigwedge(m(y) - m(x)) = x \wedge (x \to y)$. Second, let $x \leq y$ then $m(y) \subseteq m(x)$ and therefore $m(y) - m(x) = \varnothing$ and $\bigwedge(m(y) - m(x)) = \bigwedge \varnothing = 1$. If $\bigwedge(m(y) - m(x)) = 1$ then $m(y) - m(x) = \varnothing$

[27]The lattice is Heyting (complete) iff satisfies the other half of that last property iff the conditional distributes over meets iff finite meets distribute over arbitrary joins.

and therefore $m(y) \subseteq m(x)$, so $x = \bigwedge m(x) \leq \bigwedge m(y) = y$. Third, $x \to y \wedge y \to z = \bigwedge[m(y) - m(x)] \wedge \bigwedge[m(z) - m(y)] = \bigwedge[m(y) - m(x) \cup m(z) - m(y)] \leq \bigwedge(m(z) - m(x)) = x \to z$, since $m(z) - m(x) \subseteq m(z) - m(y) \cup m(y) - m(x)$. Fourth, $(x_1 \vee x_2 \vee \dots) \to y = \bigwedge(m(y) - m(x_1 \vee x_2 \vee \dots)) = \bigwedge(m(y) - m(x_1) \cap m(x_2) \cap \dots) = \bigwedge(m(y) - m(x_1) \cup m(y) - m(x_2) \cup \dots) = \bigwedge(m(y) - m(x_1)) \wedge \bigwedge(m(y) - m(x_2)) \wedge \dots = x_1 \to y \wedge x_2 \to y \wedge \dots$. The proof for the specific difference is dual. □

A complete lattice has many join-dense (meet-dense) subsets, including the lattice itself and any superset of a join-dense subset. However, sometimes a specific choice is the most natural one. In the case of concept-lattices, the set of join-irreducible elements (the infimae species) forms a join-dense subset and the set of meet-irreducible elements (the maximal genera) forms a meet-dense subset [40][28]. One can check that, in fact, these induce the specific conditional and difference operations in the concept lattice. Our join-irreducibles have the form $(ei(x), i(x))$, whereas the meet-irreducibles have the form $(e(R), ie(R))$:

Corollary 2. *Let L be a concept lattice, $K \to K' := (e(B' - B), ie(B' - B))$ and $K \setminus K' := (ei(A - A'), i(A - A'))$. Then $K \to K' = \bigwedge(m(K') - m(K))$ and $K \setminus K' = \bigvee(j(K) - j(K'))$.*

Proof. We prove it for the difference. If A is an extension, then $x \in A$ iff $ei(x) \subseteq A$. We have $\bigvee(j(K) - j(K')) = \bigvee\{\, (ei(x), i(x)) \in L \mid ei(x) \subseteq A \;\;\&\;\; ei(x) \nsubseteq A' \,\}$. Therefore, the extension of this kind is $ei(\bigcup\{\, ei(x) \mid x \in A - A' \,\})$. But $A - A' \subseteq \bigcup\{\, ei(x) \mid x \in A - A' \,\}$ therefore $ei(A - A') \subseteq ei(\bigcup\{\, ei(x) \mid x \in A - A' \,\})$. Since each $x \in A - A'$ is such that $ei(x) \subseteq ei(A - A')$, we get $\bigcup\{\, ei(x) \mid x \in A - A' \,\} \subseteq ei(A - A')$, therefore $ei(\bigcup\{\, ei(x) \mid x \in A - A' \,\}) \subseteq eiei(A - A') = ei(A - A')$. So $\bigvee(j(K) - j(K')) = K \setminus K'$. □

Let us take a closer look at the properties of these two operations.

3.4.2 Definitions in terms of Genera and Specific Differences

We can read $K' \to K$ as "the intensional difference of K with respect to K'" and $K' \setminus K$ as "the extensional difference of K' with respect to K". First of all, we have those properties related to the definition of a species in terms of the specific differences and genera that we wanted:

$$K = K' \wedge (K' \to K) \iff K \leq K' \iff K' = K \vee (K' \setminus K)$$

The formula corresponds to classical Aristotelian definitions. Take the extensional difference. A genus K' is the sum of its species K with the extensional difference of K' with respect to K. We have $(A', B') = (A, B) \vee (ei(A' - A), i(A' - A)) = (ei(A \cup ei(A' - A)), B \cap i(A' - A))$. The genus is obtained by selecting from the intension of the species those attributes shared by all the objects that belong to the genus but that do not belong to the species. For instance, the genus *Animal* is obtained by restricting the intension of *Human* to those attributes shared by non-human animals. Now consider the intensional difference.

[28]This fact is the core of the proof of the converse of the Fundamental Theorem of Concept Lattices, which says that each complete lattice is isomorphic to a concept lattice, see [40].

A species K is the overlapping of its genus K' with the intensional difference of K with respect to K'. We have $(A, B) = (A', B') \wedge (e(B - B'), ie(B - B')) = (A' \cap e(B - B'), ie(B' \cup ie(B - B')))$. The species is obtained by selecting from the objects in the genus those that have all the attributes 'specific' to the species, that is to say, those that are not shared by all the objects in the genus. The species *Human* is obtained by restricting the extension of *Animal* to all of those that are rational.

It turns out that, just as the requirement that the lattice should be a tree, the classical picture regarding the specific difference is too simple. When we subtract one kind from another there are two ways of doing so. We can define a kind intensionally by overlapping one of its genera with the intensional difference, or we can define the kind extensionally by adding to one of its species the extensional difference. These two operations are distinct and dual to each other. For instance, if K is a species of K', we get $K = K' \wedge (K' \to K)$, but generally $K \vee (K' \to K) = K' \to K \neq K$. When we get a species by overlapping the genus with the intensional difference, we cannot recover the genus by joining the intensional difference to the species. In order to do that we have to make use of the extensional difference.

In contrast, other classical theses do hold. For instance, there is a sense in which one can get every species as a 'division' of the summum genus. Since $K \leq 1$ we have $K = 1 \wedge (1 \to K) = 1 \to K$. In other words, if we were given the summum genus 1 and the intensional specific differences of each kind with respect to the former, we could obtain each kind by overlapping the summum genus with the corresponding specific difference.

Notice what happens to the specific differences if the lattice is a tree. Suppose that $K \leq K'$, then since $K = K' \wedge (K' \to K)$ we have that either $K' \leq K' \to K$ or $K' \to K \leq K'$. In the former case, we have that $K' = K' \wedge K' \to K = K' \wedge K = K$, therefore $K = K'$. In the latter case, we have that $K = K \wedge K' = K' \wedge K' \to K = K' \to K$ and therefore $K = K' \to K$. In other words, if the lattice is hierarchical and K is a proper species of K', then the intensional difference of K with respect to K' is simply K. In a tree we lose information regarding what makes the species K *intensionally* different from other species of the same genus K'. Something like this need not happen regarding the extensional difference, so this suggests that trees are in a sense *biased* towards extensional differences.

For a different special case, consider what happens when the lattice is a boolean algebra. There are several conditions under which the lattice of kinds is boolean, such as:

Corollary 3. *Let L be a concept lattice. Then L is a boolean lattice iff $x \wedge y \leq z \Rightarrow x \leq y \to z$ and $(x \to 0) \to 0 = x$.*

When the lattice is boolean, the intensional difference $K \to K'$ is the boolean conditional $\neg K \vee K'$, whereas the extensional difference $K \setminus K'$ is the boolean difference $K \wedge \neg K'$. In the boolean case, the two De Morgan laws hold for both operations. For example, $(x \wedge y) \to z = (x \to z) \vee (y \to z)$. As will be shown later, these two operations induce negations that collapse into a classical negation in the boolean case. A fortiori, this implies that this internal negation satisfies principles which are questionable if the negation is to be understood as an Aristotelian internal negation.

3.4.3 Laws of Specific Difference

Besides the one just explained, the most interesting property is monotonicity. Let K, K', J and J' be kinds. Assume that $K \leq K'$ and $J \leq J'$, then we have:

$$J' \to K \leq J \to K' \text{ and } K' \to J \leq K \to J'$$
$$K \setminus J' \leq K' \setminus J \text{ and } J \setminus K' \leq J' \setminus K$$

These reflect how certain kinds get more general or more specific depending on others. Let us consider what it means by analogy with the case of magnitudes and the arithmetical operations of division and difference[29]. Let us imagine that K, K', J and J' were determinate values of some determinable magnitudes ordered by specificity relations. Suppose that K and K' are determinates of the same determinable (e.g. pressure) and K is a smaller value than K', i.e. $K \leq K'$, whereas J and J' are determinates of another determinable (e.g. volume) and again $J \leq J'$. Let us think of the specific conditional $X \to Y$ as if it were the quotient $Y \div X$ between the two magnitudes. Monotonicity says that the quotient $J' \to K$ between two magnitudes increases to $J \to K'$ as the numerator K increases to K' and the denominator J' decreases to J. And the quotient $K \to J'$ decreases to $K' \to J$ as the numerator J' decreases to J and the denominator K increases to K'. Now we think of the difference $X \setminus Y$ as the subtraction $X - Y$. Monotonicity says that the difference $K \setminus J'$ between two magnitudes increases to $K' \setminus J$ as the magnitude being subtracted from increases from K to K' and the magnitude being subtracted decreases from J' to J. And the difference $J' \setminus K$ decreases to $J \setminus K'$ as the magnitude being subtracted from decreases from J' to J and the magnitude being subtracted increases from K to K'. What we have here in our case is a purely qualitative picture that mirrors this correspondence.

Sometimes Kant's Law was formulated as a literal proportion or quotient between extension and intension. According to one of these formulations, the quantity of the extension of a concept is *inversely proportional* to the quantity of its intension. This formulation is very suspicious. Some kinds might have either empty extension or empty intension. Even if we restricted our attention to non-trivial kinds, why would the amount of objects in the extension of an arbitrary kind (or of a concept, for Kant) be proportionally related to the amount of attributes in the intension? Unless we make additional assumptions, these cardinalities can be arbitrary. However, the idea that there is a relation between the extension and the intension that resembles a quotient can be partially preserved by following the previous remarks. Let us baptize the following laws as the:

Law of Specific Conditional the specific conditional or intensional difference between two kinds increases (decreases) as the consequent increases (decreases) and the antecedent decreases (increases).

Law of Specific Difference the specific difference or extensional difference between two kinds increases (decreases) as the kind being subtracted from increases (decreases) and the kind subtracted decreases (increases).

[29] Of course, the analogy has its limitations. It is used usually for the special case of Heyting operations.

Suppose that K is a species of K'. In Kantian terms, the intensional difference $K' \to K$, is a concept whose range of application gets enlarged as the species gets more general and the genus gets more specific, and whose content gets enriched as the species gets more specific and the genus gets more general. The extensional difference $K' \setminus K$ is a concept whose range of application gets enlarged as the genus gets more general and the species more specific, and whose content gets enriched as the genus gets more specific and the species gets more general. Thus, the extension of the intensional difference between two kinds gets bigger as the extension (intension) of the consequent gets bigger (smaller) and the extension (intension) of the antecedent gets smaller (gets bigger), and dually for the extensional difference. Despite the fact that these are very simple ordinal laws that hold between every pair of kinds, they give us more insight into the duality that holds between the extension and intension of kinds.

3.4.4 Negations and the Hexagon of Opposition

In this section it is shown that each of the specific differences induces an internal negation that behaves non-classically. The interaction between these negations and the specificity relations between kinds can be illustrated by making use of an instance of the *Hexagon of Opposition* [6], namely the *Hexagon of Inner and Outer Negations* discussed by [84].

It is well-known that Aristotle distinguished between two kinds of negation, namely *external* or *propositional negation* and *internal* or *term negation.* The proposition *that x is not mortal* involves a propositional negation, it says that the term 'mortal' does not apply to an object x. The proposition *that x is non-mortal (immortal)* involves a term negation, for it says that the negative term 'immortal' applies to the object x. If x is $non-mortal$ (or $immortal$), then it follows that x is not $mortal$. However, the converse does not hold, since there are entities to which neither the predicate 'immortal' nor the predicate 'mortal' apply (e.g. stones or machines). This distinction is present also in Kant's logic. Kant classified the forms of judgements according to quantity, quality, relation and modality. In his Jäsche lectures on logic [63], he says (in what follows the *sphere* is the extension):

> As to quality, judgements are either *affirmative* or *negative* or *infinite*. In the *affirmative* judgement the subject is thought *under* the sphere of a predicate, in the *negative* it is posited *outside* the sphere of the latter, and in the *infinite* it is posited in the sphere of a concept that lies outside the sphere of another.

Whereas the quantificational hexagon classifies judgments according to their quantity in universal and particular (singular propositions are a special case), the internal or term negations allow making distinctions corresponding to quality (see [84] for a discussion in the context of Kant's Antinomies). *Affirmative judgements* have the form $K \leq K'$, they state that the extension of one kind is included in the extension of another. *Negative judgements* have the form $K \nleq K'$, they state that the extension of one kind is not included in the extension of another. For instance, there may be a K which is not also a K'.

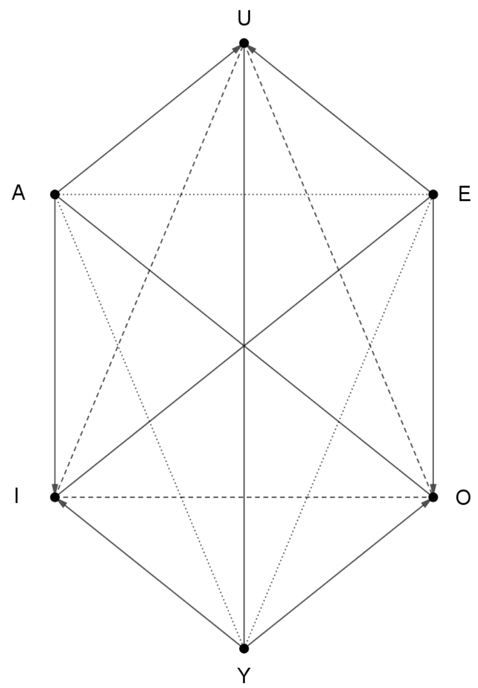

Figure 3.2: Hexagon of Opposition

Therefore, negative judgements correspond to propositional negations. Nevertheless, this does not immediately imply that K-s belong to a certain non-trivial negative kind. *Infinite judgements* have the form $K \leq \neg K'$, they state that the extension of one kind is included into the extension of a kind which is opposite or negative to it. These correspond to term negations. Due to contrariety, infinite judgements imply negative judgements, but the converse does not hold.

The interaction between the two kinds of negation and the specificity relations can be illustrated by making use of the *Hexagon of Opposition* [6].[30] Consider the six kinds of basic general propositions that form the Hexagon, as shown in Figure 3.2. Here $[A]$-$[O]$; $[E]$-$[I]$; $[U]$-$[Y]$ are contradictories. $[A]$-$[E]$-$[Y]$ are contraries, $[I]$-$[O]$-$[U]$ are subcontraries and the rest are relations of implication (subalternation). We could instantiate the hexagon in its *quantificational* form by considering non-null kinds. For instance, we could consider $[E]$ propositions to be of the form "No K-s are K'-s: $K \wedge K' = 0_{L^*}$" (see [88]). However, we will focus instead on the hexagons induced by the interaction between external and internal negations. Consider the following instance of the *Hexagon of Inner and Outer Negations* [84]:

i All K-s are either K'-s or *not*-K'-s: $K \leq K'$ or $K \leq \neg K'$. [U]

ii All K-s are K'-s: $K \leq K'$. [A]

[30] The theory behind the traditional Aristotelian Square of Opposition has been generalized and promoted mainly by J-Y Béziau [6] and it has now many applications (for recent work see [7]). Here we will only make use of the standard *Hexagon of Opposition*, which was first introduced by Blanché.

iii All K-s are *not*-K'-s: $K \leq \neg K'$. [E]

iv Some K-s are not *not*-K'-s: $K \nleq \neg K'$. [I]

v Some K-s are not K'-s: $K \nleq K'$. [O]

vi Some K-s are neither K'-s nor *not*-K'-s: $K \nleq K'$ and $K \nleq \neg K'$. [Y]

As McLaughlin and Schlaudt argue, introducing [U] and [Y] propositions gives more information regarding the behavior of the negations. Here [A] propositions (affirmative judgements) and [E] propositions (infinite judgements) are contraries. Although no K can be both a K' and a *not*-K', it might happen that some K-s are neither K'-s not *not*-K'-s, which is what [Y] propositions express. In other words, the Excluded Middle fails for term negation. In contrast, [U] propositions say that K-s are objects for which it makes sense to say that they are either K'-s or *not*-K'-s. In other words, all K-s are included among the K'-candidates. That infinite judgements imply negative ones is seen in the subalternation relation that holds between [E] propositions and [O] propositions.

It will be shown now that the operations of specific difference introduced in the previous sections induce their own internal negations. In other words, the Minimal Conception implies the existence of two sorts of *negative kinds*. The strategies followed in those sections were inspired by Wille's attempts in [142] of finding appropriate negations in a concept lattice. As he argues, there are two natural negation-like operations:

$$\neg K = (e(B^c), ie(B^c)) \qquad\qquad *K = (ei(A^c), i(A^c))$$

Whereas the one on the left takes the objects sharing those attributes that are not in K, the one on the right takes the attributes shared by all the objects that are not in K. Wille abstracts the properties of these operations to study the resulting lattices, which he calls:

Definition 24. *Let* $(L, \wedge, \vee, \neg, *, 0, 1)$ *be a bounded lattice and* $\neg, *\colon L \to L$ *two monadic operators. Then* L *is a* dicomplemented lattice *iff:*

1. $**x \leq x$. *[Intensiveness]*
2. $x \leq y \Rightarrow *y \leq *x$. *[Antitonicity]*
3. $x = (x \wedge y) \vee (x \wedge *y)$.
4. $x \leq \neg\neg x$. *[Extensiveness]*
5. $x \leq y \Rightarrow \neg y \leq \neg x$. *[Antitonicity]*
6. $x = (x \vee y) \wedge (x \vee \neg y)$.

He calls $*$ the *weak negation* and $\neg$ the *weak opposition*. We will keep the names, although a better choice would have been 'extensional negation' for $*$ and 'intensional negation' for $\neg$. Some examples are:

i. Every boolean algebra is dicomplemented, where $\neg = *$ is the boolean complement.

ii. The concept lattice of a formal context is dicomplemented, where the negation operations are $\neg K = (e(B^c), ie(B^c))$ and $*K = (ei(A^c), i(A^c))$.

Some basic properties of these negations are:

Proposition 12. *Let L be a dicomplemented lattice. Then:*

1. $\neg\neg\neg x = \neg x \leq *x = ***x$.
2. $\neg * x \leq **** x = ** x \leq x \leq \neg\neg x = \neg\neg\neg\neg x \leq *\neg x$.
3. $x \wedge \neg x = 0$ *and* $x \vee *x = 1$.
4. $\neg(x \vee y) = \neg x \wedge \neg y$ *and* $*(x \wedge y) = *x \vee *y$.
5. $\neg 0 = 1 = *0$ *and* $\neg 1 = 0 = *1$.

Now we show that these negations can be obtained from the operations of specific difference introduced in the previous sections:

Proposition 13. *Let* $(L, \rightarrow)$ *be a complete lattice with specific conditional. Then* $\neg x = x \rightarrow 0$ *satisfies:*

1. $y \leq \neg x \iff x \leq \neg y$.
2. $x \leq y \Rightarrow \neg y \leq \neg x$.
3. $x \leq \neg\neg x$.
4. $\neg x \leq x \rightarrow y$. *[Explosion]*
5. $\neg x \vee y \leq x \rightarrow y$. *[Disjunctive Syllogism]*
6. $x = (x \vee y) \wedge (x \vee \neg y)$.
7. $\neg(x_1 \vee x_2 \vee \dots) = \neg x_1 \wedge \neg x_2 \wedge \dots$. *[De Morgan Opposition]*

Proof. (1) $y \leq \neg x = x \rightarrow 0$ iff $y \rightarrow (x \rightarrow 0) = 1 = x \rightarrow (y \rightarrow 0)$ iff $x \leq y \rightarrow 0 = \neg y$ by permuting and Tautology. (2)-(3)-(4) Follow from Monotonicity. (5) By Explosion and Weakening. (6) We just need to prove half of it. By Disjunctive Syllogism, De Morgan and Identity $(x \vee \neg y) \leq y \rightarrow x = (y \rightarrow x) \wedge 1 = (y \rightarrow x) \wedge (x \rightarrow x) = (x \vee y) \rightarrow x$, therefore $(x \vee \neg y) \wedge (x \vee y) \leq x$. (7) By De Morgan. □

Therefore, if $\rightarrow$ is a specific conditional, then $\neg x = x \rightarrow 0$ is a weak opposition. Dually, if $\backslash$ is a specific difference, $*x = 1 \backslash x$ is a weak negation. In this way we obtain Wille's negations as special cases of the differences. It is easy to check that, in the case of concept lattices, the weak opposition is $\neg K = (e(B^c), ie(B^c))$ and the weak negation is $*K = (ei(A^c), i(A^c))$, as expected.

The operations of weak negation and weak opposition behave like term negations of some sort. Stating that it is not the case that K is K', $K \not\leq K'$, is not equivalent to stating that K is a *not*-K'. The latter corresponds to an infinite judgement involving a negative kind, the former does not. Whereas the weak opposition satisfies the principle of non-contradiction $K \wedge \neg K = 0$, the weak negation satisfies the excluded middle $K \vee *K = 1$. But the weak opposition usually does not satisfy the excluded middle, nor does the weak negation satisfy the non-contradiction either. Moreover, for each of these negations contraposition fails. Although if K is a species of K' then $\neg K'$ is a species of $\neg K$, the converse does not hold, because $\neg\neg K$ is usually different from K (as in the intuitionistic case). This is what one would expect from an internal negation. All mortals are not-immortals, for no entity can be both mortal and immortal. But some entities to which neither the term 'mortal' nor the term 'immortal' applies could be not-immortal.

For the weak opposition, one considers all the attributes that are not in the intension of a kind and takes the objects exemplifying all those attributes. The extensions of both kinds only overlap in those objects that belong to the null kind (which is usually empty). In the case $K \vee \neg K \neq 1$ this means that an object x may be neither a K nor a $\neg K$. The reason is that if an object does not exemplify an attribute in B, it does not follow that it exemplifies all the attributes that are not in B. Of course, it still holds that either $K_x \leq K$ or $K_x \not\leq K$. So either x belongs to the extension of K or it does not. Note that it is not that it is indeterminate whether x is a K or a $\neg K$. It is perfectly determinate that it is neither of them. The point is that x does not have all those properties needed to be a K nor all those properties needed to be a $\neg K$. To use a classical example, a stone is neither mortal nor immortal, because for something to be either of them it should be alive in the first place.

For the weak negation, one considers all the objects that are not in the extension of K and takes all the attributes that all these objects exemplify. The resulting extension may considerably overlap with that of K. In the case $0 \neq K \wedge *K = (A \cap ei(A^c), ie(B \cap i(A^c)))$ this means that we have an object that is both a K and $*K$. Object x has all the properties needed to be a K and all the properties needed to be $*K$, so one could say that x is a boundary case of K. Now the reason is that if an attribute is not exemplified by all the objects in A it does not follow that it is exemplified by all the objects not in A. Consider a folk classification of animals that includes the vernacular kind *Fish*. Since the Excluded Middle holds for $*$, every animal is either a Fish or a $*$-Fish. Some aquatic animals that we would consider to be fishes, but that shared many other attributes with animals that are not fishes, might be boundary cases of the kind *Fish*.

When the lattice is boolean, both negations collapse into the boolean negation $*K = \neg K = (e(B^c), i(A^c))$. There is still a difference between propositional and internal negations, since $K \not\leq K'$ is not equivalent to $K \leq \neg K'$. Nevertheless, the double negation and contraposition laws hold, so we have that $\neg\neg K = K$ and $K \leq K' \Leftrightarrow \neg K' \leq \neg K$.

3.4.5 Differences Between this Approach and the Traditional Aristotelian Picture

Let us recall the operations and special elements in the lattice of kinds:

i. Logical Determination: $(A, B) \wedge (A', B') = (A \cap A', ie(B \cup B'))$.

ii. Logical Abstraction: $(A, B) \vee (A', B') = (ei(A \cup A'), B \cap B')$.

iii. Specific Conditional: $(A, B) \to (A', B') = (e(B' - B), ie(B' - B))$.

iv. Specific Difference: $(A, B) \setminus (A', B') = (ei(A - A'), i(A - A'))$.

v. Weak Opposition: $\neg(A, B) = (e(B^c), ie(B^c))$.

vi. Weak Negation: $*(A, B) = (ei(A^c), i(A^c))$.

vii. Summum Genus: $1 = (S, i(S))$.

viii. Null Kind: $0 = (e(Q), Q)$.

ix. Infimae Species: $x = (ei(x), i(x))$.

x. Maximal Genera: $R = (e(R), ie(R))$

If one accepts the Minimal Conception as an accurate description of what kinds are, then one is forced to accept the view of kinds as given by the model we just saw in this chapter. However, one may object that there are some crucial differences between the classical Aristotelian picture and the one given in this chapter. According to this objection, the model would be revisionary. If some contemporary theories of kinds, such as essentialism, follow the classical Aristotelian picture, then this would seem to imply that there is a tension between the model given here and what were supposed to be some clear examples of it, namely essentialist theories.

First of all, whereas the Aristotelian picture could be thought to be committed to the hierarchy thesis, hierarchical structures have been shown to be a very special case of the approach developed in this chapter. However, it is not clear that the Aristotelian picture of kinds is really committed to the hierarchy thesis, for instance, the syllogistic logic of Corcoran and Martin does not make any assumptions regarding the hierarchy constraint. Furthermore, we have seen some reasons for why the hierarchical constraint may be too strong. If the lattice is hierarchical and finite, the extension of each kind gets partitioned (not necessarily dichotomously) into the extensions of the kinds that it covers (the only objects common to them are those that belong to the null kind). So whenever the extensions of the species of a common genus overlap, as it happens when kinds are described by several overlapping classifications of the same domain [136], the hierarchy thesis fails. In fact, for this very same reason some contemporary essentialists such as Ellis [34] have denied that the thierarchy thesis holds for kinds. Thus if the Aristotelian picture is really committed to the hierarchy condition, the picture given by the Minimal Conception of Kinds is more general and cautious, while still contains the former one as a special case.

Second, the specific differences of kinds are themselves kinds. Furthermore, the specific difference between two kinds exists even if one is not a species of the other, whereas in the classical picture the specific difference is defined only between a species and the genus that is immediately above it. Since the existence of the specific differences is guaranteed by the fact that the lattice is complete, one may object to it by saying that it is an unwelcome consequence of completeness. However, as we previously saw, every bottomed meet-semilattice becomes a complete lattice if a top element is attached to it. The assumption of the existence of a summum genus is not too strong, after all, it is plausible to say that given a domain to be classified there is a genus to which all the objects belong. A fortiori, the specific differences are defined in every finite lattice and the corresponding syllogistic logic is not far from having the resources to express Aristotelian definitions of terms by making use of the syntactic analogues of specific differences. Moreover, note that *if* one accepts the Minimal Conception of Kinds, *then* one is forced to accept the existence of these specific difference kinds. So if the Aristotelian picture rejects the existence of these kinds, it has to reject the Minimal Conception too. This seems odd considering that contemporary essentialists defend theses that are stronger than the Minimal Conception and which they take to be inspired by Aristotle's metaphysics of kinds.

Third, the specific difference (intensional) that gets subtracted from the genus to obtain the species is usually different from the specific difference (extensional) added to the species to get the genus back. At first, this may look strange. However, the special case of boolean algebras can be used to explain the difference, since the conditional and difference operations have distinct logical properties. Maybe the objection is that, granted that two such operations exist, only one of them corresponds to Aristotelian specific differences. However, which one is that?

Fourth, there are two different non-classical internal negations that induce negative and opposite kinds. Aristotelian internal negations seem to behave like weak oppositions, or even more strongly than these. However, they do not seem to be boolean negations. Due to the fact that Aristotle recognized the difference between internal and external negation, there is a controversy regarding whether the negation of terms should satisfy classical principles such as contraposition or double negation. Moreover, if these negations are to behave in the classical special case of boolean structures as the classical negation, they have to be induced by the specific differences as it was described in this chapter.

Despite these differences, the model preserves the core of the classical Aristotelian approach, which is the one stated in the Minimal Conception. One could get closer to more familiar pictures by forcing the lattice to be hierarchical or by requiring it to have such a structure (say, boolean) that the distinctions between the negations collapsed. However, what would be the point of doing that? Any attempt to make the lattice closer to such approaches will require the context to satisfy additional constraints, like the hierarchy condition, which will force the order structure to have an implausible shape or the attributes to be closed under conditions that conflict with their being sparse. Moreover, these two approaches, the tree-approach and the boolean-approach, are in fact incompatible. If a lattice is both a tree and a boolean algebra then it is either the boolean algebra of two atoms or that of one atom. For the rest of cases, we cannot have both. For instance, traditional Porphyrian trees are counterexamples to the claim that the lattice of kinds must be boolean. One may reject booleaness and ask at least for distributivity, in order that the lattice of kinds behaves like usual logics do. However, trees tend not to be distributive. Take a bottomed tree L with at least three coatoms x, y, z (elements immediately under the top element). Since these elements are incomparable, they are disjoint, a fortiori, the tree contains a copy of the diamond lattice $0 \leq x, y, z \leq 1$. It is a standard result that if this is the case then the lattice is not distributive. So if we want to keep trees as a special case we cannot assume that the lattice is distributive.

3.5 A Modal Picture

At this point, the reader may be puzzled about the absence of modality in the previous considerations. The notion of a natural kind is clearly linked to modality. The extension of a natural kind can change from one possible world to another. There could have been some natural attributes that do not actually obtain. As a consequence, there could have been natural kinds that were different from the ones we find in the actual world. Moreover, if the exemplification relation I is not that of essential exemplification, objects could have exemplified attributes different from those that they actually exemplify. The previous Chapter was devoted to introduce the basic commitments of essentialist theories of natural kinds. Every conception of natural kinds should say something about these issues, so why have I disregarded them?

There are several reasons why I have not considered modality in detail. The main one is simply that it makes the picture more complex. Chapter V will show that the connections between the similarity relations that hold between the objects, their attributes and the species-genus order structure are already complicated enough without even considering the role of modality. Nevertheless, the discussions in the previous chapter centred around modal issues and the answer to the Achilles' heel of any class nominalism, namely the coextensionality problem, requires a direct appeal to modality. For these reasons, I should give some brief remarks regarding the relation between modality and natural kinds. The way I will frame modality in the concept-lattice model takes some insights from [88].

Let us go back to our model (S, Q, I). Suppose that (S, Q, I) is the whole *modal space*, where S is the set of *possible objects*, Q is the set of *possible natural attributes* and the *relation of exemplification* I holds xIP iff "object x could have exemplified natural attribute P".

One's assumptions regarding how narrow or broad the modal space is are reflected in the exemplification relation I. For instance, under essentialist constraints relation I may be a 'small' subset of $S \times Q$, whereas anti-essentialism will require I to be very close to $S \times Q$. Possible worlds can be taken to be represented by subrelations of I (see [88]), the set of possible worlds being some $W \subseteq \wp(I)$. Nevertheless, since the domain of objects, the domain of attributes and the exemplification relation may change from one world to another we should not forget that a possible world w is better represented by the subcontext $w = (S_w, Q_w, I_w)$ of the modal space induced by such a subrelation. That is to say, we have that $S_w \subseteq S$, $Q_w \subseteq Q$ and $I_w \subseteq I$. The domains of objects and attributes are obtained by projection from I_w, in other words, $S_w := \{x \in S \mid \exists P \in Q \ (x, P) \in I_w\}$ and $Q_w := \{P \in S \mid \exists x \in S \ (x, P) \in I_w\}$[31]. Any such a world comes with its own extension and intension operators e_w and i_w, which give for any set of attributes and any set of objects their corresponding extension and intension in w, respectively. To distinguish between the extension (intension) operator of the modal space and the extension (intension) operator of a given world, we will call the former ones the *global extensions* (intensions) and the latter ones the *local w-extensions* (intensions) at world w. In the case of an object x in the modal space, its *w-intension* is the set of attributes $i_w(x)$ that x has in world w. If P is an attribute in the modal space, its *w-extension* is the set of exemplars $e_w(P)$ that P has in world w. The extensions and intensions differ from one world to another. For now on we will just assume that every possible world is trivially metaphysically accessible from any other (so that the corresponding logic turns out to be S5).

If $w \in W \subseteq \wp(I)$, then since I satisfies the No-Bare Particulars (i.e. $i(x) \neq$

[31] This seems to be at odds with some contingentist views on properties, according to which nonexistent properties may still be instantiated at worlds. I thank Bruno for this remark.

$\varnothing$) and Exemplification (i.e. $e(P) \neq \varnothing$) Principles, w satisfies them too. For if $x \in S_w$, then $(x, P) \in w$ for some $P \in Q_w$ and if $P \in Q_w$, then (x, P) for some $x \in S_w$. But if some of the attributes in Q are essentially exemplified by some objects in S (the notion of essence here might not need the modal one), then this will induce closure conditions on the possible worlds. For instance, suppose that x essentially exemplifies P, then any possible world that contains x must also contain P:

$$\text{If } x \text{ essentially exemplifies } P\text{, then } \forall w \in W\ (x \in S_w \Rightarrow P \in Q_w\ \ \&\ \ xI_wP)$$

Moreover, we may want to exclude $\varnothing$ as a possible world too. We may select one of these possible worlds as the actual one, say $@ \subseteq I$. We will not require possible worlds to be maximally specific regarding the actual one, as it is sometimes done. But if it were so required, we could consider the possible worlds to be the maximal elements in $W - \{I\}$, the worlds (as exemplification relations) being ordered by inclusion. The rest of the elements would be just parts of worlds. In that case the possible worlds would be the biggest parts of the modal space.

As we just said, a possible world w can be represented by the context induced by $I_w \in W$, namely (S_w, Q_w, I_w), where $S_w := \{x \in S \mid \exists P \in Q\ (x, P) \in I_w\}$ and $Q_w := \{P \in Q \mid \exists x \in S\ (x, P) \in I_w\}$. Our possible worlds are ordered by inclusion, which can simply be understood as the parthood relation. Let us take a closer look. We can define the following subrelations of I:

$$\Box_{re} := \{(x, P) \in I \mid \forall w \in W\ xI_wP)\}$$
$$\Diamond_{re} := \{(x, P) \in I \mid \exists w \in W\ xI_wP\}$$
$$\Box_{exre} := \{(x, P) \in I \mid \forall w \in W\ (x \in S_w \Rightarrow xI_wP)\}$$

and simply use prefix notation $\Box_{re}Px$ and $\Diamond_{re}Px$ instead of writing $x\Box_{re}P$ and $x\Diamond_{re}P$. Our de re contexts are now the binary relations "x de re necessarily exemplifies P", "x de re possibly exemplifies P" and "x exemplifies P wherever it exists". Of course, if $\Box_{re}Px$ holds for some object x then x exists in every possible world. These subrelations correspond to possible worlds. So long as each (x, P) is found in some world, we will have $I = \Diamond_{re}$ and this will be the modal space. In contrast, $\Box_{re}$ is the intersection of all the possible worlds $\Box_{re} = \bigcap W$, which may or may not be empty, depending on our additional assumptions (the definition of intersection of contexts is straightforward). It will be empty if we assume there are no necessarily existing entities. The relation $\Box_{exre}$ is the usual weakening of de re modality, $\Box_{exre}Pa$ iff aIP for every world w at which a exists.

Second, identity is a transworld-relation. The Indiscernibility of Identicals holds for global intensions:

$$\forall x, y \in S\ (x = y \Rightarrow i(x) = i(y))$$

if two objects in the modal space are identical, then they have the same possible attributes. Its restriction to local intensions of objects belonging to the same world holds too:

$$\forall x, y \in S_w\ (x = y \Rightarrow i_w(x) = i_w(y))$$

In words, if two objects in world w are identical then their w-intensions are identical. But its restriction to local intensions may fail when the same object is compared in two different worlds. In other words, the following can be false:

$$\forall x \in S_w \ \forall y \in S_{w'} \ (x = y \Rightarrow i_w(x) = i_{w'}(y))$$

the reason is of course that an object x may exemplify different attributes in different worlds.

Third, since each possible world w has its own extension and intension operators, it induces a lattice of natural kinds $\mathbf{B}(w)$. Thus natural kinds can change from one possible world to another. We have here a distinction between *local natural kinds* and *global natural kinds.* A global natural kind is simply a kind in the lattice induced by the modal space (S, Q, I), whereas a local kind is a kind in the lattice induced by a possible world (S_w, Q_w, I_w). A global natural kind has in its extension all its possible instances and in its intension all the attributes that are possibly exemplified by all its instances, whereas a local natural kind has in its extension all its instances in a given possible world w and all the attributes common to these instances in that same world w. For instance, if *global cat* is a global natural kind, then its extension includes all the possible cats and its intension includes those attributes that are common to all and only those possible cats. In our world @, all the actual cats form the extension of a local kind of cats. This distinction does the job we need to treat the coextensionality problem for nominalism in Chapter V. For example, suppose that in the actual world @, we have the kinds $K_@$ = *creature with a heart* and $K'_@$ = *creature with a kidney.* In our world they happen to be coextensional, so $A_@ = A'_@$ and therefore $K_@ = K'_@$. But these are just local kinds.

The previous paragraph may give the wrong impression that given a global kind $K = (A, B)$ and a world w, the corresponding local kind K_w is completely determined. This is not the case. For instance, given K and w there are at least two natural candidates for the corresponding local kind, namely $K_w = (ei(A \cap S_w), i(A \cap S_w))$ and $K'_w = (e(B \cap Q_w), ie(B \cap Q_w))$, depending on whether we restrict the extension or the intension of the global kind. These two local kinds may be distinct in w. We may assume that we made a choice among these options so that, for each global kind K, there is a function $f_K \colon W \to \bigcup_{w \in W} \mathbf{B}(w)$ that maps each world w to the corresponding local kind K_w of K in w. If we want to allow for some global kinds lacking a corresponding local kind in some worlds, we can restrict the domains of these functions accordingly (i.e. the domain may be some $W' \subseteq W$).

Fourth, consider essentialist theses in the model. Suppose that some of the natural attributes in Q are not essentially exemplified by any object. Let $E \subseteq I$ be the relation *essentially exemplifies*, which we will assume from now on as primitive. Informally, xEP iff *x essentially exemplifies P.* Given the essentialist theses not every subrelation of I can be considered to be a possible world. For the set $W \subseteq \wp(I)$ the essentialist will require that the following constraint is satisfied:

$$xEP \Rightarrow \forall w \in W (x \in S_w \Rightarrow x I_w P) \ [Closure]$$

or more succintly:

$$xEP \Rightarrow \Box_{exre} xP \ [Essentialism - DereModality]$$

The converse need not hold given that, as we saw in the previous Chapter, essentialists of natural kinds follow Kit Fine in this regard. From the condition it follows that $P \in Q_w$. If an object x essentially exemplifies a given attribute P, then the attribute will be exemplified by that object in each possible world in which the object exists. By contrast, the same attribute P may exist in a world w because it is (essentially or accidentally) exemplified by a given object y in w without the object x having to exist in that world. In Fine's terms, if it is in the nature of x that x is P, then if x exists, P will exist too. But it is not in the nature of P that x is P, although it is in the nature of P, due to Exemplification, that some object y exemplifying P exists whenever P exists. According to essentialism, the global natural kinds are the elements of $\mathbf{B}(S, Q^*, E)$ where $Q^* = \{P \in Q \mid (x, P) \in E \text{ for some } x \in S\}$ and not the elements of $\mathbf{B}(S, Q, I)$, since the 'kinds' of the latter structure contain some attributes in their intensions that are accidental to their instances.

So we have that $E \subseteq \Box_{exre} \subseteq \Diamond_{re} = I$. Given the definition of a natural kind, for any essentialistic global natural kind $K = (A, B) \in B(S, Q^*, E)$ and its corresponding w-local natural kinds $K_w = (A_w, B_w)$ we have the two basic essentialist theses that we discussed in Chapter II:

$$\exists P_1, P_2, \dots \in Q^* \ (x \in A \Leftrightarrow xEP_1 \ \& \ xEP_2 \ \& \ \dots) \ [General Essentialism]$$

$$x \in A_w \Rightarrow \forall w' \in W \ (x \in S_{w'} \Leftrightarrow x \in A_{w'}) \ [Individual Essentialism]$$

General Essentialism trivially follows from the definition of global kinds, since for an object x to belong to the extension of K it must essentially exemplify all the attributes in its intension. See that a local version of general essentialism follows directly, since for an object to belong to a local kind it must exist in that world and therefore it will exemplify all its essential attributes there too. *Individual Essentialism* says that if an object belongs to a local kind K_w in w, then in every possible world where it exists it also belongs to the local kinds of the same global kind as K_w. It follows from the fact that if an object x belongs to a world w' then it will exemplify in w' all the attributes that it essentially exemplifies (which therefore exist in w' too) and a fortiori will belong to the corresponding local kind. In this way, the model can deal with the fundamental principles of Essentialism that were discussed in Chapter II.

3.6 Conclusion of Chapter III

The aim of this chapter was to explore some minimal assumptions about kinds through a formal model, in order to get a clear picture of the external structure of kinds. Two assumptions were made. First, that kinds are ordered by specificity relations. Second, that kinds are two-dimensional entities. They have an extension, consisting of all the members of the kind, and an intension, consisting of all the sparse attributes of the kind. Moreover, the objects in the extension are exactly all of those having all the sparse attributes in the intension.

In order to study the external structure of kinds, we considered the thesis that kinds are hierarchically arranged in the form of a tree. Some arguments found in the literature suggest that this is too strong an assumption. In particular, crossings within a classification and crossings between classifications make the hierarchy condition fail systematically. However, Thomason's algebraic model of kinds is more general than that and could in principle be used to shed some more light on the external structure of kinds. But there are two problems with Thomason's model. The first one is that it does not give any argument for why kinds should be modelled by using complete lattices. In order to find such an argument, the contemporary approaches to Syllogistic Logic by Corcoran, Smiley and Martin were considered. It turns out that complete lattices can be used as a semantics for syllogistic logic. The second problem is that neither Thomason's model nor Corcoran-Martin's logic provide much information regarding the specificity relations between kinds and they are silent with regards to the relations between the objects and attributes of a kind. In contrast, following a proposal made by Mormann, a model based on Wille's Theory of Concept Lattices, which are complete lattices, was suggested. According to this model, kinds are modelled as pairs (A, B) where A (the extension) is the set of all the objects having all the attributes in B (the intension), and B is the set of all the attributes shared by the objects in A. Moreover, we saw that in this model, the species-genus relations between kinds satisfy Kant's Law of the duality between extension and intension and that these minimal assumptions are enough to induce an Aristotelian conception of definitions of kinds in terms of their genera and specific difference.

More specifically, the theory of concept lattices provides a good model for the Minimal Conception, and it gives us more information regarding what such a conception involves:

1. It matches the ontological assumptions:
   (a) The primitive entities assumed by the model are just a set of Objects, a set of (natural) Attributes and an Object-Attribute relation of Exemplification.
   (b) The primitive principles, namely No-Bare Particulars and Exemplification, hold. Other more controversial principles, like the Identity of Indiscernibles or Coextensionality, need not hold.
   (c) Attributes can be taken to be natural or sparse. In other words, the set of attributes is not closed under any formal operation and attributes ground non-trivial similarities between objects.
2. The description of kinds satisfies the assumptions given in the Minimal Conception:
   (a) Kinds are modelled as pairs consisting of a set of objects and a set of attributes, which correspond to the extension and intension of the kind.
   (b) The order structure of kinds satisfies the Kantian law of extension and intension.
   (c) Object-kind instantiation relations and kind-attribute exemplification relations can be explained easily.
3. The model captures the basics of an Aristotelian theory of definition and syllogistic reasoning:

(a) The model encodes the basic syllogistic reasoning for kinds, since the class of structures used forms a semantics for syllogistic logic.

(b) Meets and joins can be interpreted as operations of logical determination and logical abstraction.

(c) Two new operations of intensional and extensional difference can be introduced that correspond to the specific difference of one kind with respect to another. This leads to the explanation of definitions of species in terms of genera and specific differences.

(d) The specific differences induce two distinct non-classical negations, interpreted as term or internal negations.

4. Other additional features can be handled by the model. In particular, a hierarchical ordering for natural kinds turns out to be a very special case and a sketch of a modal approach can be given by considering subcontexts.

There are several lessons to draw from the previous considerations. First, there is something like a mathematically well-behaved natural kind calculus at our disposal. It is true that there are some differences with the classical Aristotelian approach. Nevertheless, the mathematical theory preserves many of the crucial insights of the Aristotelian logical tradition, including not only syllogistic structure but also the genus-specific difference theory of definition. Second, this very same model shows why the hierarchy condition over the external structure of natural kinds turned out to be a very naive constrain. Simply put, there are too many non-hierarchical ways in which kinds can relate to each other. Indeed, there are as many as kinds of complete lattices. In exchange, we recovered a more fundamental and explanatory principle, namely Kant's Law of extension and intension (and other related properties) and we also got more insight into the nature of Aristotelian definitions of kinds. Finally, the model assumes in its ontology two sorts of primitive entities, objects and attributes, alongside a primitive formal relation of exemplification that holds between them. These attributes behave like universal entities. For instance, each object can exemplify several attributes and each attribute can be exemplified by several objects. Moreover, attributes are assumed to be primitive entities whose identity conditions do not depend on the objects that exemplify them. In contrast, Chapters IV and V suggest how to develop an alternative resemblance nominalistic model that gives a reduction of this model of kinds under some specific circumstances. The goal is to get the concept lattice. To achieve that the nominalist will try to reconstruct a context (S, Q, I) from some primitive particular entities in resemblance relations by using moves analogous to the ones displayed by the realist.

Chapter 4

Resemblance Structure of Natural Attributes

When any objects resemble each other, the resemblance will at first strike the eye, or rather the mind, and seldom requires a second examination ...
A Treatise of Human Nature, I-3
DAVID HUME

The purpose of Chapter IV is to explore a proposal concerning the internal structure of natural kinds, that is to say, about the relations that hold among the members of a kind. According to the Minimal Conception, members of a kind share some (natural) properties. In the model that was introduced in the previous chapter, these properties were the natural attributes. Resemblance nominalists may accept this on the proviso that these properties are conceived as maximal classes of resembling objects. This chapter works in tandem with Chapter V, which is in some sense the formal counterpart of this one. The goal is to reconstruct attributes in a nominalist way [1].

The structure of the chapter is as follows. First, I explain what I will understand in the following by similarity or resemblance by explaining the main formal properties of similarity relations. Second, I will consider some objections to the categorical or binary notion of similarity. In particular, the objections by Tversky to the symmetry of similarity will be discussed in detail. Third, I will review the main sorts of resemblance nominalism, namely egalitarian, aristocratic and collectivist nominalism. In particular, Pereyra's egalitarian resemblance nominalism is discussed in detail and some criticism of his approach is given. I will suggest here that aristocratic resemblance nominalism can provide a successful answer concerning the internal structure of kinds by countering some objections to it. Fourth, I will introduce three different models for aristocratic resemblance nominalism, which form the basis for the discussions in the next chapter. The first model is the polar model, it is a topological model that comes from the work by Rumfitt and Mormann on the conceptual spaces approach to vagueness. The second model is the order-theoretic model, it is based on a suggestion by Mormann to generalize the polar model to the class of weakly-scattered spaces. The third model is the similarity model, which is a new model introduced in this thesis and the main object of study of the next chapter and is based on the theory of similarity structures. Both models involve a commitment to certain objects, to be called 'paradigms', that satisfy some of the constraints previously discussed in the chapter. I discuss the plausibility of these assumptions in detail. Finally, since aristocratic resemblance nominalism posits the existence of paradigmatic objects, I will discuss the merits of two different views on their nature. According to one of these views, paradigms are qualitatively thin objects. According to the other view, paradigms are objects chosen by a subject during his conceptual learning process.

[1] Nevertheless, the goal of Chapters IV-V will be metaphysically quite modest. I will just try to deflect some objections without fully arguing for resemblance nominalism. For a philosophically more developed and full defence of resemblance nominalism (of the egalitarian sort) the reader can consult the work of Rodriguez Pereyra [109].

4.1 Similarity from a Formal Point of View

In Chapter III we saw that realist accounts of natural kinds can account for a Minimal Conception of Kinds. These models give a semantics for syllogistic logic, explain the composition of kinds in terms of some operations that can be interpreted as logical division and abstraction (meets and joins) and subsume the hierarchical models as a special case. Furthermore, the models represent kinds as pairs of an extension and an intension, their species-genus orders as the corresponding inclusions (in such a way that the Kantian law is satisfied), the object-kind instantiation and the kind-attribute exemplification relations indirectly by the membership relation, and so on. The axioms assumed are quite minimalistic and the attributes satisfy the basic constraints for them to be considered natural, since they ground the similarities between objects and they are not closed under every familiar formal operation.

Nevertheless, according to these theories, objects share certain natural 'attributes' that behave like universal entities: each object can exemplify several attributes, each attributes can be exemplified by several objects. Usually, each attribute is exemplified by at least one object, each object exemplifies at least one attribute and two attributes can be exemplified by the same objects. Moreover, these entities are assumed as primitive. This prompts the question whether one could get at the Minimal Conception of Kinds by reducing the realist model to a nominalistically acceptable one. Instead of assuming these attributes as primitive entities, one could try to give a nominalistic explanation of them. I think that this can be done by starting where Quine [107] left, namely by (roughly speaking) constructing attributes as maximal classes of similar objects and then obtaining kinds by means analogous to the ones used in the previous chapter. But first things first, for such an approach to succeed we should grasp what similarity or resemblance relations are supposed to be. The aim of this first section is to explain what I will understand by similarity or resemblance and introduce some of its main properties[2].

4.1.1 What is Similarity?

Similarity is everywhere. We make countless qualitative comparisons in everyday life, both in literal and metaphorical discourse, by using certain expressions like "... are similar ...", "... and ... resemble each other", "... and ... are alike" or "... looks like a ..." in sentences such as "he is like his father" or "the stars look like distant lighthouses". Moreover, similarity is the ground for a big group of inferences made both in ordinary and scientific discourse, namely analogical inferences. These inferences are of the form "x is P, y is similar to x, therefore by analogy y is P" and seem to be at the core of many discoveries. Similarity is so flexible that seems to be transcategorial. Objects, parts of objects, attributes, processes, events, modes and so on can be similar to each other. In fact, entities from different ontological categories can be similar to each other too (or at least, that seems to be an appropriate way to interpret Plato's claim that concrete objects resemble their forms). Similarity comes in degrees. Entities may be approximately similar or exactly similar (duplicates) to each other. Two black

[2] Many properties to be surveyed here are very nicely explained and rigorously analyzed in the wonderful thesis on similarity by [50]. More fine-grained distinctions between kinds of similarities can be found in that reference.

shoes are similar up to a finer degree of similarity than are a black shoe and a black raven. Similarity can be intrinsic, as when we say that the twins Jim and Tim are very much alike because it is difficult to tell them apart even when they are wearing very different clothes, or it can be extrinsic, as when we say that planes fly like birds, that some key-rings can be used as bottle openers or that Jim and Tim both behave in very childish ways. The relevance of degrees of similarity easily shifts with the context. In a store full of lamps and beds, saying that the lamp in the corner is similar to the bed in the middle of the room because they are both pieces of furniture is, at the very least, to violate a couple of conversational maxims. If one is told in such a room that the lamp and the bed are similar, one expects to find some (maybe unexpected or unnoticed) finer similarity, say in their shape, in their design or in their manufacturer. Having background information or new evidence (or even a theory), also influences the choice of the relevant similarity. Once we know that tomatoes are really fruits, it feels uneasy to see them placed next to apples and bananas in a supermarket. Similarity, in its simplest formulation as a binary and categorical relation, has some nice formal properties. Necessarily, every entity is similar to itself. Necessarily, if an entity is similar to another entity, then the latter is similar to the former. More crucial to our discussion, presence of similarity relations is correlated with having common properties (what these properties really are is a different question). In fact, it seems that the following proposition is necessarily true too:

x and y are similar to each other iff they have a common property.

But as we know, this quickly leads us into trouble. If the degree of similarity is too coarse, then the previous proposition is trivially true of any pair of objects. In contrast, if the degree of similarity is too fine, then the proposition is true only when x and y are the same object. This suggests that similarity cannot escape from the distinction between naturalness and abundance either. The interesting, relevant, explanatory and informative degrees of similarity are the sparse or natural ones. The uninteresting, irrelevant and almost vacuous degrees of similarity are the abundant ones. Any attempt to ground natural kinds (or attributes) on similarities must assume that the degrees of similarity at issue are sparse. The difference between positing primitive natural similarities and positing primitive natural collections of things starts to blur, as Lewis already pointed out.

Despite its beautiful formal properties and its many potential applications, similarity is not a very popular relation among philosophers nowadays. The explanation for this fact seems to lie on the enormous amount of objections that have been made to it. Among others, it has been argued that[3]:

i. Similarity is 'logically repugnant'. [107]

ii. Similarity is scientifically sterile. [107]

iii. Similarity is not a binary categorical relation.

iv. Similarity is not necessarily reflexive. [137]

[3]Of course, there are other interesting questions regarding the indeterminacy, relativity, and so on of similarity that we will not discuss here.

v. Similarity is not necessarily symmetric. [137]

vi. Similarity is heavily context-dependent. [49]

vii. Similarity is a subjective relation.

viii. Similarity is reducible to exemplifying a common universal. [2]

ix. Similarity cannot be used to reduce universal attributes. [2]

The list of objections could go on. This is somewhat surprising considering that similarity has also been invoked as part of a solution to many philosophical problems, such as the problem of universals [109], counterfactuals [72], counterparthood [74], causality (by Hume), scientific and artistic representation, natural kinds and classification [107], and so on. I think that most of these claims are either false or gross exaggerations. The first two objections were forcefully put forward by Quine:

> "We cannot easily imagine a more familiar or fundamental notion than this, or a notion more ubiquituous in its applications. On this score it is like the notions of logic: like identity, negation, alternation and the rest. And yet, strangely, there is something logically repugnant about it. For we are baffled when we try to relate the general notion of similarity significantly to logical terms. (...) It is a mark of maturity of a branch of science that the notion of similarity or kind finally dissolves, so far as it is relevant to that branch of science." [107]

The next two sections will be devoted to objections (i)-(ii). I do not see (iii) as a real objection. There are different kinds of similarity depending on their formal features, just as there are different notions of belief (categorical, comparative and degree-like) depending on their formal features. First, we have the binary *categorical* similarity, expressed by statements of the form "x and y are similar". We also have a notion of *comparative* similarity, either triadic "x is (equally or) more similar to y than z is" or tetradic "x is (equally or) more similar to y than z is to w". We also have a notion of *collective* similarity of the form "A-s and B-s are similar to each other" or "A-s are similar to each other". Finally, there is a notion of *gradual* or *degree-like* similarity "x and y are similar up to a (fixed or not) degree d"[4]. One can think about the degree-like similarity as a collection of ordered categorical similarities. From a different point of view, one could think about categorical similarity as a similarity up to some fixed degree d (that is usually implicit). This is analogous to how belief is usually conceived of. Categorical belief can be considered to be a special case of degree-like belief, where there is a fixed degree of belief which is implicitly given by the context. One can think about categorical, comparative and degree-like

[4]Properties of categorical similarity are studied in [87], [89], [121]. Properties of some comparative similarities can be found in [143], [89] and [128]. A first approximation a notion of degree-like similarity can be found in in the structures of proximity by [128], which get represented by the kind of metric spaces used in psychology.

similarities as analogous to the classificatory, comparative and metric concepts too [59]. The interest on different notions of similarity seems to have been developed mainly due to the lack of success of the categorical approach. If the objection is that the categorical notion is useless and should be replaced by its comparative or degree-like counterparts, then an appropriate way to reply is to show that categorical similarity can do most of the work that it was supposed to do. This is left for later.

The second section will deal with (iv)-(v), which target the formal properties of categorical similarity. Most of them come from the famous experimental findings of the psychologist A. Tversky, who claimed that subjects make non-symmetric similarity judgements (for a different defence of some of these basic properties of similarity, see [109]). Regarding (vi), as [49] forcefully argued, similarity seems to be highly *contextual*. Whether two objects are similar or not may vary from one context to another depending on the aspects or respects of comparison being considered as relevant in each case. If we do not want to appeal to respects of comparison (given that this seems to commit us to the existence of attributes) we can say that similarity is intrinsically degree-like, and that by being dependent on context the degree of similarity under consideration changes from one context to another. (vii) is related to (vi). These two aspects threaten to make similarity more a contribution of the epistemic subject to the world than part of the structure of the world itself. I cannot discuss these two objections in the thesis, they are here mentioned just for completeness. The last two (viii)-(ix) are heavier and relate to the problem of universals. Although I will not give a full defence of resemblance nominalism, I will make some commitments to it and the Chapter V is devoted to give answers to the most pressing objections to resemblance nominalism.

4.1.2 Invariants of Categorical Similarity

This section could have been called 'the logical form of categorical similarity'. It has two main aims. The first one is to counter Quine's claim that categorical similarity is 'logically repugnant'. The second one is to provide us with the mathematical notions that we need in order to introduce a model for resemblance nominalism. In what follows I will deal mainly with categorical similarity, assuming that it corresponds to a degree of similarity implicitly fixed by the context. The first thing I want to say is that there is a mathematical theory of similarity. It goes under several labels: the theory of similarity structures [89], the theory of tolerance relations [121], the theory of reflexive undirected graphs, and so on. In what follows I will use the name 'theory of similarity structures', since this is closer to the term introduced by Carnap in [16] and [17]. A categorical similarity is just any reflexive and symmetric relation. This includes equivalences as a special case. Similarity has been many times thought about metaphorically in spatial terms. One imagines the objects as points in a space. The more similar these objects are, the closer they are in the space. It is no wonder then that the concept of similarity has close connections to the notions of approximation, closeness, connectedness and other geometric-topological properties. So although we will usually think of the elements of the domain as objects and the similarity as resemblance, one can think about these elements as points, about the subsets as regions and about the similarity as some sort of closeness or proximity relation between points.

A first step towards a general characterization of such a theory would be in a kleinean way. Thus I want to propose that:

Theory of Similarity The theory of similarity structures is the study of similarity invariants, the study of those properties that are invariant under similarity-preserving transformations.

We start with the following axiomatic presentation[5]:

Definition 25. *Let S be a set and $\sim \subseteq S \times S$ a binary relation. $(S, \sim)$ is a* similarity structure *iff $\forall x, y \in S$:*

i $x \sim x$. *[Reflexivity]*

ii $x \sim y \Rightarrow y \sim x$. *[Symmetry]*

The transitive similarities are the *equivalence relations*. As in the case of equivalences, we say that the degree of similarity $\sim$ is *finer* than $\sim'$ iff $\sim \subseteq \sim'$. We can interpret $x \sim y$ as "x is sufficiently similar to y", "x is close to y", "x is a neighbour of y", "x is directly connected to y", "x is analogous to y", and so on. I will use the concept of similarity in this restricted sense. Thus, it may be useful to give some examples of similarity relations:

i Let A, B be sets. Then $A \sim^* B := A \cap B \neq \varnothing$.

ii Let A, B be finite sets. Then $A \sim^* B := |A \cap B| \leq n$ for some fixed natural number n.

iii Let $f\colon A \to B$ and $g\colon A \to B$ be two functions, then $f \sim^* g := \exists x \in A\ f(x) = g(x)$.

iv Let (S, E) be a simple graph. Then (S, E^*), where $E^* := E \cup \{(x, x) \mid x \in S\}$ is a similarity structure.

v Let $(S, \leq)$ be a preordered set. Then $x \sim^* y := \exists z \in S\ z \leq x, y$.

vi Let $(S, \leq)$ be a lattice. Then $x \sim^* y := x \wedge y \neq 0$.

vii Let (S, d) be a metric space. Then $x \sim^* y := d(x, y) \leq \epsilon$ for some fixed real number $\epsilon > 0$.

viii Let (S, T) be a topological space. Then $A \sim^* B := cl(A) \cap cl(B) \neq \varnothing$.

ix Let (S, Q, I) be a formal context. Then $x \sim^* y := i(x) \cap i(y) \neq \varnothing$.

These examples show that similarities are everywhere. The fourth one is crucial, for it allows us to depict each finite similarity structure as an undirected graph. Objects will be pictured as dots and similarities as edges connecting these dots. The fundamental concepts we will make use of are first the following similarity operators:

[5]As far as I know, some of the material on similarity structures presented here and in Chapter V, like the operator *cro* with its properties and uses, the axiom systems and the proofs of the main properties from them or all the results on similarities of order 1 (e.g. gen quasianalysis and the main theorems) is new. In Chapter V I will try to show that their introduction is justified by hinting at how fruitful these concepts are.

Definition 26. *Let* $(S, \sim)$ *be a similarity structure and* $A \subseteq S$*. Then we define:*

i $co(A) := \{x \in S \mid \exists y \in A \; x \sim y\} = \{x \in S \mid co(x) \cap A \neq \varnothing\}$ *is the* similarity neighbourhood*.*

ii $int(A) := \{x \in S \mid \forall y \in A^c \; \neg(x \sim y)\} = \{x \in S \mid co(x) \subseteq A\}$ *is the* similarity interior*.*

iii $ext(A) := \{x \in S \mid \forall y \in A \; \neg(x \sim y)\} = \{x \in S \mid co(x) \subseteq A^c\}$ *is the* similarity exterior*.*

iv $out(A) := \{x \in S \mid \exists y \in A^c \; x \sim y\}$ *is the* similarity outside*.*

v $bd(A) := \{x \in S \mid \exists y \in A \; x \sim y \;\;\&\;\; \exists z \in A^c \; x \sim z\}$ *is the* similarity boundary*.*

vi $bd_{int}(A) := bd(A) \cap A$ *is the* similarity interior boundary*.*

vii $bd_{ext}(A) := bd(A) \cap A^c$ *is the* similarity exterior boundary*.*

viii $cro(A) := \{x \in S \mid \forall y \in A \; x \sim y\}$ *is the* similarity cocore*.*

Except for the last one *cro*, which is considered for the first time in this thesis, properties for the operators can be found in [13]. As usual, we define $co(x) := co(\{x\}), ext(x) := ext(\{x\})$, and so on. The operators have very intuitive interpretations. First, $co(A)$ gives us all the elements that are similar or close to some elements in A, whereas $cro(A)$ gives us the elements that are similar or close to all the elements in A. Spatially, $co(A)$ enlarges the region A by adding nearby points. Second, $int(A)$ gives us all the elements whose nearby elements are already in A, whereas $bd(A)$ includes those elements that are close to elements beyond A. The boundary can be split into an internal side $bd_{int}(A)$ and an external side $bd_{ext}(A)$. Spatially, $int(A)$ contracts the region A by eliminating those elements in A that belong to the (interior) boundary. Finally, $ext(A)$ and $out(A)$ give us elements that are outside of A in different ways. Whereas $ext(A)$ includes all those which are not similar to any element in A, $out(A)$ includes those elements that are not in the interior of A. (Mormann, REF) shows that *co* and *int* form a galois connection. The structures induced by the resulting closure operators, which involve the operators *ext* and *out*, are further explored in [13]. These operators have many interesting properties that we cannot deal with right now. An important one relates the similarity neighbourhood to the similarity interior, i.e. $co(A) = int(A^c)^c$.

We also have the following fundamental kinds of points and sets, which bring the proto-spatial nature of similarity to the fore:

Definition 27. *Let* $(S, \sim)$ *be a similarity structure and* $A \subseteq S$*. Then we define:*

i A *is a* clique $\Leftrightarrow \forall x, y \in A \; x \sim y$*.*

ii A *is* maximal $\Leftrightarrow \forall z \in S \; (\forall x \in A \; z \sim x \Rightarrow z \in A)$*.*

iii A *is a* similarity circle $\Leftrightarrow A$ *is a maximal clique.*

iv A *is* similarity dense $\Leftrightarrow co(A) = S$*.*

v x *is* similarity isolated $\Leftrightarrow co(x) = \{x\}$*.*

vi S *is* similarity connected $\Leftrightarrow (co(A) = A \Rightarrow A = S \text{ or } A = \varnothing)$.

vii S *is* totally similarity disconnected $\Leftrightarrow (x \sim y \Rightarrow x = y)$.

The *set of all the similarity circles* will be $SC(S)$. As will be shown, similarity circles play a major role in the theory, since they will be our surrogates for natural attributes. One can introduce a similarity structure by giving directly the set of all its similarity circles. For example, the similarity $x \sim y \sim z$ can be introduced as $SC(S) = \{\{x, y\}, \{y, z\}\}$. Although the reasons for this will be given in Chapter V, we will freely follow this practice to introduce examples. It is easy to check that the definition of connectedness here given is equivalent to the graph-theoretical one. In other words, a similarity is connected iff for any two points x and y there is a path $x = x_0 \sim x_1 \sim \cdots \sim x_{n-1} \sim x_n = y$ from one to the other. Connectedness is a fundamental invariant of similarity and already allows us to distinguish between two extreme cases. If the degree of similarity d is the coarsest one $\sim = S \times S$ then any two objects will be d-similar and so indiscernible to each other. If the degree of similarity is the finest one $\sim = \Delta$, then it will be coextensional to identity, the similarity will be totally disconnected and every point will be isolated. Between these two cases we have non-trivial connected similarities, that is to say, degrees of similarity which are such that one can reach any object from any other object by 'jumping' from one object to another which is sufficiently like it. The coarser the similarity is, the easier it will be to connect any two such objects. For instance, we could have a path like $a \sim b \sim c$ where a is a black raven, b is a white raven and c is a white shoe. Note that, in contrast, every equivalence relation is disconnected.

In Chapter V we will need some other properties of the operators, such as [6]:

Proposition 14. *Let* $(S, \sim)$ *be a similarity structure. Then:*

i $A \subseteq co(A)$. *[Extensiveness]*

ii $co(A) \cap B = \varnothing \Leftrightarrow A \cap co(B) = \varnothing$. *[Autoconjugation]*

iii $co(\varnothing) = \varnothing \ \& \ co(S) = S$. *[Normality]*

iv $A \subseteq B \Rightarrow co(A) \subseteq co(B)$. *[Monotonicity]*

v $co(\bigcup A_j) = \bigcup co(A_j)$. *[Union preserving]*

vi $x \in co(A) \Leftrightarrow co(x) \cap A \neq \varnothing$. *[Neighbourhood point]*

vii $co(A) = \bigcup\{co(x) \subseteq S \mid x \in A\}$. *[Algebraic]*

viii $co(x) = \{z \in S \mid x \in co(z)\} \ \& \ x \in co(y) \Leftrightarrow x \sim y \Leftrightarrow y \in co(x)$.

Proof. We only prove (iii)-(v) from (i)-(ii). (iii) Since $\varnothing \subseteq co(\varnothing) \subseteq co(co(\varnothing))$ by (ii) $\varnothing \cap co(co(\varnothing)) = \varnothing$ iff $co(\varnothing) \cap co(\varnothing) = co(\varnothing) = \varnothing$ and $S \subseteq co(S) \subseteq S$ by (i). (iv) Let $A \subseteq B$ and $x \in co(A)$, then $\{x\} \cap co(A) \neq \varnothing$ iff $A \cap co(x) \neq \varnothing$, so $B \cap co(x) \neq \varnothing$ but then $\{x\} \cap co(B) \neq \varnothing$, i.e. $x \in co(B)$. (v) If $co(\bigcup A_j) = \varnothing$, then by (i) $\bigcup A_j = \varnothing$, therefore $A_j = \varnothing$ for each j, and by (iii) $\varnothing = co(\bigcup A_j) = \bigcup \varnothing = \bigcup co(A_j)$. Now suppose $x \in co(\bigcup A_j) \neq \varnothing$, then $\{x\} \cap co(\bigcup A_j) \neq \varnothing$ iff $co(x) \cap (\bigcup A_j) \neq \varnothing$ iff $\bigcup(co(x) \cap A_j) \neq \varnothing$ iff $(co(x) \cap A_1) \neq \varnothing$ or $(co(x) \cap A_2) \neq \varnothing$ or ... iff $\{x\} \cap co(A_1) \neq \varnothing$ or $\{x\} \cap co(A_2) \neq \varnothing$ or ... iff $x \in \bigcup co(A_j)$. The converse follows from monotonicity. □

[6] That autoconjugation corresponds to symmetry can be found in (Maddux, 2006).

Although co is extensive, monotone, normal and preserves arbitrary unions, it is not idempotent. In this it differs from topological closure operators. As a consequence, the similarity neighbourhood can be iterated, enlarging more and more the original region. A different more 'algebraic' axiomatic system can be given if we start from this similarity neighbourhood operator[7]:

Definition 28. *Let (S, co) be a set S with a monary operator $co\colon \wp(S) \to \wp(S)$. Then (S, co) is a similarity structure iff it satisfies:*

i $A \subseteq co(A)$. *[Extensiveness]*

ii $co(A) \cap B = \varnothing \Leftrightarrow A \cap co(B) = \varnothing$. *[Autoconjugation]*

Define $co(x) := co(\{x\})$. A homomorphism is a function $f\colon (S, co) \to (S', co')$ such that $f(co(A)) \subseteq co'(f(A))$.

The similarity neighbourhood co has as a companion the cocore operator cro, which gives us all the elements that are similar to those of a given set:

Proposition 15. *Let $(S, \sim)$ be a similarity structure and $cro(A) := \{y \in S \mid \forall x \in A\ y \sim x\}$. Then:*

i $x \in cro(x) = co(x)$ & $cro(A) = \{x \in S \mid A \subseteq co(x)\}$.

ii $A \subseteq cro(B) \Leftrightarrow B \subseteq cro(A)$.

iii $A \subseteq B \Rightarrow cro(B) \subseteq cro(A)$. *[Antitonicity]*

iv $\bigcap cro(A_j) = cro(\bigcup A_j)$. *[De Morgan]*

v $cro(\varnothing) = S$ & $cro(S) = \varnothing \Leftrightarrow$ *there are no dense elements in* S.

vi $crocro(A)$ *is a closure operator* & $cro(A) = crocrocro(A)$.

vii $A \neq \varnothing \Rightarrow cro(A) \subseteq co(A)$ & $int(A) \neq \varnothing \Rightarrow cro(A) \subseteq A$.

viii $x \sim y \Leftrightarrow x, y \in cro(\{x, y\}) \Leftrightarrow x \in cro(y)$.

ix A *is a clique* $\Leftrightarrow A \subseteq cro(A)$ & A *is maximal* $\Leftrightarrow cro(A) \subseteq A$.

x A *is a similarity circle* $\Leftrightarrow cro(A) = A$.

xi $crocro(A) = cro(A) \Leftrightarrow cro(A)$ *is a similarity circle.*

Proof. (i) is trivial. (ii) If $A \subseteq cro(B)$, then if $x \in Bx \sim y$ for all y in A, i.e. $x \in cro(A)$. And if $B \subseteq cro(A)$ then if $x \in Ax \sim y$ for every $y \in B$. (iii) Suppose that $A \subseteq B$ and $x \in cro(B)$. Then $\{x\} \subseteq cro(B) \Rightarrow A \subseteq B \subseteq cro(x) \Rightarrow \{x\} \subseteq cro(A) \Rightarrow cro(B) \subseteq cro(A)$. (iv)-(vi) follow from the fact that (cro, cro) form the extension-intension galois connection of the context $(S, S, \sim)$, as in Chapter III. (vii) Suppose that $int(A) \neq \varnothing$. Then if $y \in cro(A)$, $y \sim z$ for every z in $int(A)$, therefore $y \in A$. (viii) Follows from (i). (ix) If A is a clique, every x in A is similar to every y in A, i.e. x is in $cro(A)$. And if $A \subseteq cro(A)$, then if x and y are in A we have $x \sim y$, i.e. A is a clique. A is maximal iff if $x \sim y$ for all $y \in A$, then $x \in A$ iff $cro(A) \subseteq A$. (x) Follows from (ix). (xi) Follows from (x). □

[7] I have not seen this axiom system anywhere, but the autoconjugation property is known to correspond to symmetry so this was to be expected. The properties of the *cro* operator are new.

A new set of axioms can be given using this operator, from which the previous properties follow:

Definition 29. *Let* (S, cro) *be a set* S *with a monary operator* $cro\colon \wp(S) \to \wp(S)$*. Then* (S, cro) *is a similarity structure iff it satisfies:*

i $x \in cro(x)$*. [Extensiveness]*

ii $A \subseteq cro(B) \Leftrightarrow B \subseteq cro(A)$*. [Autoconjugation]*

The concept of an equivalence class splits into several richer notions when equivalences are generalized to similarities. Three are the most fundamental ones. We have sets consisting of all the elements that are similar to at least of element of some set ($co(A)$), sets consisting of all the elements that are similar to all the elements of some set ($cro(A)$) and we have also maximal sets of pairwise similar elements (elements in $SC(S)$). We will make heavy use of these notions in the next Chapter. There is still a very strong link between these concepts:

Proposition 16. *Let* $(S, \sim)$ *be a similarity structure and* $A \subseteq S$*. Then:*

i $A \in SC(S) \Leftrightarrow A = \bigcap\{co(x) \subseteq S \mid x \in A\} \Leftrightarrow A = cro(A)$.

ii $A = co(x) \Leftrightarrow A = \bigcup\{T \in SC(S) \mid x \in T\}$.

If we interpret $\sim$ as an approximate similarity, we can define exact similarity or duplication as an equivalence relation that holds between those entities that are approximately similar to the same entities [87], [89]. A fortiori, two duplicates belong to the same similarity circles. Since we will take similarity circles as surrogates for natural attributes, it will follow too that two entities are duplicates iff they exemplify exactly the same natural attributes:

$$x \approx_{co} y \Leftrightarrow co(x) = co(y) \Leftrightarrow i(x) = i(y)$$

We will consider structures for which this relation coincides with identity, in other words, those structures that satisfy the following crucial axiom [89]:

Definition 30. *Let* $(S, \sim)$ *be a similarity structure. Then* S *satisfies the* Similarity Neighbourhood Indiscernibility Axiom *(SNI) iff* $co(x) = co(y) \Rightarrow x = y$*.*

This axiom is the similarity analogue of the *Identity of Indiscernibles*. It says that if two entities are similar to the same entities, then they are identical. Requiring (SNI) is harmless, after all we can always start from any similarity and obtain its (SNI)-quotient. Therefore, this simply implies that for the most part we will be considering equivalences classes of objects, or in other words, objects up-to-exact similarity.

Now it is time for a different thing. We must introduce the relevant structure-preserving maps[8]:

Definition 31. *Let* $(S, \sim)$ *and* $(S', \sim')$ *be similarity structures and* $f\colon S \to S'$ *a function. If* $x \sim y \Rightarrow f(x) \sim' f(y)$ *then* f *is a* similarity homomorphism*. If* $f(x) \sim' f(y) \Rightarrow x \sim y$*, then* f *is* faithful *and if it is also bijective then* f *is a* similarity isomorphism*. If* $S = S'$ *then* f *is said to be a* similarity automorphism*.*

[8] One can easily check that this gives us a category whose objects are similarity structures and whose morphisms are similarity homomorphisms. Mormann calls it 'SIM'.

We can think of an homomorphism f as a similarity transformation or deformation between the elements of the domain S, thus as a process of qualitative 'continuous' change. Given an object x, $f(x) = y$ is the result of deforming x in such a way that the (degree of) similarity between x and the objects to which it is similar is preserved. Some equivalent ways of presenting homomorphisms are:

Proposition 17. *Let $(S, \sim)$, $(S', \sim')$ be similarity structures, $x, y \in S$, $A \subseteq S$ and $f\colon S \to S'$ a function. Then the following conditions are equivalent:*

1. *f is a similarity homomorphism.*
2. *$x \sim y \Rightarrow f(x) \sim' f(y)$.*
3. *$f(co(A)) \subseteq co'(f(A))$.*
4. *$f(cro(A)) \subseteq cro'(f(A))$.*
5. *If A is a $\sim$-clique, then $f(A)$ is a $\sim'$-clique.*

This already gives us a hint concerning what kind of properties are relevant from the point of view of the theory of similarity: point-point similarities, 'closeness' of points to sets and classifications by sets all whose elements are similar to each other. We finally get to:

Definition 32. *Let $(S, \sim)$ and $(S', \sim')$ be similarity structures and $f\colon S \to S'$ a similarity isomorphism. Let P be a property of the similarity structure S. Then P is a* similarity invariant *iff S is $P \Rightarrow S'$ is P.*

In other words, similarity invariants are those properties common to isomorphic similarities. Some examples of similarity invariants are (some of these will be studied later on):

1. Being of cardinality n: the number of entities is n.
2. Being a clique: any two entities in the collection are similar to each other.
3. Being a similarity circle: any two entities in the collection are similar to each other and any entity which is similar to all of them is already in the collection.
4. Being the similarity neighbourhood of some set A: every entity is similar to some element in A.
5. Being an isolated point: an entity which is only similar to itself.
6. Being connected: any two entities are indirectly similar to each other.
7. Being totally disconnected: no two distinct entities are similar to each other.
8. Being (SNI): no two distinct entities are similar to the same entities.
9. Being of order 1: two similar entities are similar to a paradigmatic entity.

Therefore, the theory of similarity structures is the theory of similarity invariants. This will give us a way to detect whether a structural property is acceptable from the point of view of resemblance nominalism. If the property is not a similarity invariant, then it is not 'objectively out there' so to speak. By the time being, I hope these notions are enough to show some of the mathematical richness hidden in the deceptively simple concept of similarity. Put short, Quine was wrong.

4.2 Objections to the Properties of Categorical Similarity

4.2.1 Similarity is not Reflexive

It is a truism that every object resembles itself. But philosophers like to challenge truisms. The first objection goes as follows:

Similarity is not reflexive

A first argument for the reflexivity of similarity would appeal to the *Principle that Identity implies Similarity*:

$$x = y \Rightarrow x \sim y \ [IS]$$

The principle seems reasonable. If two objects are identical, then they must be similar to each other. Why? If these two objects were not similar to each other, for no degree of similarity, then they would be different to each other for some degree of similarity. Therefore, they would be distinct objects. There are two possible explanations. The first one appeals to properties, respects or aspects of comparison. If two objects are different, then there is a property that differentiates one from the other, and by the *Indiscernibility of Identicals* they must be distinct objects. This is what Armstrong [2] holds. This move is not available to someone who thinks that similarity is a primitive relation. But such a philosopher would simply think that the Indiscernibility of Identicals should be replaced by a different principle. Consider the *Principle that Difference implies Distinction*:

$$\exists d \in D \ \neg(x \sim_d y) \Rightarrow x \neq y \ [DD]$$

If there is a degree of similarity d in which two objects are different, then they must be distinct objects. This seems obvious. The reason is that (IS) seems to be satisfied for any degree of similarity, however coarse or fine. There is therefore a similarity-correlate for the *Indiscernibility of Identicals* which is the contrapositional of (DD):

$$\forall d \in D \ (x = y \Rightarrow x \sim_d y) \ [PII - Similarity]$$

If two objects are identical, then they are similar up to any degree of similarity whatsoever. A philosopher that accepts similarity as primitive should adopt something like this principle as fundamental. If we now 'hide' the degrees of similarity and leave them as something to be given by context, we have that for any degree of similarity:

$$x = y \Rightarrow x \sim y \; [IS]$$

And therefore the reflexivity of similarity would follow from the reflexivity of identity:

$$x \sim x \; [Reflexivity]$$

Thus the following question seems pressing: could there be an object that was not similar to itself? The nominalist has to assume that there could not be an x that was not similar to itself. So he assumes from the start that reflexivity holds. In contrast, the realist *explains* reflexivity in his framework [2]. It follows from his *Principle of no-Bare Particulars.* From the realist point of view, an object could fail to be similar to itself only if it was a bare particular and similarity only tracked sparse properties (considering properties like *being self-identical* or *being identical to x* as not sparse). But most universalists would simply assume the impossibility of bare particulars. So the realist is not in a better position, for he also has to assume that each object exemplifies at least one attribute.

Pereyra [109] counters this argument by trying to show that the formal properties of similarity follow from more basic nominalistic principles. Apart from the properties of identity, I doubt that there is anything more basic than the reflexivity or symmetry of similarity. As I said, at most reflexivity would follow from the properties of identity in conjunction with principles connecting identity and similarity. But the appropriate answer to Armstrong is, I think, to point out that the nominalist can also explain the properties that the realist assumes about the instantiation relation. As we saw in Chapter III, the realist assumes the *Principles of No-Bare Particulars* and *Exemplification.* The former one is explained in the nominalistic framework by first giving surrogates for attributes (namely sets of similar objects), then giving a surrogate for the instantiation relation (namely set-theoretic membership) and finally showing that it follows from the reflexivity of similarity that each object belongs to one of these collections of similar objects. The nominalist can also explain the *Principle of Exemplification.* The reason is that the surrogates for attributes will be maximal sets of similar objects and therefore the empty set will not be an attribute. This will be seen in Chapter V[9].

Nevertheless, some cognitive psychologists have formulated objections to reflexivity. The experiments devised by [137] seem to show that sometimes we judge some entities as being more similar to other entities than to themselves. This may seem to go against reflexivity. I will deal with this objection alongside the next one, for reasons that will become clear soon.

4.2.2 Similarity is not Symmetric

Our second truism says that necessarily, if an object is similar to another, then the latter is similar to the former one. The second objection says:

[9]There may be other reasons why reflexivity could fail. For instance, if x endures through time, then x in t may not be similar to x in t', if t' is further enough in time from t. Suppose that x is a caterpillar that undergoes metamorphosis, then x is a butterfly in t' that may not be similar to x. Therefore x is not similar to itself. This resembles the corresponding problem for identity through time.

If one explains similarity as sharing a common attribute, then symmetry follows trivially (Armstrong, 1978). It follows from the commutativity of conjunction, if we define similarity as exemplifying a common attribute $x \sim y := \exists P \in Q\ xIP\ \ \&\ \ yIP$.

But cognitive psychologists have also presented objections to symmetry [137]. The experiments seem to show that an object x could be more similar to an object y than y would be to x. As a special case, x could be more similar to y than x is to itself. The root of the problem seems to be that we may consider different salient properties when comparing the two objects and uttering a similarity comparison judgment like 'x is similar to y'. In some cases the experimental subject tends to say things like "x is more similar to y than y is to x". Tversky explained several of these phenomena that were related to metaphors and to the degree of prototypicality of objects.

I think that Tversky's criticism against symmetry can be challenged on several grounds. I will present three objections. First, theorists of spatial models of similarity have developed alternative explanations of the alleged violations of symmetry and have suggested corresponding modifications of their models to account for these facts. Second, Tversky's own contrast model makes substantive questionable assumptions, as I will discuss at length. Third, even if Tversky's criticisms are successful they only apply to either a comparative or a degree-like notion of similarity and not to the categorical version of it. This is not a cheap reply, for as I will show, a plausible comparative notion of similarity can be defined from the categorical one that explains why the violations of symmetry occur just under the assumption of the existence of paradigmatic objects in the similarity structure. The defendant of categorical similarity can explain the controversial cases without dropping the symmetry axiom.

Since the groundbreaking work of E. Rosch (see e.g. [112] and [113]), as we said, it is assumed in the cognitive psychology literature that some objects exhibit a greater degree of prototypicality than others and thus are more representative of the kind of objects they belong to. For instance, a robin is a more prototypical bird than a penguin, an orange is a more prototypical fruit than a nut and a chair is a more prototypical piece of furniture than a stove. Thus, given two birds x and y, a subject may utter a similarity comparison judgment of the form "x is more similar to y than y is to x", for example when y (a robin) is a more prototypical bird than x (a penguin). Metaphors and similes produce similar cases of directionality and may even involve a change in meaning. The classical example is the metaphor "that butcher is a surgeon", which means something quite different from "that surgeon is a butcher". Or similes like "her eyes are like pearls", which mean something different from "pearls are like her eyes". The directionality present in these cases seems to be related to the pragmatic purpose of letting our hearer understand that we are attributing to her eyes some aesthetically pleasant properties typically associated with pearls (e.g. brightness and so on), and not vice versa.

Spatial models of similarity have troubles accommodating these violations of symmetry. According to spatial models, objects are represented as points in a metric space (see [94] or [42]):

Definition 33. *Let S be a set and $d\colon S^2 \to \boldsymbol{R}$ a real valued function. Then (S,d) is a* metric space *iff $\forall x,y,z \in S$:*

i $d(x, y) \geq 0$. *[Positiveness]*

ii $d(x, y) = 0 \Leftrightarrow x = y$. *[Indiscernibility]*

iii $d(x, y) = d(y, x)$. *[Symmetry]*

iv $d(x, z) \leq d(x, y) + d(y, z)$. *[Triangle Inequality]*

The similarity or dissimilarity between two objects is represented by the distance function $d(x, y)$. We can understand the distance between two objects as their *degree of dissimilarity.* For instance, the definition says that the degree of dissimilarity of x and y is $d(x, y)$. Thus, (ii) and (iii) are requirements analogous to reflexivity and symmetry. For instance, (ii) says that two objects are maximally similar iff they are identical, which is again a version of the Identity of Indiscernibles. In contrast, (iii) says that the similarity of x to y equals the similarity of y to x. Since the metric distance is a function and by (i)-(ii), any two different objects will be similar to each other *up to some degree of similarity.* In other words, any two objects are comparable by similarity if one chooses a coarse enough degree of similarity. We will consider the *Triangle Inequality* later on, when we discuss transitivity. The crucial issue here is that spatial models assume something akin to reflexivity and symmetry. A fortiori, we have that if $x \neq y$, then $0 = d(x, x) < d(x, y)$.

What Tversky's findings showed is that we could have objects such that $d(x, y) < d(x, x)$ for two different objects x and y and that $d(x, y) \neq d(y, x)$. The answer from the spatial camp (e.g. [95]) was to conjecture that in Tversky's controversial cases, the subject that made the similarity comparisons was introducing certain biases or that he was focusing on objects in such a way that these psychological processes messed with the basic properties of similarity. The contribution of the subject would amount to introducing weights in the similarity comparison (in the distance function), or to distorting the dimensions of the space. Simply put, the violations of reflexivity and symmetry would be due not to the nature of similarity itself but to the effect of the biases of the subject that may put more weight on some or other respects of comparison. So this is my first reply, namely that the spatial models can be modified in such a way that the violation of symmetry is explained as an effect of the subject making the comparison, while keeping the axioms intact. Moreover, the proponents of spatial models have shown that this interpretation coheres with the empirical data. According to Decock and Douven Decock [27], a different answer from the spatial camp has been to account for the effects of the context (including non-symmetry) by appealing to several different relations of similarity. This is the case of Gärdenfors conceptual spaces approach. In other words, the differences in context are explained by making use of a different conceptual space in each context. Since each conceptual space makes use of a similarity relation (with possibly different formal properties), since the similarity is relative to the respects of comparison (wich are represented as dimensions of the space) and since the salience of these respects changes from one context to another, the spatial camp can provide an answer to Tversky's objections.

Now to my second point. To explain his findings Tversky introduced his *contrast model*[10]. We are given a context (S, Q, I), where $S = \{a, b, c, \ldots\}$ is

[10] Another presentation of Tversky's model, compared to the spatial accounts is to be found in Decock and Douven's paper [27].

a domain of objects to be compared by similarity and Q a set of attributes or features of the objects in S. We will write $i(a) = A$, $i(b) = B$, and so on[11]. First, we close the set $i[S] = \{i(x) \subseteq \wp(Q) \mid x \in S\}$ under intersection and difference. In other words, if $A, B \in i[S]$, then $A \cap B \in i[S]$ and $A - B \in i[S]$. We call this new set Q^*, it will contain all the interesting combinations of attributes of objects (the ones that will be used to make comparisons). Now we define:

Definition 34. *Let (S, Q, I, s, F) be a context and let Q^* be the closure of the set $i[S] = \{i(x) \subseteq \wp(Q) \mid x \in S\}$ under intersections and set-theoretic differences. Let $s \colon S \times S \to \boldsymbol{R}$ and $F \colon Q^* \times Q^* \times Q^* \to \boldsymbol{R}$ be real-valued functions. Then (S, Q) is a* contrast model *iff:*

1. *$s(a, b) = F(A \cap B, A - B, B - A)$. [Matching]*
2. *$s(a, c) \leq s(a, b) \Leftrightarrow A \cap C \subseteq A \cap B$ & $A - B \subseteq A - C$ & $B - A \subseteq C - A$. [Monotonicity]*
3. *If (a, b) and (c, d), and (a', b') and (c', d') agree on the same two components, while the pairs (a, b) and (a', b') and the pairs (c, d) and (c', d') agree on the remaining third component, then $s(a', b') \leq s(a, b) \Leftrightarrow s(c', d') \leq s(c, d)$. [Independence]*
4. *The following hold: [Solvability]*
   - *(a) For all pairs (a, b), (c, d), (e, f) there is a pair (p, q) such that $A \cap B \approx P \cap Q$ & $C - D \approx P - Q$ & $F - E \approx Q - P$, i.e. it agrees on the first, second and third components of the respective pairs.*
   - *(b) If $s(c, d) < t < s(a, b)$, then there are $e, f \in S$ such that $s(e, f) = t$ and if (a, b) and (c, d) agree on one or two components then (e, f) agrees with them on these components too.*
   - *(c) There are pairs (a, b), (c, d) that do not agree in any components.*
5. *Let V, V' and W, W' belong to ϕ_i and ϕ_k for $i, k = 1, 2, 3$. Then $(V, V') \approx (W, W')_i \Leftrightarrow (V, V') \approx (W, W')_k$. [Invariance]*

For axioms (3)-(5) we need some auxiliary definitions. They will not be discussed here, I give them just for the sake of completeness. Let $V, W \in Q^*$. Then Tversky defines $V \approx W$ as follows:

$$V \approx W := \exists X, Y, Z \in Q^* \; F(V, Y, Z) = F(W, Y, Z) \text{ or}$$
$$F(X, V, Z) = F(X, W, Z) \text{ or } F(X, Y, V) = F(X, Y, W)$$

We say that (a, b) and (c, d) *agree on the first component* iff $(A \cap B) \approx (C \cap D)$. Analogously with the other two components (arguments) of F. Now let $\phi_1 = \{X \cap Y \mid X \cap Y \in Q^*\}$, $\phi_2 = \{X - Y \mid X - Y \in Q^*\} = \{Y - X \mid Y - X \in Q^*\} = \phi_3 \subseteq Q^*$ be the sets of the corresponding first, second and

[11] Tversky did not talk about contexts, but I will change some non-essential aspects of hits formulation so that its relation to the concepts introduced in Chapter III is made clear. The domain of F is strictly speaking smaller than $Q^* \times Q^* \times Q^*$ but is difficult to make this precise without introducing the rest of auxiliary definitions. I will give all the details of Tversky's model just to be fair to his view, but the reader should skip axioms (4)-(5). Tversky himself leaves (4) and (5) to the Appendix of the paper due to their cumbersome formulation and purely technical interest.

third arguments of F, e.g. ϕ_1 is the set of all the intersections of the intensions of objects in S. Let $X, X' \in \phi_1$ and $Y, Y' \in \phi_2$, then we define:

$$(X, X')_1 \approx (Y, Y')_2 :=$$
$$\exists (a,b), (a',b')\ s(a,b) = F(X,Y,Z) = F(X',Y',Z) = s(a',b')$$

Here $(X, X')_1 = (A \cap B, C \cap D)$ somehow represents the interval between $A \cap B$ and $C \cap D$ and the relation just defined represents the matching of intervals. Finally he extends this definition as:

$$(V, V')_i \approx (W, W')_i := \exists (a,b), (a',b')\ (V, V')_i \approx (X, X')_j \approx (W, W')_i$$
$$\text{for some } (X, X')_j\ j \neq i \text{ and } i = 1, 2, 3$$

The reader is advised to ignore conditions (4)-(5) and the definitions just given, they are included here just for the sake of introducing Tversky's full model. The most easily understandable conditions are the first two. By (Matching) the similarity between two objects a and b is a function of the attributes shared by them $A \cap B$ and the attributes that differentiate one from another $A - B$ and $B - A$. By (Monotonicity) the similarity between objects a and b is greater to that between a and c iff they have more common properties or they have less different properties. Here $s(a, b)$ is an ordinal measure of the similarity between a and b. To put it simply, the similarity between two objects increases the more properties they share and the fewer their differences happen to be. These two constraints are purely ordinal, they only require mapping the degree of similarity between two objects to a real number by using a similarity scale s and for this scale to be ordinal, i.e. for its order to be monotonous with respect to the inclusions of the intersections and differences of the intensions of objects.

Nevertheless, as we see, Tversky introduces three more axioms, namely (Independence), (Solvability) and (Invariance). The axiom of (Independence) was experimentally questioned by [45]. The latter two are only explained in the Appendix of his paper [137]. Although they are not usually discussed, these axioms impose an implausibly rich structure on the domain of objects. For instance, the second condition of (Solvability) requires that if a and b are more similar to each other than c and d, then for any real value t that is among the values that represent these similarities (namely, the values $s(a, b)$ and $s(c, d)$), there will be two objects e and f which are such that their degree of similarity is exactly t (under the ordinal measure s). But of course, between any two real numbers there is an uncountable amount of real numbers, so this axiom immediately implies that there is an infinite amount of objects in the domain. And this is not just a worry about the cardinality of the domain of objects. The degrees of similarity between any two pairs of objects are just as fine-grained as the real line. That means that if you detect that cat a is more similar to cat b than dog c is to dog d, then there are as many degrees of similarity between these two (the degree of similarity between a and b and the degree of similarity between c and d) as numbers between the corresponding real numbers representing them, and as many pairs of objects as needed for these numbers to represent each of these degrees of similarity. A lucky researcher only needs to find two pairs of objects that are comparable regarding their degrees of similarity, for in doing that he can infer that for any real-numbered degree of similarity that is among these two there will be two objects in the world (possibly hidden in some remote corner of the world) that are similar to exactly that degree. It may be a demanding quest to find these, but with enough time he is guaranteed to succeed. Tversky then proves the following representation theorem:

Theorem 4 (Representation Theorem, Tversky)**.** *Let (S, Q, s, F) be a contrast model. Then there exist a similarity scale $Sim\colon S \times S \to \boldsymbol{R}$, a non-negative scale $f\colon Q^* \to \boldsymbol{R}^+$ and three positive numbers $\theta, \alpha, \beta \geq 0$ such that for all $a, b, c, d \in S$:*

i $Sim(a, b) \geq Sim(c, d) \Leftrightarrow s(a, b) \geq s(c, d)$.

ii $Sim(a, b) = \theta f(A \cap B) - \alpha f(A - B) - \beta f(B - A)$.

iii *f and Sim are interval scales.*

Here f and Sim are representing functions with range $\mathbf{R}$. Note that the 'minus' inside the parentheses is the set-theoretic difference and the minus outside the parentheses is the arithmetic difference. The parameters θ, α, β are used to give more weight, respectively, to the coincidences or differences in features between a and b. Whereas s is an ordinal similarity scale that only depends on the sets of attributes, Sim is the resulting interval similarity scale that depends on the external parameters θ, α, β. What the representation theorem shows is that if the ordering of degrees of similarity, as dependent on sets of atributes (coincidences and differences in attributes), satisfies the axioms of the contrast model, then there will exist an interval scale that measures the degree of similarity between each pair of objects as a linear combination of the measures of the attributes (not) shared by considering some additional parameters that represent the focus or attention of the subject.

The proof works by showing that the domain of F, i.e. the product of (the quotients under $\approx$ of) ϕ_1, ϕ_2 (i.e. the sets of intersections and differences of features) ordered under the (quotient) relation $\leq$, is an *additive conjoint structure*. Additive conjoint structures are a basic kind of measurement structure identified by the *Representational Theory of Measurement* and are one of the most interesting examples of measurement structures that is available to the social sciences. The basic idea is that some magnitudes (say, the degree of similarity between two objects) are the result of the combination of two or more component magnitudes that are independent of each other (there is no law that makes the value of one dependent on the value of the others). In contrast with extensive measurement structures like those found in physics, the resulting magnitudes are not simply the concatenation of their component magnitudes. So roughly speaking, for the proof to work the set Q^* of all the combinations of intensions of objects must be such that the ordering of triples of sets of attributes $(\phi_1 \times \phi_2 \times \phi_3, \leq)$ is a conjoint structure. That means that the *degree of similarity* can be decomposed into three (or better put, two) components ϕ_1 (the sets of common attributes) and $\phi_2 = \phi_3$ (the sets of differences) which are independent of each other. The resulting measurement scale Sim is an interval scale.

The general expression for the similarity between two objects in (ii) weights the similarity between objects as a contrast between the properties shared and the differences present. So if the subject pays more attention to the similarities then the coincidence of features will have greater weight. But if he focuses more on the set-theoretic differences $A - B$ than on the differences $B - A$, then this will produce cases of non-symmetry. Tversky could then explain violations of symmetry by appealing to the notion of attention:

Attention Hypothesis The direction of 'asymmetry' is determined by which stimuli (objects) are most salient, in such a way that the least salient stimuli is more similar to the most salient one than vice versa.

This can be seen in the parameters α, β and θ, which represent the degree of attention by the subject or the salience of the objects. Let us suppose that given two objects a and b, the object b is more prototypical than a. Then b-s properties are more salient than a-s. If we now ask the subject "which object is similar to which one?", we will force her to make a directional similarity judgement. It is highly probable that she will answer "a is more similar to b (than vice versa)". In the contrast model one can show [137]:

Proposition 18 (Non-symmetry)**.** *If* $\alpha > \beta$, *then* $Sim(a, b) > Sim(b, a) \Leftrightarrow f(B - A) > f(A - B)$.

Therefore, the contrast model captures the difference $Sim(a, b) > Sim(b, a)$. If the subject focuses in the direction $a \to b$, then saying that a is more similar to b than vice versa is equivalent to saying that the difference $B - A$ is bigger than the difference $A - B$. Thus a is more similar to b than vice versa because there are less differences in the direction $a \to b$ than in the direction $b \to a$. The similarity judgement is symmetric iff the differences of one object with respect to the other one have the same weight or have exactly the same value:

$$Sim(a, b) = Sim(b, a) \Leftrightarrow \alpha = \beta \text{ or } f(A - B) = f(B - A)$$

So there are two causes for the violations of symmetry. It may happen that the subject focuses more on the differences $A - B$ than on the differences $B - A$, or vice versa, i.e. $\alpha \neq \beta$. Or it may happen thar there is a bigger number of $A - B$ than of $B - A$ differences. In other words, symmetric judgements are just a special case in which either none of the objects receives more attention than the other one by the subject or the amount of differences is the same. In Tversky's contrast model similarity is not necessarily symmetric.

Let us look more closely at Tversky's answer. The solution that the researchers from the spatial camp gave to the influence of the subject is not very different from Tversky's. Since $Sim(a, b)$ and $Sim(b, a)$ are *both* a function of $A \cap B, A - B$ and $B - A$ and there is nothing else to consider (say, no other sets of attributes), in principle nothing in these sets of features by themselves points to any difference between the similarities of objects. To account for the violations of symmetry Tversky uses two strategies. First, he introduces the external parameters α, β and θ to model the focus of the subject on either the common attributes or the differences. The parameters are weights that modify the representing values, they do not affect the represented attributes (not) shared by the objects. The specific values of the parameters are not invariant under transformations that would preserve coincidences and differences between pairs of objects. The properties of the similarity relations themselves stay intact. Moreover, as we said that is something that spatial models can mimic too. Thus this part of Tversky's solution can be interpreted as imposing either the effects of the attention or biases of the subject or the effects of the context on the similarity comparisons.

Second, Tversky represents $A - B$ and $B - A$ by different real values. In the simplest case, say when the sets of features are finite, these numbers may correspond just to the different cardinalities of the sets. In that case the values may be distinct say because one of the objects has more features than the other and thus these features make the number of differences it has with respect to the other grow. The crucial question now is how much depends the similarity between two objects on the combination of the amount of the attributes shared and on the amount of their differences. Tversky is modelling similarity as if it were an interval magnitude. Thus, degrees of similarity as given by a 'similarity scale' are invariant under scalings and translations, analogously to how values of temperature as measured by a certain scale (Celsius or Fahrenheit) are represented. According to the contrast model it makes sense to consider differences between degrees of similarity, because these are grounded on amounts of properties. To change from one similarity scale to another, one could just scale equally all the degrees of similarity between pairs of objects as given by some similarity scale, and then possibly add some additional fixed quantity to each of the resulting degrees to get to the corresponding values in the other scale. What interval scales leave invariant are quotients of differences in degrees of similarity. In other words, if we have the following degrees of similarity for pairs of objects according to some similarity scale $s(a, b) = 0.5$; $s(c, d) = 0.4$; $s(e, f) = 0.3$; $s(g, h) = 0.2$, so that we have the quotient $[s(a, b) - s(c, d)]/[s(e, f) - s(g, h)] = 1$, then it makes sense to compare arithmetic differences between the degrees of similarity. The crucial discussion here is whether overall or approximate degree-like similarity is in general something more than an ordinal magnitude, whether it always makes sense to compare (arithmetic) differences in degrees of similarity between pairs of objects.

I think that one can safely say that the degree of similarity between a and b is greater than that between c and d. Thus, there does not seem to be any sort of problem with ordinal comparisons. But what is one to make of the alleged equalities, orders or ratios between the *differences* among degrees of similarity? Does it always make sense to say that the difference between degrees of similarity $s(a, b)$ and $s(c, d)$ is, say, *five times bigger* than that between $s(e, f)$ and $s(g, h)$? Remember that we are talking here about quotients of arithmetical differences. There are some contexts in which it does. If the attributes considered are magnitudes, say lengths, then we can assume that their degrees of similarity are given by the euclidean distance between the corresponding real numbers in the real line. Objects can inherit these differences in degrees of similarity. But magnitudes are highly structured attributes. What are we to say of other attributes had by objects? Consider objects exemplifying purely categorical attributes such as being red, being religious, being socialist, and so on. How many times bigger is the difference between the degrees of similarity of two pairs of people having the same nationality and the degrees of similarity of two pairs of people having the same religious beliefs?

But I find it hard to believe that we are talking about the same relation. What I suspect that is happening in these cases is that we are making simultaneous use of several degrees of similarity at the same time. This is the third point I want to make. Categorical similarity deals only with a fixed degree of similarity at a time. A proper degree-like similarity deals with several degrees of similarity at the same time. I think that it is very important to realize that the notion of similarity being discussed by these authors seems to be comparative or even degree-like. In the spatial model, the distance metric gives to each pair of objects the degree to which they are dissimilar to each other. These degrees change from one pair of objects to another. The distance is grounded on a similarity scale, which is obtained from an ordinal relation of comparative similarity. In the contrast model, the representation function gives to each pair of objects the degree to which they are similar to each other. These degrees change from one pair of objects to another. Again, these similarity relations are grounded on an ordinal similarity scale which is also thought to be interval. One way to explain what happens in the cases of violations of symmetry (and reflexivity) is to think that the subject focuses on some properties or respects of comparison instead of others. Another way to put it is to think that the whole assignment of degrees of similarity assumed for comparison is not fixed, it changes due to the effect of the focus of the subject. In any case, my point is that whatever concerns the interaction between different degrees of similarity need not affect the notion of *categorical* similarity. This is analogous to the difference between degrees of belief and categorical or absolute belief (yes-or-no belief). Categorical belief can be understood as belief under a fixed degree d which works as a threshold and which is possibly given by the context. But categorical beliefs, comparative beliefs and degrees of beliefs have different properties and therefore correspond (at least in principle) to different concepts. When we consider whether $x \sim y$ we are taking into account just an absolute or fixed degree of similarity that works for every pair of objects (there is one for every pair), not a degree of similarity that changes from one pair to another (for every pair there is one). So it is not obvious why Tversky's objections should affect our notion of similarity.

The previous remarks may at first seem like a cheap solution. We are talking about different concepts of similarity, but is that all? There is a further point I want to make. With the model I will introduce later on Tversky's typicality effects can be explained if combined with a suitable definition of a paradigmatic entity and a comparative similarity *starting from a purely categorical one.* However, this cannot be done before introducing the relevant formal machinery. The reader is directed to the last section of this Chapter, where a full explanation can be found.

Let us sum up what we discussed. First, Tversky presented some experimental findings that showed that similarity did not obey some basic axioms of metric spaces, namely minimality and symmetry. These results could be interpreted as challenges to the reflexivity and symmetry of categorical similarity. Tversky's model introduced additional parameters to account for the effect of the subject in the similarity comparisons. These allowed him to explain the controversial cases of violation of symmetry. But these can be mimicked by the spatial models too. Second, Tversky's model makes unreasonably strong assumptions. For instance, it makes similarity an interval magnitude and some of the axioms require imposing on the degrees of similarity a structure almost as fine-grained as the real line. Third, the notion of similarity discussed both by the spatial and contrast models is either comparative or degree-like, so it is not clear that it goes against the properties of categorical similarity. In contrast, the notion of similarity we are considering is that of categorical similarity, that works under a fixed degree of similarity for all objects. This similarity relation

Figure 4.1: Counterexample to Transitivity: Colour Sequence

seems to be necessarily reflexive and symmetric. If this does not seem to be so a reasonable explanation is that we are changing the topic. In other words, we are making use of a notion similarity relation that implicitly uses several degrees of similarity (or respects of comparison) at the same time and then we are ignoring the effects of the context. Moreover, a comparative similarity can be defined from the categorical one that explains the typicality effects, by assuming the existence of paradigmatic objects. Thus there are strong reasons to think that Tversky's objections do not really damage the properties of categorical similarity.

4.2.3 Similarity is not Transitive

A third objection could be:

Similarity is not transitive

This is hardly a problem if similarity is understood as *overall* or *approximate* similarity and formally explained as it is done in this Chapter. A classical example of the non-transitivity of similarity is given by the approximate similarity between hues of colours (Poincare?). In the sequence of colours of Figure 4.1, each pair of adjacent colours is similar to each other but the first one is not similar to the last:

Once more, if one explains approximate similarity as sharing at least one attribute, violations of transitivity can be explained. Simply put, $x \sim y$ and $y \sim z$ do not imply $x \sim z$, since the property shared by x and y could be distinct from the one shared by y and z [2]. Nevertheless, this definition of similarity will not explain the counterexample of colours. That is why Armstrong redefined similarity as having either a common attribute or sufficiently similar attributes. The task of explaining what is for attributes themselves to be similar without appealing to a primitive similarity relation is not an easy task, see [33].

Spatial models of similarity, such as the one used by the Conceptual Spaces approach ([42]), tend to assume that similarity satisfies some sort of property that is reminiscent of transitivity, namely the triangle inequality. Remember that:

$$d(x,z) \leq d(x,y) + d(y,z)[TriangleInequality]$$

Therefore, what the triangle inequality says is that the degree of similarity between x and z is greater than (or equal to) the sum of the degrees of dissimilarity of x and y and y and z. From a metric similarity we can define a categorical similarity as follows:

$$x \sim y := d(x,y) \leq \epsilon, \text{ for some } 0 \leq \epsilon \in \mathbf{R}$$

Reflexivity and symmetry follow from (ii) and (iii). What this bridge principle says is that x and y are ϵ–similar or "similar enough" or "similar under the degree of similarity ϵ fixed by the context" iff their degree of dissimilarity is (equal or) less than ϵ. Here the fixed value 'ϵ' can be thought of as a bound to the degree of dissimilarity of two objects. So one can think about ϵ, roughly speaking, as a degree of similarity that is coarse enough to make x and y similar to each other. The more the degree of dissimilarity between x and y, the bigger this threshold must be (the coarser the degree of similarity between objects must be selected). Let us take a fixed value ϵ and the similarity $x \sim y \sim z$. This could simply be a substructure of the similarity induced by a metric space. It follows from (iv) that $d(x, z) \leq 2\epsilon$. So in the case where we have that x is similar to y and y to z but x is not similar to z, what (iv) implies is that there will always be a coarser degree of similarity, namely one less or equal to 2ϵ, under which x is similar to z. We simply have to double the degree of similarity to get to one that is coarse enough to subsume the two dissimilar objects.

For the triangle inequality to make sense it seems that one must take seriously the idea that degrees of similarity can be added to each other (in some weak sense), and once again, Tversky argued against the triangle inequality via some experiments. But this is not the case. The empirical structures assumed by spatial models are the so-called *proximity structures* and these do not need to satisfy anything close to the triangle inequality [128]. In other words, the triangle inequality can be considered to be just an artefact of the metric representation. The similarity used by the spatial models is basically an ordinal magnitude. So long as the proponents of the spatial approach keep this in mind (and therefore, so long as the principles and theses they propose are invariant under ordinal transformations), there is no real danger here. In this thesis we will not assume that similarity satisfies anything close to the triangle inequality and therefore we can put these worries aside.

The non-transitivity of similarity is fundamental for many of the consequences that will be explored in Chapter V. To put it simply, the fact that similarity is in general not transitive allows the domain of objects to behave as if it were a space of some sort. A hint can be given by appealing to classes of similar entities. If y can be similar both to x and z without x and z being similar to each other, then y can belong to a class to which x belongs and a class to which z belongs without x belonging to that same class alongside y and z. Thus y belongs to several overlapping classes of similar objects. In a sense one could say that y belongs to the *boundaries* of these classes. Classes of similar objects have some sort of proto-spatial structure: some of their members may belong just to that class, others may belong to several classes at the same time. This will not happen if similarity is transitive because the resulting classes will be disjoint.

4.3 Egalitarian Resemblance Nominalism

4.3.1 Naive Egalitarianism

Resemblance nominalists are unsatisfied by the appeal to universal entities as a solution to the problem of universals. According to them, there are no universal entities. There is nothing literally common to the particulars being considered. For traditional resemblance nominalists these particulars are objects (not tropes). The nominalist proposes to construct surrogates for the entities posited by the realist, in such a way that whatever the realist says about the world by appealing to universals can be mimicked by a corresponding appeal by the nominalist to collections of primitively similar objects. Although one could think about more versions of resemblance nominalism, I will limit myself to reviewing the three main contenders: egalitarian, aristocratic and collectivist resemblance nominalisms. Moreover, whenever I can I will just talk about particulars instead of restricting myself to objects. The reason is that there are analogous versions of resemblance nominalism for tropes, and many of the objections to these latter views are just analogous to the objections to object nominalism[12].

Some philosophers seem to find very counterintuitive the idea that attributes are just collections of similar objects. One of the ways this objection is further developed is as the complain that whereas some attributes are intrinsic to objects, the nominalist explains all of them as extrinsic properties based on similarities among several objects [4]. I think that this is a misleading way of thinking about the nominalist view. According to object nominalists there are no attributes. There are just primitively similar objects (similar in different degrees). In principle, their alleged attributes (including relations), aspects, modes or ways of being are not entities at all. However, different nominalists explain this in different ways. Some nominalists will defend that these attributes are just logical fictions, the result of a process of logical abstraction that consists in comparing those similar objects and doing as if there was something common to them. That is why the collections of similar objects are proposed as adequate surrogates for the attributes. This view is in principle compatible with the traditional psychological conception of abstraction. Therefore, attributes could be conceived of as the result of a psychological process of abstraction that a subject could make if she detected the relevant similarities. In other words, attributes would be psychological concepts. Some of the similarities may be just imposed by her cognitive system (abundant), whereas others may be really objective or sparse. In that sense attributes would be ontologically dependent for their existence, but not necessarily for their identity (unless they are abundant), on minds[13]. Such a nominalism is a mixture of resemblance and conceptualist nominalisms. Other nominalists will argue that there are no attributes as described by the realist, that there are just primitive facts regarding which objects are similar to which, and that the psychological processes of categorization have nothing to do with the issue. Still others may defend that the issue is just that to make sense of what realists say one can take whatever entities ontologically dependent on these similarity facts turn out to fulfill the required theoretical roles. And for these purposes it happens to be the case that taking collections of similar objects is a very convenient choice.

One of the reasons why I introduced the idea of similarity invariants is due to what I think is a constraint that resemblance nominalism should meet. According to resemblance nominalism, the internal structure of natural kinds (or of properties) is based upon relations of similarity. So long as the nominalist

[12] Comparisons between the merits and difficulties of realist and nominalist approaches can be found in any introduction to the problem of universals, such as [4]. As the reader will realize, my focus in this chapter is mainly of structural issues.

[13] Such a thesis would require treating in depth the psychological literature on concept formation, so I will not delve into that.

appeals to a categorical similarity, it seems plausible to ask for the following constraint to be satisfied:

Invariants The internal structure of a kind should be determined by similarity invariants.

The resemblance nominalist should propose explications of natural kinds (or of properties in general) that make use of similarity invariants. Why? If the fundamental structure of a kind is to be explained in terms of similarity relations among its members, then being a natural kind should be invariant under similarity morphisms. If we do not ask for this, then there is a more fundamental concept than similarity that explains what the structure of a kind is. In that case, either a weaker or a stronger property would explain the structure of kinds. Each nominalistic theory may make a different choice regarding which similarity invariants finds more explanatory. We will review now the three most popular kinds of resemblance nominalism, namely egalitarian, aristocratic and collectivist resemblance nominalism [109], but we will focus mainly on the first two. I will adopt the aristocratic view.

According to (naive) *Egalitarian Resemblance Nominalism*, properties are maximal collections of pairwise similar particulars. All the members in the class fulfil the same roles as building blocks of the class. Therefore, a particular belongs to the class iff it is sufficiently similar to all its members. The roots of this position can be found in Carnap's explication of properties as similarity circles [17]. Carnap proposed to consider similarity circles as surrogates for properties. The main challenge to this position came from Goodman's coextensionality, companionship and imperfect community problems (all of which were already anticipated by Carnap). According to the simplest version of egalitarian nominalism, the truthmaker of "x is P" is the fact that $x \sim y_1$ & $x \sim y_2$ & ... where for each y_i it is also true that "y_i is P". A simple but misleading way of expressing this is by saying that "x is P" is true iff x is similar to the P-s.

The main contemporary defendant of egalitarian resemblance nominalism is Pereyra [109]. However, his specific approach is somewhat more complicated. He approaches the main problems of resemblance nominalism, imperfect community and companionship, separately. These two problems will be introduced in more detail in the next chapter. Nevertheless, in order to understand Pereyra's approach it is convenient to give a short description of them now. The nominalist suggests taking the properties to be maximal classes of pairwise similar objects. This will not work, as Carnap himself pointed out. On the one hand, properties (or their extensions) can be properly included into one another, whereas similarity circles cannot, since they are maximal by definition. This is the *companionship problem*. On the other hand, there can be a set of objects that share pairwise a property without all of them sharing any property, a fortiori, all of them will be similar to each other and will form a similarity circle. This is the *imperfect community problem*. These two problems show that the naive egalitarian approach does not work.

4.3.2 Pereyra's Egalitarianism

In order to deal with these two problems, Pereyra makes two moves. First, instead of taking as primitive a categorical similarity relation that holds just between objects, he takes the similarity to hold between objects, pairs of objects, pairs of pairs of objects, and so on. The idea is that, if the nominalist takes all these resemblance facts as primitive, that should give her information enough to filter those classes that do correspond to properties, avoiding the imperfect community. Second, in order to distinguish between some properties that are included in others, Pereyra introduces different degrees of similarity. This allows him to avoid companionship.

In order to show how his strategy works, I will reformulate some of the theses that Pereyra defends in [109]. Nevertheless, I will try to keep the notation and assumptions as closely as possible to the original ones. Recall that the aim of the nominalist is to reconstruct a realist context (S, Q, I)[14]. From now on we will assume that such a context is finite (because Pereyra assumes so), set-theoretical $Q \subseteq \wp(S)$, satisfies the non-bare particulars and exemplification principles and is such that no property is included into another.

Definition 35. *Let (S, Q, I) be a finite context such that $Q \subseteq \wp(S)$ and $I = \in$. Let us assume that:*

i If $x \in S$, then $i(x) \neq \varnothing$.

ii If $R \in Q$, then $e(R) \neq \varnothing$.

iii If $R, P \in Q$ and $P \subseteq R$ then $P = R$.

We need this last assumption because at first Pereyra only deals with the imperfect community, not with the companionship problem. The following operator will be very useful, given an arbitrary set Y and a subset $A \subseteq Y$:

$$Pair(A) := \{\{x, y\} \mid x, y \in A \ \& \ x \neq y\}$$

This operator gives all the pairs of distinct elements in A. Now we must extend the set of objects to include all the pairs, pairs of pairs, and so on.

Definition 36. *Let $(S, Q, \in)$ be a context. Then its domain of objects* is the set S^* defined as follows:*

i $S \subseteq S^*$.

ii If $a, b \in S^*$ and $a \neq b$, then $\{a, b\} \in S^*$.

The domain now contains all objects, pairs of distinct objects, pairs of pairs of objects, and so on. The domain of properties is expanded too. The idea is that pairs of objects can also be similar to each other. In fact, if binary relations are to be reconstructed too, Pereyra says, that should be the case. Consider for instance the properties $Red = \{a, b, c\}$ and $Blue = \{e, f\}$. Then although the pairs $\{a, b\}$ and $\{a, c\}$ are similar to each other in being pairs of red objects, the pairs $\{a, b\}$ and $\{e, f\}$ are not. Since the two former pairs are similar, they must have a property in common. Of course, a pair of red objects is not itself red. Pereyra suggests that the property be something like *being a pair of red objects*. So the domain of properties needs to be expanded suitably. Recall that Q is finite. All the properties of objects are supra-indexed by a 0, we have $Q = \{X_1^0, \ldots, X_i^0\}$. We now extend the domain of properties as follows:

[14] This whole reconstruction process will be considered in detail in the next chapter.

Definition 37. *Let* $(S, Q, \in)$ *be a context. Then its domain of properties* is the set* Q^* *defined as follows:*

i $Q \subseteq Q^*$.

ii If $X^n \in Q^*$*, then* X^{n+1}:=$Pair(X^n) \in Q^*$

For example, if $Red^0 = \{a, b, c\}$, then $Red^1 = \{\{a, b\}, \{b, c\}, \{a, c\}\}$, $Red^2 = \{\{\{a, b\}, \{b, c\}\}, \{\{a, b\}, \{a, c\}\}, \{\{b, c\}, \{a, c\}\}\}$, and so on. In other words, each property X^0 induces a whole hierarchy of higher-order properties that include in their corresponding extensions the pairs of pairs of ... of the members in the extension of X^0.

Then a function that maps objects* to properties* is introduced. Let us call the resulting structure a 'Pereyra context':

Definition 38. *Let* $(S, Q, \in)$ *be a context where* S^* *and* Q^* *are defined as previously. If* $f\colon S^* \to Q^*$ *is a function defined piecewise as follows, then* (S^*, Q^*, f) *is the* Pereyra context *induced by* $(S, Q, \in)$*:*

i $f(x) = i(x)$ *if* $x \in S$.

ii $f(x) = \{X_1^{n+1}, \ldots, X_i^{n+1}\}$ *if* $x = \{a, b\} \in S^* - S$ *and* $f(a) \cap f(b) = \{X_1^n, \ldots, X_i^n\}$.

iii $f(x) = \varnothing$ *otherwise.*

In words, the function extends the intension function that maps each object to its properties. In the case of pairs, it takes the pair of elements a, b and maps it to all the higher-order properties that correspond to the properties common to them. Pereyra uses this function to introduce a similarity relation:

Definition 39. *Let* (S^*, Q^*, f) *be a Pereyra context. Then* $a \sim^* b$ *iff* $f(a) \cap f(b) \neq \varnothing$ *is the* Pereyra similarity *in* S^*.

Note that this similarity relation generalizes the usual one. This similarity holds not only between objects, but also between pairs of objects, pairs of pairs and so on. This is one step where Pereyra deviates from the original strategy.

In order to distinguish between a given set of objects, the set of its pairs, and so on, Pereyra introduces some further notation. First, all subsets α of the domain of objects S are supra-indexed by a 0, the rest are defined as follows:

$$\varnothing \neq \alpha^0 \subseteq S$$
$$\alpha^{n+1} = Pair(\alpha^n)$$

So if $\alpha^0 = \{x, y, z\}$ then $\alpha^1 = \{\{x, y\}, \{x, z\}, \{y, z\}\}$. The second point where Pereyra deviates from the original strategy is in his definition of the nominalistic properties. Instead of taking similarity circles, he proposes:

Definition 40. *Let* (S^*, Q^*, f) *be a Pereyra context and* $\alpha^0 \subseteq S$. *Then* α^0 *is a* perfect community *iff* $\forall n \in \boldsymbol{N} \quad \forall a, b \in \alpha^n \quad a \sim^* b$.

In other words, a perfect community is something like a stratified clique. It is not difficult to see that any subset of a perfect community is also a perfect community. Therefore, although all the properties are perfect communities, not all the perfect communities are properties. What Pereyra claims is that perfect communities are exactly the subsets of the original properties in Q. In other words, that they are exactly the sets of objects which are such that all their members share a property. Since we assumed that no property in the context is properly included into another, the maximal perfect communities are exactly the properties. Thus the following would hold (recall that through this section a 'context' is finite, set-theoretic, has no property included into any other and so on):

Theorem 5. *Let $(S, Q, \in)$ be a context. Then its Pereyra context (S^*, Q^*, f) induces a similarity structure $(S^*, \sim^*)$ which is such that the family of maximal perfect communities is just Q.*

As an example, take the context $Q = \{\{x,y\},\{x,z\},\{x,w\},\{y,z\},\{y,w\},\{z,w\},\{x,y,z\},\{x,w,z\},\{x,y,w\},\{y,w,z\}\}$. Let us rename the properties as XY^0, XZ^0, XW^0 and so on. In this example, the set $\alpha^0 = \{x,y,z,w\}$ is an imperfect community, since there is no property common to its four members. All its members are similar to each other, and so it forms a clique. However, it is not a perfect community, since the set $Pair(\alpha^0) = \alpha^1 = \{\{x,y\},\{z,w\},\dots\}$ is not a clique. The reason is that the pairs $\{x,y\},\{z,w\}$ are not similar to each other. Whereas $f(\{x,y\}) = \{XY^1, XYZ^1, XYW^1\}$, we have that $f(\{z,w\}) = \{ZW^1, XZW^1, YZW^1\}$, so the two sets of properties are disjoint.

In sum, in order to reconstruct the properties, the nominalist only needs to take as primitive the similarity relation $\sim^*$. She starts from the structure $(S^*, \sim^*)$. It is important to notice, as Pereyra says, that it will not be enough to consider only similarities between pairs of objects. One has to consider similarities between pairs of pairs, between pairs of pairs of pairs, and so on.

In order to prove the former theorem, something like the following lemma is needed:

Lemma 4. *Let $(S, Q, \in)$ be a context and (S^*, Q^*, f) its Pereyra context. Then $\alpha^0 \subseteq S$ is a perfect community iff there is a $X^0 \in Q$ such that $\alpha^0 \subseteq X^0$.*

The right-to-left direction can be proven by induction. One just needs to show that if $\alpha^0 \subseteq X^0$, for some property $X^0 \in Q$, then $\alpha^n \subseteq X^n$ for every natural number n. So every subset of a property is a perfect community.

However, Pereyra does not give a proof for the left-to-right direction of this lemma, which is the crucial one. Although the proposition seems to be true, showing that it is so requires some work. In page 168 of [109] he says:

> "First, if α^0 is a perfect community then, for every n, α^n is a community, while if α^0 is an imperfect community then there is some n such that α^n is a non-community. This can be easily proved since (1), of section 9.4, conjoined with the assumption that the members of a class α^0 form a perfect community allows us to reach by a sort of induction that, for every n, α^n is a community."

Here 'community' means a subset of a property. The assumptions Pereyra refers to are these (page 165):

> "(1) If certain properties are shared by certain entities then the properties shared by their pairs are the corresponding higher-order

properties. Thus if F^n is shared by x, y and z, F^{n+1} is shared by $\{x, y\}, \{x, z\}$ and $\{y, z\}$."

"(2) If an nth-order pair has a property F^n then its bases share the property F^0."

The first assumption is true by definition and does not help giving a proof, therefore Pereyra seems to be referring to the second one. In fact, this proposition (2) can help proving the result, but it itself is still unproven. For this we must establish a correspondence between each set of pairs of pairs of ... and its 'basis', the set of elements from which it has been obtained. Then (2) can be proven. Let us define the basis of a set of pairs of pairs ... by the following function $base\colon S^* \to \wp(S)$:

$$base(x) = \{x\}, \text{for } x \in S$$
$$base(A) = \bigcup\{base(a) \mid a \in A\}, \text{for } A \subseteq S^* - S$$

In particular, $base(\{a, b\}) = base(a) \cup base(b)$. For example, $base(\{\{1, 2\}, \{2, 3\}\}) = \{1, 2, 3\}$, as expected.

Lemma 5. *Let base: $S^* \to \wp(S)$ be the function just defined. Then:*

i Let α^n be induced by $\alpha^0 \subseteq S$. Then $base(\alpha^n) = \alpha^0$.

ii $base(X^n) = X^0$, for each property $X^0 \in Q$.

iii If $A \subseteq B$, then $base(A) \subseteq base(B)$.

iv $X^n \in f(\{a, b\}) \Leftrightarrow \{a, b\} \subseteq X^n$.

v If $\{a, b\} \in \alpha^n$ is such that $X^n \in f(\{a, b\})$, then $base(\{a, b\}) \subseteq X^0$.

Proof. (i) By induction. (ii) By (i). (iii) If $A \subseteq B$ then $base(A) = \bigcup\{base(a) \mid a \in A\} \subseteq \bigcup\{base(b) \mid b \in B\} = base(B)$. (iv) By induction: Let $X^n \in f(\{a, b\})$. If $n = 1$, then $X^0 \in f(a) \cap f(b)$, therefore $a, b \in X^0$. Suppose that it holds for n. If $X^{n+1} \in f(\{a, b\})$, then $X^n \in f(a) \cap f(b) = f(\{c, d\}) \cap f(\{e, f\})$ for $a = \{c, d\}, b = \{e, f\} \subseteq X^n$, by hypothesis. So $a, b \in X^{n+1}$. Conversely, let $\{a, b\} \subseteq X^n$. If $n = 1$, then $\{a, b\} \subseteq Pair(X^0)$ so $a, b \in X^0$ therefore $X^1 \in f(\{a, b\})$. Suppose that it holds for n. If $\{a, b\} \subseteq X^{n+1}$ then $a = \{c, d\}, b = \{e, f\} \subseteq X^n$, therefore $X^n \in f(a) \cap f(b)$ and so $X^{n+1} \in f(\{a, b\})$. (v) If $\{a, b\} \in \alpha^n$ is such that $X^n \in f(\{a, b\})$, by (iv) we have $\{a, b\} \subseteq X^n$, and by (ii), (iii) we have $base(\{a, b\}) \subseteq base(X^n) = X^0$. □

The contrapositional of (v), which is the assumption that Pereyra was referring to, is:

$$base(\{a, b\}) \nsubseteq X^0 \Rightarrow X^n \notin f(\{a, b\}), for\{a, b\} \in \alpha^n$$

Therefore, if the basis of a given 'n-th order pair' has no properties, then the corresponding pair has no properties either, in other words:

$$f(base\{a, b\}) = \varnothing \Rightarrow f(\{a, b\}) = \varnothing$$

Let α^0 be a perfect community and suppose by reductio that it is not included in any property $X^0 \in Q$. One has to show that there is a number n such that α^n is not a clique anymore. The simplest way to do so would be to find a natural number n which is big enough so that α^n contains an element $\{a, b\}$ whose basis is α^0. This means that $\{a, b\}$ has to include each of the elements of α^0 in one of its pairs of pairs of pairs ... How can we find such an n?

We need to find a way to represent a given set, like α^0, by a set of pairs of pairs of ... in such a way that all the elements in the set occur in one of these pairs. That such a representation must exist is plausible, given that the set α^0 is finite. Moreover, there are many ways to do it. I will just give one example on how one could do it, since filling the details can be a bit tricky.

Let N^* be the set of all natural numbers, all pairs of natural numbers, all pairs of pairs, and so on, as we defined before. Let $s\colon N^* \to N^*$ be the following auxiliary function:

$$s(n) = n + 1, \text{for } n \in N$$
$$s(\{a, b\}) = \{s(b), ss(b)\}, \text{for } \{a, b\} \in N^* - N$$

Next we define the following sequence of sets:

$$A_1 := \{1, 2\}$$
$$A_{n+1} = \{A_n, s(A_n)\}$$

The first sets in this sequence are:

$$A_1 = \{1, 2\}$$
$$A_2 = \{\{1, 2\}, \{3, 4\}\}$$
$$A_3 = \{\{\{1, 2\}, \{3, 4\}\}, \{\{5, 6\}, \{7, 8\}\}\}$$
$$A_4 = \{\{\{\{1, 2\}, \{3, 4\}\}, \{\{5, 6\}, \{7, 8\}\}\}, \{\{\{9, 10\}, \{11, 12\}\}, \{\{13, 14\}, \{15, 16\}\}\}\}$$
$$\ldots$$

Let $A = \{x_1, \ldots, x_n\}$ be a finite set of at least two elements, say, $A = \alpha^0$. The previous sequence of sets gives a way of representing A as a set of pairs of pairs of ... in such a way that every element of A appears in one of these pairs. In order to force only elements of A to appear in those pairs, some elements must appear twice. Each of the previous sets A_m has as basis 2^m, for m a natural number. For any A with cardinality $|A| = n$, one should choose the simplest representation, namely A_m where m is the smallest natural number such that $|A| = n \leq 2^m$. For example, if $A = \{x_1, \ldots, x_5\}$ has cardinality 5 one chooses $m = 3$, and so A gets represented as the set $A_3 = \{\{\{x_1, x_2\}, \{x_3, x_4\}\}, \{\{x_5, x_1\}, \{x_2, x_3\}\}\}$ whose basis is 2^3. Thus, if $|A| = n$ the function that sends A_m to A, or what is the same, 2^m to A, simply maps the number (or position) p to x_p if $p \leq n$ and to x_{m-n} otherwise (to allow for repetitions). It is clear that such a representation will contain all and only elements in A (some of them occur at most twice) as elements in pairs of distinct pairs of distinct pairs of ... Thus, the basis of such a pair is simply A. In any case, once such an A_m is found, it must belong to some α^k constructed from α^0, and so it follows from the previous results that if the α^0 lack a property, $f(\alpha^0) = f(base(A_m)) = \varnothing$, then A_m lack properties too, $f(A_m) = \varnothing$. Therefore, A_m cannot be similar to any other element in α^k and so α^k is not a clique, which contradicts the assumption that α^0 is a perfect community and finishes the proof of the lemma required.

Pereyra argues that the nominalist has to assume as primitive similarity relations between ordered pairs anyways if he is to reconstruct binary relations (as sets of ordered pairs). This is indeed the case. However, the similarity relations that need to be assumed as primitive among ordered pairs will differ from the ones needed to solve the imperfect community for monadic properties. In order to reconstruct binary relations, the set Q of the context must contain sets of ordered pairs, whereas S must contain ordered pairs. Then the induced 0-similarities hold between these ordered pairs, the 1-similarities hold between pairs of these ordered pairs, and so on. Consider the case where S also contains some of the members of these pairs and Q contains some monadic property. Then the similarities that hold between these objects and the ordered pairs and that are used to reconstruct relations are different from the similarities that hold between the pairs for reconstructing the monadic properties. In other words, there are two different kinds of similarities among ordered pairs. Some of are used to reconstruct monadic properties from the elements and the others are used to reconstruct binary relations from these ordered pairs.

There are other problems with Pereyra's solution[15]. The first one is formal, it is that half of the interesting result is missing. In other words, the previous result only gives half of the nominalistic reduction. Even if there is a way to define a similarity relation between objects, pairs of objects, pairs of pairs of objects and so on, which is such that the perfect communities are exactly the original properties, this does not say what the original similarity by the nominalist should look like. A fortiori, it is not shown either that there is indeed a unique correspondence between such a similarity and the one induced following the previous recipe. The similarity structure that the nominalist starts from must contain objects, all pairs of objects, all pairs of pairs of objects, and so on. But what other conditions must satisfy? Clearly, it is not enough to take any similarity structure $(S, \sim)$ and then extend the similarity to all the pairs in S^*. This will not do, one has to assume as primitive the whole similarity structure that includes all the pairs. The point is that some further axioms have to be given concerning the relation between the similarities between pairs and the similarities between objects.

There is another problem which concerns adequacy. The structure $(S^*, \sim^*)$ is much more complex than it seems. It is not clear to me how different Pereyra's egalitarianism is supposed to be from collectivist approaches. This strategy assumes from the outset primitive similarities between pairs of pairs of pairs of ... of objects. This is not very different from assuming similarities between arbitrary sets of objects, as several authors (like Busse [15]) have noted. In fact, one can collapse any such pair of pairs of ... of objects onto the set of objects from which it was constituted by using the *base* function. Such a move is needed to actually prove the main result. If we take a closer look at how this works, we can see that what is doing the job is the following relation:

$$X^n \in f(\{a, b\}) \Leftrightarrow \{a, b\} \subseteq X^n$$

[15] I will ignore the finiteness constrain, but it is also problematic given that the strategy should work independently of the cardinalities of the domains of objects and properties involved.

In other words, what we have here is the resemblance $\{a,b\} \sim \{c,d\}$ iff $\{a,b\},\{c,d\} \subseteq X^n$ for some n, that holds among sets. It is true that the relation is not reflexive, but one can easily restrict the attention to the sets that are similar to themselves (those whose bases are included in a property). This is in fact Goodman's collectivist solution to the problem of imperfect community in disguise. What Pereyra's solution is doing is representing each base as a set of pairs, but a resemblance relation between the bases themselves could do the job equally well. The resemblance is $A \sim B$ iff $A, B \subseteq X$ for some property $X \in Q$. This Goodmanian strategy is used by Rooij and Schulz [111] too to give a set-theoretic version of Pereyra's approach (their formulations are somewhat different though). Starting with sets allows getting rid of the cumbersome representation in terms of pairs of pairs of ... There is no much difference since similarities between pairs, between pairs of pairs, and so on had to be assumed. Moreover, the representation is also simpler in terms of cardinality, since one only assumes a resemblance relation on $\wp(S)$ and thus the domain has cardinality 2^m.

Things become even more complex with respect to Pereyra's solution of the companionship problem. The problems that this solution has are analogous to the ones just mentioned. Furthermore, this solution expands on the previous one and so inherits the cumbersomeness of the representation in terms of pairs of pairs of pairs of ... In sum, Pereyra gives an interesting way to approach the nominalist reconstruction of universal (natural) attributes. However, the strategy has several flaws. First, it is incomplete. We have been shown how to define a similarity relation starting from a realist structure (a context) in such a way that the properties are certain subsets (perfect communities) of the domain of objects. But the conditions that the primitive similarity must satisfy for the converse to hold are not given. Second, assuming as primitive similarity relations among pairs of pairs of ... seems to make this version of resemblance nominalism really a variant of collectivist nominalism, as several authors have already noted. In fact, a simpler version of the strategy could be given just by starting with a resemblance relation between sets.

4.3.3 Collective Resemblance Nominalism

Our second kind of nominalism is *Collective Resemblance Nominalism*. It appeared as a different answer to Goodman's problems for nominalism, first suggested by Goodman himself in a mereological fashion in [48] and then again by Lewis [73]. The idea is that properties are collections of collectively similar particulars. Collective similarity has some special features that make it irreducible to binary or pairwise similarity. The point of collectivism can be made by comparing the following two principles that link pairwise similarity to collective similarity (see Guigon's [50]). We will read $A \sim^* A$ as "the A-s are similar to each other collectively":

Rdistributivity If the A-s resemble each other, then for any x and y such that x and y are among the A-s, x resembles y. Equivalently, if the A-s share a property, then each pair of A-s shares a property.

$$A \sim^* A \Rightarrow \forall x, y \in A \; x \sim y$$
$$i(A) \neq \varnothing \Rightarrow \forall x, y \in A \; i(x) \cap i(y) \neq \varnothing$$

Rcumulativity If any x and y such that x and y are among the A-s resemble each other, the A-s resemble each other. Equivalently, if each pair of A-s shares a property, the A-s share a property.

$$(\forall x, y \in A \; x \sim y) \Rightarrow A \sim^* A$$
$$(\forall x, y \in A \; i(x) \cap i(y) \neq \varnothing) \Rightarrow i(A) \neq \varnothing$$

According to the collectivist, what imperfect communities show is that approximate binary similarity is cumulative. But it should not be, given that members of an imperfect community do not share a natural property. Therefore, they argue, we should assume as primitive a collective resemblance relation that is not cumulative and forget about binary similarity. As several authors like Busse [15] or Lewis have suggested, resemblance nominalism is very close to what is known as *Natural Class Nominalism*, which just takes some basic collections of particulars as fundamental, thus bypassing the problem of constructing them from a similarity relation. The boundaries are even more difficult to draw regarding collectivist approaches. We will not consider Collectivism in this thesis, but since it is one of the main competitors it is important to keep it in mind.

4.4 Aristocratic Resemblance Nominalism

4.4.1 Naive Aristocraticism

In contrast to Egalitarianism, *Aristocratic Resemblance Nominalism* says that properties are collections of particulars that are sufficiently similar to some special particulars which behave as paradigmatic members of the collection. So there are some particulars, the paradigms, that occupy a special role in the unification of the collection. A particular belongs to the class iff it is sufficiently similar to the paradigm(s) of the class. For instance, the truthmaker of "x is P" is the fact that $x \sim p$, where p is a paradigm. The truthmaker of "All P-s are R-s" is the fact that if $x \sim p$ then $x \sim r$, where p and r are some paradigms of the corresponding classess. Once again, a simple but misleading way of expressing this is by saying that "x is P" is true iff x is similar to the P-paradigm p. The first one to defend an aristocratic nominalism was Price [102], who says:

> "It is agreed by both parties that there is a class of red objects. The question is, what sort of a structure does a class have? That is where the two philosophies differ. According to the Philosophy of Universals, a class is so to speak a promiscuous or equalitarian assemblage. All its members have, as it were, the same status in it. All of them are instances of the same universal, and no more can be said. But in the Philosophy of Resemblances a class has a more complex structure than this; not equalitarian, but aristocratic. Every class has, as it were, a nucleus, an inner ring of key-members, consisting of a small group of standard objects or exemplars. The exemplars for the class of red things might be a certain tomato, a certain brick and a certain British post-box. Let us call them A, B and C for short. Then a red object is any object which resembles A, B and C as closely as they resemble one another. The resemblance

> between the exemplars need not itself be a very close one, though it is of course pretty close in the example just given. What is required is only that every other member of the class should resemble the class exemplars as closely as they resemble one another."[102]

Nowadays this position is not very popular due to some objections raised by Armstrong [2] and Pereyra [109], [110]. The main ones are:

1. There are some technical problems regarding the power of aristocratic nominalism to reconstruct properties from the paradigmatic elements.
2. The motivation for introducing paradigmatic elements seems suspicious. That some elements are more paradigmatic than others seems to be the result of a conventional choice or of a psychological mechanism that does not have so much to do with the objective structure of the world.

I will soon develop a formally adequate model of aristocratic resemblance nominalism. This should be enough to answer the first objection. Nevertheless, one of the objections launched by [110] appears to go against the formal adequacy of the model. I think that it just goes against the assumptions regarding the paradigmatic objects that such a model would make, but in any case it is worth considering:

> "A further problem is that Cargile's paradigms cannot do the work they are supposed to do, namely to collect all and only F-things. Cargile's examples involve a single paradigm. But a single paradigm cannot collect all and only the things it is supposed to collect. To see this consider the white paradigm. If it collects white things because they resemble it, it also collects non-white things, for many of these also resemble it. This is because the white paradigm will have other properties apart from being white." [110]

This is a serious problem. Let us grant that each property has at least one paradigmatic exemplar. In addition, let us suppose that each property has *at most one* paradigmatic instance (this is the sort of model I will introduce in Chapter V). Is not this assumption blatantly false? Why should each property have a unique paradigm? This sounds unrealistic. As Pereyra says, each object, including the paradigms, will have more than one property. Therefore it will not be similar just to the objects that have the property P of which it is a paradigm. Since it will be similar to objects having a different property R, it will follow that P-s and R-s are similar to each other. But this is precisely what the non-transitivity of similarity is supposed to avoid.

Let us take a closer look. First, we should not forget about the different degrees of similarity. Regarding degree d, object p may be a paradigm of property P and therefore p is not an instance of any other property P'. But if the degree of similarity d' is finer than d, p may be a paradigm of a different property P''. Or if d' is coarser than d, p may not be a paradigm at all. So although p may only be d-similar to the white things, it can be d'-similar to white and square things. The point is that whether an object is a paradigm or not may depend on the relevant degree of similarity. In short, being a paradigmatic element need not be an invariant property under some similarity transformations (in fact, according to the model to be given below, it is not invariant under similarity homomorphisms, just under polar continuity). I grant that the idea that each property has at most one paradigmatic exemplar may sound somewhat suspicious or at least like a high idealization. In fact, something like a generalized aristocratic nominalism with more paradigms can be developed, as it will be shown in Chapter V. Although an explanation for this will be given in the last sections of the next chapter, developing such a project fully falls outside of the scope of this thesis.

Regarding the second objection, the worry seems to be that it is implausible to think that nature would be structured in such a way that some members of a kind K or an attribute P would be more K-s or more P-s than others. Paradigms may have some role to play in allowing us to successfully refer to all the members of a collection of objects, and they may also have some other psychologically useful roles to fullfil (making conceptual categorisations possible, for instance). But if objects are primitively objectively similar to each other, then what does it mean to say that some objects are more paradigmatically P-s than others? The problem is that commitment to paradigmatic objects seems to introduce some subjective components into the picture. After all, it seems that the choice of some objects as paradigms instead of others would be motivated by several psychological factors, like the fact that the subject founds some objects to be paradigmatic simply because they resemble more those exemplars that he has already found before. But which objects the subject has encountered first is just a purely contingent feature of his learning process. If anything, nature is egalitarian. All the instances of a property or a kind are on a par. Any alleged difference between the members of the kind is to be attributed to the idiosyncrasies of the subject making the similarity comparisons.

But why should we think that nature has such a high level of symmetry? We might be wrong after all. Consider the structure of space. One may think that space is necessarily homogeneous: a god could permute any space-time point by any other while the overall structure of space remained intact. *Nobody would notice.* This seems prima facie reasonable, after all, what could possibly distinguish one point or location of space from another? The space should look the same from any of its locations. But it turns out that, at least for physical space, this may not be the case. Whether physical space is homogeneous or not is a substantive empirical question. Analogously, nature may be non-homogeneous too in having some exemplars of its natural properties (or kinds) clustered more closely around a 'nucleus' or 'core' and others around its 'boundary'. This may happen due to the fact that objects may belong to several classes and therefore may occur in the boundaries of these. Think about the case of the colour wheel. Some hues seem to be more red-like than others and some hues are at the boundaries of several colours. The spatial analogy is crucial here for reasons that were hinted at in Chapter III while we discussed the Hierarchy constraint. The traditional conception of classification is non-spatial and ignores the possibility of the classes being internally structured in different ways. Moreover, some of our concepts may have aristocratic structure because they successfully track the structure of the corresponding attributes or kinds. If so, then the fact that a given object is a paradigm or not is fixed by the world, not by us.

Nevertheless, the previous objections cannot be answered in more detail unless we describe a specific instance of aristocratic resemblance nominalism. In the following sections I will consider three models for aristocratic resemblance nominalism and I will argue that they are materially adequate. This will provide us a starting point for the reduction of the realist model.

4.4.2 A Topological Model

Let us introduce our first model for aristocratic nominalism. Recall the conceptual spaces approach, which was introduced in Chapter II. Conceptual spaces are spatial models of conceptual categorisation that represent objects as points, natural properties (simple concepts) as spatial regions and degrees of dissimilarity as distances. Prototypical objects are represented as certain chosen points in the space, the attributes being the collections of points which are sufficiently close to these. We could use conceptual spaces to give a model for aristocratic nominalism, by taking paradigms to be the prototypical objects that classify the rest. However, these models make use of *degrees* of similarity. What we need here is a model that appeals to a *categorical* notion of similarity. Thus, it would be convenient if there was a conceptual spaces-like approach which was closer to this categorical conception and did not appeal to degrees as represented by the distances between points. Such an approach can be found, I think, in the polar model proposed by Rumfitt in [114] and developed by Mormann in [90].

Rumfitt proposes in [114] a topological model of vagueness as a way to deal with the Sorites paradox. The crucial point is the introduction of a structure of objects that are "similar" to certain paradigmatic objects. These structures are called 'polar distributions', paradigms being called 'poles'. A simple example is the colour circle used to represent the different hues. The more paradigmatic exemplars of red, yellow, orange and so on are taken to be the poles. Any other colour in between, such as an orangish red or a yellowish green, gets mapped to the poles to which it is similar. Rumfitt maps each pole to a linguistic predicate, in such a way that the former ones determine the extension of the latter ones, thus providing a semantics for the language. His approach is inspired by the *prototype* and *exemplar* models of psychological categorisation that we mentioned before. Rumfitt's proposal has been developed by Mormann [90] as a qualitative model for Gärdenfors theory of conceptual spaces. I will consider it here as a model for aristocratic nominalism.

I will use the formulation in [90] of some of the basic notions. I introduce condition 'PII' (Identity of Indiscernibles again) for reasons that will become evident in the next chapter:

Definition 41. *Let S be a non-empty set and $P \subseteq S$. A* polar distribution *over S is a function $m\colon S \to \wp(P)$ that satisfies (1)-(2). A distribution is (PII) iff it satisfies in addition (3):*

1. *$\forall x \in S\ m(x) \neq \varnothing$.*
2. *$\forall x \in S\ \forall p \in P\ m(x) = \{p\} \Leftrightarrow x = p$.*
3. *$\forall x, y \in S\ m(x) = m(y) \Rightarrow x = y$. (PII)*

The polar distribution is denoted by (S, P, m). Elements in P are *poles* or *paradigms*. The crucial insight by Rumfitt was that the role played by poles in the polar distributions induces literally a spatial structure in the domain of objects. To explain this we need to introduce some basic notions of topology. The concepts are standard and can be found in any textbook on topology, say [141]:

Definition 42. *Let S be a set and $O(S) \subseteq \wp(S)$. Then $(S, O(S))$ is a* topological space *iff:*

1. *$S \in O(S)$ and $\varnothing \in O(S)$.*
2. *$A, B \in O(S) \Rightarrow A \cap B \in O(S)$.*
3. *$A_1, A_2, \dots \in O(S) \Rightarrow \bigcup_i A_i \in O(S)$.*

Members of $O(S)$ are called *open sets*. A set $B \subseteq S$ is *closed* iff B^c is open. $C(S)$ is the family of closed sets of S. A set $C \subseteq S$ is *clopen* iff is open and closed. Given an element x, an *open neighbourhood of x* is an open set $N(x) \in O(S)$ which is such that $x \in N(x)$. In particular, we will say that a point x is open (closed) iff $\{x\}$ is an open (closed) set. Loosely put, one can think about a topological space as a set of points S and some families (open, closed, ...) of regions closed under familiar set-theoretic operations.

Definition 43. *Let $(S, O(S))$ be a topological space and $A \subseteq S$. Then:*

1. *$Cl(A) := \bigcap\{B \in C(S) \mid A \subseteq B\}$ is the* closure *of A.*
2. *$Int(A) := \bigcup\{B \in O(S) \mid B \subseteq A\}$ is the* interior *of A.*
3. *$Bd(A) := Cl(A) \cap Cl(A^c)$ is the* boundary *of A.*
4. *$Ext(A) := Int(A^c)$ is the* exterior *of A.*

To determine whether two topological spaces are structurally similar we need the notion of continuity:

Definition 44. *Let $(S, O(S))$ and $(S', O(S'))$ be topological spaces and $f\colon S \to S'$ a function. Then:*

1. *f is* continuous *$\Leftrightarrow \forall A \in O(S')\ f^{-1}(A) \in O(S)$.*
2. *f is a* homeomorphism *$\Leftrightarrow$ f is bijective, f is continuous and f^{-1} is continuous.*

Examples:

i. $O(S) = \{S, \varnothing\}$ is the *indiscrete space* and $O(S) = \wp(S)$ is the *discrete space.*

ii. Let $S = \{0, 1\}$, then $O(S) = \{\varnothing, \{1\}, \{0, 1\}\}$ is the *Sierpinski space.*

iii. Let $(S, \leq)$ be a preordered set. Then $O(S) = \{A \subseteq S \mid A \text{ is an up-set}\}$ is the *Alexandroff topology* of S.

The indiscrete and the discrete are the two trivial or extreme examples of spaces. The indiscrete space is so coarse that one cannot use regions to distinguish between the points. In the other extreme, the discrete space is so fine-grained that each collection of points is a region. The Alexandroff topologies are the ones we are interested in, they can be equivalently introduced as follows:

Definition 45. *Let $(S, O(S))$ be a topological space. Then S is an* Alexandroff space *iff $O(S)$ is closed under arbitrary intersections.*

Alexandroff topologies correspond to preorders. As the previous example shows, every preorder induces an Alexandroff topology by taking the open sets to be the up sets. Conversely, every Alexandroff topology induces a preorder over its points called the *specialization preorder* and defined as $x \leq y \Leftrightarrow x \in Cl(y) \Leftrightarrow \forall A \in O(S)(x \in A \Rightarrow y \in A)$. Moreover, a function from one Alexandroff topology to another is continuous iff it is monotone with respect to the corresponding preorders.

The indiscrete and discrete spaces already hint at a minimal condition that any interesting space should satisfy. Namely, the space should give us enough resources to distinguish the points by making use of the (open) regions. Topologists have studied many of the so-called 'separation axioms'. For our purposes, the weakest of these axioms is enough:

Proposition 19. *Let $(S, O(S))$ be a topological space. The following conditions are equivalent:*

1. *$O(S)$ is a T_0 space.*
2. *For all x, y in S, there is an open neighbourhood of x which is not of y, or viceversa.*
3. *The specialization preorder is a partial order.*

In other words, the T_0 Alexandroff spaces are the ones corresponding to posets. These brief topological remarks are enough to consider the fruitfulness of the polar approach. It turns out that a polar distribution induces an Alexandroff topological space as follows [114], [90]:

Proposition 20. *Let (S, P, m) be a polar distribution. Let $O(S) := \{A \subseteq S \mid \forall x \in A\ (p \in m(x) \Rightarrow p \in A)\}$. Then $O(S)$ is a T_0 Alexandroff topology over S called the* polar topology.

The topological closure of a paradigm p is $Cl(p) := \{x \in S \mid p \in m(x)\}$ [114]. The specialization order is $x \leq^* y$ iff $x = y$ or $y \in m(x)$. The smallest open set for each x is $N_x = \{x\} \cup m(x)$. There is a property that characterizes this sort of spaces. First note that the set of open points in the space is exactly P, the set of poles or paradigms.

Definition 46. *Let $(S, O(S))$ be a topological space and $A \subseteq S$. Then A is* dense *iff $Cl(A) = S$. Moreover, $O(S)$ is* weakly-scattered *iff the set of open points is dense.*

A dense region in a space is a set of points which are 'everywhere', so to speak. Whichever point in the space we choose, we will always be able to find one of these points arbitrarily close to it. The most important fact about polar spaces is the following one, as shown by Mormann [90]:

Proposition 21. *Let $(S, P, O(S))$ be a polar topology. Then $O(S)$ is weakly-scattered.*

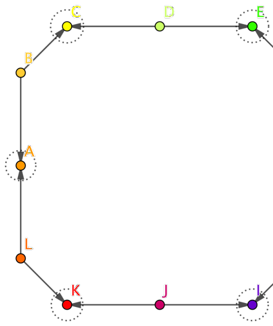

Figure 4.2: Example of Polar Space: Colour Circle

So a polar distribution is a space whose fundamental regions, the open sets, are 'centred around' the paradigms. Moreover, the paradigms, which are the open points, form a dense subset. In other words, whichever object we choose, we will always be able to find a paradigm that is arbitrarily close to it.

Let us put an example of a polar distribution. We will take the classic example of the colour wheel or colour circle, but in a discretized fashion. Let $S = \{A, B, C, D, E, F, G, H, I, J, K, L\}$ be a set of spots and $P = \{A, C, E, G, I, K\}$, where the paradigms are A (orange), C (yellow), E (green), G (blue,) I (purple), K (red) and the rest of items are coloured spots that are in intermediate position between these, e.g. B (orangish yellow) and J (purplish red). The assignment m is shown by arrows in Figure 4.2 (paradigms are circled). For instance, the smallest neighbourhood of J is $N_J = \{J, K, I\}$ and $Cl(A) = \{A, L, B\}$. This example satisfies (PII).

The model preserves the ontological commitments of aristocratic resemblance nominalism. There is just one basic category of entities, objects, and some of them (the poles) have a special property. In his proposal Rumfitt argues by assuming an implicit comparative relation of similarity of the sort "pole p is maximally similar to object x". Some of these comparisons are captured explicitly by the specialization order and the mapping that classifies objects according to their poles. Each object is mapped to a set of poles, pre-theoretically understood to be those poles to which the object is similar. The axioms impose very weak constrains. First, each object is mapped at least to one pole which is maximally similar to it. Second, an object is a pole iff the only pole it is mapped to is itself. This implies that every non-paradigmatic objects gets mapped to at least two distinct poles. If an object is a pole then no other pole is maximally similar to it, that seems plausible. The converse may seem less plausible, because it involves an indiscernibility constraint. If an object is maximally similar just to one pole, then it is in no way different to that pole. Thus it could be used as a pole for that same property too. The polar model identifies all these objects that are indiscernibles or duplicates from a given pole[16].

Moreover, the model has spatial content given by its topology. Given any object in an open region, one can always find a region which is small enough to contain all the objects that are sufficiently close to it, namely the region formed by the object and its poles. One can get as close as wanted to any object in the domain by selecting the appropriate pole. Given a pole, the set of all those objects mapped to it form a closed region. Thus any object that is mapped to a pole is spatially arbitrarily close to it. This closed region can be taken to be a surrogate for the corresponding universal attribute[17]. In other words, a property is the set of all those objects that are arbitrarily close to a given pole or paradigm. Thus, the interior of a property is just its pole, whereas the boundary contains all those non-paradigmatic instances of the property. Two properties overlap exactly at their boundaries, which contain no poles. The model has other nice topological properties (for instance, related to boundaries) that we will not deal with. The interested reader is referred to [90]. It is to be said that since polar spaces are only a special case of a more general class of spaces, that of weakly-scattered spaces, Mormann takes the latter to be the appropriate models for conceptual spaces.

In any case, the only drawback this model has is that it does not mention similarity relations at all. If the model is to be a model for resemblance nominalism then similarity should figure somewhere. The distribution suggests that objects are to be mapped to the poles to which they are similar, but the formal properties of categorical similarity are absent.

4.4.3 An Order-Theoretic Model

The second model I will consider is order-theoretic. In [90] and [91] Mormann argues for replacing polar spaces by the more general class of weakly-scattered Alexandroff spaces. One of the reasons he gives is that this allows replacing the categorical notion of a prototype by a comparative notion. Instead of saying simply that an object is a prototype, objects are ordered according to *how prototypical* they are. The poles of the polar model are then the most prototypical elements in the model (which are guaranteed to exist). This order of prototypicality is the specialization order induced by the topology. Mormann takes the class of weakly-scattered Alexandroff spaces to be adequate for his purposes of giving a model for conceptual spaces. However, my main purpose here is to give a model for aristocratic nominalism. For this reason, I will put one further constraint in order to get a model for aristocraticism that will turn out to be equivalent to the other two models.

The basic idea is this. Objects are ordered by how qualitatively rich they are. We say that $x \leq y$ iff $i(x) \subseteq i(y)$ iff y has all the properties that x

[16]One could have an even more general model where the latter requirement is dropped and there are many objects indiscernible from a given pole. In fact, Rumfitt seems to suggest this.

[17]I take the properties to be the closures of poles. In contrast, [114] and [90] take the properties to be the open sets, more specifically, the open regular sets. The reason why I choose the closures will be explained in the next chapter. Roughly speaking, it will follow that all the instances of the property are similar to each other. Moreover, the closures of poles will be exactly certain class of similarity circles. Rumfitt and Mormann choose the open sets because they want the polar structure to be a semantics for *vague predicates* and very plausibly the extensions of vague predicates are open sets. Moreover, Mormann proposes the more general weakly-scattered Alexandroff spaces as the more appropriate framework.

has and possibly more. The dual order can be interpreted as the order of *prototypicality*, $x \leq y$ iff x is more prototypical than y. In other words, the order of prototypicality is inverse to the order of qualitative richness. The more properties an object has, the less prototypical it is, and vice-versa. The order of prototypicality has some minimal elements, the objects which are such that no other object is more prototypical than they are. These elements will be called 'paradigms' and they form the set of elements for which the order of prototypicality collapses into the categorical notion of prototypicality. Up to now, what we have done is this: we took the specialization order and we put it upside-down (we took the dual order).

The new constraints I will demand are two. First, I will force each object to be a 'qualitative sum' of the paradigms which are more prototypical than it is. This condition can be interpreted as some sort of 'qualitative atomism'. Second, I will take attributes to be surrogated by ultrafilters (this will be explained below). The reasons for demanding these two features will become clearer in the next chapter (in fact, they were suggested by the results I obtained there, not vice versa).

The former idea can be explained in terms of the standard notion of *atomistic lattices* [26]. However, some posets that are not lattices are already structured enough to look like atomistic lattices, let us introduce them:

Definition 47. *Let L be a poset. Then L is* atomic *iff for each element x there is an element z such that z is minimal and $z \leq x$. If L is atomic, then L is an* atomistic poset *iff each element in L is the join of its minimal elements.*

We get the dual notions of *co-atomic* and *co-atomistic*. Atomic posets are also called *(DCC) posets*, since they are the posets that satisfy the *Descending Chain Condition*, which prevents the poset from having infinitely descending chains of elements $\cdots < x < x' < x''$. Any such chain will sooner or later hit a minimal element. The dual notion is that of (ACC) or *Ascending Chain Condition.*

Note that in an atomistic poset not every pair of elements needs to have a join, and that if a poset has a bottom element, it is atomistic iff it has just one element. So atomistic posets do not have bottoms. The following is an easy reformulation of the condition:

Proposition 22. *Let L be a poset. Then the following conditions are equivalent:*

1. *S is an atomistic poset.*
2. $\uparrow x = \bigcap\{\uparrow p \mid p \in \downarrow x \ \ \& \ \ p \textit{ is minimal}\}$ *for every x in S.*

The second formulation is the corresponding topological separation condition and it will get an interpretation right now. Let $Min(S)$ be the set of minimal elements in S and $min(x)$ the set of minimal elements below x. We have that if $x \leq y$ then $min(x) \subseteq min(y)$ by transitivity and conversely, if $min(x) \subseteq min(y)$ then $x = \vee min(x) \leq \vee min(y) = y$. From this fact the previous proposition follows.

In this model, attributes will be represented as certain sets of objects named 'fixed ultrafilters'. Let us now introduce the following notion [141]:

Definition 48. *Let L be a poset and $F \subseteq L$. Then:*

1. *F is a* filter *iff F is an up-set such that every pair of elements in F has a lower bound in F.*

2. *F is an* ultrafilter *iff F is a maximal filter.*

3. *F is a* fixed ultrafilter *iff F is a non-trivial ultrafilter and $\bigwedge F \neq 0$ exists.*

The *trivial* ultrafilter is $\uparrow 0$. When L is a lattice, a filter is closed under meets. If the lattice is complete then $\bigwedge F$ is of course guaranteed to exist. Thus, attributes have the following properties:

1. If x is F and y is qualitatively richer than x (x is more prototypical than y), then y is also F.

2. If x and y are F, then there is a z which is also F and which is more prototypical than x and y (z is qualitatively poorer than x and y).

3. Every attribute F is maximal, in the sense of not being properly included into another attribute.

4. There is an element $\bigwedge F$ which is the qualitatively richest element among all those objects that are more prototypical than every exemplar of F.

5. The set of all objects is not an attribute.

The third condition prevents companionship and corresponds to the maximality of similarity circles, whereas the last condition excludes the trivial property shared by all objects. The first, second and fourth conditions say that a property is a collection of objects that can be 'refined' until one reaches an exemplar that is more prototypical than any other exemplar of the property. This element is precisely a paradigmatic object, namely, the paradigm of the property:

Lemma 6. *Let L be an atomistic poset and $A \subseteq L$. Then A is a* fixed ultrafilter *iff $\exists p \in Minimal(L)$ $A = \uparrow p$.*

Proof. ($\leftarrow$) Let p be a minimal element. Then $\uparrow p$ is a filter and $\bigwedge \uparrow p = p$. Suppose that $\uparrow p \subseteq A$ and A is a non-trivial filter. Let $x \in A$, then there is a $z \in A$ such that $z \leq x$. Since $p \in A$ it follows that there is an element $pz \leq p, z$ in A. Since p is minimal, $pz = p \leq z \leq x$. Therefore $x \in \uparrow p = A$, which is an ultrafilter. ($\rightarrow$) Let A be a fixed ultrafilter. We prove that $\bigwedge A \neq 0$, which exists by assumption, is a minimal element. Suppose that $q \leq \bigwedge A$. Consider $A \cup \uparrow q = \uparrow A \cup \uparrow q = \uparrow (A \cup \{q\})$, which is an up-set. If $x, y \in A \cup \uparrow q$, then either $x, y \in A$ or $x, y \in \uparrow q$ or $x \in A$ and $q \leq y$. In the first case, since A is a filter, there is a $z \leq x, y$ in $A \subseteq A \cup \uparrow q$. Analogously in the second case. In the third case, since $q \leq \bigwedge A \leq x$ we have that $q \leq x, y$ and so this shows that $A \cup \uparrow q$ is a filter. Since A is a non-trivial ultrafilter, $q \in A = A \cup \uparrow q$, therefore $\bigwedge A \leq q$, so $\bigwedge A$ is a minimal element. By the other direction we just proved, $\uparrow \bigwedge A$ is a non-trivial ultrafilter. Now, $A \subseteq \uparrow \bigwedge A$ and since A is an ultrafilter, $A = \uparrow \bigwedge A$. □

So a property is the set of all the objects that are qualitatively richer (less prototypical) than a given paradigmatic object. This lemma shows that paradigms and properties are in bijective correspondence, which is a very important fact that we will make use of later on when we deal with quasianalysis.

Let us come back to the topology once more. Recall that Alexandroff spaces and preorders are mathematically equivalent. The weakly-scattered Alexandroff spaces of Mormann's model [90] are exactly the co-atomic posets, in other words, the posets that satisfy the Ascending Chain Condition. So we can consider a more specific class of weakly-scattered Alexandroff spaces, the ones that correspond to the co-atomistic posets. We just need the topological separation condition that is dual to the one we mentioned before, namely:

Definition 49. *Let $(S, O(S))$ be a weakly-scattered Alexandroff topological space. Then S is* co-atomistic *iff $Cl(x) = \bigcap\{Cl(p) \mid p \in N_x \text{ and } p \text{ is open}\}$ for every $x \in S$.*

What the previous condition says is that the set of objects that have the same properties (and possibly more) of a given object x is exactly the set of objects that belong to all the properties induced by the paradigms of x. Topologically, this says that every object z that is sufficiently close to x is sufficiently close to every paradigm of x. To put it shortly, whereas the order-theoretic model describes the order of qualitative richness, this more specific Alexandroff model describes the order of prototypicality. Nevertheless, both models are equivalent. In the next chapter it will be shown that, under an indiscernibility constraint ((SNI), (PII) or anti-simmetry) the three models we are discussing are indeed equivalent. This is I think a surprising result, for the following reason: despite the fact that weakly-scattered spaces are richer in their order, the coatomistic weakly-scattered spaces turn out to be dually equivalent to the (PII) polar spaces. Thus both models can serve the same purposes.

Of course, there are some differences between the use of the order I make and Mormann's approach. For instance, Mormann wants to deal with vagueness, and therefore takes properties to be open (regular) sets. Since I put the model upside-down, ultrafilters are closed properties and therefore will not be useful for his purposes. However, this is essential for the applications I will give of this model, since this order will make ultrafilters to be just similarity circles and therefore attributes in the nominalist sense. For another difference, Mormann suggests the class of weakly-scattered Alexandroff spaces, but he does not restrict this class, so he might reject the axiom just introduced.

4.4.4 A Similarity Model

Regrettably, the previous models make no mention at all of similarity relations between objects. In contrast, I want to propose a new model for aristocratic resemblance nominalism that makes heavy use of the resources provided by the theory of similarity structures. In [89] Mormann considers an interesting result by Brockhaus on quasianalysis [14]. We will review these results later on in chapter V. The crucial point I want to mention here is that Mormann makes a distinction between certain kinds of similarity structures for which Brockhaus' Theorem holds. This distinction makes use of the notion of the *order* of a similarity structure. The class of similarities for which the theorem holds are called 'similarity structures of order $n \leq 2$'. This class of similarity structures is linked to the existence of certain elements in the domain that have very special features regarding their relations to similarity circles. In a sense, there is enough information in the relations these elements have to other elements in the structure to 'generate' the similarity circles. For these reasons, Mormann calls them 'generators'. The class of similarity structures that, I think, provides a successful model for aristocratic resemblance nominalism is that of similarity structures of order 1. For reasons of conceptual clarity, I propose to extend the notion of order 1 both to similarity circles and to elements (to generators). This move will turn out to be very useful in the reduction steps taken in the next chapter. Moreover, it will help us to clarify the introduction of new concepts that explain the role that these special elements have in the structure. It should be clear by now that these special elements will be precisely our paradigmatic objects.

In the following definitions one can take the second condition as defining the concept appearing in the first condition, the rest are trivially proven to be equivalent conditions:

Definition 50. *Let $(S, \sim)$ be a similarity structure and $x, y, p \in S$. The following conditions are equivalent:*

i. *p is a* generator of order 1*.*

ii. $x \sim p \sim y \Rightarrow x \sim y$.

iii. $x \sim p \Leftrightarrow p \leq_{co} x$.

iv. $co(p) \in SC(S)$.

v. *$p \in int(T)$ for some $T \in SC(S)$.*

We define as $Gen(S) := \{p \in S \mid \forall x, y \in S \ x \sim p \sim y \Rightarrow x \sim y\}$ the *set of generators of order 1.* This set is defined independently of the order of the similarity. We could have called them 'transitive elements' too. I will usually drop the adjective 'of order 1' when referring to a single generator.

Definition 51. *Let $(S, \sim)$ be a similarity structure and $T \subseteq S$. The following conditions are equivalent:*

i. *T is a* similarity circle of order 1*.*

ii. *$T = co(p)$ for some $p \in S$.*

iii. *T is a clique and $int(T) \neq \varnothing$.*

Proof. (i)-(ii) by definition. (i)-(iii) Suppose (iii), let $p \in int(T)$. If $z \in S$ and $z \sim x$ for all $x \in T$, we have $z \sim p$, so $z \in T$. Therefore, T is a similarity circle of order 1. Converse is obvious. □

We define as $SC_1(S) := \{T \in SC(S) \mid \exists p \in S \ co(p) = T\}$ the *set of similarity circles of order 1.* This set is also defined independently of the order of the similarity. Similarity circles of order 1 are fundamental to the purposes of this PhD thesis. In the following chapter I will propose them as surrogates of universal attributes. In other words, an attribute will be reconstructed as the collection of all the objects that are similar enough to a paradigm.

Definition 52. *Let $(S, \sim)$ be a similarity structure. The following conditions are equivalent:*

i. $\sim$ *is a* similarity of order 1*.*

ii. $x \sim y \Rightarrow \exists p \in Gen(S)\ x \sim p \sim y$.

iii. $x \sim y \Rightarrow \exists T \in SC_1(S)\ x, y \in T$.

We will sometimes assume that our similarities satisfy the following separation axiom introduced by Mormann [89]:

Definition 53. *Let* $(S, \sim)$ *be a similarity structure. Then* S *is* (SNI) *or satisfies the* Similarity Neighbourhood Identity Axiom *iff* $\forall x, y \in S\ co(x) = co(y) \Rightarrow x = y$.

In an (SNI) similarity structure of order 1, each similarity circle of order 1 has *exactly one generator*. But for most of our purposes we do not need the full power of this indiscernibility axiom. It is enough to just assume that each similarity circle of order 1 has exactly one generator. In other words, we just identify duplicate paradigms:

Definition 54. *Let* $(S, \sim)$ *be a similarity structure of order 1,* $p, q \in Gen(S)$ *and* $T \in SC_1(S)$*. The following conditions are equivalent:*

i. S *is* pure*.*

ii. $p \sim q \Rightarrow p = q$.

iii. T *has a unique generator* p*.*

Every (SNI) similarity of order 1 is pure, but the converse is false. For instance, the similarity $y \sim p \sim x\ \ \&\ \ y \sim q \sim x$ is pure, but is not (SNI) since x and y are indiscernible. Adopting this axiom is ontologically innocuous, since given any similarity one can always force paradigms to be identical by taking the quotient of the structure. In such cases we will be dealing just with equivalence classes of indiscernible paradigms.

As we will see later on, similarity structures of order 1 have very interesting features that make them relatively simple, at least compared to their siblings of higher order. This family is easily seen to be infinite. Take an n-polygon and attach to each edge a triangle in such a way that the base of the triangle is identified with the edge. As is easily checked, the resulting similarity structure is (SNI) and is of order 1. Since this construction works for any number n of edges, the family of similarity structures of order 1 is infinite. An even simpler construction starts from a point x and attaches edges to x, forming stars (the domain can be uncountable, e.g. take the unit circle alongside the origin in the Cartesian plane and define the paradigms to be the points in the circle and the only non-paradigm to be the origin). Informative and concrete examples of similarities of order 1 will have to wait until we explore the role of quasianalysis in the next chapter. Nevertheless, we can already consider some examples:

i. $p \sim x \sim q$ is the smallest (SNI) similarity of order 1 [87].

ii. $x \sim p \sim y\ \ \&\ \ x \sim q \sim y$ is the smallest pure similarity of order 1 which is not (SNI).

iii. Let the context be $S = \{x, y, z\}$ and $Q = \{\{x, y\}, \{y, z\}, \{x, z\}\}$. The similarity, which is of order 1 but not (SNI), is the so-called 'Goodman Triangle' $x \sim y \sim z \sim x$ [87]. Here $Gen(S) = \{x, y, z\}$ is an imperfect community.

To sum up, the structures we will be concerned with are the following ones:

Definition 55. *Let* $(S, \sim)$ *be a set* S *with a binary relation* $\sim \subseteq S \times S$*. Then* S *is a* pure similarity structure of order 1 *iff* $\forall x, y \in S\ \forall p, q \in Gen(S)$*:*

i. $x \sim x$. *[Reflexivity]*

ii. $x \sim y \Rightarrow y \sim x$. *[Symmetry]*

iii. $p \sim q \Rightarrow p = q$. *[Pure]*

iv. $x \sim y \Rightarrow \exists p \in Gen(S)\ x \sim p \sim y$. *[Order 1]*

For this class of similarity structures, I will introduce the following representation function $gen\colon S \to \wp(Gen(S))$:

$$gen(x) := \{p \in Gen(S) \mid x \sim p\} = co(x) \cap Gen(S)$$

The function $gen(x)$ will turn out to be very important soon. In the next chapter it will be seen that $gen(x)$ is in fact a *non-standard quasianalysis*. This function maps each object to the set of its generators (its paradigms).

It might not be completely obvious now how Goodman's problem of imperfect community will be solved by making use of these notions, this will be explained in the next chapter in detail. Nevertheless, a hint can be given. Consider those similarity circles induced by paradigms. Any such similarity circle contains a paradigm, which is unique by purity. An imperfect community is a collection of objects that pairwise share a property but that do not have any single property in common collectively. If we assume that each property is paired with a unique paradigmatic instance, then it is easy to see that no imperfect community will correspond to a similarity circle of order 1.

In the previous sections I mentioned that in this model a comparative similarity relation could be defined in such a way that Tversky's non-symmetry effects are explained. Let us explain this as an application of the model. Let me take an example from Goldstone [46]. Suppose that the subjects rate, given the three countries Russia, Cuba and Jamaica, that their similarities are given by the following ordinal similarity scale $S(Russia, Cuba) = 7, S(Jamaica, Cuba) = 8, S(Russia, Jamaica) = 1$. Using Multidimensional Scaling Techniques psychologists show that one can infer that at least two dimensions are needed to give an adequate metric model for these similarities, for instance *political affiliation* and *climate*. In the first dimension, Russia is closer to Cuba than to Jamaica while in the second dimension Cuba is closer to Jamaica than to Russia. Now let us reduce this to the relational model by defining the following categorical similarity $x \sim y := S(x, y) \geq 5$. Two countries will be similar enough iff they have a score higher than (or equal to) 5 in the similarity scale. This results in the similarity structure $Russia \sim Cuba \sim Jamaica$, where the properties are *political affiliation* $= \{Russia, Cuba\}$ and *climate* $= \{Cuba, Jamaica\}$. Here

Russia and Jamaica are the paradigmatic countries regarding, respectively, political affiliation and climate. Now define the following comparative similarity taken from [89]:

$$T(x, y, z) := co(x) \cap co(z) \subseteq co(y) \cap co(z)$$

In words, y is more similar to z than x iff y and z are similar to all the objects to which x and z are similar (and possibly more). For example, we have that $T(Russia, Cuba, Jamaica)$ and $T(Jamaica, Cuba, Russia)$. Now we can explain what happens in cases of typicality effects of the sort "Cuba is more similar to Russia than Russia is to Cuba", Russia being the paradigmatic country here. The reason is that although it is true that $T(Russia, Cuba, Russia)$ for $co(Russia) \subseteq co(Cuba)$, it is not the case that $T(Cuba, Russia, Cuba)$ since we do not have $Jamaica \in co(Cuba) - co(Russia)$. In words, Cuba is more similar to Russia than Russia is to Cuba, because we are considering other objects to which Cuba is similar (and so other properties that Cuba has and Russia lacks) that 'interfere' so to speak in our similarity judgement. In fact, in this sense Cuba is just as similar to Russia as Russia is to itself since we have both $T(Russia, Cuba, Russia)$ and $T(Russia, Russia, Russia)$. Even more, so long as one of the objects is paradigmatic while the other is not (both being similar to each other) in the sense to be explained when we introduce the model below, this is guaranteed to happen:

Corollary 4 (Typicality Effects)**.** *Let $(S, \sim)$ be a similarity structure of order 1. Let $p \in Gen(S)$, $x \in S - Gen(S)$ and $p \sim x$. Then $T(p, x, p)$ but $\neg T(x, p, x)$.*

Proof. $T(p, x, p)$ iff $co(p) \subseteq co(x) \cap co(p) = co(p)$ but $T(x, p, x)$ iff $co(x) \subseteq co(x) \cap co(p)$ iff $co(x) \subseteq co(p)$ iff $co(x) = co(p)$. So $x \in Gen(S)$, which contradicts the assumption that x is not paradigmatic. □

Now that the model is in place, I claim that it is plausibly materially adequate, since it preserves directly the ontological commitments of aristocratic resemblance nominalism. First of all, the fundamental ontology consists of objects in relations of resemblance. All entities belong to the same category, although some of them have special features, since they act as paradigmatic objects. Second, there are just four axioms.

1. Every object is similar to itself.
2. If an object is similar to another, the latter is similar to the former.
3. Two similar paradigms are indiscernible (identical in the model).
4. If two objects are similar, then they are both similar to a common paradigm.

The first two are the basic logical properties of a binary categorical similarity that have already been argued for. The third one is an indiscernibility constraint that identifies in the model any two similar paradigmatic objects. It implies two things. First, that we identify any two duplicate paradigms. From this point of view the condition is purely technical, it requires only that for the most applications that we will make of the model, whenever two duplicate paradigms occur we will consider the class of all objects duplicate to these ones and we will only choose one as a 'representative' of the class. This means simply that the model cannot distinguish between two similar paradigms, because these are considered to be indiscernible. But it also implies a second, and perhaps more controversial, constraint: that each paradigmatic object exemplifies a unique attribute. A fortiori, no paradigmatic object is similar to another paradigmatic object. I will try to give some arguments for it in the following sections. Nevertheless, if one thinks about the categorical similarity as a degree of similarity implicitly given by the context, the condition only says that no paradigm is sufficiently similar to another paradigm. This seems plausible. In a given context, if p is a paradigm of *Red* and q is a paradigm of *Round*, then it makes sense to consider the degree of similarity as being high enough to prevent p from being similar to q. A coarser degree of similarity according to which p and q are similar would defeat the very purpose of appealing to paradigms to classify the rest of the objects in this context. If p is also round, then if there is another object x which is red but not round, then x and not p is more plausibly the paradigm of *Red*.

The fourth condition is the one that does all the work. It says that there are enough paradigms to explain the similarity between objects. This seems to be the core requirement of aristocratic nominalism. Any similarity between a pair of objects will correspond, according to the realist, to their having a common attribute. Any attribute will have at least one paradigmatic exemplar and exemplifying the attribute will be equivalent to being similar to that paradigmatic exemplar. Therefore if two objects are similar, then there will be at least one paradigmatic object to which both are similar (namely, the paradigm of the attribute they will have in common). Although a property may have several paradigms, all of them will be exactly similar and therefore by (3) the uniqueness of the paradigm will follow.

In addition to the previous axioms, the model uses the following definition of a paradigmatic object:

Paradigm An object is a paradigm iff any two objects that are similar to it are similar to each other.

Finally, in the next Chapter the model will be used to reconstruct attributes as similarity circles of order 1. Thus it is convenient to state already that:

Attribute An attribute is the collection of all the objects that are similar to a given paradigm.

There are two comments to make. One is that in contrast to the Pricean view that introduces paradigms comparatively, this one is a purely categorical notion. Objects are not more or less similar to the paradigms than to other objects. Either they are similar to them or they are not. The second comment is that, as we said, it will follow in the model that (relative to a fixed degree of similarity) a paradigm exemplifies exactly one attribute. This implies that Pereyra's worry concerning transitivity, that a paradigm may collect exemplars of two different attributes, will not arise. Paradigms are the objects that act transitively by definition. Given that they are to fulfil the role of grounding the unity of the class, I think that this definition of a paradigm is reasonable. If an object is paradigmatic (relative to a fixed degree of similarity), then any two objects that are similar to it will be similar to each other. That is in fact the

	P	R	W
p	X	X	X
x	X		X
y	X	X	
z		X	X

Table 4.1: Imperfect Community Example

role of a paradigmatic object according to aristocratic resemblance nominalism. The converse may sound more suspicious. May not be the case that any two objects that are similar to a certain given object p are similar to each other *by coincidence*, without p being a paradigm? If there are imperfect communities, something like this may seem to happen. An example is given by the Table 4.1.

The similarity induced by this property assignment will be the blob. Every pair of objects are similar to each other. A fortiori, all of them satisfy the right-hand side of the definition of a paradigm. But is any of them a paradigm? Since we will take certain similarity circles as surrogates for attributes, the whole domain $S = \{x, y, z, p\}$ will be a similarity circle and thus any of the elements could in some sense be considered as a paradigmatic exemplar of S (as a property). But S is an imperfect community that does not correspond to any of the previously given ones. Thus one could object to the definition by appealing to a pre-theoretical notion of a paradigmatic object and pointing to this imperfect community. Note that the similarity induced by this counterexample does not satisfy purity (the third condition), since all the objects are paradigms. So if we had strong reasons to accept that axiom, it would not be a counterexample at all.

However, let us grant that one could reject such an assumption. Do we have such a pre-theoretical notion of a paradigm? I am not sure that this is so. In any case, this prompts an interesting question: what is it to be a paradigm? I will consider two possible answers. According to the former one, paradigms are ontologically special entities, they are limiting cases of property bearers or qualitative atoms. Thus properties and kinds are really aristocratically structured. According to the latter one, paradigms are not ontologically different from the other objects. The difference is a difference in their epistemological roles that it is brought up by the processes of conceptual categorization that occur during the learning period of a given epistemic subject. Thus the aristocratic structure of properties and kinds is 'put there' by the subject. This makes aristocratic positions a mix of resemblance nominalism and conceptualism. These answers give two completely different pictures of what these paradigms are supposed to be and will be considered later on.

4.5 What Paradigms could Be

I have suggested that an aristocratic version of resemblance nominalism is still defensible and I have given two formal models for it. Nevertheless, a crucial issue concerns the nature of these paradigmatic entities. I will consider two proposals. According to the first one, the difference between paradigmatic objects and non-paradigmatic objects is based on the ways in which a given epistemic subject may learn about the world (and about the kinds in the world). According to the second one, the difference between paradigmatic objects and non-paradigmatic objects is based on the ways that the entities themselves are related to each other. Whereas both approaches are compatible with the idea that the degrees of similarity are given by the world, the former one puts more weight on the role of an agent and introduces conceptualistic components into the theory. Whereas the former one puts resemblance nominalism closer to conceptualism, the latter one makes it closer to trope theory.

4.5.1 Paradigms as Constituents of Concepts

The first approach to the nature of paradigms basically follows the psychological exemplar theory of concepts that was explained before. According to it, paradigms are just some objects that a given epistemic subject has found in his concept-learning process. The subject stores mental token copies of these objects in his memory and uses the latter ones as fixed reference points to build his concepts around them. Thus when he finds two new objects x and y, he compares how similar they are to a mental copy of an object p that he has previously found. If they are similar enough to it then the extension of the concept induced by p is enlarged so that x and y fall under it. For any two objects x and y that appear to be similar enough, he will be able to find out a stored paradigm p to which both of them are similar.

The main advantages of this position are epistemological. First, it can be more easily explained how could one ever get and store knowledge of a given attribute or kind. Both attributes and kinds are concepts (kinds are complex concepts), and these are just mental entities developed by the subject based on certain objects he has already found. One does not know the whole extension of the attribute or kind, only some of its members. To know whether a newly encountered object belongs to the kind he just has to compare it to the paradigms instead of comparing it to all the members of the kind (which would be an impossibly time-consuming task). As the subject learns more about the world the corresponding extensions of the concepts grow, shrink or get modified in different ways. The subject gets to generalizations that purport to be about all the members of the attribute or kind by induction from this limited sample and then tests these against newly encountered objects. Second, one can explain why some attributes or kinds seem to overlap by considering these as cases of vagueness, as has been done using conceptual spaces (see [146]). The same object falls under several concepts because it is similar to various paradigms. Overlappings of attributes and kinds are the rule rather than the exception and explain in which sense concepts themselves can be vague. Vagueness would then be explained as a mental phenomenon. Third, the position is more compatible with naturalism and in particular, it is very close to the psychological accounts of conceptual categorization in terms of prototypical instances or exemplars as depicted in [42], and has lead to empirically testable hypotheses. The conceptual spaces approach seems to assume something like this conception of paradigmatic entities and has shown it can do some fruitful work. Evidence for this can be found in the fast development of the conceptual spaces approach (see again [146]).

In fact, Douven and Gärdenfors [30] have argued for several 'design principles' that natural concepts should satisfy. They reason by analogy as follows.

Suppose that we had to develop a conceptual scheme to allow for a system to make correct, sufficiently fine-grained and successful classifications. Of course, such a system would be limited in several ways, for instance, it would have a limited memory, limited perceptual capacities (which restricts the fine-grainedness of the classifications it would make), it would be forced to make decisions quickly, it should be able to successfully communicate with other systems, and so on. According to the authors, in order to design such an *optimal* conceptual scheme, it would be plausible to follow these principles, which we quote directly from [30][18]:

1. Parsimony: The conceptual structure should not overload the system's memory.
2. Informativeness: The concepts should be informative, meaning that they should jointly offer good and roughly equal coverage of the domain of classification cases.
3. Representation: The conceptual structure should be such that it allows the system to choose for each concept a prototype that is a good representative of all items falling under the concept.
4. Contrast: The conceptual structure should be such that prototypes of different concepts can be so chosen that they are easy to tell apart.
5. Learnability: The conceptual structure should be learnable, ideally from a small number of instances.

At the level of the whole system of concepts, *parsimony* and *informativeness* pull in opposite directions, as the authors rightly point out. The more concepts we have, the more informative and less parsimonious the system of concepts will be, and vice versa. A system with many concepts can be used to make many inferences about the objects classified, but this makes all the concepts involved harder to remember. Dually, a system wich requires remembering fewer concepts is easier to use, but it supports fewer interesting inferences and thus gives less information about the items classified.

As we have seen several times in this thesis, this dual trade-off is present in the different models of kinds. In the case of the similarity models discussed in this chapter (or the special cases of equivalences and partitions), parsimony and informativeness concern how fine-grained the similarity relation chosen is. The more fine-grained the similarity is, the greater the amount of attributes (similarity circles) and also of paradigms will be, and dually. In the case of the concept lattice model of kinds, this fact is well illustrated by Kant's Law. In order to achieve a parsimonious system, the extensions of concepts have to be large. But the larger the extensions of the concepts are, the smaller their intensions will be, and a fortiori, the concepts will be less informative. Dually, in order to achieve an informative system, the intensions of concepts have to be large. But the larger the intensions of the concepts are, the smaller their extensions will be, and a fortiori, the system of concepts will be less parsimonious.

Similarly, at the level of the paradigms chosen, *representation* and *contrast* pull in opposite directions too. Whereas representation requires maximizing

[18]See [91] for arguments by Mormann that the class of weakly-scattered spaces also satisfy the design principles.

the similarities among the instances of a concept, contrast requires maximizing the dissimilarities between paradigms of different concepts. In the case of the similarity model, consider again the basic axioms and definitions, now restated in terms of concepts and prototypes:

- No two prototypes are similar to each other.
- An item is a prototype iff any two items that are similar to it are similar to each other.
- If two items are similar, then they are both similar to a common prototype.
- A (simple) concept is the collection of all the items similar to a given prototype.

On the one hand, representativeness is guaranteed by combining the axiom of order 1, that forces two similar objects to be similar to a common paradigm, with the definition of a paradigm and the definition of attributes as similarity circles of order 1, which makes all the objects similar to a common paradigm similar to each other and the resulting collection to be maximal. On the other hand, contrast is assumed by the axiom of Purity, which states that no two paradigms are sufficiently similar to each other. In other words, one can interpret the assumptions made by the similarity model directly as the requirements that the conditions of representation and contrast be satisfied.

Finally, parsimony and *learnability* are very tightly linked to each other. However, whereas the former involves the whole system of concepts, the latter is concerned with the amount of prototypes required. A parsimonious system requires fewer concepts. Since each concept corresponds to a (set of) paradigm(s), a fewer amount of paradigms will make concept learning easier for the agent. In the case of the models proposed here, there is a bijection between attributes and paradigms. This means that the amount of paradigms posited is the minimal one that guarantees such connection between parsimony and learnability: in order to learn a concept one has to store just one paradigm and then compare every new instance of the concept to this paradigm. And in order to remember a concept, one only has to remember the corresponding paradigm. From a conceptualist point of view though, a unique paradigm for each concept is a quite strong idealization though.

The main disadvantages are ontological. The main problem is that the choice of paradigms seems to be completely arbitrary. If the subject had had different experiences or had lived in a different environment, he would have encountered different objects. Therefore, the objects that would have given rise to the mental paradigms would have been different. In other words, paradigms are not 'chosen by nature'. The difference between a paradigm and a non-paradigmatic object is a difference simply in which of the two objects the subject finds first in his learning process. This objection is again the one made by Pereyra [109]. In other words, the choice of an object as a paradigm may have more to do with psychological biases, social and cultural prejudices or the scarcity of exemplars of certain kinds than with tracking real similarities.

Of course, it is highly unlikely that a Japanese person develops her concept of BIRD by storing in her memory the same exemplars that I did for my concept of BIRD. The paradigms chosen by each of us are distinct. However, this does not imply that the choice of paradigms is *arbitrary*. It is not so clear that every object found by the subject could do as a paradigm for a concept. If the exemplar is badly chosen and the degrees of similarity are fixed by the world, the subject may not find out enough new exemplars to develop a sufficiently stable concept. The objects he finds may not be similar enough to the stored one and the resulting concept will be unfruitful enough to be left apart. Or he may fix some new objects as paradigms and find out that what he took as a paradigm was in fact similar enough to the new ones to fall under the newly developed concept. In such a case, the subject would modify his system of concepts. Moreover, if the choice of an object as paradigm of a concept was completely arbitrary, this would not explain how it is that psychologists have found stability among subjects regarding which objects are considered to be more or less paradigmatic or prototypical for each kind. Furthermore, this stability does not only hold for the members of a given community, it holds across different cultures too. This debate between *conceptual universalists* and *conceptual relativists* is not solved yet. It could be that relativists are right after all and there are no universal concepts shared by the members of all the different cultures. Nevertheless, the point is that the thesis that paradigms are prototypical instances of concepts is compatible with the idea that the subject is tracking real similarities. Some of the similarity relations he finds between a new exemplar and the mental copy of the object (chosen to be paradigmatic) may be natural or sparse. Others may be abundant. To have some hints regarding whether they are more likely to be sparse than abundant the concept will have to be used for serious research and shown to be fruitful. But the subject does make a choice regarding the relevant degree of similarity, for any two paradigms will not be similar to each other under that same degree he is choosing as 'yardstick'. Such a position will be a hybrid view between resemblance nominalism and conceptualism. It is a resemblance nominalism, since it accepts that objects are in objective resemblance relations to each other and that the similarity satisfies necessarily (de dicto) the formal properties of being reflexive and symmetric. But it also has conceptualistic components, since it considers attributes and kinds to be extensions of those concepts (as collections of similar objects) developed by a subject after making some choices regarding which objects should be taken as paradigmatic with respect to some given degree of similarity.

Moreover, the fact that objects are only paradigms relative to the choice of some degree of similarity may receive some support from the model. *Being a paradigm* is a similarity invariant, in the sense that if p is a paradigm of $(S, \sim)$ and $(S', \sim')$ is f-isomorphic to S, then $f(p)$ will be a paradigm in S'. But finer or coarser similarities may not preserve paradigms. Thus, being a paradigm is relative to the degree of similarity that is being considered. For instance, as we will see, the similarity $\sim$ induced by $SC(S) = \{\{p, x\}, \{x, y, z\}, \{q, y\}\}$ is (SNI) of order 1 and here r is a paradigm (whereas x and y are not). But now consider a finer degree of similarity $\sim' \subseteq \sim$ given by $SC'(S) = \{\{p\}, \{x, r\}, \{r, y\}, \{q\}\}$. This new similarity is also (SNI) and of order 1. Here the object r is no longer a paradigm, whereas y and x have become paradigms. Objects are paradigms relative to a certain degree of similarity. If that is the case, then no object is a paradigm in an absolute sense. One has first to choose the degree of similarity to be considered.

Under this conception of paradigms the model does not (at least directly) reflect the ontological structure of the world, but the epistemological states of a given subject that purports to know about it. Which are the objects is given by the world, alongside all the degrees of similarity. But it is the subject who has chosen some degree of similarity as relevant and some of the objects as paradigmatic relative to that degree. To learn about these objects, at a given time t the subject chooses a degree of similarity and some of these objects as paradigms. The choice of the degree of similarity may have different motivations. The subject may have some theory concerning the objects that he wants to test, and chooses that degree of similarity as a sorting criterion. Or what may be more usually the case, the subject being psychologically disposed to make such a choice. For any new exemplars he finds, he compares it to mental copies (stored in his memory) of the objects he chose as paradigms. If it is similar enough to one or more of them (that is to say, similar up to the chosen degree of similarity), then it will fall under the corresponding concept(s). If not he may forget about it or form a new concept by taking it as a paradigm. Thus the subject goes on enlarging the concepts or modifying them in view of the new evidence gathered. From some of these concepts, say of attributes, he may develop more complex concepts of kinds. He tests the usefulness of these concepts by comparing them to the world. Some degrees of similarity will be highly theoretically loaded and may result in very fruitful (e.g. predictive) concepts. These concepts may hint at natural attributes and kinds. Others may only be tracking too coarse degrees of similarity that correspond to more abundant attributes and kinds. The subject does not have access to all the exemplars of the attribute or kind, just to a limited sample of them (those from which he has developed his concepts). To put it shortly, the model does not check (directly) whether the subject is tracking real kinds, but how the corresponding concepts of the epistemic subject are developing[19].

4.5.2 Paradigms as Qualitative Atoms or Tropers

We just saw one way to interpret paradigms. But is there any other view? What could these paradigmatic objects be, if not objects selected by an epistemic subject? Let us take the description of a paradigm *literally.* At first glance, these entities look like tropes. A particular instance of the property *Crimson* looks like the sort of entity that could be a paradigm of crimson: for a given fixed degree of similarity, such an entity only has the property of being *Crimson* (it is also red, say, for a coarser degree of similarity) and any two entities that are sufficiently similar to it (up to the same degree) will be crimson and therefore will be similar to each other. However, tropes are not objects, they are properties. Since I want to remain as closely as possible to object nominalism, let us see whether there could be objects that could fulfill such description.

In order to understand what these paradigmatic objects would be, one can appeal to the distinction developed by [75] and [41] between *tropes* and *tropers* or *modifier tropes* and *module tropes.* According to these authors, the notion of trope is ambiguous. On the one hand, tropes may be considered to be particular characterizers. On this account tropes are properties, they are called 'modifier tropes'. On the other hand, tropes can be considered to be very special characterizable particulars, called 'tropers' or 'module tropes'. Tropers are

[19]Moreover, it gives a frozen picture of the knowledge of the subject for some time t (and some implicitly conveyed restricted spatial location.

qualitatively thin particular objects. Pre-theoretically speaking, a troper is an object that has a unique property. Thus, the notion of a troper is dual to the notion of *haecceity*. Consider the following example used by García. Take two balls a and b. Ball a has a 'trope' *sphericity-a* which exactly resembles another trope *sphericity-b* that b has. According to a trope theorist, *sphericity-a* is a property of the ball a and is not itself spherical. According to a troper theorist, *sphericity-a* is an object and is itself spherical (although the troper theorist will posit neither a trope *sphericity-sa* nor a universal *sphericity* to account for this fact of course). Whereas according to a troper theorist *sphericity-a* is an object, and plausibly is a mereologically ordinary part of the object a, according to a trope theorist *sphericity-a* is a property and it is therefore either a mode of a or a constituent of a, but not an ordinary part of a. What links a troper with an object is the ordinary mereological relation of parthood, after all, tropers are just objects. What links a trope with an object is a relation of characterization, which can be either primitive or surrogated as membership to a class of compresent or existentially dependent tropes. Thus a troper resemblance nominalist is an objectual resemblance nominalist that assumes a further distinction between qualitatively coarser or thinner objects. Suppose that Alice points at the ball a and says "a is round". According to the trope theorist, Alice is pointing at a property, the trope of roundness that characterizes a. According to the troper theorist, Alice is pointing at an object which is part of a, a qualitatively thin part of a which only has one property, namely, that of being round. One could say that tropers are *qualitative slices* of ordinary objects.

Thus the object resemblance nominalist can make sense of the existence of tropers by assuming that objects are ordered by qualitative thinness. By considering the previous explanation, I propose to render this ordering as follows:

> x is qualitatively thinner than y iff if x has a property P, then y has P too.

Which corresponds in the model to the order induced by the intension or quasianalysis:

$$x \leq_q y \Leftrightarrow q(x) \subseteq q(y) \Leftrightarrow i(x) \subseteq i(y)$$

A fortiori, it makes sense to consider those objects which are *maximally qualitatively thin* (equivalently, minimally qualitatively thick). As I will show later, such a distinction need not be taken as a primitive fact, it follows from the resemblance relations that hold between objects. Indeed, one can define the order of qualitative thisness as follows. Let x and y be objects, then:

> x is qualitatively thinner than y iff if z is similar to x, then z is similar to y.

Which corresponds in the model to the order induced by the similarity neighbourhood:

$$x \leq_{co} y \Leftrightarrow co(x) \subseteq co(y)$$

The tropers will be qualitatively minimal objects, that is to say, objects which are such that no other object is qualitatively thinner than they are. The consequences of this fact will be seen in the next Chapter.

Once the classical resemblance nominalist acknowledges this distinction, he obtains the benefits of the trope theoretic positions. First of all, think about what it is for an object to exemplify a property. It is for it to be similar to one of its parts, which is a troper. If two objects x and y have a common property P, then each x and y have as parts tropers t_x and t_y which are duplicates. Thus what makes x and y be similar to each other is that they are both similar to tropers t_x and t_y. Any such a troper is as good as any other as a candidate for a paradigmatic instance of property P. Another interesting feature of tropers is that they solve the coextensionality problem without requiring any appeal to modal realism, just as tropes do. Suppose there were two coextensional properties F and G. Then F and G are both exemplified by at least one common object x. This object x has two tropers t_F, t_G as parts, the former one has the property F and the latter one has the property G. Due to their qualitative thisness, t_F is not G (and t_G is not F). But t_F is in the extension of F, because t_F *is an object* that exemplifies F. Since the extension of F is identical to the extension of G, this implies a contradiction. Therefore, there cannot be two coextensional properties. The existence of qualitatively thin objects prevents the coextensionality problem from even starting.

In the next chapter I will show in more detail how the appeal to paradigmatic entities also gets rid of the imperfect community in a new and simple way. Nevertheless, some hints of the solution can be given now informally by interpreting paradigms as tropers. Consider again our previous assumptions, now rephrased:

- Two tropers which are similar to each other are indiscernible (identical in the model).
- An object is a troper iff any two objects that are similar to it are similar to each other.
- If two objects are similar, then they are both similar to a common troper.

As I said, the first one is not a real commitment, it is a purely technical condition. Whenever it is not satisfied it can be forced by considering instead equivalence classes of duplicate objects. If there are tropers, then when we say "an object is similar to a paradigm" we mean "any object from this class of duplicate objects is similar to a troper from this class of duplicates" which is equivalent to "all the objects in this class of duplicates are similar to all the tropers in this class of duplicates". For instance, if we have two tropers of redness, we consider the class of all the tropers of redness and take any of each elements as the paradigm. The second one is the definition of a paradigmatic object. Suppose that p is a troper. If x and y are approximately similar to p, since p 'only has one property P' (pre-theoretically speaking), x and y must be similar to each other. Now we can show that if we take tropers as candidates, the converse plausibly holds. Suppose that any two objects x and y that are similar to p are similar to each other. Let p have two (possibly identical) properties P and R. Then p has two tropers as parts, t_P and t_R, one for each of these properties. Since p and t_p are P, they are similar to each other. Analogously, p and t_R are similar to each other by being R. By the assumption, t_P and t_R are similar to each other. Since the only property that t_P has is P and the only property that t_R has is R, P must be identical to R. Therefore p is a troper, identical to t_P and t_R. Therefore, tropers are exactly the objects satisfying the definition of a paradigmatic object. Lastly, if two objects x and y are similar then 'they have a common property P'. We can find in x a troper which is a duplicate of a troper in y. These tropers are objects that only have that property P. Thus x and y are both similar to these two tropers, which are paradigms by the previous principle. Positing paradigms and interpreting them as tropers gives us an aristocratic objectual resemblance nominalism that can deal with Goodman's objections. This can already be taken as a good reason to accepting their existence:

Fruitfulness Argument for Paradigms

i. If assuming the existence of certain entities solves the difficulties faced by a theory by explaining why the latter do not arise, then plausibly these entities exist.

ii. Assuming the existence of paradigmatic-like entities such as tropers solves Goodman's objections.

iii. Therefore (plausibly), paradigmatic-like entities such as tropers exist.

Furthermore, if paradigms are taken to be tropers, since tropers are objects this does not force the object resemblance nominalist to go beyond its original ontological commitments.

Is there any other argument that may back up the thesis that such entities exist? Here a known quote by Bertrand Russell may be helpful:

> "When I speak of 'simples', I ought to explain that I am speaking of something not experienced as such, but known only inferentially as the limit of analysis. It is quite possible that, by greater logical skill, the need for assuming them may be avoided."

One way is to think about these entities as *limiting cases* of property bearers. They are what it would be left if we deprived an object of almost all of its properties except for one. It will be useful to argue by analogy here. Consider space. Spatial entities, like regions, can be thought to be ordered by their being contained inside one another by their extension. We can conceive of a spatial entity which is such that no other spatial entity is contained in it (it lacks proper subregions). How do we do it? We conceive first of an extended region and we go on slowly 'shrinking it' by 'deleting its extension' until we get to an atomic spatial region. These atomic regions are spatial entities which have a minimal extension (or no extension at all), they are called *points*. Now consider time. Temporal entities, like intervals, can be thought to be ordered by their being 'contained' inside one another by duration. We can conceive of a temporal entity which is such that no other temporal entity is contained in it (it lacks proper subintervals). How do we do it? We conceive first of an interval with duration and we go on slowly 'shrinking it' by 'deleting its duration' until we get to an atomic temporal interval. These atomic intervals are temporal entities which have a minimal duration (or no duration at all), they are called *instants*. Next, consider the parthood relation. Composable or mereological entities, say parts or wholes, can be thought to be ordered by their being 'contained' inside one another by parthood relations. We can conceive of a mereological whole which is such that no other mereological whole is part of it (it lacks proper parts). How do we do it? We conceive first of a mereological whole and we go on slowly 'deleting its parts' until we get to an atomic whole. These atomic wholes are the minimal parts, they are called *atomic parts* or *mereological atoms.* Finally, consider properties or similarity relations. Entities that have properties or are in similarity relations, say objects, can be thought to be ordered by their being 'contained' inside one another by these properties or similarity relations. More precisely, they can be thought to be ordered by their qualitative thinness. An object x is qualitatively thicker than y iff x has all the properties that y has (and possibly others) iff x is similar to all the objects to which y is similar. In a sense, x is qualitatively richer than y, any adequate description of y will be an adequate description of x, and there will be richer descriptions of x that do not apply to y. We can conceive of an object which is such that no other object is qualitatively thinner than it (no other object has less properties). It is an object which is qualitatively impoverished. How do we do it? We conceive first of an object and we go on slowly 'deleting its properties' until only one of them remains. These atomic objects are the ones which are minimal with respect to similarity or having properties, we can call them *paradigms*, *tropers*, *module tropes* or *qualitative atoms*[20]

Thus, paradigms or qualitative atoms are the similarity analogues of points, instants and atomic parts. They are the building blocks of the similarity structure, just as points, instants and atomic parts are the building blocks of spatial, temporal and mereological structures. Just as a region can be thought to be spatially composed of points, an interval temporally composed of instants and a whole mereologically composed of its parts, an object can be thought to be qualitatively composed of its qualitative atoms.

Analogical Argument for Paradigms

i. Spatial, temporal and mereological entities are the bearers of spatial, temporal and mereological relations respectively.

ii. Spatial, temporal and mereological entities are spatially, temporally and mereologically composed of spatial, temporal and mereological atomic entities, namely points, instants and atomic parts, respectively.

iii. Objects are the bearers of properties or similarity relations.

iv. Therefore (analogically), objects are qualitatively composed of similarity or property-like atomic entities, namely paradigms or qualitative atoms.

I grant that an argument by analogy is usually not a very powerful argument. The success of the analogy depends on how similar the two situations are in those respects that are relevant to make the comparison. What is relevant here is the common formal ontological, plausibly compositional, structure. But one may worry that, even though we can make sense of points, instants and atomic

[20]There are more examples of this pattern, for instance singletons, point-events, basic actions, logical atoms, and so on. I am tempted to suggest that all of them are, in some sense, cases of mereological composition.

parts as being 'inside' or 'contained' in regions, intervals and wholes, that is not the case for objects. Objects are not 'qualitatively contained' in one another. What is this *qualitative composition*? The first thing to say is that this idea is not completely alien to philosophers. It is, after all, the basic intuition lurking behind (universal or trope) bundle theories of objects. Moreover, the idea makes sense under some substance-attribute theories too. For example, according to the immanent realist, objects do literally overlap due to their properties being in them. So if y has all the properties that x has, i.e. $i(x) \subseteq i(y)$, then in some sense many of the 'qualitative parts' of x are literally contained in y. Qualitative composition is in principle not objectual mereological composition. An object y may have all the properties that an object x has, without x being literally a part of y in the ordinary sense. For instance, take two duplicates x and y and add a property to y. This does not make x a part of y. But one may wonder also whether the way in which a point or an instant is contained in a region or an interval can be literally thought of as ordinary mereological composition. Since ordinary objects are concrete and standard mereology is modelled after them, we may find it obvious that points are parts of regions. But what about instants? In what sense is an instant literally a part of an interval? An instant is not spatially extended, it is not a concrete object. And to say that an instant is abstract is simply to miss the point (how could a temporal entity be atemporal?). But the relation between instants and intervals closely resembles that between points and regions or parts and wholes. The former resembles the latter one close enough for us to legitimately talk about compositional structure. And the same happens in our case, qualitative composition behaves sufficiently enough like mereological composition.

If qualitative atoms or paradigms are just tropers, then they are indeed parts of ordinary objects. In that case it is no wonder that the qualitative order of thinness closely follows the mereological one. Given that tropers are such that no other object is qualitatively thinner than they are, a plausible conjecture is that they are *atomic parts* of ordinary objects. I will show that this is indeed the case. Let us suppose that our objects are mereologically structured by some partial order. For the parthood and similarity relations to be compatible, they have to satisfy some constrain. I suggest the following:

Definition 56. *Let S be a set, $\leq$ a partial order and $\sim$ a similarity of order 1 over S. Then $(S, \leq, \sim)$ is a* partial similarity poset of order 1 *iff* $\forall x, y \in S\ (\exists z, w \in S\ z \leq x\ \&\ w \leq y\ \&\ z \sim w \Rightarrow x \sim y)$.

The condition requires that two objects are similar if they have similar parts. The converse is inmediate due to the reflexivity of the partial order. If two objects do not have similar parts, since each object is part of itself then they cannot be similar. I think it is a reasonable constraint. If two objects have similar parts, then since the similarity is approximate and not exact the objects will be similar to each other (the similarity is inherited 'upwards'). If we assume that there are paradigms and that the indiscernibility is satisfied, then we immediately have that (the converse of (1) does not hold):

Proposition 23. *Let $(S, \leq, \sim)$ a partial similarity poset of order 1, $x, y \in S$ and $p \in Gen(S)$. Then:*

1. $x \leq y \Rightarrow co(x) \subseteq co(y)$.

2. $x \leq p \Rightarrow co(x) = co(p)$.

3. *S is (SNI) or pure* $\Rightarrow$ *p is minimal.*

Proof. (1) Let $x \leq y$. If $z \sim x$, since $z \leq z$ we have $z \sim y$. Therefore $co(x) \subseteq co(y)$. (2) If $x \leq p$ then by (1) $co(x) \subseteq co(p) \in SC_1(S)$. Therefore $co(p) \subseteq co(x)$. (3) If S is (SNI) or pure, from (2) follows that $x \sim p$ and therefore $p = x$, p is minimal. □

(1) says that if object x is part of object y, then x is qualitatively thinner than y. What (3) shows is that under the assumption that objects are mereologically structured by a partial order which is compatible with the similarity, that the similarity satisfies the indiscernibility axiom and that there are paradigmatic objects, it follows that these paradigmatic objects must be mereologically minimal. In other words, if there are paradigms, they must be atomic parts (the converse is easily seen to be false though). If the indiscernibility is not satisfied, then not every paradigm will be mereologically atomic. But all of them will be indiscernible from their parts, which will also be paradigms. This means that if there are mereologically minimal objects, they will also be paradigms and if there are none then we will have infinite descending chains of paradigms. In both cases paradigms will turn out to be 'quite small' mereological parts. The argument does not show that paradigms exist, but it shows that if they exist then they will be mereologically atomic (or qualitatively indiscernible from mereologically smaller parts):

Mereological Argument for Paradigms

i. There are paradigmatic objects. [Similarity is of order 1]

ii. Objects are mereologically structured. [Poset]

iii. Two objects are similar iff they have similar parts. [Partial Similarity]

iv. Any two paradigms are indiscernible (identical in the model). [Pure]

v. Therefore (deductively), paradigms are atomic parts.

The position may look metaphysically extravagant. We may dislike positing points, instants and atomic parts. They seem to be suspicious limiting cases we get at just by a process of abstraction applied over bona fide entities. It may be the case that, at the end of the day, points, instants, atomic parts and paradigms are just the result of such limiting processes of abstraction. But if so one would have to show how this process works. The development of techniques that would allow us to dispense with them usually takes more time and knowledge[21], and for the time being they do a sucessful job as building blocks of the corresponding composed entities. The claim I want to make is that we should be as justified on positing the existence of qualitative atoms as we are on the existence of their spatial, temporal and mereological analogues[22].

[21] Witness the amount of non-trivial mathematical results that are needed to get rid of points in topology.

[22] Of course, one is not mathematically forced to accept the existence of paradigms, just as one is not mathematically forced to accept the existence of points or mereological atoms. Just as there are gunky spaces, there are also similarity structures that lack this sort of atoms. One example is a cyclic structure $x \sim y \sim z \sim w \sim x$. Another example is an infinite path $\cdots \sim x \sim y \sim z \sim w \sim \ldots$

4.6 Conclusion of Chapter IV

The purpose of this chapter was mainly defensive, namely to clear the ground for a formal model of natural kinds based on aristocratic resemblance nominalism.

First, I have given some explanations regarding the concept of categorical (all-or-nothing) similarity or resemblance, including some of its most basic formal properties. The properties shown should be enough to convince the reader that Quine's animadversions regarding the logical status of categorical similarity are ill-founded.

Second, I have discussed some objections against the fundamental properties of categorical similarity. In particular, the objections against symmetry stemming from the psychological literature and Tversky's attribute model have been discussed in detail. Tversky objects to the symmetry of similarity by providing evidence of non-symmetric similarity judgements. However, I have given three reasons to resist the argument. First, spatial models can give alternative explanations for this fact without dropping the symmetry axiom. Second, Tversky's own alternative model, the attribute model, makes several questionable assumptions regarding the nature of similarity. Third, the discussions seems to concern the notion of degrees of similarity, not categorical relations of similarity. This is important, for a comparative similarity relation can be introduced from a categorical one which allows explaining the violations of symmetry in terms of the paradigmatic entities involved.

Third, I briefly reviewed the three main forms of resemblance nominalism, namely egalitarian, aristocratic and collectivist nominalism. I have discussed Pereyra's egalitarian approach in detail and make some objections to it. Although Pereyra's egalitarianism provides a way to recover properties as certain 'stratified cliques', it has two problems. On the one hand, only half of the required result is provided. He does not give the axioms that a similarity between pairs of pairs of pairs of ... have to satisfy in order to reconstruct any realist context. Thus, the approach is incomplete. On the other hand, the approach is too close to collectivism, since the similarity relations between pairs of pairs of pairs of ... can be replaced by the similarity relations between the corresponding bases (which are sets). These similarities can be plausibly interpreted to be collective. Thus, the approach risks being inadequate. Moreover, this solution only works for the imperfect community problem. In order to deal with the companionship problem, more assumptions have to be made.

Fourth, I have considered three different models for aristocratic resemblance nominalism. One of them makes use of polar models introduced by Rumfitt and Mormann, which provide a formal framework for conceptual spaces approach. The second one is based on Mormann's suggestion that the specialization order of weakly-scattered spaces is a good model for the prototypicality order among objects. I extended this model by (putting it upside-down and) adding an axiom to it and I made a proposal regarding what the attributes would be (fixed ultrafilters). The disadvantage of these models is that they do not mention similarity at all. The third one is introduced for the first time in this thesis and is based on similarity structures of a special kind. According to this model, no two paradigms are sufficiently similar to each other and any two similar objects are similar to a common paradigm. This model will allow for reconstructing attributes as collections of objects similar to a given paradigm. This one preserves both the spirit and the letter of aristocratic resemblance nominalism.

Finally, I have considered two different conceptions of what paradigms are. According to the former conception, paradigms are some exemplars found (and stored in memory) by an epistemic subject during her concept learning process. In contrast, this solution puts resemblance nominalism closer to conceptualism. The notion of a paradigm appealed to by the similarity model was shown to satisfy the criteria that proponents of the conceptual spaces approach have given. According to the latter one, paradigms are tropers, qualitatively thin objects that are mereologically related to ordinary objects. I argued that this choice solves the coextensionality problem and makes plausible the assumptions made in the model. I also considered three more arguments for the existence of tropers. First, one can argue for the existence of tropers by showing how positing them solves Goodman's problems. Second, one can argue for the existence of tropers as qualitative atoms by analogy with space, time and composition. Third, if one accepts the basic assumptions of the model then if objects are mereologically structured it follows that tropers will be mereological atoms (or at least will form an infinite chain of smaller and smaller duplicate parts). This solution moves resemblance nominalism closer to trope theory.

The upshot is that the notion of categorical similarity stands the test and that a version of aristocratic resemblance nominalism can at least be presented in a systematic fashion. The purpose of Chapter V is more constructive, its aim is to show that the three models just introduced are equivalent and can give an answer to the infamous Goodman's objections to similarity and quasianalysis, towards which we now turn.

Chapter 5

Resemblance Structure of Natural Kinds

The supreme maxim in scientific philosophizing is this: wherever possible, logical constructions are to be substituted for inferred entities.
Our Knowledge of the External World
BERTRAND RUSSELL

In chapter IV I chose resemblance nominalism as an account of the internal structure of natural kinds. There I presumed that resemblance nominalism could in fact give an adequate answer to the problem of kinds. This chapter works in tandem with the previous one. First, I briefly discuss the methodological question concerning what a successful nominalist reduction should be like. I propose to reformulate the formal side of the problem of universals as the quest for certain structural representation between the corresponding models. Next, I deal with the second pack of objections to a resemblance nominalism of kinds, namely Goodman's companionship and imperfect community objections to similarity. I consider them alongside several replies by Leitgeb and Mormann. Third, I show that the polar and similarity models for aristocratic resemblance nominalism introduced in the previous chapter are equivalent and can indeed answer Goodman's objections, so long as we restrict our attention to a special class of realist structures. Finally, I show what the resulting nominalist lattice of natural kinds looks like, by proving several theorems about its structure. This will give us a picture of the external structure of kinds according to the nominalist. In particular, it is shown that kinds form a complete (co)atomistic lattice.

5.1 Adequacy Conditions for a Nominalist Reconstruction

This section is devoted to discussing methodological matters concerning adequate nominalist reconstructions. Many of these ideas can already be found in the discussions about the possibility of Constructive or Constitutional Systems of Concepts by Carnap [17] and Goodman [48], and others come from the discussions of philosophers of science concerning the nature of scientific representation and the notions of theoretical reduction and theoretical equivalence. The contemporary popular Quinean paraphrasing techniques seem to me to be a drawback from the approaches assumed in these areas.

Philosophers have many times tried to avoid commitment to entities that they considered to be ontologically suspicious, or have tried to show that some entities are more fundamental than others. To argue for their views, they have attempted to reconstruct the undesirable entities from the favoured ones. Consider for instance a philosopher who gave the following quick arguments for the existence of certain entities called 'absences'. First, absences seem to be causes and having causal power is a sufficient condition for existence. To take a standard example, the lack of water killed my plant. Second, we seem to quantify over absences in natural language. As an answer to the question "what killed my plant" I could say "the lack of water killed it", from which one may suggest that it informally follows that there exists something that killed my plant. Third, some specific domains, like that of classical propositions, Russellian facts, Platonic Forms or standard parts and wholes are closed under negation-like operations that produce 'negative' entities. Based on these considerations the (Heideggerian) philosopher may argue that absences are entities in their own right. A cautious philosopher will balk at this ontological inflationist attitude by suggesting that less suspicious entities could fulfill some of the tasks that absences are posited for. To explain away the apparent need to posit absences, the latter philosopher will suggest some sort of 'revision' or 'reconstruction' of what the former philosopher is saying. A more interesting (and non-exhaustive) list of well-known examples could be the following one, where many pairs can be read in both directions:

- Logical Atoms-Logical Complexes.
- Particulars-Bundles of Universals.
- Universals-Collections/Sums/Pluralities of Particulars.
- Instants-Intervals.
- Points-Regions—Points-Lines—Points-Solids.
- Propositions-Possible Worlds.
- Constituents-States of Affairs/Facts.
- Experiences-Qualitative atoms.

The aim of such a reduction is not always clear, but I would say that it purports to be *explanatory*. The nominalist or anti-realist about X-s wants to explain the same phenomenon that the realist about X-s wants without committing himself to the existence of X-s. As a first step, the nominalist will suggest that a different ontology will be sufficient to give an explanation of the aforementioned phenomenon. However, it will usually not be enough for the nominalist to give an account of the phenomenon in his terms by appealing to his basic ontology, for the realist will always be able to reply that there are some features of the phenomenon that he is capturing by appealing to his posited entities that are not accounted for by the nominalist. So long as the nominalist finds these features worth explaining, he will have to find a way to explain them by appealing just to the sort of entities he accepts. Since the ontology of the nominalist will be sparser than that of the realist, he may not be able to give such an explanation by appealing to the basic entities of his system. How is he then to perform his task? He may construct, constitute or derive some entities from the more basic ones, by performing certain nominalistically acceptable operations over the latter, and then show that these derived entities can indeed fulfill the same roles that the entities posited by the realist do. The morale of the story is that the nominalist can dispense with the entities posited by the realist, since he can find other constructed or derived entities to do the same job.

For instance, the nominalist about universals wants to give an explanation of the same phenomena that the realist wants, be it resemblance among objects, naturalness, having a common nature, the truthmaking relations between some entities and some specific statements of the form "a is F", the categorial structure of the world, or whatever. But the nominalist wants to give this explanation without committing himself to the existence of universals. So he proposes a sparser ontology, say objects or tropes, and suggests that it will be enough to perform the desired explanation. The realist will reply that the nominalist is still not able to explain in a systematic fashion some facts like those that make statements like "a and b are F" true. It is not clear how the nominalist can give a *systematic* answer that guarantees that the same 'recipe' will be followed to give a nominalistically acceptable explanation in all the cases that are similar to the one just mentioned. Here is where the reductive move is to play its role. The nominalist proposes some entities derived from the fundamental ones in his system, say mereological sums, pluralities, sets or collections, predicates, concepts, and so on, and shows that they can perform the same tasks that the universals do. The fundamental methodological question that should concern us is how to make such a nominalist reduction.

It is currently customary in metaphysics to assume that the task of such nominalist reconstructions is one of making *logical paraphrases* for any (true) sentence that implies the existence of the undesired entities. This method was popularized in its current form by Quine, particularly in [105], inspired by the success of the Russell's use of definite descriptions for similar purposes. It consists in making a translation of the sentence, formulated first in a natural language, to the language of some formal logic. This logic is usually classical first order logic suitably augmented (e.g. with identity). Once in this logic, it is shown that the sentence really does not commit oneself to the existence of the purported entities, since these entities are not the values of the existentially quantified variables that occur in the w.f. formula that translates the original sentence. So the basic procedure consists in mapping each sentence in natural language that contains the controversial term (e.g. 'shadow') to a w.f. formula in the logic chosen, so that all the truth values are preserved. Then one looks for the entities at the domain of values for the variables that are quantified in the formula. But there are some differences between current uses of neoquinean methodology and Quine's own approach. On the one hand, Quine would only consider as worthy of translation those sentences found in our most successful scientific theories. On the other hand, in most of the examples that one comes across in the literature, the paraphrases are given in some sort of hybrid language of natural and formal languages. A problem with this practice is that usually the authors give only a rough and sketchy presentation of the 'translation' by a couple of examples, instead of giving a proof that the paraphrase works for any sentence that implies the existence of the entities in question. It is not enough to say that one has a paraphrase for one or two kinds of sentences, as for instance for "there are opaque objects that project shadows" [62]. The reason is that we could always find a different sentence that may imply the existence of shadows and did not adjust to the paraphrase template given.

I think that the problems can be circumvented if one turns to a more interesting version of "paraphrasing" that comes from mathematics and that philosophers of science concerned with the problem of scientific representation know well. The paradigmatic example of this method is given by the Representational Theory of Measurement (from now on, 'RTM') developed by [128][1]. The

[1] My suggestion is inspired by the approach developed by [87] and [89], who reformulates

idea is that paraphrases should be understood as structural representations, in the sense explained by Swoyer [129]. According to RTM, measurement consists in the establishment of structural mappings from qualitative structures to numerical structures. This explains why real numbers are so useful for making measurements, since they can be used to make what Swoyer calls 'surrogative reasoning'. We go from the represented entities in the world to the numbers in our representation and we make calculations with these numbers. Then we infer something about the represented entities using the representing entities and we come back to the world to check that the conclusions hold. What guarantees that our 'prediction' of the behaviour of the entities in the world is more or less right is precisely the structural correspondence between both domains. This correspondence preserves the relations and operations found in the qualitative domain. If for instance the concatenation of bodies, when considering their lengths, was not associative then there could not be much that we could infer from the numbers we use to measure them. Or consider the possibility that our mapping did not respect the order between lengths. For instance, suppose that $z < x < y$, but the corresponding numerical representations are arbitrarily 'ordered' e.g. $f(z) < f(x) > f(y)$. Even if we ignore the psychological impulse to say that the numbers do tell us that x must be greater than y, we are simply not able to make any sense of this situation. What the theory establishes is something like the *conditions of possibility of measurement* itself. It also explains why the choice of a particular scale is irrelevant, showing that the choice of scale is unique up to an equivalence relation for all those representations that leave invariant the relevant structure. Moreover, it also explains the differences between those structures more often encountered in physics (the 'extensive-like' structures) and those found in psychology and other social sciences. It explains them as differences between ways in which entities in the world may be structured.

The classical simplest example is that of an extensive structure. Extensive structures are empirical structures of the form $(S, \leq, \star)$, where S is a domain of objects to be measured (e.g. a collection of rigid rods), $\leq$ is a linear order that holds between those objects (e.g. $x \leq y$ iff y is longer than x) and $\star$ is a concatenation operation under which the domain is closed (e.g. $x \star y$ is the rod obtained by juxtaposing the rod y with the rod x). A scale is a function $s\colon (S, \leq, \star) \to (\mathbf{R}^+, \leq, +)$ that maps each object to a real number representing the attribute that is being measured (e.g. length) in such a way that the empirical structure gets preserved. In other words, $x \leq y \Leftrightarrow s(x) \leq s(y)$ and $s(x \star y) = s(x) + s(y)$. Note that not all the structure of the real numbers is needed for these purposes, we only require the order and additive structure of the real line. Moreover, any other scale would do so long as it preserves the same structure. In that sense, the scale for an extensive structure is unique up to a given notion of equality between the scales (in this case, the scale has to be a ratio scale). A proof of the existence and uniqueness (up to equivalence) of a scale for a given structure (say, extensive structures) shows which conditions are necessary and sufficient for certain numerical statements about it to be meaningful.

The results of this theory can be interpreted as nominalist paraphrases that show that real numbers are dispensable for measurement. An ontology of particulars, sets and (extensional) relations and operations over them (which can

---

the quasianalytic approach as a case of structural representation (see section 3 below).

be interpreted as ideal actions) suffices. Of course, the results themselves are neutral regarding the realism-nominalism disputes. The theory does not say whether the entities in the qualitative domain are ordinary substantial particulars, tropes or even magnitudes taken as universal properties. But it does show that it is not necessary to assume that numbers themselves are in some way 'in' the entities to be measured. Entities which behave structurally in a sufficiently and relevantly similar way to numbers are enough to explain why measurement works. Moreover, this goes hand in hand with the modern development of mathematics. Mathematicians have developed many different ways to get structures that are isomorphic to numeric structures just by starting with 'qualitative' structures and adding more and more assumptions. From the point of view of these theories, number structures are just highly complex qualitative structures that happen to have interesting order-theoretic, algebraic or geometric properties which are related in very nice ways to each other.

Why should one consider these intermediate structures introduced by RTM? After all, if number structures are the ones doing all the work, why not simply assume that the world is just as highly structured as these are? Because we would be making unnecessarily stronger assumptions about how (some portion of) the world is structured. We may be even imposing structure where there isn't any. These features will be 'representational artefacts' [129]. But then we could fall into the trap and take too seriously some features of the representation which lack a worldly correlate. In contrast, if we assume just those features that are needed to do the job (those that are conditions of possibility of the corresponding phenomenon), then at least we know that such a structure will be embeddable in any other structure that should do the same job. This can be taken as an Ockhamist argument in favour of any RTM-like project. So we can think about the nominalist project roughly along the lines of the RTM approach. For a given structure posited by the realist, we will look for the minimal conditions that a structure formed from the entities acceptable to the nominalist needs to satisfy for the realist representation to be meaningful (in his terms).

What the existence of such a structural mapping shows is that we may be able to avoid commitment to some entities by considering them simply a metaphorical way of speaking about other entities. But this mapping by itself does not show that the disputed entities do not exist. The nominalist usually requires additional arguments. Traditionally, these have been arguments of ontological parsimony or epistemological (e.g. empiricist) arguments. Nevertheless, the existence of some mapping of this sort seems to be a necessary condition, for if such a 'translation' is impossible in principle, then the nominalist thesis is doom to failure. An additional reason for the structural mapping not being enough by itself is that philosophers usually look for identities and not only isomorphisms (that explains the heated discussions concerning the work by Kripke on identity statements). Whatever this really amounts to, if there is no such an isomorphism then there is no identity, and therefore it makes sense to ask first for the existence of such structural mappings. This does not mean that it is easy to find such correspondences. It may be a mathematical problem too difficult to solve for philosophers. Furthermore, this does not mean either that we should always seek for identities or for isomorphisms. The example of the Representational Theory of Measurement already shows that we may be content with an embedding. This is reasonable, for the nominalist will surely reject some features of the real numbers as mere representational artefacts that do not correspond to anything in the world. If this is so, there is no reason why we could not argue for weaker structural correspondences.

Nevertheless, I will impose stronger constraints than those required by the Theory of Measurement. The additional condition required is that the *surrogate or representing entities themselves* will have to be constructed from the primitive nominalist ontology, which is an idea that can already be found in the works of Carnap and Goodman[2]. So the methodology that I am proposing combines two steps. First, surrogates for the undesired entities are constructed by making use only of primitive entities in the system by applying logico-mathematical operations to them. Second, the existence of a (unique) structural mapping from the surrogate structure to the structure posited by the realist is proven. This shows that the surrogates do behave as desired and can actually fulfil the tasks they were introduced for. In other words, it is not enough to establish the existence of a certain structural mapping, the surrogates have to be constructed from the primitive entities. Therefore, the procedure is even more demanding than in the measurement case. An analogous move in RTM would be to construct *the real line itself* from the extensive structure. With all this in mind, we can try to give some requirements for an *adequate* nominalist reconstruction of universal entities along these lines:

i The ontology must be nominalist. In our case, it should include particulars, sets of particulars and one (or at least few) relation which can be reasonably interpreted as a resemblance relation. Moreover, this relation should have some expected formal properties, depending on whether it represents a categorical, comparative, degree-like or collective resemblance relation.

ii The structure of universal attributes should be reconstructed from the ontology in (i) as previously discussed. The correspondence between the structures should be an isomorphism (or some weaker structural relationship, if suitably argued).

iii Ideally the reconstruction should satisfy some additional constraints. It should be as simple and general as possible. It should have some unification power, so it should at least imply as special cases ones already known. It should have some explanatory power, so it should at least explain why other alternatives work or do not work, hint at why it seems pretheoretically plausible to posit the entities that are being reconstructed and explain in some way the features that realists usually attribute to universals.

Conditions (i)-(iii) are inspired by Carnap's notion of *explication* [18]. (ii) requires the explicatum to be formally adequate or *precise*, whereas (iii) requires it to be *simple* and *fruitful*. As we know, some of these conditions, such as adequacy and fruitfulness, often pull in opposite directions. Regarding (i) which concerns *material adequacy*, the following moves seem to be cheating or at least suspicious:

[2]This makes the proposal closer to the requirement that there is definitional equivalence between the two classes of models than just requiring the existence of an isomorphism from each of the models in one class to another model in the other class.

- Starting from one similarity relation for each of the universal properties. Many philosophers have claimed that one would be smuggling in universals as 'respects' of comparison.
- Introducing external entities which are not obtained as a result of applying logico-mathematical operations over entities taken from the basic ontology.
- Obtaining the reconstruction by 'trivial' moves.

Concerning the first case, one could appeal to degrees of similarity without necessarily committing himself to universals, this is for instance what [109] does. As long as the degrees of similarity are not interpreted as 'respects' of comparison, this is a legitimate strategy. Regarding the second case, I think it is a very important point. Suppose I claim that the basic ontology consists in particulars, sets and a degree-like similarity relation. Then I introduce the euclidean space $\mathbf{R}^3$, with its metric, vector-space structure and even cartesian coordinates. Finally I say that the surrogates for universal properties are to be found in some regions of this space (say, convex regions) that are to be described using the coordinates in some specific way. In this case, I would owe the nominalist an explanation of where does this euclidean space come from, why are all its properties needed and what does it have to do with the original ontology. So the first problem is that the entities introduced could be at odds with the nominalist ontology. Second, the more structure we put the greater the risk is of taking too seriously features that may simply be (at least for the nominalist) representational artefacts. We could then start wondering about the worldly correlate of the inner product, about the ontological status of vectors or about the number of dimensions that our space should have. Of course, one could argue that it is legitimate to appeal to certain entities in a fictionalist fashion, leaving the reconstruction of these entities as a task for further inquiriy. But in that case one should specify why. The third case is difficult to make more precise. The following would be a clear example of a violation of this condition. We start from a context (S, Q, I) and we define a collective resemblance relation over the power set of S as follows:

$$A \sim B := \exists P \in Q \ A = B = e(P)$$

The reflexive nodes in the graph $(\wp(S), \sim)$ are exactly the extensions of the original properties. This case does not violate the two previous conditions, since there is only one resemblance relation, it only appeals to particulars and sets of particulars and no extra-systematic entities are introduced. But it is clearly trivial. The similarity is "the A-s and the B-s are similar iff they are exactly the extension of a property P". This definition has some awkward consequences too: it makes the similarity necessarily transitive and just as fine as identity, proper subsets of extensions of properties are not similar to themselves (even if all the members clearly share some property and therefore should be similar to each other), etc. At most it could be a way of subsuming naturalness nominalism as a trivial case of collective resemblance nominalism.

5.2 Goodman's Problems for Resemblance Nominalism

5.2.1 Companionship and Imperfect Community

As it is well known, Goodman [48] highlighted some problems for Carnap's quasianalytical account, as developed in the Aufbau, that were later interpreted as objections against resemblance nominalism. These problems had already been identified by Carnap in [16] and [17], but Goodman did not find Carnap's answers convincing. The discussion of how Goodman's problems block the attempt at reconstructing a realist context is neatly reconstructed by Leitgeb [70]. Although I will make some changes in terminology to adjust it both to the terms that we have been using so far and to those of the realist-nominalist controversy[3], I follow his presentation closely.

Suppose that the realist asks the nominalist to give a model of the world. The nominalist chooses a similarity structure $(S, \sim)$. The realist then claims that he can choose a given context (S, Q, I) to show that a similarity $\sim^*$ can be defined in this context which is equal to the one chosen by the nominalist. The realist has several options. The simplest one is to choose a context (S, Q, I) which is such that $P \in Q$ iff $e(P) \in SC(S)$. To put it simply, he chooses attributes whose extensions coincide with the similarity circles. Then he defines the new similarity among the objects as $x \sim^* y := i(x) \cap i(y) \neq \varnothing$, which turns out to be equal to $\sim$. Following this recipe, the realist can reformulate whatever the nominalist says in his own terms. The realist gives a reduction of the nominalist model.

Let us go back, now from the point of view of the nominalist. The realist starts from a context (S, Q, I), which satisfies at least the following condition:

$$\forall x \in S \; \exists P \in Q \; xIP \text{ [Instantiation/Exhaustiveness]}$$

This is not a problem since the realist we are considering accepts the *Principle of No-Bare Particulars*, that is, that necessarily any particular object must instantiate some universal property (the principle will follow from the reflexivity of similarity). The nominalist is willing to accept a counterpart of this principle in his terms. Namely, the reflexivity of categorical similarity. Then the realist advances the following definition of similarity:

$$\forall x, y \in S \; x \sim y := \exists P \in Q \; (xIP \;\; \& \;\; yIP) \text{ [Similarity]}$$

This is a crucial step, since the resemblance nominalist could simply reject this definition and ask for a different one. For instance, he could argue that this is not a convincing definition of similarity, since properties themselves can be more or less similar to each other. He could ask for something like Carnap's part-similarity (see [17]) or for an overall similarity of a different sort. Nevertheless, the resemblance nominalist grants this definition and starts from the similarity structure. Notice that now the answers that the nominalist can give are constrained. Under the previous definition of similarity, and given that the surrogates for attributes must be sets, we have that any attribute P in the

[3] For example, what here is called a 'formal context' that satisfies some additional conditions, using the terminology of concept lattices, is called 'property structure' by Leitgeb, and so on.

context is a clique. That is to say, for any $x, y \in P$ it trivially holds that $x \sim y$. Therefore, the nominalist knows that he must choose cliques as surrogates for attributes. In fact, all the attributes the realist is asking for are among the cliques of the nominalist. Is the problem of the nominalist solved then? Not really. If we start from a context (S, Q, I) and define a similarity structure $(S, \sim)$ in terms of it, which cliques should the nominalist choose to surrogate the attributes in Q? It will not do to simply say that they must be the extensions of the attributes in Q. That would be tantamount to cheating. Such a move would be like starting from the beginning with a chosen set of cliques $Q \subseteq \wp(S)$. A naturalness class nominalist may be allowed to make that move, our resemblance nominalist is not.

So as a second crucial step, the nominalist proposes to follow Carnap's suggestion. He will take as surrogates for the universal attributes the similarity circles $SC(S)$ of the structure $(S, \sim)$. That is to say, universals are maximal collections of pairwise similar particulars. He finally offers the context $(S, SC(S), \in)$ as a surrogate of the structure (S, Q, I) given by the realist, arguing as we previously explained that they are isomorphic (not *identical* of course, the nominalist rejects the existence of universals!). To summarize:

Naive Egalitarian Resemblance Nominalism $(L, \leq)$ is the concept lattice of the exhaustive set-theoretic context $(S, SC(S), \in)$ induced by the similarity structure $(S, \sim)$.

Unfortunately, this will not do. Not every context can be uniquely reconstructed from a similarity structure following the recipe given by resemblance nominalism. The classical objections of coextensionality, companionship and imperfect community can be seen as different ways in which the required isomorphism fails. The first objection is the *coextensionality problem.* There could be two different but coextensional properties. Therefore, there are more properties than sets and no possible bijection. The classical example is that of the different but coextensional predicates 'creature with a kidney' and 'creature with a heart'. Pereyra [109] argues that this particular counterexample is wrong but in any case, he acknowledges the possibility of there being coextensional predicates that correspond to different sparse properties. In other words, the nominalist can reconstruct at most those contexts that satisfy the *condition of extensionality*:

$$\forall P, R \in Q \; e(P) = e(R) \Rightarrow P = R \text{ [Extensionality]}$$

Which is *dual* to the (unqualified) *Principle of Identity of Indiscernibles*:

$$\forall x, y \in S \; i(x) = i(y) \Rightarrow x = y \text{ [Identity of Indiscernibles]}$$

By 'dual' I mean that we can obtain one of the principles from the other one by substituting the variables for a certain kind of entity (e.g. particulars) for the variables of another kind of entity (e.g. universals) and the corresponding operator (e.g. intension) for its dual (e.g. extension). This gives us some clues, since the problems that they give rise to are structurally the same. In other words, they impose the same representational limits to the reconstructions given by those who only accept one kind of entities, be they particulars or universals. For instance, the nominalist must accept [Extensionality], and therefore cannot distinguish between coextensional universals. Analogously, the bundle theorist must accept [Identity of Indiscernibles], and therefore cannot distinguish between indiscernible particulars, i.e. particulars that instantiate exactly the same universals[4]. If the bundle theorist uses for the reduction a similarity relation of coinstantiation, i.e. being instantiated by the same particulars, then he may face analogous problems to the companionship and imperfect community. Moreover, the nominalist may construct new surrogates for universals choosing appropriate extensions. The bundle theorist may construct new particulars choosing appropriate intensions.

As we know, the nominalist can appeal to other strategies to guarantee that the context to be reconstruct must satisfy Extensionality. For instance, he can accept some version of modal realism (say [73]), or he can appeal to possibilia and then show how to reconstruct possibilia themselves from other entities. I will not discuss these strategies and for the time being I will simply assume that both parties agree on restricting themselves to extensional contexts. These contexts are in unique correspondence with those contexts whose set of properties Q is a subset of the power-set of S, i.e. whose properties are just sets. So the nominalist can simply restrict his attention to the set-theoretical contexts of the form (S, Q, I) where $Q \subseteq \wp(S)$ and $I = \in$. So instead of making a distinction between a property P and its extension $e(P)$, both parties agree now to talk about the property P, taken as an extensional entity. We simply have to remember that, strictly speaking, the nominalist is giving as a surrogate a structure (S, Q^*, I^*) which is isomorphic to the one proposed by the realist, who does not accept that properties are sets.

The second objection is the *companionship problem*. Given (S, Q, I), it can happen that for two different properties one is properly included into the other one. In other words, a pair $A, B \subseteq S$ forms a companionship iff

$$A \subset B \text{ [Companionship]}$$

For instance, take $S = \{x, z\}$ and $Q = \{\{x\}, \{x, z\}\}$, then the pair $\{x\}, \{x, z\}$ is a companionship. $\{x\}$ is not a similarity circle because it is not maximal. Since similarity circles are maximal, neither of them can be properly included into the other one, and so the included property cannot be recovered. The method fails because one of the properties in Q is not constructed. In other words, the similarity structure is not 'full' [70].

One of the classical examples of companionship is as follows:

- Let x be a red scarlet object and y a red crimson object. Suppose that $x \sim y$ iff x and y are of the same colour. Then if one intends to recover all the colours from a sufficiently general degree of similarity, i.e. such that $x \sim y$, then one will only recover the colour red. If one restricts the degree of similarity such that x is not similar to y, then one can only recover the most specific determinates red scarlet and red crimson, but not the colour red.

The nominalist realizes that the properties to be recovered can be properly included into one another. So instead of the set of similarity circles, he can suggest the set $SC^*(S)$ of all the cliques of S. But then the realist points

[4]However, Pereyra has recently argued that this is not indeed the case.

out that now there are too many properties, and that some of them (e.g. the singletons, the empty set, some edges, ...) do not correspond to any of the properties in the original context. He could take some of these cliques, but without any clue that helped to distinguish those cliques that are properties from those that are not, he could not do much more.

The nominalist can complain that the realist is asking for too much. He obviously cannot recover all the properties from *a single degree of similarity*, since some properties are more specific than others. For that he would require a distinction between degrees of the same relation of similarity (e.g. [109] uses this strategy). But the problem now lies in giving a plausible account of such degrees of similarity and a general recipe that shows how to choose an adequate family of them for each context. For instance, choosing a degree of similarity for each attribute would be cheating. This is not an easy thing to do. The nominalist can also insist on the fact that the properties to be reconstructed are of the lowest level of specificity, and that the rest simply supervene on them (once more [109] seems to suggest this solution). The properties to be reconstructed are the lowest determinate natural properties. Therefore, some of the counterexamples given by the realist are not counterexamples at all.

The third objection is the infamous *imperfect community problem*. Given (S, Q, I), it may happen that for at least three particulars $x, y, z \in S$, each pair of them instantiate a common property whereas there is no property which is instantiated by all of them. In other words, a set of objects $A \subseteq S$ forms an *imperfect community* iff

$$\forall x, y \in A \ i(x) \cap i(y) \neq \varnothing \ \ \& \ \ i(A) = \varnothing \text{ [Imperfect Community]}$$

Imperfect communities are cliques and some of them can be similarity circles. For instance, suppose the context has the form of a Goodman triangle $S = \{x, y, z\}$ and $Q = \{\{x, y\}, \{y, z\}, \{x, z\}\}$. Then $\{x, y, z\}$ is an imperfect community, more specifically a similarity circle, which does not correspond to a property in Q. Now the method fails because the property constructed was not there. In other words, the similarity structure is not 'faithful' [70].

Let us put some examples of imperfect communities for different domains of entities:

- Let x, y be organisms and $x \sim y$ iff x and y can interbreed. As it is known to the philosophers of biology, this relation is indeed non-transitive (this phenomenon is called 'ring species'). If one intends to reconstruct species as similarity circles, the method may fail since we could have a Goodman triangle.

- Let x, y be (social) agents and $x \sim y$ iff x and y act jointly. If one intends to reconstruct collective actions as similarity circles, the method may fail since we could have a Goodman triangle.

- Let x, y be epistemic agents and $x \sim y$ iff x and y communicate with each other. If one intends to reconstruct shared or common knowledge (i.e. shared propositions) as similarity circles, the method may fail since we could have a Goodman triangle.

- Let f, g be functions with the same domain and range, and $f \sim g$ iff there is an x such that $f(x) = g(x)$. If one intends to reconstruct points as similarity circles from the overlappings of functions, the method may fail since we could have a Goodman triangle.

It is important to realize that not all similarity relations, taken in the purely formal sense of reflexive and symmetric relations, produce imperfect communities. It depends on the interpretation of the relation and of the entities to be reconstructed. For instance consider:

- Take entities which have continuous-like structure, like spatial regions, temporal intervals, processes, colors in the wheel, and so on. For a concrete example, let x, y be spatial regions, and $x \sim y$ iff x and y overlap. Let us say that we want to reconstruct the smallest region that contains the overlapping of x, y and z, i.e. the region common to x, y, z. It is plausible to say that if the space was modelled as a similarity structure, the corresponding region to be reconstructed must exist. The similarity circle $x \sim y \sim z \sim x$ will correspond to that region.

- Let x, y, z be people, and $x \sim y$ iff x and y are friends (or acquaintances). Let us say that we want to reconstruct (maximal) groups of friends. Then we cannot have imperfect communities. It is nonsense to say that x and y are friends, y and z are friends, x and z are friends, but there is no group of friends to which x, y and z belong. The similarity circle is this group of friends, by definition[5].

The point of these concrete examples is to hint at the fact that there is nothing defective in the formal definition of the relation of similarity. We get into trouble as soon as similarity is given certain interpretations, for instance when it is defined as "x and y have a property in common", and when certain entities are to be recovered as abstracted from similarity circles. The reason for this is to be found in the additional structure that the entities to be recovered have, as will be seen in the following sections.

It would be a mistake to think that the formal problems related to categorical similarity are in some sense rooted in its being qualitative. For instance, quantitative similarities are also subject to imperfect communities. Consider the similarity $x \sim y := |i(x) \cap i(y)| \geq n$, for some $n \in \mathbf{N}$. For instance, we may require that $|i(x) \cap i(y)| \geq \frac{|i(x) \cup i(y)|}{2}$. Then we construct an imperfect community as follows. Let x, y, z be such that x and y share n properties, y and z share n properties and x and z share n properties, but x, y, z share less than n properties. For example, for $n = 1$, a Goodman triangle suffices. Thus, although x, y, z are pairwise similar to each other, they are not collectively similar. Moreover, this is not so surprising considering that since $i(x) \cap i(y) \cap i(z) \subseteq i(x) \cap i(y), i(y) \cap i(z), i(x) \cap i(z)$, we will have that $|i(x) \cap i(y) \cap i(z)| \leq |i(x) \cap i(y)|, |i(y) \cap i(z)|, |i(x) \cap i(z)|$. To sum up, making use of a similarity defined in terms of the number of shared properties does not solve the problem.

There are several answers one can give to these problems. But before seeing the details, we must distinguish the different theoretical proposals that will be discussed. On the one hand, we have Carnap's original Aufbau approach [17]. This is the development of a very general formal procedure or principle

[5]I think these are called 'cliques' in sociology too.

of abstraction, the quasianalysis, for constituting some structures of entities from other more basic ones. This formal device was introduced for the task of building what he called "constitutional systems of concepts". This is in fact the research line that Mormann [89] and Leitgeb [70] continue in different ways. This approach is much more general than the problem we are going to discuss here, and encompasses the constitution of various formal structures[6]. I will delve into this project later on, since the the nominalist reduction can be seen as a specific application of those tools.

On the other hand, we have the proposal of the nominalist, say Pereyra [109]. This consists in assuming as primitives some particular entities in some resemblance relation (binary, comparative, degree-like, collective or whatever). From this, some suitable collections of entities are chosen, in such a way that they uniquely correspond to the extensions of any structure of universal properties, i.e. any context (¡!), that the realist may have chosen. This construction is similar both for the resemblance nominalist that starts from substantial particulars and for the trope nominalist[7]. Of course these two projects overlap, but their aims are different and we should try to keep this in mind.

5.2.2 A Summary of the Puzzle

Let us sum up the puzzle:

1. We start from a given arbitrary exhaustive set-theoretical context (S, Q, I) which acts as a realist model.
2. We define a binary categorical similarity relation in the domain of objects S as sharing a common property.
3. We select certain class of cliques in the similarity structure to act as surrogates for the original properties.
4. The class of cliques selected must be identical to the original set Q of properties.
5. We select as the class of cliques the class of all similarity circles.

We know that if we follow these steps the strategy fails. Thus, in order to answer to Goodman's problems, at least one of these steps has to be rejected. Since we want to preserve the correspondence between being similar and having a common property, the second step will not be challenged. Since all the properties are cliques in the induced similarity and the purpose is to recover some of the original properties, the third step will not be rejected. This leaves the first, fourth and fifth steps. I will consider two approaches found in the literature. Leitgeb [70] preserves the fourth and fifth steps but rejects the first one. He proposes to reconstruct just a restricted class of contexts, not the class

[6]These are structures that are specially useful for doing formal ontology, such as mereological and topological structures

[7]Trope theorists usually appeal to the transitive relation of exact resemblance, so that the maximal classes of resembling tropes (which are the surrogates for universal properties) end up being equivalence classes. This avoids imperfect communities. I am not sure that this strategy works. After all, it is plausible to argue that tropes are also more or less approximately similar to each other (consider colour or magnitude tropes). If this is the case, then the threat posed by the imperfect community will still be there.

of all such contexts. Mormann [89] suggests several solutions. According to his main solution, steps one, four and five are rejected. Not every context should be reconstructed, because we need not have access to a given such context (since all we know are the similarity relations). Since we do not have access to such a context, the cliques chosen need not be identical to some given properties. The class of cliques chosen is in fact a restricted class of similarity circles (namely those of order $n \leq 2$). Since we are not to reconstruct an arbitrary context, we are free to choose the similarity structure from which to start. This allows us to choose a certain class of similarities (similarities of order $n \leq 2$) for which there is indeed a unique corresponding context such that the properties of the former are the similarity circles just mentioned[8].

My proposal contains a bit of Leitgeb's and a bit of Mormann's. I will reject steps one and five. I will preserve Leitgeb's requirement that a previously given realist context should be reconstructed, in such a way that the class of cliques selected is identical to the class of properties to be reconstructed. But I will also follow Mormann in selecting only a certain class of similarity structures (namely the pure similarities of order 1) and of similarity circles (namely the circles of order 1, and more generally the simple circles), so that a plausible model for nominalism results. The reason for this move is to be found in the philosophical purposes of the model in this thesis: in order to fulfil the nominalist reduction requirement and obtain a unique correspondence between some given realist models and the nominalist ones, I need to follow Leitgeb's path and restrict my attention to a subclass of such contexts. But in order to give a model for aristocratic nominalism that involves paradigms I need to follow Mormann's path and restrict the attention to a certain class of similarities and of similarity circles.

5.2.3 Property-First Approaches

One answer is to accept Goodman's objections, but to reply that the task of reconstructing *any* possible (exhaustive and set-theoretical) formal context was too much to ask for. This strategy rejects the first step of the puzzle. A reason could be that, after all, our task is to reconstruct structures of *natural properties*, not of abundant ones. Not every possible set-theoretical context is guaranteed to be a suitable context of sparse properties. For instance, as we have discussed, natural properties are not necessarily closed under the usual boolean operations. One could argue that the structure of properties should satisfy some additional structural constraints for it to be a structure of natural properties. However, neither Goodman nor the realists have given us any clues regarding what these would be. In the absence of such conditions, the nominalist is free to propose them. Let us call this the *property-first approach*[9].

Suppose that one thinks that, for a given domain of entities, the natural properties are fundamental properties. In such a domain no property implies the others, all are 'implicationally independent' from each other, so to speak. Then one can suggest the conditions given in Leitgeb's [70]:

[8] According to another solution by Mormann, step four is rejected. The class of cliques need not be identical to the one given, it can be equivalent in some specified weaker sense.

[9] Leitgeb himself proposes to consider the following contexts as contexts of natural properties. He then rejects this option as implausible.

Definition 57. *Let (S, Q, I) be a finite exhaustive set-theoretical context. Then it is a* Gilmore context *iff:*

i $\forall P, R \in Q \; (P \subseteq R \Rightarrow P = R)$. *[Maximality]*

ii $\forall P, R, T \in Q \; \exists N \in Q \; (P \cap R) \cup (P \cap T) \cup (R \cap T) \subseteq N$. *[Gilmore]*

Another way to understand what (i) says is to define as usual the boolean implication in the power set of S (in Q is undefined!) as $P \to_B R := P^c \cup R$, so that $P \subseteq R$ iff $P \to_B R = S$. Therefore, what (i) says is that no two distinct properties imply each other, they are independent. There is another condition equivalent to (Gilmore) which was proposed by Hazen and Humberstone in [58]:

$$\forall A \subseteq S \; (\forall x, y \in A \; \exists K \in Q \; x, y \in K \Rightarrow \exists K' \in Q \; A \subseteq K') \text{ [H-H]}$$

If any two objects in A share a property, then all of them share a property. In other words: there are no imperfect communities. Any Gilmore context defines a similarity structure over S as $x \sim' y := i(x) \cap i(y) \neq \varnothing$. What Leitgeb [70] shows is that this correspondence is unique[10]:

Theorem 6. *Let $(S, \sim)$ be a finite similarity structure. Then $(S, SC(S), \in)$, where $SC(S)$ is the set of all similarity circles of S, is a Gilmore context whose similarity structure $(S, \sim')$ is such that $\sim = \sim'$. Conversely, let (S, Q, I) be a Gilmore context. Then $(S, \sim')$ is a similarity structure which is such that $(S, SC'(S), \in)$ where $SC'(S) = Q$ and $\in = I$.*

To be clear, the aim of Leitgeb [70] is to assess the adequacy of Carnap's method of quasianalysis, that will be introduced soon. His conclusion is precisely that, due to Goodman's objections, the quasianalysis only works adequately under the conditions stated by the previous theorem. The version of quasianalysis that Leitgeb discusses is the one from Carnap's *Aufbau*, we call it here *the standard weak quasianalysis*[11]. In this thesis (following Mormann) we identify the method of quasianalysis itself with the function which maps each object to the set of *all* the similarity circles to which it belongs (this is important, for we will see now several variations of this notion):

$$q \colon S \to \wp(SC(S)) \; q(x) := \{T \in SC(S) \mid x \in T\}$$

So q is the intension function of the context induced by the similarity. For Leitgeb it seems that the method of quasianalysis is the correspondence (described in the previous theorem) between the similarity structures and the contexts of properties. In any case, this is a minor difference. According to Leitgeb, under the assumptions made so far, Carnap's method of quasianalysis of reconstructing a given structure of properties from a similarity relation fails. In other words, not every context of properties can be reconstructed from a similarity

[10]Leitgeb gives still another equivalent formulation using notions from hypergraph theory. In the case of finite structures, he also gives a combinatorial result that shows that the correspondence cannot obtain because there many more property structures than similarities. Thus, in the finite case the correspondence fails *for cardinality reasons*.

[11]Later on we will mean by 'standard' any quasianalysis that assigns similarity circles, independently of whether these are *all* of them or just some. The reason is that other entities apart from cliques can be assigned to objects too, e.g. paradigms.

following the previously listed steps. This is indeed the case, as the previous theorem shows, *so long as we do not drop other assumptions of the puzzle.*

Realists will agree with this conclusion, and they will answer in the following way. There are formal contexts which are plausibly structures of natural properties, but whose similarity faces the companionship problem, e.g. these context do not satisfy (Maximality). Leitgeb himself argues that it would be implausible to consider the previously mentioned contexts as the contexts of natural properties. For instance, take these examples of contexts from the standard book on concept lattices [40]. These are used typically as nominal and ordinal scales in social sciences. The corresponding properties are determinates of a common determinable:

Ordinal Context a context where the extensions of properties in Q form a chain. E.g. *extremely loud* $\leq$ *very loud* $\leq$ *loud.*

Bi-ordinal Context a context where the extensions of properties in Q satisfy the hierarchy condition (it may not have a top), i.e. they form a disjoint union of chains. E.g. *very low* $\leq$ *low* & *very loud* $\leq$ *loud.*

Interordinal Context a context where some extensions of properties in Q form a chain whose complements are also extensions of properties in Q. E.g. *very hot* $\leq$ *hot* $\leq$ *warm* $\geq$ *cold* $\geq$ *very cold.*

For example, consider the interordinal context. If x is very hot, then is hot and therefore (plausibly) warm, but it is not cold. If it is very cold then it is cold and therefore (plausibly) warm, but not hot. The realist will point out that it is embarrassing for the nominalist that he cannot reconstruct such common examples of ordered properties. Thus, the discussion will again revolve around the plausibility of the thesis that natural properties are implicationally independent from each other. Plausible cases of imperfect communities are even easier to come up with, as Goodman's classic example shows. Even assuming a distinction between natural properties and natural kinds, imperfect communities can also happen for the latter ones, as can be seen from the example of ring species. So the point is that the realist will insist on the fact that there are legitimate contexts that should be reconstructed by the nominalist. Nevertheless, the property-first approach does offer a partial solution to the problem. At least for Gilmore contexts of universals, the nominalist strategy succeeds.

My own assessment of Goodman's problems is similar to that made by Leitgeb. Since I take the nominalist to be concerned with a given realist structure of properties to be reconstructed, I agree with Leitgeb that not every such structure can be reconstructed. The main difference with the approach taken in this thesis is that I will make use of some of the results by Mormann (to be described in the next section) in order to select a specific subclass of realist structures to be reconstructed. This class overlaps the one described by Leitgeb, but neither contains the other. This choice will also force me to select specific subclasses of similarity structures and of similarity circles, namely the pure similarities of order 1 from the similarity model in the previous chapter. The reason for this is to be found in some results by Brockhaus, that will be considered now.

5.2.4 Similarity-First Approaches

Another answer is to reject that the nominalist has any access to a previously given structure of properties that has to be reconstructed. If so, the nominalist has to develop some sort of theory or hypothesis that explains the similarities between the objects by positing *quasi*properties as theoretical representational constructs. This interpretation is, roughly speaking, what one gets by applying the answer that Thomas Mormann proposed in [87] and [89] to the point of view of the nominalist. In other words, this strategy rejects steps one and four (and possibly five too). The nominalist can start from the formal properties of the similarity and try to argue that several other axioms should be accepted. From this point of view, some of the axioms that [89] proposes, concerning the indiscernibility of the similarity neighbourhood (C3) or the principle or parsimony (C4), are plausible. Let us call this the *similarity-first approach*. This approach makes sense in those cases where we start from a similarity relation and we want to discover interesting kinds of objects[12]. But before considering this approach in more detail, we have to understand what Carnap's main formal innovation in his *Aufbau* project was. This is a formal tool called 'quasianalysis'.

What is a quasianalysis? From a formal point of view, it is a function that represents a similarity structure by a set-theoretical structure, preserving the structure of the former into the latter. Thus, a quasianalysis is a representation function just as those found in the set-theoretic representations of other structures, such as lattices. From an ontological point of view, things are a bit more complicated. The quasianalysis can be seen as a generalization of the *method of abstraction* of Russell and Whitehead, by which a set of objects in equivalence relations are transformed into a set of equivalence classes, thus collapsing the identity relation into the coarser equivalence relation and moving one level up to a layer of higher abstraction. Quine was well aware that this method could be useful for making ontological reductions. So a quasianalysis is a principle of abstraction. Put more interestingly, a quasianalysis is a general method by which one can constitute or synthesize new entities from more basic or fundamental ones in such a way that the original structure is preserved. The method is a way of enriching the content of the original structure. If the elements of the structure are particulars, the quasianalysis gives us a *bundle representation* of these particulars. In other words, it represents them as bundles of some previously chosen properties that Carnap called 'quasiproperties'. In the most interesting case, these quasiproperties themselves are constructed from the particulars too, say as similarity circles.

Carnap's proposal was to use these principles of abstraction as *synthetic* methods of constitution of new entities from some previously given ones that were to be taken as more basic or fundamental[13]. These principles of abstraction could allegedly be applied again and again to jump from a structure of basic entities to even more complex structures, constructing what pre-theoretically

[12] As far as I understand, this is roughly speaking what is done with the spatial models of categorisation used in cognitive psychology that were briefly mentioned in Chapter IV.

[13] The choice of a *base* of entities was largely a pragmatic issue for Carnap, depending on what piece of science one wanted to reconstruct and why. So the fact that Carnap chose a phenomenalistic base to give an example of constitution of entities is irrelevant. This feature has been repeatedly pointed out by Carnap scholars. There are several proposals of bases in the Aufbau, some of them are physicalistic, some are phenomenalistic and Carnap even plays with the idea of starting from a cultural basis.

could be considered to be the parts, properties or kinds to which these basic entities belonged. This partially explains what Carnap meant by 'synthetic'. Some of the entities that one could construct could, pre-theoretically speaking, be interpreted as *constituents* of the objects. An analytic procedure would consist in dividing an entity into its logical constituents, be they parts or properties. In contrast, a synthetic procedure would consist in starting from some pre-theoretically (logically) complex or extended entities *as if they were atomic* and constructing their logical constituents from them, in such a way that one could later on give an enriched structural description of the basic entities as complex. In this way one does not need to assume the constituents as if they were basic entities in the system. So, taking the previous example by Goodman, imagine that one starts with lines instead of points as basic entities. Although pre-theoretically we consider the points to be constituents of the lines, we can construct them as overlappings of lines, i.e. as sets of lines. Then we can re-describe the lines as bundles of points, that is to say, as sets of sets of overlapping lines. In the simplest case, one represents an object by an equivalence class. This is not unrelated to the methods of reduction we are used to. For instance, we are familiar with the representation (or reduction) of possible worlds by (to) equivalence classes of propositions. But we know that, as Lewis suggested, we could start from the possible worlds and represent (reduce) propositions as (to) sets of possible worlds. Carnap's method does the same job for any non-transitive similarity.

It is time to take a closer look to the formal features of quasianalysis. We start from [89]:

Definition 58. *Let $(S, \sim)$ be a similarity structure, Q a non-empty set, $\sim^*$ the similarity relation $A \sim^* B := A \cap B \neq \varnothing$ on $\wp(Q)$, and $q\colon S \to \wp(Q)$ a function. Then q is a* standard quasianalysis *iff $Q \subseteq SC(S)$ and $q(x) := \{T \in Q \mid x \in T\}$. Moreover, q is a* weak quasianalysis *iff satisfies (C1-C2),* strong *iff satisfies (C1,C2,C3,C4) for any $q'\colon S \to \wp(Q')$ defined as follows:*

i $x \sim y \Leftrightarrow q(x) \sim^* q(y)$. *[C1-C2]*

ii $co(x) = co(y) \Leftrightarrow q(x) = q(y)$. *[C3]*

iii $co(x) \subseteq co(y) \Leftrightarrow q(x) \subseteq q(y)$. *[C3*]*

iv If $Q' \subseteq Q$ is such that $q'\colon S \to \wp(Q')$, defined as $q'(x) := \{T \in Q' \mid x \in T\}$ satisfies (i)-(ii), then $Q' = Q$. [C4]

Here Q is a set of *quasiproperties*, surrogates for attributes. Recall that $co(A) := \{z \in S \mid z \sim x\}$ is the similarity neighbourhood of A and the similarity interior is $int(A) := co(A^c)^c$. By (i) a quasianalysis is a similarity homomorphism that by (ii) preserves the equivalence relation (and even the order (iii)) among neighbourhoods. (iv) is a principle of parsimony[14]. It says that we cannot replace the quasianalysis by one using fewer quasiproperties without violating one of the previous conditions. (C3*) is stronger than (C3). A quasianalysis is injective iff $\sim$ satisfies the indiscernibility axiom (SNI). What

[14]We can restrict the range of q by introducing the overlapping similarity relation $\sim'$ in Q taking as range of q the set of cliques of $(Q, \sim')$. So, $(SC(S), \sim^*)$ is the *intersection graph* of $\sim$ and the elements of S are represented by cliques in this graph. But we will not make use of this. For more on these intersections graphs, see [70].

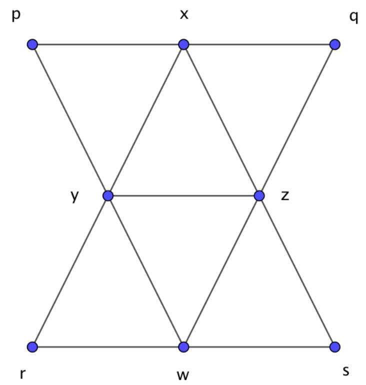

Figure 5.1: Counterexample to Uniqueness

Mormann does here is to subsume the quasianalytic account as a special case of the RTM approach. What the function q does is precisely to represent each object as a bundle of its quasiproperties.

In this section we will focus on the strong standard quasianalysis. Two fundamental questions concerning any representation are (1) whether there are any of them and (2) if so whether it is unique. The first question has a positive answer [89]:

Theorem 7 (Existence)**.** *Let $(S, \sim)$ be a similarity structure. Then it has a strong standard quasianalysis.*

The second question, however, has a negative answer. There are similarity structures which have several strong standard quasianalysis, the simplest example is given by [89] and is depicted in Figure 5.1. It has two different strong standard quasianalysis, one of them includes the similarity circle $\{x, y, z\}$ and the other one the circle $\{y, z, w\}$.

But there is a family of similarity structures which do have a unique strong standard quasianalysis. There are several ways to describe it. One of them goes back to a theorem by Brockhaus [14]. To explain this, we have to introduce several concepts. The first one is [87], [89]:

Definition 59. *Let $(S, \sim)$ be a similarity structure. Then $T \in SC(S)$ is a* similarity circle of order n *iff $\exists x_1, \ldots, x_n \in T$ such that $co(x_1) \cap \cdots \cap co(x_n) = T$.*

The elements $x_1, \ldots, x_n$ are called the *generators* of the similarity circle. The result by Brockhaus says is that if we restrict to similarity circles of order $n \leq 2$ then there is a unique quasianalysis. That is to say, the relevant class of structures is:

Definition 60. *Let $(S, \sim)$ be a similarity structure. Then S is a* similarity of order 2 *iff there is a family $F \subseteq SC(S)$ of similarity circles of order 2 which is such that $\forall x, y \in S\ x \sim y \Rightarrow \exists T \in F\ x, y \in T$.*

To prove the theorem we need to introduce some additional notions discussed by Mormann [87] and [121].

Definition 61. *Let $(S, \sim)$ be a similarity structure and $F \subseteq SC(S)$. Then F is a* similarity covering *iff* $\forall x, y \in S \; x \sim y \Rightarrow \exists T \in F \; x, y \in T$.

As usual, if F is a similarity covering and $F' \subseteq F$, then F' is a *similarity subcovering* iff F' is still a similarity covering. The family of similarity coverings of $\sim$ ordered under inclusions is upwards closed and therefore closed under arbitrary non-empty unions, the top being $SC(S)$. By reflexivity, F is also a covering in the usual sense. Observe that it is not enough for F to be a covering in the usual sense. What we must cover is not the domain but the similarity, because pairwise similarities have to be explained as sharing a common property. For example:

> Let $SC(S) = \{\{p, x\}, \{q, y\}, \{r, z\}, \{x, y, z\}\}$. Then $Q := SC(S) - \{x, y, z\}$ is a covering of S, but it is not a similarity covering. The similarities $x \sim y \sim z \sim x$ are unexplained by Q.

This is important for the following reason. Take Mormann's [89] counterexample of a similarity of order 3:

> Let $SC(S) = \{\{p, x, y\}, \{q, x, z\}, \{x, y, z\}, \{y, z, w\}, \{r, y, w\}, \{s, w, z\}\}$ and define the families $Q := SC_1(S) = SC(S) - \{\{x, y, z\}, \{y, z, w\}\}$, $Q' := SC(S)–\{\{x, y, z\}\}$ and $Q'' := SC(S) - \{\{y, z, w\}\}$.

This counterexample is crucial. As we saw, the similarity has two strong standard quasianalysis that can be constructed by using the minimal similarity coverings Q' and Q''. However, Q is a family of similarity circles of order 1 that covers the domain. Nevertheless, Q is not a similarity covering for it cannot account for the similarity $y \sim z$. That the condition must be so strengthened should be expected, after all what we need is that each pairwise similarity corresponds to having a common property. To prove the theorem we just need to introduce the relevant notion, that of a similarity covering which has no irredundant circles. This condition is the one corresponding to axiom (C4), just as the notion of similarity covering corresponds to (C1)-(C2):

Definition 62. *Let $(S, \sim)$ a similarity structure and $F \subseteq SC(S)$ a similarity covering. Then F is* minimal *iff* $\forall A \in F \; \exists x, y \in A \; (x \sim y \;\; \& \;\; \forall B \in F \; (x, y \in B \;\Rightarrow B = A))$.

This notion was introduced by Schreider under the name 'base'. If $x, y \in A$ are such that $\forall B \in F \; x, y \in B \Rightarrow B = A$ then we say that x, y are *F-indispensable* for A. A similarity covering F is minimal iff each of its circles has a pair of F-indispensables. In other words, a pair of F-indispensables belongs exclusively to one of the circles in the covering. Minimal Covering requires that each circle in the covering has such a pair of exclusive elements (note that what is exclusive is the pair, not just one of the elements).

Proposition 24. *Let $(S, \sim)$ be a similarity structure and $F, F' \subseteq SC(S)$ be similarity coverings. Then the following conditions are equivalent:*

i F is minimal.

ii *F has no proper similarity subcoverings.*

iii *$F - \{A\}$ is not a similarity covering of S, for each $A \in F$.*

iv *Every set A in F has a pair of F-indispensable elements.*

Proof. (ii)-(iii) Suppose (ii), then since it is properly contained in F, $F - \{A\}$ is not a similarity covering of S. (iii)-(iv) Suppose (iii) and, by reductio, that there is an $A \in F$ such that if $x \sim y$ then x and y belong to a $B \in F$ distinct from A. Define $F' := F - \{A\}$. If $a \sim b$, then there is a $C \in F$ such that $x, y \in C$, and if $C = A$ we can replace C by a B distinct from A. Therefore, F' is a similarity covering of S, contradicting the assumption. (iv)-(ii) Suppose that $F \subseteq F'$ and $A \in F'$. Then there are two $x, y \in A$ such that $\forall B \in F'\ \{x, y\} \subseteq B \Rightarrow B = A$. But since $x \sim y$, $\{x, y\} \subseteq C$ for some C in F, and since C belongs to F' too, we have $A = C$ and therefore $F = F'$. □

A minimal similarity covering is non-redundant, each of its members is needed to cover the similarity. If $x \sim y$ then x and y may not be indispensable for any of the sets. For example:

> Let $SC(S) = \{\{p, x, y\}, \{q, x, y\}\}$. The only minimal similarity covering of this similarity is $SC(S)$, but x and y are not indispensable for they appear in each circle.

Definition 63. *Let $(S, \sim)$ be a similarity structure. Then S is* finite-like *iff every similarity covering $F \subseteq SC(S)$ has a minimal similarity subcovering.*

Every finite similarity structure is finite-like. Moreover:

Corollary 5. *Let $(S, \sim)$ be a similarity structure of order $n \leq 2$. Then S is finite-like.*

Proof. Let $F \subseteq SC(S)$ be a similarity covering. If T is a similarity circle of order $n \leq 2$ then it has two generators x, y. Now, if $x, y \in T'$ then $T' \subseteq co(x) \cap co(y) = T$ therefore $T' = T$ for any circle T'. So T is the only circle containing x, y and since $x \sim y$ and F covers the similarity, $T \in F$. Therefore $SC_2(S) \subseteq F$. Since S is of order $n \leq 2$, $SC_2(S)$ covers the similarity and given that each pair of generators in each circle of order $n \leq 2$ are indispensable to the circle, $SC_2(S)$ is a minimal similarity covering. Therefore, S is finite-like. □

I do not know whether every similarity structure is finite-like. A straightforward application of Zorn's Lemma is not enough, the family of subcovers of a similarity covering ordered by inclusion may still have infinitely descending chains. But if there is a counterexample it is quite complex. With this notions at hand, we obtain Brockhaus's result from the following lemmas:

Lemma 7. *Let $(S, \sim)$ be a finite-like similarity structure. Then S has a unique minimal similarity covering $F \subseteq SC(S)$ iff F is a similarity covering all of whose members are similarity circles of order $n \leq 2$.*

Proof. Let S have a unique minimal similarity covering F. Let $T \in F$, we prove it is of order $n \leq 2$. Suppose for reductio that T is not of order $n \leq 2$. Since F is minimal, there are at least two elements $x, y \in T$ which are F-indispensable for T. Since x and y do not generate T, there is a $z \in co(x) \cap co(y) - T$

such that $\{x, y, z\} \subseteq T_{xy}$ for some $T_{xy} \in SC(S)$ since $\{x, y, z\}$ is a clique. Besides, T_{xy} is not in F by indispensability. Let $Ind(T) := \{\{a, b\} \subseteq S \mid a \text{ and } b \text{ are F-indispensable for } T\}$. For each $\{a, b\}$, since $a \sim b$ by the previous reasoning they are included in at least one circle T_{ab} which is not in F. Assign one such circle T_{ab} to each pair $\{a, b\} \in Ind(T)$ and take the family of all these similarity circles $SInd(T)$. Define $F^* := (F - \{T\}) \cup SInd(T)$. By stipulation, T does not belong to $SInd(T)$. For any $a, b \in S$ such that $a \sim b$, if they are F-indispensable for T, then there is a $T_{ab} \in SInd(T) \subseteq F^*$ such that $a, b \in T_{ab}$. If they are not F-indispensable for T, then there is a $T' \in F - \{T\} \subseteq F^*$ which is such that $a, b \in T'$. Therefore, F^* is a similarity covering of S. By finite-likeness, it has a minimal similarity subcovering $F^{**} \subseteq F^*$ and clearly T is not in F^{**}. Therefore, by uniqueness $F = F^{**}$, which contradicts the fact that T is not in F^{**}. So T is of order $n \leq 2$.

Conversely, let F be a similarity covering of circles of order $n \leq 2$. We prove F is minimal and unique. Let $F' \subseteq F$ be a similarity subcovering and suppose there is a $T \in F - F'$. Then T is of order $n \leq 2$. Let x, y be two (possibly identical) generators of T, then $cro(\{x, y\}) = cro(\{x\} \cup \{y\}) = cro(x) \cap cro(y) = co(x) \cap co(y) = T$. Since $x \sim y$, there is a $T' \in F'$ such that $x, y \in T'$ and therefore $T' = cro(T') \subseteq cro(\{x, y\}) = T$, i.e. $T' = T$, which contradicts the assumption. Therefore $F = F'$ and F is minimal. Let $F^* \subseteq SC(S)$ be a minimal similarity covering of S. Let $T \in F$, since it is of order $n \leq 2$ take two (possibly identical) generators x, y of T. Since $x \sim y$, again there is a $T' \in F'$ such that $x, y \in T'$ and therefore $T' = cro(T') \subseteq cro(\{x, y\}) = T$, it follows that $T = T' \in F'$ and thus $F \subseteq F^*$. Since F^* is minimal, $F = F^*$ and therefore F is unique. □

Equivalently put, S has a unique minimal similarity covering by circles iff it is of order $n \leq 2$. In principle, being the unique minimal similarity covering does not immediately imply being the *smallest* one. The poset of similarity coverings could have the following shape. There are at least two chains stemming from the top. One is an infinitely descending chain of coverings. The other one, infinite or not, ends in the unique minimal covering.

Lemma 8. *Let $(S, \sim)$ be a similarity structure. Then S has a unique strong standard quasianalysis $q\colon S \to \wp(Q)$ iff S has a unique minimal similarity covering $F \subseteq SC(S)$.*

Proof. Let q be the unique strong standard quasianalysis of S, we prove Q is a unique minimal similarity covering of S. First, $Q \subseteq SC(S)$ and by (C1-C2) is a similarity covering, for if $x \sim y$ then there is a $T \in q(x) \cap q(y) \neq \varnothing$ and by standarness $x, y \in T$. Let $Q' \subseteq Q$ be a similarity covering of S and define $q^*(x) := \{T \in Q' \mid x \in T\}$. Since Q' covers the similarity, $x \sim y \Rightarrow q^*(x) \cap q^*(y) \neq \varnothing$. The converse holds because circles are cliques, thus q^* satisfies (C1-C2). If $co(x) = co(y)$, then $T \in q^*(x)$ iff $x \in T$ iff $T \subseteq co(x) = co(y)$ iff $y \in T$ iff $T \in q^*(y)$. Conversely, if $q^*(x) = q^*(y)$ then $z \sim x$ iff $q^*(z) \cap q^*(x) \neq \varnothing$ iff $q^*(z) \cap q^*(y) \neq \varnothing$ iff $z \sim y$ by (C1-C2). Thus q^* satisfies (C3). But q is a strong quasianalysis, therefore $Q' = Q$ and so Q is a minimal similarity covering. Suppose that Q is not unique, there is a $Q^* \subseteq SC(S)$ such that Q^* is a minimal similarity covering. Then $q^*(x) := \{T \in Q^* \mid x \in T\}$ is a standard quasianalysis satisfying (C1-C4) by previous reasoning and since the quasianalysis is unique,

$q = q^*$. Therefore, $Q^* = Q$ and Q is the unique minimal similarity covering of S.

Conversely, define $q^*(x) := \{T \in F \mid x \in T\}$. It is standard and satisfies (C1-C3) by previous reasoning. Let $F' \subseteq F$ be such that $q^{'}(x) := \{T \in F' \mid x \in T\}$ satisfies (C1-C3). If $F' \neq F$, then since F is a minimal similarity covering and $F' = F - \{T_k\}$ for some T_k, F' does not cover the similarity anymore, i.e. $q^{'}$ does not satisfy (C1). Therefore $F = F'$ and q^* is a strong standard quasianalysis. We now show that q^* is unique. Let q' be a strong standard quasianalysis of the form $q' \colon S \to \wp(Q')$. We just proved that $Q' \subseteq SC(S)$ must be a minimal similarity covering. Since S has a unique minimal similarity covering, we have $Q' = F$. □

As a corollary we obtain the following result. I propose to call it the *Brockhaus-Mormann-Schreider Theorem* (or *BMS-Theorem*, for short):

Theorem 8 (BMS Uniqueness)**.** *Let $(S, \sim)$ be a finite-like similarity structure. Then the following conditions are equivalent:*

i S is a similarity structure of order $n \leq 2$.

ii S has a unique strong standard quasianalysis.

iii S has a unique minimal similarity covering by similarity circles.

What the BMS guarantees is that similarity structures of order 1 or 2 have a unique strong standard quasianalysis. The previous counterexample is a structure of order 3, where all elements in $\{x, y, z\}$ or in $\{y, z, w\}$ are generators.

Now recall from Chapter IV that the structures we are concerned with are:

Definition 64. *Let $(S, \sim)$ be a set S with a binary relation $\sim \subseteq S \times S$. Then S is a* pure similarity structure of order 1 *iff $\forall x, y \in S$ $\forall p, q \in Gen(S)$:*

i $x \sim x$. [Reflexivity]

ii $x \sim y \Rightarrow y \sim x$. [Symmetry]

iii $p \sim q \Rightarrow p = q$. [Pure]

iv $x \sim y \Rightarrow \exists p \in Gen(S)$ $x \sim p \sim y$. [Order 1]

Since every similarity of order 1 is finite-like and of order 2, this class satisfies the conditions required by the BMS theorem:

Corollary 6. *Let S be a similarity structure of order 1. Then it satisfies BMS.*

Now we can give some concrete examples of similarities of different orders:

i $(S, S \times S), (S, \Delta), \varnothing, (\{x\}, \sim_x)$ and $(\{x, y\}, \sim_{xy})$ where $x \sim_{xy} y$, are of order 1.

ii $p \sim x \sim q$ is the smallest (SNI) similarity of order 1 [87].

iii $x \sim p \sim y$ & $x \sim q \sim y$ is the smallest pure similarity of order 1 that is not (SNI).

iv Every cycle and every path with at least 4 points is a (SNI) similarity of order 2. The square $x \sim y \sim z \sim p \sim x$ is one of the smallest (SNI) similarities of order 2.

v Let the context be $S = \{x, y, z\}$ and $Q = \{\{x, y\}, \{y, z\}, \{x, z\}\}$. The similarity, which is of order 1 but not (SNI), is the so-called 'Goodman Triangle' $x \sim y \sim z \sim x$ [87]. Here $Gen(S) = \{x, y, z\}$ is also the imperfect community.

vi Let the context be $S = \{x, y, z, p, q, r\}$ and $Q = \{\{x, y, p\}, \{y, z, q\}, \{x, z, r\}\}$. The similarity, which is of order 1 and (SNI), is given by $SC(S) = \{\{x, y, p\}, \{y, z, q\}, \{x, z, r\}, \{x, y, z\}\}$. Although all the properties in Q are similarity circles of order 1, the imperfect community $\{x, y, z\}$ is not of order 1 but of order 2. Here $Gen(S) = \{p, q, r\}$.

vii Let $SC(S) = \{\{p, x\}, \{q, y\}, \{r, z\}, \{x, y, z\}\}$, then it is (SNI) of order 2.

The example (v) shows that some contexts that induce similarities with imperfect communities can still be recovered from them *if we take as properties only the similarity circles of order 1.* The existence of generators gives us an additional piece of information that we can use to discard some similarity circles. There lies the importance of Brockhaus Theorem. The examples (v)-(vi) show some important features of our definition. First, some similarities of order 1 like (v) contain similarity circles of higher order. Second, the crucial point is that there must be enough similarity circles to cover the similarities between the objects, and not just to cover the objects. If that is satisfied, then some circles may be redundant, as $\{x, y, z\}$ is in (v). But if that is not satisfied, then some circles will be indispensable to cover these similarities, as $\{x, y, z\}$ is in (vi).

For the case of similarity structures of order 1 remember that I introduced a new function $q' \colon S \to \wp(Gen(S))$. It turns out that it is equivalent (in an interesting sense) to the unique strong standard quasianalysis q:

$$q(x) := \{T \in SC_1(S) \mid x \in T\}$$
$$gen(x) := \{p \in Gen(S) \mid x \sim p\} = co(x) \cap Gen(S)$$

They are interdefinable, i.e. $q(x) = \{T_p \in SC_1(S) \mid p \in q'(x)\}$ and $gen(x) = \{p \in Gen(S) \mid T_p \in q(x)\}$. The reason is that in a pure similarity structure of order 1, each similarity circle has a unique generator. The function $gen(x)$ will turn out to be very important soon. The sense in which both functions are equivalent is the following one, introduced by [89]:

Definition 65. *Let $(S, \sim)$ be a similarity structure, Q and Q' sets of quasiproperties, $f \colon S \to \wp(Q)$ and $g \colon S \to \wp(Q')$ mappings. Define for $P \in Q$ as $e_f(P) := \{x \in S \mid P \in f(x)\}$, the* f-extension of P. *Define now the following quasianalysis called the* extensionalization of f *as $f^*(x) := \{e_f(P) \mid P \in f(x)\}$. Then f and g are* extensionally equivalent *$f =_{EXT} g$ iff $f^* = g^*$.*

We first check that gen is indeed a non-standard strong quasianalysis:

Proposition 25. *Let $(S, \sim)$ be a pure similarity structure of order 1. Then $gen \colon S \to \wp(Gen(S))$ defined as $gen(x) := \{p \in Gen(S) \mid p \sim x\}$ is a strong non-standard quasianalysis.*

Proof. (C1)-(C2) and $(C3^*)$ are immediate by S being of order 1. For (C4), take $Q^* \subset Gen(S)$ and a function $gen^*: S \to \wp(Gen(S))$ defined as $gen^*(x) := \{p \in Q^* \mid p \sim x\}$ satisfying (C1)-(C3). Since there is a $p \in Gen(S) - Q^*$, by purity we have that $gen(p) = \{p\}$, and therefore we have that $p \sim p$ but $gen^*(p) \cap gen^*(p) = \varnothing$, which violates (C1-C2). ☐

We check that the previously defined functions q and gen are indeed extensionally equivalent:

Proposition 26. *Let $(S, \sim)$ be a (SNI) pure similarity structure of order 1, where q and gen are defined as above. Then $q =_{EXT} gen$.*

So every element x in S can be represented either by the set of its similarity circles of order 1 $q(x)$ or by its set of generators $gen(x)$. Since the structure is (SNI), gen is injective and therefore $gen(x) = gen(y)$ implies $x = y$. The representation by gen will acquire an interpretation later on.

Now we come back to the dispute between nominalism and realism. As an answer to the questions concerning the feasibleness of the constitutional systems programme, and if we agree that Carnap's project does not depend on the similarity structure having to reconstruct a previously given arbitrary context[15], there is no objection to make to the approach so far. But as a nominalist proposal, to the similarity-first approach the realist will answer as follows. *If* the aim is to reduce or paraphrase *any* talk concerning properties to talk about similarity, *then* supposing that this requires us to give a direct reconstruction of any formal context, this still does not solve the problem, since we have uncontentious examples of contexts of properties which should be given some nominalist reconstruction and do not correspond to these similarity structures. As an example, take any ordinal context where the attributes are chained.

There are several other strategies. For instance, [89] shows that imperfect communities can be considered as transitory phenomena that disappear if one enlarges the domain in question. The author also shows that there are different notions of equivalence between quasianalytic mappings. So even if some reconstructions may not be identical to the one we started with, they may still be equivalent to it in some sufficiently relevant sense. The nominalist can give *approximate* reconstructions. Moreover, different answers result from dropping the requirement that the similarity be defined as coincidence in some property or the requirement that the surrogates for properties be similarity circles. As I said, it is up to the nominalist to propose a different notion of similarity to be defined from (S, Q, I) or a different surrogate for the universals once the similarity structure is assumed. Concerning the former case, several authors have suggested to start from triadic or tetradic comparative relations of similarity, collective relations of similarity or other different relations. Of course, the difficult part of these proposals rests on reconstructing the properties. Concerning the latter case, one could try to argue that any subset of a property is in some weak sense a property too, or at least a part of a property. Taking all the cliques would solve the companionship problem in this way (but not the imperfect community). Or one could use this to try to argue that it is enough to show that any property can be found as a clique in a similarity structure, in other words, that is enough to show that the realist context can be order-embedded into the

[15] For example, Leitgeb disagrees on this point. I will not enter into this dispute, which concerns Carnap's project in the Aufbau.

context determined by all the cliques of the similarity. In any case, in the following sections I will use the formal tools that we have gathered to discuss how aristocratic resemblance nominalism can deal with Goodman's problems.

5.3 A Nominalist Reduction of Universals

Given the difficulty of Goodman's problems, a piecemeal approach may be more enriching and explanatory. The idea is that which nominalist structure we should choose may depend on the additional structure present in the realist context. Here I want to tackle what I think it is a very promising approach and which was introduced in the previous chapter: *Aristocratic Resemblance Nominalism*. Remember that according to it, there are some particulars, the paradigms, that occupy a special role, they are the ground of the class. A particular belongs to the class iff it is sufficiently similar to the paradigm(s) of the class. What I want to show is that aristocratic nominalism succeeds in reconstructing at least some interesting realist contexts.

5.3.1 Equivalence between Polar and Similarity Models

There is a close conceptual affinity between the polar and similarity models for resemblance nominalism. In this section I will show that they are strictly speaking equivalent. Moreover, I will also describe the class of realist contexts to which they are equivalent. This will provide a partial reduction of the realist ontology. The contexts to be reduced are the following ones:

Definition 66. *Let (S, Q, I) be an exhaustive set-theoretical context and $P \subseteq S$. Then (S, P, Q, I) is a* polar context *iff it satisfies:*

1. $\forall R \in Q\ \exists p \in P\ (p \in R\ \ \&\ \ \forall r \in P\ (r \in R \Rightarrow p = r))$.
2. $\forall p \in P\ \forall R, T \in Q\ (p \in R \cap T \Rightarrow R = T)$.

Members of P are once again *paradigms*. So (1)-(2) say that in polar contexts there is a bijection between properties and paradigms. Each property is in some sense 'generated' by a unique element, which is the paradigm of the property. Of course, this amounts to some idealizations, usually several paradigms are assigned to each property. But since each paradigm belongs at most to one property, one can simply think about the paradigm of a property R as if it were a representative of all the paradigms in R. We can now prove the first main result:

Theorem 9. *Let (S, P, m) be a polar distribution. Then $(S, P, Q^*, \in)$, where $Q^* := \{Cl(p) \subseteq S \mid p \in P\}$, is a polar context whose polar distribution (S, P, n) is such that $n = m$. Conversely, let $(S, P, Q, \in)$ be a polar context. Then (S, P, n), where $n(x) := \{p \in P \mid i(x) \cap i(p) \neq \varnothing\}$, is a polar distribution whose polar context $(S, P, Q^*, \in)$ is such that $Q^* = Q$.*

Proof. $(S, P, Q^*, \in)$ is a realist set-theoretical structure, since $p \in m(x) \neq \varnothing$ implies $x \in Cl(p) \neq \varnothing$. If $R \in Q^*$, then $p \in m(p) = \{p\}$ iff $p \in Cl(p) = R$. Suppose $p, r \in P \cap R = P \cap Cl(q)$. Then $q \in m(p) \cap m(r) = \{p\} \cap \{r\}$, so $p = q = r$. Let $p \in P \cap R \cap T = P \cap Cl(r) \cap Cl(t)$, then $r, t \in m(p) = \{p\}$, so $R = Cl(r) = Cl(p) = Cl(t) = T$, which proves that the structure is polar.

We prove n is polar. Since $p \in m(x) \neq \varnothing$ for all x, $x, p \in Cl(p)$. Let $p \in P$ and $q \in n(p)$, then $p, q \in Cl(r)$, therefore $r \in m(p) \cap m(q) = \{p\} \cap \{q\}$ and so $n(p) = \{p\}$. We prove $n = m$. Let $p \in n(x)$, then $x, p \in Cl(q)$ for some $Cl(q) \in Q^*$. Therefore $q \in m(p) = \{p\}$, which implies $p \in m(x)$.

Conversely, if $p \in m(x)$ then $x, p \in Cl(p)$ so $p \in n(x)$. Let $n \colon S \to \wp(P)$ be defined as $n(x) := \{p \in P \mid \exists R \in Q\ x, p \in R\}$. By assumption, for every x we have $x, p \in R$ for some $p \in P, R \in Q$ and thus $p \in n(x)$. Let $p \in P$ and $q \in n(p)$. Then $p, q \in R$ so by polarity $p = q$, which makes n polar. It follows that $(S, P, Q^*, \in)$ is polar. We prove now that $Q = Q^*$. Let $Cl(p) \in Q^*$. Since p is a paradigm, it corresponds to a unique $R \in Q$. If $x \in Cl(p)$, then $p \in n(x)$, so there is a $T \in Q$ such that $x, p \in T = R$. If $x, p \in R$ then $p \in n(x)$ so $x \in Cl(p)$. So $Cl(p) = R$. It follows that $Q^* \subseteq Q$. Let $R \in Q$, then it corresponds to a unique paradigm $p \in R$, we analogously prove that $Cl(p) = R$. □

So, some of the closed sets in the polar topology are the surrogates for universals. We can extend our correspondences to pure similarity structures of order 1:

Theorem 10. *Let (S, P, m) be a polar distribution. Then $(S, \sim)$ defined as $x \sim y := \exists p \in P\ p \in m(x) \cap m(y)$ is a pure order 1 similarity. It induces a polar distribution $(S, Gen(S), gen)$ which is such that $P = Gen(S)$ and $m = gen$. Conversely, let $(S, \sim)$ be a pure order 1 similarity. Then $(S, Gen(S), gen)$ is a polar distribution where $(S, \sim')$ is such that $\sim = \sim'$.*

Proof. Let (S, P, m) be polar and define $x \sim y := \exists p \in P\ p \in m(x) \cap m(y)$, which is symmetric. By polarity, $m(x) \neq \varnothing$ and reflexivity follows. Let $p \in P$, if $w \sim p \sim z$, then there are $r, s \in P$ such that $r \in m(w) \cap m(p)$ and $s \in m(z) \cap m(p)$, so by polarity $r = p = s$. Therefore $p \in m(w) \cap m(z)$ and $w \sim z$. So p is a generator. If $x \sim y$ then there is a $p \in P \cap m(x) \cap m(y)$ and so the similarity is of order 1. Suppose that $p, r \in P$ then if $p \sim r$ we have $m(p) \cap m(r) = \{p\} \cap \{r\} \neq \varnothing$, so $p = r$ and the similarity is pure.

Now let $(S, Gen(S), gen)$. We have $Gen(S) \subseteq S$ and $gen \colon S \to \wp(Gen(S))$ which satisfies $gen(x) \neq \varnothing$ by order 1. If $p \in Gen(S)$, then $gen(p) = \{p\}$ because if $q \in Gen(S) \cap gen(p)$ we have $q \sim p$ and by purity $p = q$. So it is a polar distribution. We already showed that $P \subseteq Gen(S)$. Let $p \in Gen(S)$ we prove $m(p) = \{p\}$. If $q \in m(p)$ then $q \in m(p) \cap m(q)$ and so $p \sim q$. Since $p, q \in Gen(S)$ by purity $p = q$. Therefore $m(p) = \{p\}$ and so $p \in P$. It follows that $P = Gen(S)$. Therefore, $p \in m(x) \Leftrightarrow p \in P\ \ \&\ \ p \sim x \Leftrightarrow p \in Gen(S)\ \ \&\ \ p \sim x \Leftrightarrow p \in gen(x)$, because if $p \in P$ and $p \sim x$ then there is a $q \in m(p) \cap m(x)$ such that $q = p$.

Conversely, let $(S, \sim)$ be a pure similarity of order 1. We already proved that $(S, Gen(S), gen)$ is a polar distribution and that $(S, \sim')$ is pure of order 1. We show $\sim = \sim'$. By order 1, $x \sim y \Leftrightarrow \exists p \in Gen(S)\ p \in gen(x) \cap gen(y) \Leftrightarrow x \sim' y$. □

Thus, there is a correspondence between polar distributions and the similarity structures of order 1. Therefore:

Corollary 7. *Let $(S, P, Q, \in)$ be a polar context. Then its similarity structure $(S, \sim)$, where $x \sim y := \exists R \in Q\ x, y \in R$, is pure of order 1. It induces the polar context $(S, Gen(S), SC_1(S), \in)$ which is such that $SC_1(S) = Q$ and $P = Gen(S)$. Conversely, let $(S, \sim)$ be a pure similarity structure of order 1.*

Then $(S, Gen(S), SC_1(S), \in)$ *is a polar context whose similarity* $(S, \sim')$ *is such that* $\sim = \sim'$.

Corollary 8. *Polar distribution is (PII)* $\Leftrightarrow$ *Polar context is (PII)* $\Leftrightarrow$ *Similarity is (SNI).*

Proof. (i)-(ii) $m(x) = m(y) \Leftrightarrow \forall p \in P\ (p \in m(x) \Leftrightarrow p \in m(y)) \Leftrightarrow \forall Cl(p) \in Q^*\ (x \in Cl(p) \Leftrightarrow y \in Cl(p)) \Leftrightarrow i^*(x) = i^*(y)$.

(i)-(iii) If (S, P, m) is (PII) and $co(x) = co(y)$, then $p \in m(x) \Leftrightarrow p \in P\ \&\ p \sim x \Leftrightarrow p \in P\ \&\ p \sim y \Leftrightarrow p \in m(y)$. Conversely, if $(S, \sim)$ is (SNI) and $m(x) = m(y)$, then $x \sim z \Leftrightarrow \exists p \in Gen(S)\ x \sim p \sim z \Leftrightarrow \exists p \in P\ p \in m(x) \cap m(z) \Leftrightarrow \exists p \in P\ p \in m(y) \cap m(z) \Leftrightarrow \exists p \in Gen(S)\ y \sim p \sim z \Leftrightarrow y \sim z$. Thus $co(x) = co(y)$ and by (SNI) we have $x = y$. □

So polar distributions, polar contexts and pure similarity structures of order 1 are all mathematically equivalent structures. Paradigms are exactly the generators. Closures of paradigms are some polar-closed sets which are exactly the properties and which are also exactly the similarity circles of order 1. A fortiori, exemplification now has spatial content: to say that a particular x has a property P is just to say that x is arbitrarily close to the paradigm p of P. Polar distributions give us a new axiomatization of pure similarities of order 1. These correspondences extend to some of the structure preserving mappings, we will skip the proofs for readability:

Lemma 9. *Let* (X, P_x, m_x, T_x), (Y, P_y, m_y, T_y) *be polar distributions. If* f *is a homeomorphism from* X *to* Y*, then* $f(P_x) = P_y$ *and* $f(m_x(z)) = m_y(f(z))$.

Proposition 27. *Let* $(X, \sim_X, gen_X), (Y, \sim_Y, gen_Y)$ *be (SNI) similarity structures of order 1 and* $f\colon X \to Y$ *a function. Then:*

1. *If* f *is polar continuous, then it is a similarity homomorphism.*
2. *If* f *is a similarity isomorphism, then* $f(Gen(X)) = Gen(f(X))$.
3. f *is a similarity isomorphism* $\Leftrightarrow$ f *is a polar homeomorphism.*

From the allegedly anaemic similarity we got an enriched similarity structure with its topology and polar structures. The paradigms induce a classification over the domain of objects, the classes being the similarity circles of order 1, which are closed sets. Each similarity circle T can be partitioned into two parts, the similarity interior $int(T) = \{p\}$ that contains the generator of the class, and the similarity interior boundary $bd_{int}(T) = T{-}p$ that contains the rest of elements in the class (recall $bd(A) := co(A) \cap co(A^c)$ and $bd_{int}(A) := bd(A) \cap A$). As we will see, the similarity interior boundary of the class coincides with its (polar) topological boundary $Bd(T)$ (again, we will skip the proofs for readability):

Corollary 9. *Let* $(S, \sim)$ *be an (SNI) similarity of order 1 with its polar topology and similarity order. Let* $T, T' \in SC_1(S)$*, then:*

i $bd_{int}(T) = Bd(T)\ \&\ int(T) = Int(T) = \{p\}\ \&\ cro(T) = T = Cl(T)$.

ii If T *and* T' *are distinct, then* $T \cap T' = Bd(T) \cap Bd(T')$.

iii If $\sim$ *is not the identity, then there are two overlapping circles of order 1.*

iv $x \sim y \Leftrightarrow N_x \cap N_y \neq \varnothing \Leftrightarrow \exists p \in Gen(S)\ x, y \in Cl(p)$.

v $co(x)$ *is polar-closed.*

vi $co(p) = T$ *for* $p \in Gen(S)$ *is an atom regular-closed set. (see [90])*

Point (i) says that for similarity circles (of order 1), the topological and similarity operations coincide. (ii) says that kinds corresponding to similarity circles overlap at their boundaries. (iii) says that (for non-trivial cases) at least two kinds must overlap, since the structure is (SNI) and therefore each non-paradigmatic element must belong to several classes. This implies that the corresponding classifications, in contrast with the kind of traditional classifications discussed in Chapter III, have many 'criss-crossings'. (viii) says that the similarity corresponds to the overlapping of the corresponding minimal open neighbourhoods. This allows us to give a spatial reading of kinds: two objects are similar iff they are arbitrarily close to a common paradigm, and if an object is arbitrarily close to the extension of a kind then it belongs to that kind (i.e. kinds are stable).

5.3.2 Equivalence between Order and Similarity Models

I will now show that the order-theoretic and similarity models are also equivalent to each other. Let us consider the following pre-order from Mormann's [89], it can be defined in any similarity structure:

Definition 67. *Let* $(S, \sim)$ *be a similarity structure. Then* $(S, \leq_{co})$ *where* $x \leq_{co} y := co(x) \subseteq co(y)$ *is the* similarity preorder *over* S.

It is easy to see that $(S, \sim)$ satisfies (SNI) iff $\leq_{co}$ is a partial order. For this reason, from now on we will only deal with (SNI) structures. To distinguish it from the polar order $x \leq^* y$ iff $x = y$ or $y \in gen(x)$, we will call this order the *similarity order* or *similarity poset*. One can think about it as a *qualitative ordering*. We can say that x is *qualitatively thinner than* y iff $x \leq_{co} y$. This is what I have been calling qualitative thinness up to now.

We want to establish a correspondence between this order and the similarity. Let us introduce some auxiliary definitions. An element x is a *dense element* $\Leftrightarrow co(x) = S \Leftrightarrow x$ belongs to every similarity circle. Equivalently in order-theoretic terms x is dense iff it is the top element. Since $\sim$ is (SNI), the dense element is unique. See that the similarity order cannot have a bottom. If $x = 0$ then $co(x)$ would be included in every other neighbourhood. But if there is a y similar to x, then $co(y) = S$, and by (SNI) $y = x$. Therefore, x is similar only to itself, and if the structure has at least two points, x cannot be the bottom. Nevertheless, for similarities of order 1 some interesting meets and joins do exist:

Proposition 28. *Let* $(S, \sim, \leq_{co})$ *be an* (SNI) *structure of order* 1*. Then:*

i p *is minimal* $\Leftrightarrow p$ *is a generator of order 1.*

ii $x = \bigvee gen(x)$.

iii $x = \bigwedge(\bigcap\{T \in SC_1(S) \mid x \in T\}) = \bigwedge(\bigcap q(x))$.

iv If p *generates the similarity circle* T_p *of order* 1*, then* $p = \bigwedge T_p$.

v *x is maximal* $\Leftrightarrow \{x\} = \bigcap q(x)$.

Proof. We only prove (i)-(ii). (i) Let p be minimal. It belongs to some similarity circle T of order 1 with unique generator z. Then $T = co(z) \subseteq co(p)$ and by minimality $p = z$. Conversely, let p be the generator of an order 1 circle T and $y \leq p$. Then $y \in co(y) \subseteq co(p) = T$ and therefore $co(y) = T$. It follows that $p = y$ and so p is minimal. (ii) x is greater than each of its generators. Let $p \leq y \leq x$ for each generator p of x. Then $gen(x) \subseteq gen(y) \subseteq gen(x)$ and therefore $x = \bigvee gen(x)$. □

What (i) says is that paradigms are such that no other object can be qualitatively thinner than they are. By (ii) we see now that elements are represented by their generators in a very strong sense. Indeed, they can be *directly constructed* from them. This is very important, as we will see now.

The first thing to notice is that two isomorphic similarity structures have isomorphic posets [89]:

Proposition 29. *Let* $(S, \sim)$, $(S', \sim')$ *be* (SNI) *isomorphic similarity structures. Then their posets* $(S, \leq_{co})$, $(S', \leq'_{co})$ *are isomorphic.*

Proof. Let f be the similarity isomorphism. If $x \leq y$ and $f(p) \sim' f(x)$, it follows that $p \sim x$, so $p \sim y$ and $f(p) \sim' f(y)$, i.e. $f(x) \leq' f(y)$. Conversely, if $f(x) \leq' f(y)$, then if $p \sim x$, $f(p) \sim' f(x)$, so $f(p) \sim' f(y)$ and $p \sim y$, i.e. $x \leq y$. □

However, the converse is false, as Mormann [89] shows. A simpler counterexample is this. Take the square $x \sim y \sim p \sim q \sim x$ and the similarity structure consisting of four isolated elements. These two similarities are T_1 and determine the same poset, the antichain of four elements, and therefore the same completion, the diamond of four points. Notice that the square is of order 2. This matters, for if the similarities are of order 1 then it is easy to show that the correspondence holds:

Proposition 30. *Let* $(S, \leq_{co}), (S', \leq'_{co})$ *be the isomorphic posets of two similarity structures* $(S, \sim), (S, \sim')$ (SNI) *of order* 1*. Then* $(S, \sim), (S, \sim')$ *are isomorphic.*

This suggests that there is an interesting correspondence between the similarity and order in similarities of order 1. This is indeed the case. The following is one of the main new results that grounds the rest of results of the chapter:

Theorem 11. *Let* $(S, \sim)$ *be a (SNI) similarity structure of order 1. Then* $(S, \leq_{co})$ *is an atomistic poset where the minimal elements are* $Gen(S)$*. Moreover,* $(S, \sim^*)$*, defined as* $x \sim^* y := \exists z \in Min(S)\ z \leq x, y$*, is identical to* $(S, \sim)$*. Conversely, if* $(S, \leq)$ *is an atomistic poset, then* $(S, \sim^*)$ *is a (SNI) similarity structure of order 1 such that* $\leq = \leq_{co^*}$*.*

Proof. Let $(S, \sim)$ be the similarity. We already proved that the generators are exactly the minimal elements and that every element is the join of its generators. Now $x \sim y$ iff $gen(x) \cap gen(y) \neq \varnothing$ iff $x \sim^* y$. Conversely, let $(S, \leq)$ be an atomistic poset. If p is minimal and $x \sim^* p \sim^* y$ then $z \leq x, p$ and $w \leq y, p$ and by minimality $z = p = w$ and so $x \sim^* y$. Conversely, if p is a generator and $x \leq p$ then $min(x) \subseteq min(p)$. If $q \in min(p)$ and $y \sim^* q$ then $y \sim^* p$ therefore

$co^*(q) \subseteq co^*(p)$. Since p is a generator of order 1, $x \in co^*(p) = co^*(q)$ and therefore $q \sim^* x$, from which it follows that $q \leq x$ by minimality and therefore $min(x) = min(p)$. Thus $x = p$ and so p is minimal. So the minimals are exactly the generators and by definition S is of order 1. Let $co^*(x) = co^*(y)$, then by order 1 $min(x) = gen^*(x) = gen^*(y) = min(y)$ and therefore $x = y$. Finally, $x \leq y$ iff $min(x) \subseteq min(y)$ iff $gen^*(x) \subseteq gen^*(y)$ iff $x \leq_{co^*} y$. □

In other words, the class of (SNI) similarities of order 1 is the class of atomistic posets. The generators of order 1 are exactly the minimal elements and the similarity circles of order 1 are the principal filters $\uparrow p$ of the minimal elements. In other words, the similarity circles of order 1 are exactly the non-trivial fixed ultrafilters.

Note that the similarity corresponds to having a common lower bound, to 'overlapping non-trivially'. In a sense it is old news to say that the similarity circles of order 1 are the fixed ultrafilters. We are requiring not only that any two elements in the circle overlap (are similar to each other), but that all of them *collectively overlap*. In other words (surprise!): there are no imperfect communities.

This result has a very interesting consequence. Recall that Mormann [91] proposes to take the general class of weakly-scattered Alexandroff spaces as a model for conceptual spaces. I suggested that we added one more axiom in order to restrict this class (co-atomistic spaces), so that its dual corresponds to atomistic posets. By combining the two results of these sections we obtain the following corollary, that establishes the equivalence of the two topological models:

Corollary 10. *Let $(S, O(S))$ be a co-atomistic weakly-scattered Alexandroff space. Then $(S, Max(S), m')$ where $m'(x) := \{p \in S \mid x \in Cl(p)$ and p is open$\}$, is a (PII) polar distribution such that $(S, \leq)$ defined as $x \leq y \Leftrightarrow m'(y) \subseteq m'(x)$ is the specialization order of the original space. Conversely, if $(S, P, m$ is a (PII) polar distribution, then $(S, \leq)$ is the specialization order of a co-atomistic weakly-scattered Alexandroff space which is such that $Max(S) = P$ and $m' = m$.*

This is an interesting result for the following reason. Although the class of weakly-scattered spaces is more general than that of polar spaces, an additional axiom makes them equivalent, so that the comparative and categorical notions of prototypicality can be defined in terms of each other.

5.3.3 Relevance and Limitations of the Previous Result

To sum up, we had several different models of aristocratic nominalism: the polar topology induced by a polar distribution, polar contexts and similarity structures of order 1. These three models are equivalent. Let us summarize this result roughly as:

Theorem I Polar contexts, polar topologies and pure similarity structures of order 1 are equivalent.

Moreover, under the further assumption of indiscernibility, we got:

Theorem I* Atomistic posets, co-atomistic topologies and (SNI) similarity structures of order 1 are equivalent.

The result establishes a correspondence between all these models. But what is the philosophical relevance of this result? There are two things to consider. On the one hand, how does it deal with Goodman's problems? For instance, so long as we take attributes to be represented by similarity circles, the companionship problem cannot be dealt with. How is this a reply to Goodman? The result might seem quite limited. It gives the resources to 'translate' only a certain class of realist structures.

However, what the nominalist can question is whether the extensions of attributes can be nested *at all*. The nominalist assumes that the realist structure to be reconstructed has, for each property, a unique object that only has that property. In such a structure, no property can be properly included into others. Consider the interpretation of paradigms as tropers, qualitatively thin entities, that we made in the previous chapter. If there is, for each property, a unique object which is such that it only has that property, then it is no mystery that no companionship problems arise. If all R-s are T-s, then the paradigm of R is also a T and since such an object has a unique property, $R = T$. Analogously, there cannot be imperfect community problems. If there is a (maximal) imperfect community, it is not an attribute and therefore there is no paradigm corresponding to it. The members of the imperfect community are pairwise similar to each other because they are pairwise similar to a paradigmatic object (the paradigm of the property that the two objects share). But there is no paradigm to which all the objects are similar. Equivalently put, similarity circles of order 1 cannot be imperfect communities.

For the strategy to work in every case, the realist would have to accept the correspondence between attributes and paradigms. To many this will sound unreasonable. The realist can reject the existence of objects that have a unique property as suspicious entities. If the realist accepted the existence of tropes, then that would be an argument for assuming this kind of contexts. But most realist of universals do not accept tropes. Why would they accept qualitatively thin objects?

Nevertheless, it is not clear to me that the aristocratic nominalist cannot reply again. Similarity circles cannot be nested. However, all the objects that one paradigm is similar to can also be similar to another paradigm. In other words, if we delete the paradigms from the similarity circles, the resulting collections of non-paradigmatic objects can be nested. For example, recall the example of an ordinal context where the properties in Q form a chain: *extremely loud* $\leq$ *very loud* $\leq$ *loud*. Let us grant that there is a unique paradigm p, q, r for each such property. Then, every non-paradigmatic object that is extremely loud is also very loud and loud, and every non-paradigmatic object that is very loud is also loud. If one thinks about the degree of similarity as being strict enough, the paradigm of extreme loudness is not itself very loud nor loud. Equivalently, we have $co(p) - \{p\} \subseteq co(q) - \{q\} \subseteq co(r) - \{r\}$.

Now consider the following move. The realist gives a context. The nominalist replies and says that each of the properties of the realist that lacks a paradigm should have one. So, he *introduces new paradigms into the context* which will look to the realist like entities constructed from his own ontology. In this way the nominalist 'completes' the realist structure by adding what he thinks it should be there, namely the remaining paradigms. The resulting context is polar. Then the nominalist defines a similarity as usual in this new context and recovers the properties by making use of the previous result. However, the properties recovered contain entities that the realist did not accept, namely the paradigms introduced by the nominalist. Accordingly, the nominalist deletes these paradigms and nests the remaining collections, which are the original realist properties. In other words, the new strategy amounts to introducing new paradigms ('quasi-paradigms') into the old context and then preserving this information while passing to the similarity structure. In particular, if there are no paradigms at all, according to the realist, then the set of polar properties is always empty. If so the nominalist can introduce these new paradigms (which for the realist are idealizations of some sort) and then recover the original properties from the similarity induced by deleting the paradigm of each similarity circle of order 1. I will not consider the formal details of such an approach, although with the previous result at hand this is straightforward (the move is analogous to some of the results obtained below, regarding order 1 completions).

There is another way to reply. In the last sections of this chapter I will generalize the similarity model to one where each object can be a paradigm of an arbitrary number of properties and each property can have an arbitrary number of paradigms. Although I will not specify the class of contexts that can be reconstructed, it will be clear that this class is huge and forms a more plausible model for the realist contexts. This will be done by focusing on a generalization of similarity circles of order 1 (which will be called 'simple circles'). I think that such an answer will make more difficult for the realist to reject that the contexts should have such structure. After all, the thesis that each property has exactly one paradigm might be too strong, but the thesis that, for every property, there is a smallest set of objects whose only common property is that property (a set of paradigms of the property) does not look so strong. In fact, these sets of paradigms that 'generate' the properties will be defined in such a way to allow for the property (as a set of objects) to be a set of paradigms for itself (so long as there is no smaller subset that is enough to generate the property). Thus, trivially every property can have a set of paradigms that generate it. The crucial question will be whether there is always a *smallest one*. Unfortunately, these questions lead us beyond the limits of this PhD thesis.

On the other hand, one can question how plausible these models are and what sort of 'bridge principles' we get from them. I argued for the plausibility of the similarity model in the previous chapter and I briefly suggested in what sense the topological model is plausible too. The fact that different plausible models of the same phenomenon turn out to be mathematically equivalent strongly suggests their material adequacy. In the previous chapter I suggested that the polar and similarity models were plausibly materially adequate accounts of aristocratic resemblance nominalism. Thus, Theorem I already suggests an argument in favour of similarity structures as an adequate model for aristocratic resemblance nominalism. In other words, I give the following argument for the material adequacy of similarity structures as a model of aristocratic resemblance nominalism:

Argument for Adequacy as Convergence of Models

1. If several different independently intuitively plausible models of the same phenomenon turn out to be mathematically equivalent, this suggests that they are materially adequate.

2. Pure similarity structures of order 1, polar contexts and polar topolo-

gies (and atomistic posets, under (SNI)) are mathematically equivalent and independently intuitively plausible models of aristocratic resemblance nominalism.

3. Therefore (plausibly), similarity structures give a materially adequate model of aristocratic resemblance nominalism.

This one is a 'convergence argument'. The first premiss says that if different independently intuitively plausible models that purport to be models of the same phenomenon turn out to be mathematically equivalent, this suggests that they are adequately tracking the fundamental features of the phenomenon in question. To mention a famous example, the celebrated Church-Turing thesis says that any of the models of computability (Turing machines, λ-calculus or recursive functions) adequately represents the phenomenon of computability. The Church-Turing thesis is not something that can be proven to be true, since it is a thesis concerning the correspondence between a mathematical model and a phenomenon to be modelled. Nevertheless, the fact that these three different models turn out to be mathematically equivalent strongly suggests that they are adequately tracking the phenomenon in question. The second premiss appeals to the First Theorem and to the arguments given in Chapter IV in favour of the adequacy of the model in terms of similarity structures to reach the conclusion.

The idea is this. In order to explain the assumptions of this sort of resemblance nominalism one could give different models. According to the similarity model, objects are in categorical similarity relations to each other, paradigms are objects which are such that any two objects similar to them are similar to each other, no two paradigms are sufficiently similar to each other and every pair of similar objects are similar to a common paradigm. According to this picture, objects get classified by selecting those paradigms to which they are similar and then taking the classes consisting of all the objects similar to each of those paradigms. According to the polar context model, objects belong to certain extensional properties (sets) which are in one-one correspondence with some of the objects, the paradigms. Paradigms are objects which have a unique property, a fortiori, no two distinc paradigms can share a property. Therefore, since every object has a property, every object shares a property with at least one paradigm. According to this picture, objects come already classified by the properties they share with each other. Thirdly, according to the topological model, objects are thought of as points in a space. The open points, the paradigms, are located through the space in such a way that they form a dense region, in the sense that any object we choose will be found to be arbitrarily close to at least one of these paradigms. In other words, the paradigms structure the space around them. According to this picture, objects are classified by taking for each open point all the objects that are arbitrarily close to it. Finally (once indiscernibilty axioms are added), according to the order-theoretic model, objects are related by qualitative richness or prototypicality relations, paradigms are the qualitatively poorest (most prototypical) objects and attributes are ultrafilters. These pictures are independently plausible ways to represent the commitments of the resemblance nominalist. Since they are mathematically equivalent, this suggest that they offer us an adequate model for the facts involved.

Regarding the connections between these different pictures, we have some interesting correspondences. The previous theorem establishes some sort of 'translations' among them. Let us assume the indiscernibility axioms to get the whole picture. First, similarity consists of sharing a common property, being mapped to a common paradigm, being arbitrarily close to a common open point and having overlapping bundles of attributes (this will be explained later on):

x is similar to y iff

x and y are similar to a common paradigm iff

x and y have a common property iff

x and y are arbitrarily close to a common open point.

x and y, as bundles of attributes, overlap.

Indiscernibility corresponds to being similar to the same objects, sharing the same properties and being mapped to the same paradigms:

x and y are indiscernible iff

x and y are similar to the same objects iff

x and y are similar to the same paradigms iff

x and y have the same properties iff

x and y are arbitrarily close to the same open points iff

x and y, as bundles of attributes, are identical.

Paradigms correspond to open points and to objects that have a unique property:

p is a paradigm iff

Any two objects similar to p are similar to each other iff

p has a unique property iff

p is an open point iff

p is a qualitatively minimal or most prototypical object.

Attributes are special similarity circles generated by paradigms, the collections of objects similar to a paradigm:

T is an attribute iff

T is the collection of all the objects similar to a paradigm iff

T is a maximal collection of pairwise similar objects generated by a unique object iff

T is the closure of an open point.

T is a fixed ultrafilter.

Finally, we have the correspondences of the qualitative order:

x is qualitatively thinner than y iff

y is similar to every object to which x is similar iff

y is similar to every paradigm of x iff

y has all the properties that x has iff

x is more prototypical than y is.

In any case, it is now time to get to the nominalist picture of kinds.

5.4 External Structure of Nominalist Kinds

In the previous section I offered a partial reduction of realist structures by showing first that the similarity and polar models of aristocratic resemblance nominalism were equivalent, and then proving that there is a unique one-one correspondence with a certain class of realist structures. In this section I want to consider the main application of these results, which is the core of this phd thesis: to get a picture of the resemblance nominalism approach to natural kinds. As argued in Chapter III, we need at least a complete lattice of kinds. Through the next sections, there are several questions I want to consider:

1. Which one is the most plausible candidate for the nominalist lattice of natural kinds?
2. What are the properties of this lattice? What does it look like?
3. What more specific kinds of similarities are there?

In other words, the first question asks for the most plausible nominalist model of kinds obtained from the similarity model and the second and third questions are about the structure of this lattice of kinds. All these questions are entangled and they lead to interesting results. In particular, they will give us a picture of the nominalist lattice of kinds. We will deal with them in the order just given.

5.4.1 Lattice of Nominalist Kinds

Let us start with the first question. Since we chose similarity circles of order 1 as surrogates for attributes, the most plausible surrogate for the realist account of natural kinds is the lattice $\mathbf{B}(S, SC_1(S), \in)$. That is to say, we just take the nominalist surrogates for realist attributes and we generate the lattice of natural kinds following the steps described in Chapter III. Nominalist kinds are pairs (A, B) where A is the overlapping of all the similarity circles (of order 1) in B. In other words, the extension of a kind is a set of all those objects that are similar to some set of paradigms. Its intension is a family of sets of similar objects. For the resemblance nominalist, membership to a given kind is equivalent to being similar to all the paradigms of those attributes common to the members of the kind. There is no circularity here of course, since the attributes themselves are taken to be the sets of objects similar to a given paradigm. It is important to notice that although Goodman's companionship problem did not allow us to get attributes that are implied by (or whose extension is included in) other attributes, the resulting nominalist kinds can be proper species of one another. Since the resulting lattice is obtained following the recipe in Chapter III, we immediately have nominalist surrogates for the species-genus relations and Kant's Law; the logical operations of abstraction, specification, specific difference and negation; Aristotelian syllogistics; and so on. The nominalist just takes advantage of all the resources we outlined.

Corollary 11. *Let* $(S, \sim)$ *be an (SNI) similarity of order 1. Let* $T \in SC_1(S)$ *and* $(A, B) \in \boldsymbol{B}(S, SC_1(S), \in)$*, then:*

i $A \in \boldsymbol{B}_S \Leftrightarrow A = \bigcap(\{T \in SC_1(S) | A \subseteq T\})$.

ii $B \in \boldsymbol{B}_Q \Leftrightarrow B = \{T \in SC_1(S) \mid \bigcap B \subseteq T\}$.

iii $A \in \boldsymbol{B}_S \Rightarrow A$ *is polar-closed.*

iv $A \in \boldsymbol{B}_S - SC_1(S) \Rightarrow Int(A) = \varnothing$.

Proof. (i) $A = ei(A)$ iff $A = \bigcap i(A) = \bigcap\{T \in SC_1(S) \mid A \subseteq T\}$. (ii) Idem. (iii) If A is an extension, then $A = \bigcap q(A)$, but $q(A) \subseteq SC_1(S) \subseteq C(S)$ where $C(S)$ is the family of the polar-closed sets. So A is polar-closed. (iv) If $A \in \mathbf{B}_S - SC_1(S)$, then if $x \in Int(A)$ it would follow that $\varnothing \neq gen(x) \subseteq A$. But then $A \cap Gen(S) \neq \varnothing$ and $A \in SC_1(S)$, contradicting the assumption. Therefore $Int(A) = \varnothing$. □

Point (i) says that every extension is an intersection of similarity circles. (iii) says that the extensions of natural kinds are topologically well-behaved sets in the polar topology, all of them are closed sets. Since no paradigm can have more than one attribute, the complex kinds that result from the overlapping of other kinds contain no paradigms in their extension. In other words, attributes are closed regions that are structured around a given paradigmatic object, they contain all the objects that are similar (arbitrarily close) to that paradigm. In contrast, kinds have as extensions closed regions which are the overlapping of several attributes and therefore contain the objects that are similar to the paradigms of these attributes.

Let us consider some examples. If the kind $Crow$ has as intension $B = \{Blackness, Featheredness, \dots\}$ then the extension of $Crow$ is the set of objects (the crows) which are similar to the paradigm of *Blackness*, the paradigm of *Featheredness*, and so on. The kind of crows does not include any of these paradigms, since no paradigm can have more than one attribute. Furthermore, *Blackness*, *Featheredness* and so on are maximal classes of objects that are similar to given paradigms. If we say that "all crows are birds", $Crows \leq Birds$, then the nominalist explains this claim equivalently as "if an object is similar to all the objects in the extension of *Blackness* and of *Featheredness* and of ... then it is similar to all the objects in the extension of featherness and of ...", which is also equivalent to "if an object is similar to the paradigm p_{black} and to the paradigm $q_{feather}$ and to ..., then it is similar to $q_{feather}$ and to ...".

As another example, let us consider once again the colour wheel. Our similarity structure is given by a domain of particular spots $S = \{A, B, C, D, E, F, G, H, I, J, K, L\}$ where the paradigmatic spots are $Gen(S) = \{A, C, E, G, I, K\}$ and the similarity induces the similarity circles of order 1 $SC_1(S) = \{\{A, B, L\},$

	Orange	Yellow	Green	Blue	Purple	Red
A	X					
B	X	X				
C		X				
D		X	X			
E			X			
F			X	X		
G				X		
H				X	X	
I					X	
J					X	X
K						X
L	X					X

Table 5.1: Context of the Colour Circle

$\{B, C, D\}, \{D, E, F\}, \{F, G, H\}, \{H, I, J\}, \{J, K, L\}\}$. These attributes can be renamed as $SC_1(S) = \{Orange, Yellow, Green, Blue, Purple, Red\}$. Now that we have the natural attributes, we have the context $(S, SC_1(S), \in)$ as shown by the table 5.1.

The context induces the lattice of natural kinds, whose members are the kinds $Orange$, $Orangish - Yellow$ ($Turmeric$), $Yellow$, $Yellowish - Green$ ($Lime$), $Green$, $Greenish-Blue$ ($Turquoise$), $Blue$, $Bluish-Purple$ ($Indigo$), $Purple$, $Purplish-Red$ ($Magenta$), Red, $Reddish-Orange$ ($Vermellion$), e.g. $Lime = (\{D\}, \{Yellow, Green\}) = (\{D\}, \{\{B, C, D\}, \{D, E, F\}\})$.

Let us take stock of what we have done. According to aristocratic resemblance nominalism:

Aristocratic Resemblance Nominalism $(L, \leq)$ is the concept lattice induced by the polar context $(S, SC_1(S), \in)$, which is induced by a pure similarity structure of order 1 $(S, \sim)$.

The aim of the following sections is to get more information regarding this lattice of kinds and connect this topic to that of quasi-analysis.

5.4.2 Quasianalytic Representations of Objects

The introduction of paradigmatic objects is just half of the story. Recall that the discussion about the reducibility of properties arose from Carnap's project of developing a theory of systems of constitution of concepts. Carnap considered the reconstruction of properties as similarity circles just as an application of his quasianalytic method. Mormann reformulated the quasianalysis as a function that allows us to represent objects set-theoretically as bundles of (quasi)properties. Representing entities as bundles of properties is a very popular strategy. A clear example is the way spatial models (like Carnap's or Gärdenfors's) use n-tuples of determinate properties to represent objects. Other examples could be the use of spatial and temporal coordinates or other indexes such as possible worlds to represent objects. Moreover, such representations immediately give tools to construct new surrogate entities. To put a specific example of philosophical relevance, if one starts with possible worlds and constructs propositions as similarity circles, then possible worlds can be represented as sets of propositions without circularity. The class of similarities we have been dealing with is a very rich source of these representations. In this section I will explore at least two interesting ways to represent objects quasianalytically by using similarity structures of order 1:

- Objects can be represented as the sets of those paradigms to which they are similar.
- Objects can be represented as the sets of their attributes, which in turn are just sets of similar objects.

As a consequence, whenever we are dealing with what seems to be just one similarity structure and its corresponding lattice of nominalist kinds, many other structures will result. The lattice of kinds will be translated to other isomorphic lattices where the objects are replaced by sets of paradigms or sets of properties. *Everything that holds for objects will be translated as something that holds for bundles of paradigms or bundles of properties.* This strategy allows us to get as close as we want to the ontological pictures of spatial models without having to assume as primitive or basic any properties whatsoever. Moreover, by using the set-theoretic representations we can enrich our domain of objects in such a way that operations that were only partially defined for them are now fully defined.

Our first goal will be to construct a 'small' complete lattice of individuals out of this order. Later on we will extend the original similarity relation to this lattice. This will be our small 'calculus of individuals'. The device we will use for this is the well-known method of 'completion by cuts' or 'Dedekind-MacNeille completion'. Later we will explore how this structure is related to the lattice of kinds.

Now we will construct our lattice of individuals. Let us recall the operator $cro(A) := \{y \in S \mid \forall x \in A\ y \sim x\}$ introduced in Chapter IV. It is analogous to the extension and intension functions. This is the alternative generalization for sets of the similarity neighbourhood co[16]. It is linked to the two following fundamental operations:

Definition 68. *Let* $(S, \leq)$ *be a poset and* $A \subseteq S$*. Then the* set of upper bounds of A *is* $\Uparrow A := \{x \in S \mid \forall y \in A\ y \leq x\}$ *and the* set of lower bounds of A *is* $\Downarrow A := \{x \in S \mid \forall y \in A\ x \leq y\}$.

As it is known, the set of upper bounds and the set of lower bounds form a Galois connection [26]. Therefore, their compositions are closure operators. The set of lower-upper-closed sets, i.e. $A = \Downarrow\Uparrow A$, is the complete lattice of closed sets called the *Dedekind-MacNeille completion* or ***DM**-completion* of S. We take from [26] the following well-known result:

Theorem 12. *Let* $(S, \leq)$ *be a poset. Then the function* $\boldsymbol{DM}(x) := \downarrow x$ *is an order-embedding of* $(S, \leq)$ *in the complete lattice* $\boldsymbol{DM}(S) = (\{A \subseteq S \mid A = \Downarrow\Uparrow A\}, \subseteq)$ *called its* Dedekind-MacNeille completion *or* completion by cuts. *Moreover:*

*i **DM** preserves every existing join and meet.*

[16]Its lattice of closed sets is $\mathbf{B}(S, S, \sim)$.

*ii If S is a lattice, **DM** is a lattice embedding.*

iii If S is a complete lattice, then it is isomorphic to $\boldsymbol{DM}(S)$.

iv If S is embedded in some lattice L, then $\boldsymbol{DM}(S)$ is also embedded in L.

v If L is a complete lattice, then it is the completion of any of its join-dense and meet-dense subsets.

*vi **DM** is the concept lattice of the context $(L, L, \leq)$.*

According to these properties, the completion by cuts is the smallest complete lattice in which the poset can be embedded. In that sense, it is the optimal completion. The completion constructs new objects as sets from the original particular objects. We can think about them as some sort of 'ideal particulars'. They are added to the original poset to guarantee that the join and meet operations are well defined. They are 'quasi-individuals', so to speak.

Now we simply use this fact on the similarity poset:

Definition 69. *Let $(S, \sim)$ be a (SNI) similarity structure of order 1. Then its completion $\boldsymbol{DM}(S)$ is the Dedekind-Completion of its similarity order $(S, \leq_{co})$.*

Since the posets of two isomorphic order 1 similarities are isomorphic, their completions are isomorphic too. Of course, the correspondence between the similarities of order 1 and the **DM**-completions is not unique, since several non-isomorphic posets have the same **DM**-completion.

It is interesting to note that the construction can be carried on combining the similarity operations *int*, *cro* and *co*. So in some sense we are still on similarity grounds:

Proposition 31. *Let $(S, \sim)$ be a (SNI) similarity structure of order 1, $(S, \leq_{co})$ its poset, $x \in S$, and $A \subseteq S$. Then:*

i $co(A) = \uparrow\downarrow A$

ii $\Uparrow A = croco(A)$.

iii $\Downarrow A = intcro(A)$.

iv x *is maximal* $\Leftrightarrow croco(\{x\}) = \{x\}$.

v x *is minimal* $\Leftrightarrow intcro(\{x\}) = \{x\}$.

In any case, we have now several lattices at work. Some of them are given by these structures:

- The similarity structure $(S, \sim)$. It induces the Dedekind-MacNeille completion $\mathbf{DM}(S)$. We will call $\mathbf{DM}(S)$ the *completion of the similarity* or the *lattice of individuals*.
- The polar context $(S, SC_1(S), \in)$. It induces three lattices. The most interesting ones are its lattice of intensions $\mathbf{B}_{SC_1(S)}$, which we will call the *lattice of bundles of properties*, and its *lattice of natural kinds* $\mathbf{B}(S, SC_1(S), \in)$ which we will call simply the *nominalist lattice of natural kinds.*

- The context $(Gen(S), S, \sim)$. It also induces three lattices. The most interesting one is its lattice of extensions $\mathbf{B}_{Gen(S)}$, which we will call the *lattice of bundles of paradigms*.

Each of these structures offers a unique picture of the entities we have been discussing. Take the similarity structure $(S, \sim)$. The lattice $\mathbf{DM}(S)$ represents objects in terms of their qualitative thinness relations to other objects.

Take the context $(Gen(S), S, \sim)$. The lattice $\mathbf{B}_{Gen(S)}$ represents quasianalytically the objects as bundles of paradigms. Let $A \subseteq Gen(S)$ and $B \subseteq S$. The extension operator is $e(B) = gen(B) = \{p \in Gen(S) \mid \forall x \in B\ p \in gen(x)\}$. The intension is $i(A) = cro(A)$. Therefore, $ei(A) = \{p \in Gen(S) \mid cro(A) \subseteq co(p)\}$ and $ie(B) = \{x \in S \mid gen(B) \subseteq co(x)\}$. Here gen generalizes the corresponding quasianalytic representations for sets. In other words, each object is represented as the bundle of the paradigms to which it is similar. Two objects are similar iff their bundles overlap iff they are similar to a common paradigm.

Take the polar context $(S, SC_1(S), \in)$. The lattice $\mathbf{B}_{SC_1(S)}$ represents quasianalytically the objects as bundles of properties. The extension operator is $e(B) = \{x \in S \mid \forall T \in B\ x \in T\} = \{x \in S \mid B \subseteq q(x)\} = \bigcap B$. The intension operator is $i(A) = q(A) = \{T \in SC_1(S) \mid \forall x \in A\ x \in T\} = \bigcap q(x)$. Here q generalizes the corresponding quasianalytic representations for sets. In other words, each object is represented as the bundle of its properties. Two objects are similar iff their bundles overlap iff they have a common property.

The main result of this section is that the three structures $\mathbf{DM}(S)$, $\mathbf{B}_{SC_1(S)}$ and $\mathbf{B}_{Gen(S)}$ are isomorphic:

Theorem 13. *Let $(S, \sim)$ be a (SNI) similarity structure of order 1. Then the following lattices are isomorphic:*

1. *The Dedekind-MacNeille completion $\mathbf{DM}(S)$ of the similarity.*
2. *The lattice of intensions $\mathbf{B}_{SC_1(S)}$ of its polar context $(S, SC_1(S), \in)$.*
3. *The lattice of extensions $\mathbf{B}_{Gen(S)}$ of the context $(Gen(S), S, \sim)$.*

Proof. (1)-(3): We shall prove that the function $f\colon \mathbf{DM}(S) \to \mathbf{B}_{Gen(S)}$ defined as $f(A) := A \cap Gen(S)$ is an order isomorphism. Its inverse is $g(B) = \{x \in S \mid gen(x) \subseteq B\}$. Let us recall that $A \in \mathbf{DM}(S)$ iff $A = \Downarrow\Uparrow A$, where the order in S is induced by the similarity neighbourhood co.

First, f is well-defined: we show that $gencro(A \cap Gen(S)) = A \cap Gen(S)$ in $\mathbf{B}_{Gen(S)}$. Let $p \in gencro(A \cap Gen(S)) \subseteq Gen(S)$ and $x \in \Uparrow A$. We have $A \cap Gen(S) \subseteq A \subseteq co(A)$, therefore $x \in \Uparrow A = croco(A) \subseteq cro(A \cap Gen(S))$ and therefore $p \sim x$, so $p \leq_{co} x$, i.e. $p \in \Downarrow\Uparrow A = A$.

Second, f preserves order: if $A \subseteq B$, then $f(A) = A \cap Gen(S) \subseteq B \cap Gen(S) = f(B)$. Conversely, suppose that $f(A) = A \cap Gen(S) \subseteq B \cap Gen(S) = f(B)$. Let $x \in A$ and $y \in \Uparrow B$. If $q \in gen(x)$, then for every $z \in \Uparrow A$ we have $q \leq_{co} x \leq_{co} z$, i.e. $q \in \Downarrow\Uparrow A = A$, so $gen(x) \subseteq A$. Therefore, $gen(x) \subseteq A \cap Gen(S) \subseteq B \cap Gen(S) \subseteq B$. But then $q \leq_{co} y$, i.e. $gen(x) \subseteq gen(y)$ and we have $x \leq_{co} y$. Thus, $x \in \Downarrow\Uparrow B = B$. So $A \subseteq B$ and it follows that f is injective.

Finally, f is surjective: let $B \in \mathbf{B}_{Gen(S)}$, we show that $A := \{x \in S \mid gen(x) \subseteq B\}$ is such that $f(A) = A \cap Gen(S) = B$, which gives the inverse. First, if $gen(p) = \{p\} \subseteq B \subseteq Gen(S)$, then $p \in f(A)$. And if $p \in A \cap Gen(S)$ then $\{p\} = gen(p) \subseteq B$. Second, we show that $A \in \mathbf{DM}(S)$. Let $x \in \Downarrow\Uparrow A$.

Since $B = gencro(B)$, let $y \in cro(B)$, we show that $y \in \Uparrow A$. Let $z \in A$, then $gen(z) \subseteq B$, therefore $gen(z) \subseteq gen(y)$ since $y \in cro(B)$, i.e. $z \leq_{co} y$. But $x \in \Downarrow\Uparrow A$, so $x \leq_{co} y$, i.e. $gen(x) \subseteq gen(y)$. So if $p \in gen(x)$, $p \sim y$. A fortiori, $p \in gencro(B) = B$, i.e. $x \in A \in \mathbf{DM}(S)$.

(2)-(3): The context $(S, Gen(S), \sim)$ is dual to $(Gen(S), S, \sim)$, by the symmetry of similarity. So the lattice of extensions of $(Gen(S), S, \sim)$ is isomorphic to the lattice of intensions of $(S, Gen(S), \sim)$. Now, $(S, Gen(S), \sim)$ is isomorphic to $(S, SC_1(S), \in)$, under the context isomorphism (f, g), where $f: S \to S$ $f(x) := id(x)$ and $g: Gen(S) \to SC_1(S)$ $g(p) := co(p)$. Both f, g are bijective and $x \sim p$ iff $x \in co(p)$ iff $f(x) \in g(x)$. Therefore the lattice of extensions $B_{Gen(S)}$ of $(Gen(S), S, \sim)$ is isomorphic to the lattice of intensions $B_{SC1(S)}$ of $(S, SC_1(S), \in)$. □

So the lattice of individuals $\mathbf{DM}(S)$ is isomorphic to the dual of the lattice of natural kinds induced by its polar context. This gives us several mathematically equivalent ways to represent objects, by using sets of qualitatively ordered objects, sets of nominalist properties or sets of paradigms. A fortiori, this gives also several ways to represent kinds as well. This establishes a correspondence between different ways in which the three models, the topological model, the order-theoretic model and the similarity model, can be used to represent objects *by making use of structure that is induced by these objects themselves.* One can represent an object as the set of its paradigms, as the set of its properties or as the set of all the objects that are ordered with respect to it according to qualitative richness (or prototypicality). For example, one can think about a specific crow as the set consisting of a specific paradigm of blackness, a specific paradigm of featheredness, and so on. Equivalently, one can think about the crow as the set consisting of the property of blackness, the property of featheredness, and so on. One can also think about the crow as the set of all those objects that are qualitatively richer (or less prototypical) than the paradigms to which the crow is similar. One of the uses Carnap suggested for the quasianalysis, as a synthetic procedure, was that of 'translating' a relational description of objects to a class-based or property-based description of them. In other words, the quasianalysis was supposed to be a method of summarizing the relations that hold between objects by transforming them into entities that behave like properties. Something like this is still the case: similarities between objects and paradigms, between different objects and qualitative orderings can all of them be used to represent objects in new ways as bundles of other entities.

5.4.3 Qualitative Atomism

Recall that we left several questions unanswered. The second one was "what are the properties of this lattice?". Our nominalist lattice is not an arbitrary complete lattice. The presence of paradigms constrains its structure. Since every object can be 'qualitatively composed by' its paradigms, each kind can be obtained as the overlapping of its maximal genera. In formal terms, the lattice of bundles of paradigms is atomistic and therefore the lattice of natural kinds is coatomistic. In this section we will show that the converse holds too, in other words, that every (co)atomistic complete lattice can be obtained from a similarity structure of order 1. We already show that the correspondence holds between posets and similarities. However, the correspondence can be extended by making use of the Dedekind completion.

The notion of (co)atomism for lattices is standard [26]:

Definition 70. *Let L be a complete lattice and $p, x \in L$. Then p is an* atom *iff* $0 < p \ \& \ \forall y \in L\ (0 \leq y \leq p \Rightarrow y = p)$. *$L$ is* atomistic *iff* $x = \bigvee\{p \in L \mid p \textit{ is an atom } \& \ p \leq x\}$.

We define $Atom(S) := \{p \in L \mid p \text{ is an atom}\}$ and $atom(x) :=\downarrow x \cap Atom(S)$. The notions of *coatom* and *coatomistic* are defined dually. A lattice is atomistic iff its dual is coatomistic.

What is the relation between atomistic posets and atomistic lattices? It sounds like a plausible conjecture to say that the Dedekind completion of an atomistic poset should be an atomistic lattice. We will prove this now and in the process we will get a more abstract picture of what was shown in the section before:

Lemma 10. *Let L be an atomistic poset. Let the functions be $min\colon \boldsymbol{DM}(L) \to \wp(Min(L))$, defined as $min(A) := \{p \in Min(L) \mid \exists x \in A\ p \in min(x)\}$, and $\Downarrow\Uparrow_{Min}\colon \wp(Min(L)) \to \boldsymbol{DM}(L)$, which is the restriction of the domain of the DM-closure to the set of minimal elements. Then the pair $(min, \Downarrow\Uparrow_{Min})$ is a Galois connection.*

Proof. Let $B, B' \in \wp(Min(L))$ and $A, A' \in \mathbf{DM}(L)$. Both functions are obviously monotone. We now prove that $\Downarrow\Uparrow_{Min}(B) \subseteq A$ iff $B \subseteq min(A)$. Suppose $\Downarrow\Uparrow_{Min}(B) \subseteq A$. If $p \in B \subseteq \Downarrow\Uparrow_{Min}(B) \subseteq A$, then by reflexivity $p \in min(A)$. Conversely, suppose $B \subseteq min(A)$ and let $x \in \Downarrow\Uparrow_{Min}(B)$. If $y \in \Uparrow A$ and $q \in B \subseteq min(A)$ then there is a $z \in A$ such that $q \leq z \leq y$ and therefore $y \in \Uparrow B$. So $x \leq y$ and thus $x \in A = \Downarrow\Uparrow A$. □

A fortiori, the composition $cl_{Min} = min \Downarrow\Uparrow_{Min}\colon \wp(Min(L)) \to \wp(Min(L))$ is a closure operator on the set of minimal elements of L. In contrast, the function $\Downarrow\Uparrow_{Min} min\colon \mathbf{DM}(L) \to \mathbf{DM}(L)$ is a kernel. Let $CL(Min(L))$ be the lattice of closed sets. We can now restrict the range and domain of the previous functions (respectively), by replacing the whole power set of $Min(L)$ by the smaller lattice of closed sets. We then have that the pair $(min, \Downarrow\Uparrow_{Min})$ is an order isomorphism:

Lemma 11. *Let L be an atomistic poset. Let the functions $min\colon \boldsymbol{DM}(L) \to CL(Min(L))$, $\Downarrow\Uparrow_{Min}\colon CL(Min(L)) \to \boldsymbol{DM}(L)$ be defined as before. Then:*

i The lattice $CL(Min(L))$ is atomistic.

ii The poset L is embedded into $CL(Min(L))$ by the function $min(x)$.

iii $CL(Min(L))$ and $\boldsymbol{DM}(L)$ are isomorphic under $(min, \Downarrow\Uparrow_{Min})$.

Proof. L has a bottom element iff it only has one element, so the lattice is trivially atomistic. Let us suppose that L has no bottom. (i) Since $min \Downarrow\Uparrow \varnothing = min \Downarrow L = min\{0\} = min\varnothing = \varnothing$, the operator is normal. And since p is a minimal element, then $min \Downarrow\Uparrow \{p\} = min\{p\} = \{p\}$ the closure $T1$, thus the lattice of closed sets is atomistic (see the sections below). (ii) Since $min(x) = min(\{x\})$, by Galois $min \Downarrow\Uparrow min(x) = min(x)$ so $min(x)$ is closed.

The function is monotone and is injective by atomism. (iii) We will show that the functions are inverses, in other words, $min \Downarrow\Uparrow = id_{CL(Min(L))}$ and $\Downarrow\Uparrow min = id_{\mathbf{DM}(L)}$. The former is trivial and since $\Downarrow\Uparrow min$ is kernel, we already have $\Downarrow\Uparrow minA \subseteq A$ for $A \in \mathbf{DM}(L)$. Let $x \in A = \Downarrow\Uparrow A$ and $y \in \Uparrow min(A)$. Since $min(x) \subseteq min(A) \subseteq min(y)$ we have $x \leq y$, so $A \subseteq \Downarrow\Uparrow minA$. □

In words, what happens in the case of atomistic posets is that the Dedekind completion can be simplified by representing each element in the original poset as the set of its minimal elements. From the point of view of similarity, we are representing each object by the set of its paradigms. This result is the abstract counterpart of the one given in the previous section.

Corollary 12. *Let L be an atomistic poset. Then its Dedekind completion $\boldsymbol{DM}(L)$ is atomistic.*

Since the minimal elements are in bijective correspondence with the fixed ultrafilters, we get an equivalent representation of objects as closed sets of ultrafilters.

Since the completion $\mathbf{DM}(S)$ is isomorphic to the lattice of generators $\mathbf{B}_{Gen(S)}$, which is dually isomorphic to the lattice of kinds $\mathbf{B}(S, SC_1(S), \in)$, we can give now an answer to the second question:

Corollary 13. *Let $(S, \sim)$ be a (SNI) similarity structure of order 1. Then:*

1. *$\boldsymbol{DM}(q(S))$ is isomorphic to $\boldsymbol{B}_{SC_1(S)}$ and $\boldsymbol{DM}(gen(S))$ is isomorphic to $\boldsymbol{B}_{Gen(S)}$.*
2. *$\boldsymbol{B}_{Gen(S)}$ is an atomistic lattice, the atoms are the singletons of paradigms.*
3. *$\boldsymbol{B}(S, SC_1(S), \in)$ is a coatomistic lattice.*

In other words, the lattice of nominalist kinds is coatomistic.

The natural question to ask is whether every atomistic lattice is the completion of an (SNI) similarity structure of order 1, and if so, what kind of correspondence this is. To answer this question, we extend the similarity relation to atomistic lattices. We just have to be careful with the bottom and with the corresponding similarity circle $\{0\}$[17]:

Proposition 32. *Let $(L, \leq)$ be a complete atomistic lattice. Define in L the relation $x \sim y := \exists p \in Atom(S)\ p \leq x, y$ or $x = 0 = y$. Then $\forall x, y, w, r \in L$:*

i *$\sim$ is a similarity.*

ii $co(0) = \{0\}\ \&\ co(1) = L - \{0\}$.

iii $\forall p \in Atom(S)\ p \leq x \Leftrightarrow p \sim x$.

iv *$\sim$ is of order 1, where $Atom(S) \cup \{0\} = Gen(S)$.*

[17]Notice that if $L = \mathbf{DM}(S)$ there cannot be a similarity $\sim_{DM}$ over L that both extends the original similarity and is also compatible with the meets, i.e. that $x \sim y\ \&\ w \sim r \Rightarrow x \wedge w \sim y \wedge r$. Take as a counterexample the similarity $x \sim z \sim y$ which is of order 1 and (SNI). Therefore $x, y \leq z$. So $\mathbf{DM}(S)$ is the boolean structure $\mathbf{B}^2$. Suppose that $x \sim y$ iff $x \sim_{DM} y$. If $\sim_{DM}$ is compatible with meets, then since $x \sim_{DM} z\ \&\ z \sim_{DM} y$, we have $x = x \wedge z \sim_{DM} y \wedge z = y$, so $x \sim y$. This contradicts the assumption. Therefore $\sim_{DM}$ is not compatible with meets.

v $\sim$ *is* (SNI).

vi $x \sim y \Leftrightarrow \exists T_p \in SC_1(L)\ \{x, y, x \wedge y, x \vee y\} \subseteq T_p$.

vii $x \sim y\ \ \&\ \ w \sim r \Rightarrow x \vee w \sim y \vee r$. *[Compatible with Joins]*

viii *If* $x \neq 0 \neq y$, *then* $x \leq y \Leftrightarrow co(x) \subseteq co(y) \Leftrightarrow x \leq_{co} y$.

ix $T \neq \{0\}$ *is a similarity circle of order* $1 \Leftrightarrow T$ *is a fixed ultrafilter.*

x *If* $x \neq 0 \neq y$, *then* $x \sim y \Leftrightarrow x \wedge y \neq 0$.

Proof. Since $L - \{\ 0\}$ is an atomistic poset, (i)-(v) and (viii) follow immediately, by treating the bottom separately. Note that since 0 is isolated it is the generator of the similarity circle 1 $T_0 = \{0\}$. In particular, the similarity order is isomorphic to the order of $L - \{\ 0\}$, not to that of L. For (ix) we use the previous Lemma. For (vi) If $x \sim y$ there is a generator p such that $p \leq x \wedge y \leq x, y \leq x \vee y$, therefore $\{x, y, x \wedge y, x \vee y\} \subseteq T_p$. The converse is obvious. (vii) If $x \sim y$, $w \sim r$ then x and y have a common atom $p \leq x \vee w, y \vee r$, so $x \vee w \sim y \vee r$. (x) Let $x \sim y$, so there is an atom p such that $0 < p \leq x \wedge y \leq x, y$. Conversely, if $0 < x \wedge y$, by atomism we have an atom p such that $0 < p \leq x \wedge y \leq x, y$ so $x \sim y$. □

Property (ix) hints at what is happening. Whereas the quasianalysis *gen* represents each element as the set of its atoms (its generators), the quasianalysis q represents each element as the set of its fixed ultrafilters (its similarity circles of order 1).

Corollary 14. *Let* $(S, \sim)$ *be an* (SNI) *similarity structure of order* 1. *Then* $(S, \sim, \leq)$ *is order-embedded and similarity-embedded by* $\Downarrow\Uparrow$ *in* $(\boldsymbol{DM}(S), \sim_{DM}, \subseteq)$, *where for* $a, b \in \boldsymbol{DM}(S)$ $a \sim_{DM} b := \exists c \in Atom(\boldsymbol{DM}(S))\ c \leq a, b$ *or* $a = 0 = b$.

Proof. Since $\sim$ is of order 1, $x \sim y$ iff $\exists p \in Gen(S)\ x \sim p \sim y$ iff $p \leq_{co} x, y$ iff $\Downarrow\Uparrow p \subseteq \Downarrow\Uparrow x, \Downarrow\Uparrow y$ iff $\Downarrow\Uparrow x \sim_{DM} \Downarrow\Uparrow y$. □

Let **2** be the boolean lattice $\{0, 1\}$. This atomistic lattice cannot be the completion of a similarity structure for the following reason. Since a poset is embedded in its completion, it has a smaller cardinality than the latter one. But we have shown that the similarity poset necessarily lacks a bottom. So the cardinality of the poset, and therefore of the similarity structure, would be $n \leq 1$. If $n = 0$ it is the empty set. If $n = 1$ then the poset is the lattice $\mathbf{1} = \{0\}$ and therefore it is isomorphic to its own completion. So **2** is not the completion of a similarity poset. From the similarity point of view, the reason is that (SNI) prevents the similarity poset being a chain (although we can get two disjoint copies of **2**, just take the poset of the similarity of order 2 $p \sim x \sim y \sim q$).

As we have seen, if we stick to general (SNI) similarities of order 1 the correspondence will not be unique. Each lattice corresponds to all those similarity structures that have the same completion. But if we ask for the original similarity to already be a complete join-semilattice (thus having a dense element 1), then we can guarantee the bijective correspondence (uniqueness up to isomorphism). The reason is that adding a bottom element to this similarity makes it a complete lattice, and it is therefore isomorphic to its completion. We can go back by starting from an atomistic complete lattice, imposing a similarity and deleting the bottom element. The problematic cases are the boolean lattices **0**, **1** and **2**. Although **0** and **1** do correspond uniquely to similarities, we will treat them separately since they do no fit the general construction. Both **1** and **2** are special in the sense that if we eliminated their bottoms, then we could not recovered them by the **DM**-completion since the resulting posets (**0** and **1**) would already be complete lattices. These are the only exceptions. Remember that a complete join-semilattice is a poset where every non-empty subset has a join:

Definition 71. *Let* $(S, \sim)$ *be a (SNI) similarity structure. Then* S *is* join-similarity *iff* $(S, \leq_{co})$ *is a complete join-semilattice, i.e. every non-empty subset* $A \subseteq S$ *has a join.*

There is an important detail to consider. See that S can be a join-similarity even if it is not closed under all the joins of the order induced by the similarity neighbourhood, for example:

> $SC(S) = \{\{p, x, 1\}, \{q, y, 1\}, \{r, x, y, 1\}\}$ is a (SNI) similarity of order 1. Then S is a join-similarity, for $p, r \leq x \leq 1$ & $q, r \leq y \leq 1$ forms a complete join-semilattice. Attaching a 0 we get a lattice isomorphic to the **DM**-completion of S. Nevertheless, not all the joins induced by *co* are in S, since these correspond to unions of neighbourhoods. For instance, $p \vee_{co} q$ would be such that $co(p \vee_{co} q) = co(p) \cup co(q) = \{p, x, y, q\}$, but no such element exists in S.

In other words, although arbitrary joins exist in $\leq_{co}$, these joins need not be those corresponding to the lattice induced by the similarity neighbourhood. Note that assuming that every non-empty subset has a join is a harmless assumption. We have already shown how to get these joins by using the **DM**-completion, so the reader can think about them as logical constructs.

We will consider also the bottomed $S^* := (S^*, \sim^*, \leq^*)$ where $S^* = S \cup \{0\}$, $x \sim^* y := x \sim y$ or $x = 0 = y$ and $x \leq^* y := x \leq y$ or $x = 0$, which is isomorphic to the **DM**-completion of S[18]. It is rather obvious that:

Lemma 12. *Let* $(L, \leq) \neq$ ***2*** *be a complete atomistic lattice where* $|L| \geq 2$. *Then* ***DM***$(L - \{0\}) \cong L$.

Proof. By assumption, $L \neq$ **0, 1, 2**. We know that L–$\{0\}$ is embedded into L by the identity function, that it must have at least two distinct minimal elements and that L–$\{0\}$ has no bottom (minimals are disjoint). Let L' be a complete lattice and $g\colon (L–\{0\}) \to L'$ an embedding. So we can simply extend the embedding g to $g'\colon L \to L'$ as $g'(x) = g(x)$ iff $x \neq 0$ and $g'(x) = 0'$ iff $x = 0$. Therefore L is up to isomorphism the smallest complete lattice in which $(L–\{0\})$ is embedded, **DM**$(L - \{0\}) \cong L$. □

This Lemma would be false if L was a chain, which is what happens for instance in cases **1** and **2**.

[18] The 0 is introduced conventionally. But again, if one is worried about this, one can think about it just as the set of all those objects that are less than all those objects that are greater than all those objects that are not identical to themselves, i.e. as $\Downarrow\Uparrow(\varnothing) = \varnothing$.

Theorem 14. *Let $(S, \sim)$ be a join-similarity of order* 1 *where $|S| \geq 2$. Then $\boldsymbol{DM}(S)$ is a complete atomistic lattice distinct from $\boldsymbol{2}$ which is such that $|L| \geq 2$ and $\boldsymbol{DM}(S) - \{0\} \cong (S, \sim)$. Conversely, let $(L, \leq)$ be a complete atomistic lattice distinct from $\boldsymbol{2}$ such that $|L| \geq 2$. Then $(L, \sim^*)$ is a join-similarity of order* 1 *which is such that $\boldsymbol{DM}(L - \{0\}) \cong L$.*

Proof. Let S be a join-similarity of order 1 with at least two elements. Then $\mathbf{DM}$(S) is a complete atomistic lattice distinct from $\mathbf{2}$. We show $\mathbf{DM}(S)$–$0 \cong (S, \sim)$. By the previous corollary, $(S, \sim, \leq_{co})$ is embedded into $(\mathbf{DM}(S), \sim_{DM}, \subseteq)$ under $\Downarrow\Uparrow$, since the latter preserves similarity. Since S has at least two elements, it is neither $\mathbf{1}$ nor $\mathbf{2}$. If we add a bottom to S, $S_0 := (S \cup \{0\}, \leq^*)$ where $x \leq^* y$ iff $x \leq_{co} y$ or $x = 0$, then S_0 is a complete lattice. We extend the original similarity to S_0 as $x \sim^* y := x \sim y$ or $x = 0 = y$, so $(\mathbf{DM}(S), \sim_{DM}, \subseteq) \cong (S, \sim^*, \leq_{co})_0$ and $\mathbf{DM}(S)$–$\{0\}$ is isomorphic to $(S, \sim)$.

Conversely, let $(L, \leq)$ be a complete atomistic lattice distinct from $\mathbf{1}$ and $\mathbf{2}$ with is similarity defined as before. By the previous proposition, $(L, \sim^*)$ is a join-similarity of order 1 with a dense element and at least two elements, and if $a \neq 0 \neq b$ it follows that $a \leq b$ iff $a \leq_{co} b$. Therefore, $(L$–$\{0\}, \sim^*, \leq_{co}) \cong (L$–$\{0\}, \sim^*, \leq)$, where the identity function is the similarity and order isomorphism. By the previous lemma, $(\mathbf{DM}(L$–$\{0\}), \subseteq) \cong (L, \leq)$ since $L \neq 2$, so $(\mathbf{DM}(L$–$\{0\}), \sim^*, \subseteq) \cong (L, \sim^*, \leq)$. □

The trivial lattices $\mathbf{0}$ and $\mathbf{1}$ uniquely correspond to the similarities $S = \varnothing$ and $S = \{x\}$. As said, $\mathbf{2}$ cannot be recovered from a similarity. Therefore, every complete atomistic lattice except for $\mathbf{2}$ is the completion of the similarity poset of some similarity structure of order 1. Let us summarize the results of this section roughly as:

Corollary Every (SNI) join-similarity structure of order 1 with at least two elements uniquely corresponds to a complete atomistic lattice, and conversely.

5.4.4 Kinds of Nominalist Worlds

As was suggested in Chapter III, kinds can be ordered in many different ways and thus are not necessarily hierarchically arranged. But what other specific orderings could be? I will now make use of the previous results to obtain new species of similarity structures of order 1. In this way we will get new axioms that will give us aristocratic structures that may hold in more specific domains. This will answer our last question, which was "can we give some sort of classification of the similarity structures of order 1 in terms of its completions?". To simplify the results we will assume that our similarity poset is already a join-semilattice and that the domain has at least two elements. This allows us to greatly simplify the formulation of the following axioms:

Definition 72. *Let $(S, \sim)$ be a (SNI) join-similarity of order* 1 *and such that $|S| \geq 2$. Let $p \in Gen(S)$, $x, y \in S$ and $\varnothing \subset A \subseteq S$, then:*

i *S is an* algebraic similarity *iff $p \sim \bigvee A \Rightarrow p \sim \bigvee B$ for some finite $B \subseteq A$.*

ii *S is a* topological similarity *iff $p \sim x \vee y \Rightarrow p \sim x$ or $p \sim y$.*

iii S *is a* boolean similarity *iff* $\forall\varnothing \neq A \subseteq Gen(S) \exists x \in S\ co(x) = co(A)$.

These properties are similarity invariants. To 'classify' the different kinds of similarities we just need to look at the different properties that the closure operator *gencro* may have:

Definition 73. *Let* X *be a set and* $cl\colon \wp(X) \to \wp(X)$ *a closure operator. Then* cl *may satisfy:*

i $cl(\varnothing) = \varnothing$. *[Normal]*

ii $cl(\{x\}) = \{x\}$. *[Closed Points/T1]*

iii $cl(A) = \bigcup\{cl(B) \mid B \subseteq A\ \&\ B \text{ is finite}\}$. *[Finite]*

iv $cl(A \cup B) = cl(A) \cup cl(B)$. *[Preserves Union]*

We say that a closure operator is *algebraic* iff it is normal and finite. It is *topological* T_1 iff it is topological and T_1. Note that complete atomistic lattices of at least two elements are exactly those corresponding to normal T_1 closure operators (this is known). It can be easily checked:

Lemma 13. *Let* $(S, \sim)$ *be a* (SNI) *similarity of order* 1. *Then (i) holds and if* $|S| \geq 2$, *(ii) holds too:*

i *gencro is* T_1.

ii *gencro is normal.*

Proof. (i) Let $p \in Gen(S)$. Then if $q \in gencro(p) = genco(p)$, $co(p) \subseteq co(q)$ and so $p = q$. (ii) Let $|S| \geq 2$. Then if $\varnothing \neq gencro(\varnothing)$, there is a generator $p \in cro(\varnothing) = S$. Therefore, $S \subseteq co(p)$ and so S is a clique. By (SNI), $S = \{p\}$ and therefore $|S| \leq 1$, which contradicts the assumption. So gencro is normal. □

When S has exactly one point, $gencro(\varnothing) = gen(\{p\}) = \{p\}$ and normality fails. We will now show that the following result holds:

Theorem 15. *Let* $(S, \sim)$ *be a* (SNI) *join-closed similarity of order* 1 *and such that* $|S| \geq 2$. *Then:*

i S *is an algebraic similarity* $\Leftrightarrow$ *gencro is an algebraic closure.*

ii S *is a topological similarity* $\Leftrightarrow$ *gencro is a* T_1 *topological closure.*

Let us introduce some notions from lattice theory, we apply them to similarities:

Definition 74. *Let* $(S, \sim)$ *be a join-closed similarity,* $x \in S, A \subseteq S$ *and* $T \in SC(S)$. *Then* x *is* compact *iff* $x \leq \bigvee A \Rightarrow x \leq \bigvee B$ *for some finite* $B \subseteq A$.

In the special case of a generator p, p is compact iff $p \sim \bigvee A \Rightarrow p \sim \bigvee B$ for some finite $B \subseteq A$.

Proposition 33. *Let* $(S, \sim)$ *be a join-closed* (SNI) *similarity of order* 1, *where* $|S| \geq 2$. *The following conditions are equivalent:*

i S *is an algebraic similarity.*

ii Every generator of order 1 *is compact.*

iii The gencro closure operator is algebraic.

As it is well known, there is a correspondence between algebraic complete lattices and convex spaces given by the convex or hull operators, which are just the algebraic closures. In short, the closed sets are to be understood as convex regions in a space. In our case the points are convex regions too. So if the similarity is algebraic, we can get spaces like those advocated by the theory of conceptual spaces [42], [43]. But every finite lattice is algebraic. So every finite join-closed similarity (of two points) is algebraic. Moreover, due to the correspondence between $\mathbf{B}_{Gen(S)}$ and $\mathbf{B}_{SC_1}(S)$, we can take the points to be properties and the convex regions to be objects, represented quasianalytically as bundles of properties. If we strengthen the previous condition we can even get a space which has *geometric* structure. So under stronger conditions over a similarity structure, the generators behave as points and the rest of elements, now represented as sets of paradigms, behave as lines, planes, and so on (generally, as the 'flats' of a geometry). However, we will not deal with this case here. To sum up, S is an algebraic similarity $\Leftrightarrow \mathbf{B}_{Gen(S)}$ is the lattice of convex sets of a T_1 convexity.

But we can also go in a different direction. Let us suppose that our domain contains an infinite amount of paradigms. For complete atomistic lattices to be the lattices of closed sets of certain topological spaces, the T_1 ones, it is known that the only requirement is that the atoms should be join-prime elements[19]:

Definition 75. *Let* $(S, \sim)$ *be a join-closed similarity,* $x \in S, T \in SC(S)$*. Then* x *is a* join-prime element *iff* $x \leq y \vee z \Rightarrow x \leq y$ *or* $x \leq z$*.* T *is a* join-prime similarity circle *iff* $x \vee y \in T \Rightarrow x \in T$ *or* $y \in T$*.*

In the special case of a generator p, p is join-prime iff $x \sim y \vee z \Rightarrow x \sim y$ or $x \sim z$.

Proposition 34. *Let* $(S, \sim)$ *be a join-closed* (SNI) *similarity of order* 1*, where* $|S| \geq 2$*. The following conditions are equivalent:*

i S *is a topological similarity.*

ii Every generator of order 1 *is join-prime.*

iii Every similarity circle of order 1 *is join-prime.*

iv The gencro closure operator is topological.

Thus the extensions of the context $(Gen(S), S, \sim)$ are the closed sets of a T_1 topology whose points are the generators. Of course, this topology is only interesting when the number of generators is infinite (otherwise it is discrete). Since $\mathbf{DM}(S)$ (or equivalently, the bottomed S) is isomorphic to $\mathbf{B}_{Gen(S)}$, we have obtained a topological representation of the elements in S: each element is represented as a closed set of paradigms in a T_1 topological space. Moreover, since $\mathbf{B}_{Gen(S)}$ is isomorphic to the lattice of intensions $\mathbf{B}_{SC_1}(S)$ of $(S, SC_1(S), \in)$, this means that the ie-closure of the latter context over $SC_1(S)$ is also topological. What we have here is a T_1 topological space, homeomorphic to the

[19] This forces the lattice to be co-Heyting.

former one, where the points are properties (similarity circles of order 1) and where the closed regions are objects represented quasianalytically as bundles of properties. Thus, these are very close to Carnap's attribute spaces. To sum up, S is a topological similarity $\Leftrightarrow \mathbf{B}_{Gen(S)}$ is the lattice of closed sets of a T_1 topology.

The third example is that of the boolean algebras:

Proposition 35. *Let $(S, \sim)$ be a (SNI) similarity of order 1, where $|S| \geq 2$. Let S be a boolean similarity. Then $\boldsymbol{B}_{Gen(S)}$ is a complete atomistic boolean algebra.*

Proof. We prove that $(S, \leq_{co})$ is a bottomless complete atomistic boolean algebra. Therefore, its completion will be a boolean algebra isomorphic to $\mathbf{B}_{Gen(S)}$. Consider the power set $\wp(Gen(S) - \{\varnothing\}$, we show that the function $gen \colon S \to \wp(Gen(S) - \{\varnothing\}$ is an order isomorphism. The function is well-defined by order 1, monotonous and it is injective by (SNI). The boolean similarity axiom guarantees surjectivity. The inverse is monotonous too, since if $gen(x) \subseteq gen(y)$ then by order 1 $x \leq_{co} y$. □

Note that the similarity need not be join-closed, since the axiom is so strong that it immediately implies all the boolean structure. The atoms of the boolean algebra are the paradigms, and every other object can be obtained as a boolean join from its paradigms. In fact, every possible combination of paradigms corresponds exactly to an object[20].

Due to the correspondence between atomistic lattices and similarities of order 1, the converses of these results also hold for those structures with more than two elements. For example, since the lattice of closed sets of any T_1 topology is an atomistic lattice, it (all classical topological spaces of more than two points) can be obtained from a similarity of order 1.

The reader may wonder why such an investment of time and effort in making these abstract distinctions between kinds of similarities should be of any philosophical interest. I think that the objection is fair, so let me give a couple of examples of applications. Suppose that one wants to give a nominalist reconstruction of classical propositions. Let us suppose that the realist introduces them as entities that satisfy the laws of a complete atomistic boolean algebra. Then the nominalist can say that, necessarily, there will be some similarity structure whose lattice of kinds corresponds to such an algebra, and therefore that he will be able to give an appropriate surrogate for it (maybe by using possible worlds or token-sentences). But this will not be very informative. After all, the nominalist still owes the realist some sort of story regarding how boolean operations (which for the realist in fact generate new entities) arise from these nominalistically acceptable entities. But now the nominalist has an answer: propositions are kinds (or the extensions of these kinds) induced by some boolean similarity structure of order 1.

More specifically, suppose one wants to reduce propositions to possible worlds. One may start with a set W of possible worlds and a relation of similarity which can be pre-theoretically understood as "$w \sim w'$ iff there is an atomic proposition p that is true in both w and w'". Of course, for this condition not to be circular one has to introduce it as a primitive relation. So one considers 'atomic worlds',

[20]Note what this means from the point of view of classical propositional logic: its algebraic structure can be obtained from the similarity relation of consistency, i.e. $\phi \sim \psi$ iff $\phi \wedge \psi \neq \bot$.

which are pre-theoretically those worlds in which just one atomic proposition p is true, and one understands similarity as closeness to a common atomic world. The boolean axiom says that for any set of atomic worlds, there is exactly one world which is close exactly to the former ones. In pretheoretic terms, for each set of atomic propositions, there is exactly one world in which exactly these propositions are true. Then atomic propositions are constructed as similarity circles. Lastly, we get the rest of complex propositions as the elements of the lattice of kinds. The nominalist need not defend that the whole world is so structured, it may be that just a certain substructure of it satisfies the boolean axiom.

As a second example, consider a realist about magnitudes. We have given a nominalist reconstruction of universals, but we have not touched on issues related to the additional structure that these universals may have. Suppose that the realist asks for some nominalist story regarding the composition and order of these magnitudes. How will the nominalist answer? The answer will start by considering a piece of the world that happens to be a geometric similarity structure. The more conditions we impose on this part of the world, the closer it will resemble the real line. Of course, there is quite a long (and exciting) road from merely geometric similarities to the specific geometry we may be interested in. But the point is that such a road exists.

As a third example, suppose that the realist insists that the natural kinds are hierarchically arranged as a tree [135]. Then the nominalist can point at a corresponding tree similarity as an adequate surrogate. The axiom will require now that if two objects are similar to a common paradigm, then all the objects similar to one of them are already similar to the other. The paradigms will be the leaves of the tree. Consider for instance the following relation between species as conceived under the phylogenetic concept: "$K \sim K'$ iff K and K' have a common contemporary descendant species p", the contemporary descendant species being the leaves of the tree. Recall that each species induces a unique clade, the set of all its descendant species. Then each species is represented by the contemporary species that form a subset of this clade. An axiom for trees can be given by translating the conditions discussed in Chapter III.

Lastly, suppose that one argues that natural kinds are the convex regions of a conceptual space (e.g. [42], [43]) or closed regions in a T_1 topological space (e.g. a metric space, which coheres with Carnap's attribute spaces). If the points in this space are convex or closed too, then the nominalist can argue that these convex or closed regions are induced by an algebraic or topological similarity relation, respectively. These convex or closed regions are sets of paradigms. Or equivalently, they are the natural kinds determined by the similarities. Given the quasianalysis and the correspondence between paradigms and properties, these can be seen also as spaces whose points are properties and whose convex or closed regions are objects, represented as bundles of properties.

I would like to add a final methodological remark regarding classifications that will connect this chapter with Chapter III. If similarities of order 1 are taken to be plausible models of the structure of non-classical classifications we are quickly lead to a plurality of kinds of classifications: some classifications are partitions, others are more genral similarities of order 1, others are topologies, still others geometries, and so on. This should not be a surprise, after all, there is a plurality of kinds of orderings and a plurality of kinds of structures for measurement. Moreover, the unification power of the theory of similarity structures of order 1 is quite remarkable. For instance (except for domains with less than two entities) all atomistic trees, all T_1 topologies, all the concept lattices of polar contexts and the partition lattices can be obtained as (completions of) the orders induced by a similarity relation of order 1. I think these are highly non-trivial facts to be considered by any general theory of classifications and they show how narrow a conception of classifications as a (hierarchically ordered) family(s) of exhaustive and exclusive sets is. Nevertheless, this goes far beyond our purposes here, so I will leave this point for future discussion.

5.5 A World Full of Paradigms

There is a last question I want to consider. The model of pure similarities of order 1 makes a strong assumption, namely, that each attribute is generated by exactly one paradigmatic exemplar. This requires us to posit qualitatively thin objects that fulfill this special role of being the unique paradigmatic exemplar of a property. This has two drawbacks. First, although some reasons for positing such entities were given in the previous chapter, it is still questionable whether there are qualitatively thin entities such as these ones. Second, it forces us to assume that each attribute can be generated by a unique such paradigm, whereas for attributes in general it seems that there are several possible candidates for being their paradigms. One may wonder whether a more complex version of aristocratic resemblance nominalism could be developed according to which each attribute has several paradigmatic exemplars and each such a paradigm may exemplify several different attributes. In other words, the question is whether there is a more general class of similarities that allows us to reconstruct attributes with plenty of qualitatively enriched paradigmatic exemplars. Moreover, we would like this class of similarity structures to be as closely related as possible to our pure similarities of order 1. For instance, we might want the class of pure similarities of order 1 to be just a special case of the new class. Therefore, the last question to consider is:

> Can we generalize the aristocratic model to deal with properties having several paradigms and paradigms having several properties?

Recall that according to Brockhaus' Theorem, the similarity covering must be made up by similarity circles that can be generated by two (or less) members. [89] generalized this notion in order to classify similarity structures in terms of the number of generators each similarity circle had. The idea is that a similarity structure is of order n iff each similarity circle can be generated by n members (or less). Let us recall that:

Definition 76. *Let $(S, \sim)$ be a similarity structure. Then $T \in SC(S)$ is a* similarity circle of order n *iff $\exists x_1, \dots, x_n \in T$ such that $co(x_1) \cap \cdots \cap co(x_n) = T$. Here the x_i are the* generators *of T.*

One can consider $x_1, \dots, x_n$ jointly as paradigms of the property T. Now what we need to do is to generalize this notion to similarity circles that might be generated by an arbitrary number of elements. To do that it is convenient to lift the notion of generator from single elements to sets. The idea is that the generator of a circle is a subset, the set of all the generating elements *taken collectively.* The choice I propose is this:

Definition 77. *Let $(S, \sim)$ be a similarity structure, $x, y \in S$ and $G \subseteq S$ a clique. Then the following conditions are equivalent:*

i *G is a* generator.

ii *$cro(G)$ is a clique.*

iii *G generates a similarity circle, i.e. $cro(G) \in SC(S)$.*

iv *$cro(G) = crocro(G)$.*

v *$(\forall p \in G\ x \sim p \sim y) \Rightarrow x \sim y$.*

vi *$cro(G) = \bigcap\{co(p) \mid p \in G\} \in SC(S)$.*

Proof. Assume (i)-(ii) as a definition, (v) and (vi) are just reformulations of (ii) and (iii), respectively. (ii)-(iii) Let G be a generator, so $cro(G)$ is a clique. Since G is also clique, $G \subseteq cro(G)$. If $z \in crocro(G) \subseteq cro(G)$, so $cro(G)$ is a similarity circle. (iii)-(iv) Recall that cro(G) is a similarity circle iff crocro(G)=cro(G). □

When $|G| \leq n$ we get the original notion of generator. Thus a clique G generates a circle T iff $cro(G) = T$. The operator cro moves us from the generators to their circles. In the case of similarities of order 1 to avoid clutterness we can simply replace the generator $\{p\}$ by the element p itself. This allows us to avoid an unnecessary higher layer in the set-theoretic hierarchy when representing an element by the set of its generators.

Now, if G generates a similarity circle T and $G \subseteq H \subseteq T$ then H also generates T. In fact, every similarity circle trivially generates itself. Similarity circles always have at least one generator (themselves), and usually they have many redundant generators. We are interested in those generators which are such that no smaller set generates the same circle:

Definition 78. *Let $(S, \sim)$ be a similarity structure and $G \subseteq S$ a generator. Then G is a* minimal generator *iff if $H \subseteq G$ and $cro(H)$ is a clique, then $H = G$.*

This hints at the usefulness of the generalization made. We need to conceive of generators as sets not just to consider arbitrary cardinalities, but to make possible simpler comparisons among different minimal sets of generating elements of the same circle.

Lemma 14. *Let $(S, \sim)$ a similarity structure, $G \subseteq S$ a set of generators and $T = cro(G)$ the circle it generates. Then:*

i *If T' is a similarity circle and $G \subseteq T'$, then $T' = T$.*

ii *If $H \subseteq G$ and H is a set of generators, then $cro(H) = T$.*

Proof. (i) If T' is a circle and $G \subseteq T'$, then $cro(T') = T' \subseteq cro(G) = T$ and therefore $T = T'$. (ii) If $H \subseteq G$ and H is a generator, then $cro(H)$ is a clique and $cro(G) = T \subseteq cro(H)$, i.e. $T = cro(H)$. □

(i) says that generators are exclusive to the similarity circles they generate. Now, the same similarity circle may have several different minimal generators. The last move is to consider only those similarity circle that have a smallest generator:

Definition 79. *Let* $(S, \sim)$ *be a similarity structure and* $T \in SC(S)$*. Then* T *is a* simple similarity circle *iff* T *has a unique minimal generator.*

In other words, the simple circles are the ones in bijective correspondence with their smallest generators. Each simple similarity circle has a unique small generator. A similarity covering is simple iff all its members are simple similarity circles.

Examples are:

- The similarity $SC(S) = \{\{p, q, x\}, \{x, r\}\}$ is of order 1 but $\{p, q, x\}$ is not simple. It can be generated either by $\{p\}$ or by $\{q\}$.
- The similarity $SC(S) = \{\{p, x\}, \{q, y\}, \{r, z\}, \{x, y, z\}\}$ is of order 2 but $\{x, y, z\}$ is not simple. It can be generated by $\{x, y\}$, $\{y, z\}$ or $\{x, z\}$.

By taking into account these two features, the class of similarity structures we should take a look at is the following one:

Definition 80. *Let* $(S, \sim)$ *be a similarity structure. Then* S *is* simple *iff* S *has a similarity covering by simple similarity circles.*

We define the *set of simple similarity circles* as $Simple(S) := \{T \in SC(S) \mid T \text{ is simple}\}$ and the *set of smallest generators* as $SGen(S) := \{G \subseteq S \mid G \text{ is the smallest generator of a simple circle}\}$. Every pure similarity of order 1 is simple, because its similarity is covered by $SC_1(S)$. See that if a similarity is simple then $Simple(S)$ also covers it. Of course, if all similarity circles are simple, then the similarity is simple.

Examples:

- The classical example in [89] of a similarity $SC(S) = \{\{p, x, y\}, \{x, z, q\}, \{r, y, w\}, \{w, z, s\}, \{x, y, z\}, \{y, z, w\}\}$ of order 3 has two minimal similarity coverings by simple circles, one of them includes $\{x, y, z\}$ and the other one includes $\{y, z, w\}$.
- The similarity $SC(S) = \{\{p, x\}, \{q, y\}, \{r, z\}, \{x, y, z\}\}$ is of order 2 but it has no similarity covering by simple similarity circles, since $\{x, y, z\}$ is not simple. It can be generated by $\{x, y\}$, $\{y, z\}$ or $\{x, z\}$.
- The (SNI) similarity of order 1 $SC(S) = \{\{p, x, y\}, \{y, z, q\}, \{x, z, r\}, \{x, y, z\}\}$ is such that all circles are simple, since $\{x, y, z\}$ is simple (its generator is itself). However, the covering by all the simple similarity circles is different from $SC_1(S)$, because $\{x, y, z\}$ is of order 3.
- Let the simple similarity of order 1 be $SC(S) = \{\{a, b, p\}, \{a, c, q\}, \{a, d, r\}, \{b, c, w\}, \{b, d, m\}, \{c, d, n\}, \{a, b, c, d\}\}$. Then $\{a, b, c, d\}$ is a non-simple circle of order 3, for instance it can be generated by $\{a, b, c\}$ or $\{b, c, d\}$. Here $Gen(S) = \{p, q, r, w, m, n\}$. Nevertheless, the family $SC_1(S) = SC(S) - \{\{a, b, c, d\}\}$ is a similarity covering by simple similarity circles.

The first example shows that a simple similarity can be covered by two minimal but distinct simple similarity coverings. The second example shows that some similarities of order 2 are not simple. The third example shows that in a similarity of order n, not all the simple similarity circles may be of order

n, they can be of higher order. The last example shows that a simple similarity can be such that not every similarity circle is simple. Of course, although a similarity structure may have several minimal simple coverings there is only one which is the biggest one, namely $Simple(S)$.

From now on we focus on simple similarity structures. Recall that $Simple(S)$ and $SGen(S)$ are the family of simple similarity circles and the family of smallest generators of simple circles, respectively. We now have two very similar weak quasianalysis:

$$gen\colon S \to \wp(\wp(S)) \;\; gen(x) := \{G \in SGen(S) \mid x \in cro(G)\}$$
$$q\colon S \to \wp(\wp(S)) \;\; q(x) := \{T \in Simple(S) \mid x \in T\}$$

Whereas q is standard, gen is not. These generalize to sets:

$$gen\colon \wp(S) \to \wp(\wp(S)) \;\; gen(A) := \{G \in SGen(S) \mid \forall x \in A \; x \in cro(G)\}$$
$$q\colon \wp(S) \to \wp(\wp(S)) \;\; q(A) := \{T \in Simple(S) \mid \forall x \in A \; x \in T\}$$

Although here we will always choose $Simple(S)$, all the results hold if we replace this family by a possibly smaller family $Q^* \subseteq Simple(S)$ that still covers the similarity (which must exist, by assumption). There may be several such families, although in some contexts a natural one suggests itself. For example, for pure similarities of order 1 one takes $Q^* = SC_1(S) \subseteq Simple(S)$. In such cases we just replace accordingly the family $SGen(S)$ by the corresponding family $SGen(S)^*$ of generators of the circles in Q^*. Without further constrains there is in general no unique weak quasianalysis of simple similarity circles (unless one requires it to contain *all the simple circles*).

Let us take a closer look now at the lattices induced by the contexts of simple similarities. We have at least the following two:

i The context $(SGen(S), S, I)$, where GIx iff $x \in cro(G)$, of bundles of paradigms. Consider its lattice of extensions $\mathbf{B}_{SGen(S)}$, which represents each object as a bundle of generators. Let $A \subseteq SGen(S)$ and $B \subseteq S$:

   (a) $e(B) = \{G \in SGen(S) \mid \forall x \in B \; x \in cro(G)\} = gen(B)$.

   (b) $i(A) = \{x \in S \mid \forall G \in A \; x \in cro(G)\} = \bigcap\{cro(G) \mid G \in A\} = cro(\bigcup A)$.

ii The context $(S, Simple(S), \in)$ of bundles of properties. Consider its lattice of intensions $\mathbf{B}_{Simple(S)}$, which represents each object as a bundle of properties. Let $A \subseteq S$ and $B \subseteq Simple(S)$:

   (a) $e(B) = \{x \in S \mid \forall T \in B \; x \in T\} = \bigcap B$.

   (b) $i(A) = \{T \in Simple(S) \mid \forall x \in A \; x \in T\} = q(A)$.

Furthermore, if the simple similarity is (SNI) then $\leq_{co}$ is a partial order and we can consider its Dedekind-MacNeille Completion too:

$\mathbf{DM}(S)$ is the Dedekind-MacNeille completion of $\leq_{co}$.

We know that for (SNI) similarities of order 1, $\mathbf{DM}(S)$, $\mathbf{B}_{SGen(S)}$ and $\mathbf{B}_{Simple(S)}$ are isomorphic. This does not hold in general though, as the next example shows.

Let $SC(S) = \{\{p, x\}, \{q, x, w\}, \{r, y, w\}, \{x, y, w\}\}$. This similarity is (SNI), simple and of order 2 since $\{x, y, w\}$ is of order 2 and its only generator is $\{x, y\}$. The rest of circles are of order 1, where $Gen(S) = \{p, q, r\}$. The order induced by the similarity neighbourhood is $p, q \leq_{co} x \;\;\&\;\; q \leq_{co} w \;\;\&\;\; r \leq_{co} y \leq_{co} w$. Its Dedekind completion $\mathbf{DM}(S)$ simply adds a top 1 and a bottom 0. This lattice is not atomistic. For instance, y corresponds to a join-irreducible element. But the lattice of extensions $\mathbf{B}_{SGen(S)}$ is not isomorphic to it. It contains an element that does not correspond to any element in $\mathbf{DM}(S)$. This element emerges as a representation of the generator $\{x, y\}$ of the circle $\{x, y, w\}$, by overlapping the bundles of x, y and w. So in $\mathbf{B}_{SGen(S)}$ this element is the meet of the images of x and y. However, in the $\mathbf{DM}(S)$ the meet of the images of x and y is the bottom 0.

For similarities of order higher than 1, some information encoded by the similarity gets lost when we go to the similarity order. This is easily seen to be the case for x and y if we modify slightly the preceding example. Take the similarity of order 2 $SC(S) = \{\{p, x\}, \{s, y\}, \{q, x, w\}, \{r, y, w\}, \{x, y, w\}\}$. Nothing in the $\leq_{co}$ order tells us that x and y are similar. A fortiori, the same order is induced by the different similarity of order 1 $SC'(S) = \{\{p, x\}, \{s, y\}, \{q, x, w\}, \{r, y, w\}\}$, where x and y are not similar to each other. Once we are in $\leq_{co}$, we cannot know whether the original similarity was that of $SC(S)$ or that of $SC'(S)$. We lost the relevant information.

The previous example is also interesting because it already hints at what happens in the concept lattices of simple similarities: although for similarity circles of order higher than 1 there is no element whose bundle of generators just includes the generator of that circle (because all members of such a circle have more than one property), such an 'ideal' element is produced by the lattice of extensions.

Nevertheless, due to the bijective correspondence between simple similarity circles and their generators the other two lattices are dually isomorphic:

Proposition 36. *Let* $(S, \sim)$ *be simple. Take the contexts* $(SGen(S), S, I)$ *and* $(S, Simple(S), \in)$*. Then* $\boldsymbol{B}_{SGen(S)}$ *is isomorphic to* $\boldsymbol{B}_{Simple(S)}$*.*

Proof. $(SGen(S), S, I)$ is dual to $(S, SGen(S), I)$, and the latter one is isomorphic to $(S, Simple(S), \in)$ under the function $cro\colon SGen(S) \to Simple(S)$, for if G is a smallest generator then $cro(G)$ is a simple circle. If $cro(G) = cro(G')$, given that the similarity circle has a unique smallest generator, injectivity follows $G = G'$. Since every simple circle has at least a smallest generator, the function is surjective too. Let us prove that the pair of functions (id, cro) where $id\colon S \to S$ and $cro\colon SGen(S) \to Simple(S)$ is an isomorphism of contexts. If $x \in S$ and $G \in SGen(S)$ then xIG iff $x \in cro(G)$. Therefore, the lattice of extensions of $(SGen(S), S, I)$ is isomorphic to the lattice of intensions of $(S, Simple(S), \in)$. $\square$

Now, the interesting thing is that just as in the case of similarities of order 1, $\mathbf{B}_{SGen(S)}$ is a complete atomistic lattice. Recall that the blob is the trivial similarity $\sim = S \times S$ (i.e. S is a similarity circle):

Proposition 37. *Let $(S, \sim)$ be simple, distinct from the blob and $(SGen(S), S, I)$. Then $\boldsymbol{B}_{SGen(S)}$ is a complete atomistic lattice where $Atom(\boldsymbol{B}_{SGen(S)}) = \{\{G\} \mid G \in SGen(S)\}$.*

Proof. Since ei is a closure operator on $(\wp(SGen(S)), \subseteq)$, in order to prove that the lattice is atomistic we only need to show that the closure is T_1 and normal. Since $i(\varnothing) = S$ vacuously, $ei(\varnothing) = \varnothing$ iff $e(S) = \varnothing$ iff there is no $G \in SGen(S)$ such that $S \subseteq cro(G)$ iff $\sim$ is not the blob, which holds by assumption. Let $G, G' \in SGen(S)$, then since $i(\{G\}) = i(G) = cro(G)$, we have $G' \in ei(G)$ iff $\forall x \in i(G)\, x \in cro(G')$ iff $cro(G) \subseteq cro(G')$ iff $T_G = T_{G'}$ iff $G = G'$, therefore $ei(G) = \{G\}$. □

As a corollary, if the similarity is pure of order 1 and distinct from the blob the concept-lattices are atomistic even if the similarity is not (SNI) (i.e. purity is enough). However, to get the correspondence with the DM-completion we still need (SNI) (if it is order 1 (SNI) then it is distinct from the blob iff it has at least two distinct elements).

What happens was suggested earlier: since $\mathbf{B}_{SGen(S)}$ is closed under arbitrary intersections, for each simple circle T of order $n \geq 2$, it contains the singleton whose only member is the generator G of T. In other words, the closure fills the gaps by adding to each simple similarity circle T of order $n \geq 2$ an ideal element that behaves as a generator of order 1 for that circle. That element is the one having as bundle of paradigms $\{G\}$, where G is the generator of T. A fortiori, $\mathbf{B}_{SGen(S)}$ behaves as if it had been induced by a similarity of order 1. This has a negative and a positive side. From the negative side, simple similarities do not move us beyond similarities of order 1. Once we get to their lattices it is as if we had just come back to similarities of order 1, thus forgetting the complexity carried by the different orders of the generators. To put it differently, the lattice-theoretical representations collapse simple similarities to order 1 similarities and therefore they will give us no new information. But it has a positive side too: we can use similarities of order 1 to understand what simple similarities look like.

Since $\mathbf{B}_{SGen(S)}$ is complete atomistic, it is also a (SNI) similarity of order 1. Thus, we can 'complete' our original simple similarity by transforming it into the order 1 similarity $(\mathbf{B}_{SGen(S)}, \sim^*)$. The completion procedure inserts in each simple circle of order $n \geq 2$ a new ideal element that behaves as an order 1 generator for the circle (of course, it adds many other elements too). An analogous completion works by using $\mathbf{B}_{Simple(S)}$.

Proposition 38. *Let $(S, \sim)$ be a (SNI) simple non-blob similarity. Take its* simple 1-completion *to be the similarity $(\boldsymbol{B}_{SGen(S)}, \sim^*)$ where $A \sim^* B := (A \cap B \neq \varnothing$ or $A = \varnothing = B)$, with the corresponding function $gen(x)\colon S \to \boldsymbol{B}_{SGen(S)}$. Then:*

i $\boldsymbol{B}_{SGen(S)}$ *is a (SNI) similarity structure of order 1 and for $A, B \neq 0$, $A \subseteq B \Leftrightarrow co^*(A) \subseteq co^*(B)$.*

ii gen is a similarity-embedding.

iii gen is a $\subseteq$-order-embedding and a $\leq_{co^*}$-order embedding.

iv gen preserves every existing (so non-empty) $\leq_{co}$-join in S, $gen(\bigvee A) = \bigvee_{co^*} gen[A]$.

Proof. (i) Since S is not the blob, it has at least two distinct elements. Therefore, $\mathbf{B}_{SGen(S)}$ is a complete atomistic lattice and then the similarity $A \sim B := A \cap B \neq \varnothing$ iff there is a $G \in \mathbf{B}_{SGen(S)} G \in A \cap B$ iff there is a $G \in \mathbf{B}_{SGen(S)} \{G\} \subseteq A \cap B$, for $A \neq \varnothing \neq B$, is (SNI) of order 1. We already proved that for $a, b \neq 0$, $a \leq b$ iff $co(a) \subseteq co(b)$. (ii) Since gen is a weak quasianalysis, it is a faithful similarity homomorphism. If $gen(x) = gen(y)$ then $z \sim x$ iff $\exists T \in Simple(S)\ x, z \in T$ iff $\exists G \in \mathbf{B}_{SGen(S)}\ x, z \in cro(G)$ iff $\exists G \in \mathbf{B}_{SGen(S)} y, z \in cro(G)$ iff $\exists T \in Simple(S) x, y \in T$ iff $y \sim z$, by (SNI) we have $x = y$. Therefore, gen is injective. (iii) By (SNI) $\leq_{co}$ is a partial order. If G is minimal generator, then $\{G\} = gen(T)$ for some simple circle T, for if G' is a minimal generator in $gen(T)$ it follows that $T \subseteq cro(G') = T'$, and therefore $G' = G$ and the function is well-defined. If $x \leq y$ and $x \in cro(G)$ then $crocro(G) = cro(G) \subseteq cro(x) = co(x) \subseteq co(y)$. Therefore, $y \in cro(G)$ and so $gen(x) \subseteq gen(y)$. And if $gen(x) \subseteq gen(y)$, if $z \sim x$ then there is a simple circle $x, z \in T$ and so T has a generator $G_T \in gen(x) \subseteq gen(y)$. It follows that $y \in cro(G_T) = T$ and $y \sim z$. Since gen is injective, it is an order-embedding. Since $gen(x) \subseteq gen(y)$ iff $co^*(gen(x)) \subseteq co^*(gen(y))$ iff $gen(x) \leq_{co^*} gen(y)$, we are done. (iv) Let $A \subseteq S$ suppose that $\bigvee A$ exists in S, where the join is that of $\leq_{co}$. Since the join exists and S is not the blob we have $A \neq \varnothing$, if not we would have $\bigvee A = 0$, which is impossible, so $gen(\bigvee A) \neq \varnothing$. We prove that $gen(\bigvee A) = \bigvee_{co^*} gen[A]$ for $gen[A] = \{gen(x) \mid x \in A\}$. If $gen(x) \in gen[A]$, then $x \in A$ and since $x \leq_{co} \bigvee A$, we have $gen(x) \leq_{co^*} gen(\bigvee A)$ by monotonicity. Suppose there is a $X \in \mathbf{B}_{SGen(S)}$ such that $gen(x) \leq_{co^*} X \leq_{co^*} gen(\bigvee A)$ for every $x \in A$. Since X is an extension, there is a $Y \subseteq S$ such that $X = \bigcap\{gen(y) \mid y \in Y\}$. Since $gen(x) \leq_{co^*} X$ we have $gen(x) \subseteq X$ therefore $X \neq \varnothing$, and finally $Y \neq \varnothing$. It follows that $\forall x \in A\ gen(x) \leq_{co^*} \bigcap\{gen(y) \mid y \in Y\} = X$ iff $\forall x \in A\ gen(x) \subseteq \bigcap\{gen(y) \mid y \in Y\} = X$ iff $\forall x \in A\ \forall y \in Y\ gen(x) \subseteq gen(y)$ iff $\forall x \in A\ \forall y \in Y\ x \leq_{co} y$ iff $\forall y \in A\ \bigvee A \leq_{co} y$ iff $\forall y \in A\ gen(\bigvee A) \subseteq gen(y)$ iff $gen(\bigvee A) \subseteq X$ iff $gen(\bigvee A) \leq_{co^*} X$. Therefore $gen(\bigvee A) = \bigvee_{co^*} gen[A]$. □

We must be careful, in this specific context we denote by $gen[A] = \{gen(x) \mid x \in A\}$, which is distinct from $gen(A) = \{G \in SGen(S) \mid \forall x \in A\ x \in cro(G)\}$. Although gen still preserves the similarity and the preorder in both directions, we need (SNI) for the preorder to be partial and gen to be injective. Thus, for most purposes the completion procedure will require the simple similarity to be (SNI). Since S is order-embedded in the lattice of extensions, it follows by the fundamental properties of the DM-completion that the latter one is so embedded too:

Corollary 15. *Let $(S, \sim)$ be a simple (SNI) similarity. Then $\boldsymbol{DM}(S)$ is order-embedded in $\boldsymbol{B}_{SGen(S)}$.*

In the special case of (SNI) similarities of order 1 the converse holds too, gen being surjective. In other words (modulo the 1-completion), the theory of simple similarity structures is that of (SNI) similarity structures of order 1.

Recall what our original aim was. We wanted to generalize the class of pure similarities of order 1 to get a model according to which each property is generated by a subset of its members. This gives us a generalized aristocratic resemblance nominalism (a generalized polar model) where each property has several paradigms. Take any simple similarity circle T and its generator G. The cardinality of any such G can be as wished, so long as the resulting similarity is still simple. How well does G correspond to our notion of a set of paradigms for T? First, G is a clique, so all the objects in G are similar to each other (as should be, all of them share property T). However, each paradigm in G may have more than one property, in contrast to the paradigms of similarity circles of order 1. Our paradigms can be qualitatively enriched. Moreover, each member in G may share with other members in G properties distinct from T. Even more interestingly, the same object can be a paradigm for different properties. The only requirement is that *taken collectively* all of the members of G are paradigmatic of T. A fortiori, there is no other property T' different from T which is such that all members of G are jointly instances of T'. The only property common to all the members in G is T.

How does this relate to Goodman's problems? In simple similarities there is no unique cover by simple circles. This guarantees that the approach generalizes directly the case of pure similarities of order 1. In exchange, we lose the unique correspondence with a class of realist contexts. However, we could select, for each simple similarity, the class of all simple similarity circles. This will carve a class of the contexts which correspond to the collections of all the simple circles of a simple similarity. I have not described such family of contexts, but one can guess what it would look like. Each property has a smallest generating set of paradigms, but each paradigm can have several properties and even be a paradigm of these properties. Such set of paradigms need not be finite. From a realist point of view, such contexts do not pose the problems that pure order 1 similarities have. Moreover, this does not sound so implausible. Is not there, for each property R, a smallest set of exemplars of R that only have the property R in common? This is then the question that is left open for future work: could a world full of paradigms be used as a reply to Goodman's objections?

5.6 Conclusion of Chapter V

At the beginning of this chapter, the formal side of the problem of universals has been formulated as the task of finding a suitable mapping from a surrogate structure of properties (constructed from a resemblance nominalist structure) into a realist structure of universals. Then Goodman's problems of coextensionality, imperfect community and companionship have been recasted as objections to the existence of such a mapping. After reviewing the basic results on quasi-analysis, the aristocratic models were proposed as a solution to these problems. The three models were shown to be equivalent. One starts from a similarity structure which contains some paradigmatic objects, and constructs the surrogates for universal attributes. From objects and attributes, we get the surrogate for the lattice of natural kinds. Natural kinds are reconstructed as pairs of extensions and intensions, where the extensions are certain collections were each pair of objects are similar to each other and where the intensions are the quasi-analytic representations of particular objects (i.e. bundles of objects). We can sum up the results as follows:

1. By the first main theorem, the models of aristocratic resemblance nominalism in terms of polar distributions (and topologies), polar contexts and pure similarity structures of order 1 turn out to be equivalent. This correspondence extends to atomistic posets by adding the indiscernibility constraint. Moreover:

    (a) That several intuitively correct models of the same phenomenon turn out to be mathematically equivalent supports the claim that the original model is materially adequate.
    (b) The equivalence between similarity structures and polar contexts gives a way to surrogate a big class of reductivist realist models, as introduced in Chapter III. This reduction satisfies the constraints imposed in the beginning of the chapter, namely, it is a structure-preserving map (iso) which is such that the basic entities of the realist model are reconstructed from the nominalist one.
    (c) Given the correspondence with polar topologies, we get for free a spatial model where both attributes and extensions of kinds turn out to be topologically well-behaved regions (namely, closed sets).

2. By the second main theorem, the resulting nominalist lattice of natural kinds turns out to be coatomistic, the converse (that every complete coatomistic lattice can be obtained from a similarity structure) being true too. Moreover:
    (a) There are two other lattices that are isomorphic to the (dual of the) lattice of kinds, namely the lattice of bundles of paradigms and the lattice of bundles of properties. These lattices are given by the two extensionally equivalent quasianalysis that are defined over the similarity relation. Thus, each object can be quasianalytically represented either as its (nominalist) infimae species, or as a bundle of its (extensional) attributes or as a bundle of its paradigms.

3. By the third main theorem, different properties of the lattice of kinds correspond to different axioms that similarities of order 1 may satisfy. This allows us to construct possible worlds where more specific conditions hold and therefore where different kinds of entities can be reconstructed by nominalist means. Moreover:
    (a) The conceptual spaces approach, according to which natural properties are convex regions in a space, can be considered as a special case.
    (b) Boolean approaches, according to which properties are abundant, can be considered as a special case.
    (c) Atomistic trees, as used in models for classification, can be considered as a special case.

4. By the last main theorem, some of the notions used in this chapter can be generalized to a model where each property is generated by a set of paradigmatic entities and each paradigmatic entity may have several properties. This provides the nominalist with a general aristocratic model.

These results hint at the picture of natural kinds one gets by assuming the theses of the aristocratic resemblance nominalist. Objects are fundamentally similar or dissimilar to each other. Some of these objects, the paradigms, behave as bridges that make similar any two objects that are similar to them. The collections of all the objects similar to a given paradigm work as attributes. The nominalist can then mimic the Minimal Conception by taking these collections to form the intensions of kinds. The resulting order structure satisfies some sort of qualitative co-atomism: every kind is the result of the overlapping of its maximal genera, which are the kinds induced by the attributes, which are those collections of objects gathered around a given paradigmatic object.

Chapter 6

Conclusion

The present study is an attempt
to apply the theory of relations to
the task of analyzing reality.

The Logical Structure of the World
RUDOLF CARNAP

6.1 What was Done

Let us recall what the aim of this thesis was. In Chapter II, we introduced the topic of kinds and considered some of the reasons that philosophers give to posit natural kinds. We also presented one of the main theories of kinds, namely scientific essentialism, cluster theories and conceptualism. Several objections to these positions were reviewed. The main diagnosis was that the discussions prompted by these theories (except possibly for conceptualism) were too general. They have not lead us to specific principles about the structure of kinds that may help us improve our understanding about them.

In order to better understand what kinds are supposed to be, I suggested a different strategy. Instead of defending a specific theory of kinds, I proposed looking at two structural features of kinds by using formal models of kinds:

i The *external structure* of natural kinds, that is to say, the different ways in which kinds are ordered by species-genus specificity relations.

ii The *internal structure* of natural kinds, that is to say, the fact that the members of a given kind have several common properties or are sufficiently similar to each other.

The background idea was that most contemporary theories of kinds assume at least that kinds have as members some objects that belong to the kind because they share some natural properties (or are in similarity relations to each other) and that these kinds are ordered by specificity relations. By looking at these two aspects, I conjectured that we could find more specific principles about the behaviour of kinds that we could focus the discussion on. However, in order to get at specific principles, some minimal assumptions had to be made. These were introduced in Chapter III.

About the external structure of kinds, a principle discussed in the literature was considered, namely the hierarchy constraint. According to this principle, kinds are ordered in a tree-like fashion. It was argued by appealing to current discussions in the literature that the principle is too strong, and that kinds can be ordered in many other different ways. I looked for weaker conditions. According to Thomason, kinds form an ordered structure called a 'complete lattice' that is not necessarily hierarchical. The problem with Thomason's approach was that it gave no further reasons for why this should be a plausible model for kinds. According to Martin's version of Corcoran's standard syllogistic logic, if kinds are to satisfy at least the basic relations described by syllogistic propositions, kinds must form a lattice too. Given that kinds satisfy these relations, this suggested the hypothesis that the external structure of kinds is at least that of a (complete) lattice. In other words, that kinds are ordered by specificity relations in such a way that the overlapping of kinds is a kind and that there is a summum genus that includes all the other kinds as species.

Nevertheless, the models just mentioned were not very informative regarding the way the external structure of kinds is linked to the internal structure. In other words, they do not say anything regarding how the species-genus relations depend on the fact that kinds have certain objects as members that share certain properties. For the purpose of explaining this relation, I introduced a Minimal Conception of Kinds:

Minimal Conception of Kinds Every kind has as members some objects (the extension) sharing certain sparse attributes (the intension). More strongly, all the objects share all these attributes, and these attributes are all those sparse attributes shared by these objects.

The Minimal Conception is somewhat stronger than what some theories of kinds are committed to, but it is still minimal since it does not mention other philosophically loaded concepts such as that of essences, causality or natural laws. However such a conception appeals to sparse attributes without saying what these are. I wondered whether a nominalist explanation for these attributes could be given. Accordingly, this fixed the more specific aim of this thesis, which was:

Aim of the Thesis To develop a formal model for kinds that satisfies the Minimal Conception and is based on the ontological assumptions of resemblance nominalism.

A fortiori, the project was divided into two different tasks. The first one was to provide a model for the Minimal Conception and use it to get more information regarding the external structure of kinds. This was done in Chapter III. The second one was to reconstruct such a model by starting from purely nominalist assumptions. This was done in Chapters IV and V. Overall, this resulted into the three main results of the thesis:

i I selected a class of mathematical structures, that of concept lattices, to provide a model for the Minimal Conception of Kinds. Such a model is a *realist model*.

ii I selected a class of mathematical structures to provide a model for a specific conception of resemblance nominalism, namely aristocratic resemblance nominalism. Such a model is a *nominalist model*.

iii I showed which subclasses of realist models can be reconstructed from the nominalist ones. In other words, the models used by the realist were 'translated' to models used by the nominalist.

Concerning the model for the Minimal Conception, I chose the theory of concept lattices. A concept lattice is the lattice induced by a context, which is a structure consisting in some objects that have some attributes. The elements of this lattice are pairs of sets, which were used to represent kinds. One of the sets, the extension of the kind, includes all and only those objects that have the attributes in the other set. The other set, the intension of the kind, includes all and only those attributes had by the objects in the extension. Thus, the representation of a kind in the model follows the Minimal Conception. Since concept lattices are complete lattices, the model is a special case of Thomason's and Martin's, and therefore it gives a semantics for syllogistic logic. Moreover, the connection between objects, attributes and the species-genus relation reveals a classical principle of logic, namely Kant's Law of the duality between extension and intension. According to it, the extension of a kind is inversely related to its intension. This was shown to hold in the concept-lattice model of kinds. In contrast, the hierarchy condition was shown to be just a special case of the model.

As an application of the model, I showed that two new operations of specific difference are defined in any such concept lattice. The properties of these operations were studied. They allow for definitions of kinds in terms of genera and specific differences. Moreover, each specific difference was shown to induce a non-classical internal negation of kinds, whose properties were also studied. A comparison between the picture given by the model and the classical Aristotelian one was made. Overall, these results gave more insight into the species-genus specificity relations according to which kinds are ordered and provided a realist model of kinds to be reconstructed by nominalist means in the remaining chapters.

Concerning the nominalist approach to the internal structure of kinds, the nominalist reconstruction worked in two steps (developed in Chapters IV and V, respectively):

1. First, the natural attributes were reconstructed as collections of objects similar to paradigmatic objects.

2. Second, the natural kinds were reconstructed as pairs consisting of a set of objects and a set of attributes (where these attributes were the ones obtained in the first step).

In Chapter IV, the first step was discussed. In order to reconstruct the attributes, a version of resemblance nominalism was selected and an adequate model for it was provided. Since similarity relations have a bad philosophical reputation, I answered to some of the objections that have been raised against them. In particular, Tversky's criticism of symmetry was considered at length. Then the three main versions of resemblance nominalism, namely egalitarian, aristocratic and collectivist resemblance nominalisms were reviewed. The shortcomings of the standard egalitarian approach by Pereyra were discussed in detail. In contrast, an aristocratic approach was selected. According to aristocratic resemblance nominalism, there is a distinction between paradigmatic and non-paradigmatic objects. The former are the ground for the resemblance relations that hold among all objects and thus structure the corresponding classes of objects.

The core of the chapter consisted in the introduction of two different models for aristocratic nominalism. The first one was the topological polar model by Mormann and Rumffitt. According to this model, attributes are reconstructed as certain topologically closed sets, sets containing all the objects that are arbitrarily close to a given paradigmatic object. The second one is the similarity model based on pure similarities of order 1, which was the new model introduced in this thesis. This class of structures was based on previous results by Brockhaus and Mormann on similarities and quasianalysis. According to this latter model, attributes are reconstructed as collections of objects that are similar to a given paradigmatic object. These collections are called 'similarity circles of order 1'. Despite the apparent difference, these two models were later shown to be mathematically equivalent.

Nevertheless, the reconstruction of the attributes was not proven yet. This was done in Chapter V. The goal now was to get to the Minimal Conception of Kinds just by starting from the sparser ontology of aristocratic resemblance nominalism. This required making first use of the similarity model of Chapter IV to reconstruct natural attributes and then obtaining the lattice of kinds from the resulting context (just as it was done for realist contexts in Chapter III). However, the main obstacles for reconstructing the context were Goodman's infamous companionship and imperfect community problems.

Goodman's problems required us to narrow down the class of realist contexts that can be reconstructed. This class was later called the class of 'polar contexts'. The similarity model introduced in Chapter IV was shown to be able to reconstruct such a class of contexts, thus providing nominalistic surrogates for attributes. Moreover, in the process, the two nominalistic models introduced in Chapter IV (the polar and similarity models) were shown to be equivalent. Once the context was obtained, the next step was to generate the corresponding concept lattice, which was to represent the nominalistic lattice of kinds:

Aristocratic Resemblance Nominalism The lattice of nominalist kinds is the concept lattice induced by the context induced by a similarity model, where the attributes are the similarity circles of order 1.

I proved that the concept lattices that can be obtained starting from a similarity model are all (co)atomistic. Moreover, I also proved that every such complete lattice can be obtained uniquely in that way by starting from a similarity model. This gave us a description of the class of specificity orders between kinds that the nominalist can mimic and also more specific orders that can be reconstructed just by adding the corresponding axioms to the similarity structures. Furthermore, the quasianalysis was shown to give several equivalent representations of objects: the same object can be represented as the set of its paradigms or as the set of its attributes (which are classes of similar objects).

However, there remained some doubts concerning the paradigmatic objects. The last question I considered was whether a generalized picture could be given, according to which paradigms may have several properties and each property may be generated by several paradigms. In fact, I generalized several concepts and results concerning pure similarities of order 1 to similarity structures (called 'simple' similarities) where there is an arbitrary number of paradigms for each property and each paradigm can have an arbitrary number of properties. This lead to a more plausible model for aristocratic nominalism.

In this way an answer to the questions that motivated this thesis, namely a search for a better understanding of both the external and internal structure of kinds, was given: under nominalist ontological assumptions regarding the nature of objects and resemblance relations, the internal structure of kinds is determined by the similarities among these objects and their paradigms, whereas the external structure of kinds is that of a complete (co)atomistic lattice that satisfies the requirements of the Minimal Conception of Kinds, including Kant's Law.

6.2 What is Yet to be Done

In this thesis I have just hinted at how a Minimal Conception of natural kinds can be developed by nominalistic means formally. But there is a lot of work to do for such an analysis to be shown to be fruitful for other aspects of the topic of kinds.

Consider again the arguments surveyed in Chapter II in favour of natural kinds, namely the epistemological argument based on projectability and induction, the semantic argument based on the reference of natural kind terms, the metaphysical argument based on naturalness and universals and the naturalistic argument based on the role of classifications in scientific theories. In this thesis I have only explored issues having to do with some parts of the metaphysical side of the problem. I have not said much regarding how the models for the Minimal Conception of kinds may be used to deal with the epistemological, semantic and naturalistic issues. Furthermore, I have barely touched on issues regarding how natural kinds are related to time and qualitative change, ontological dependence relations and other main metaphysical problems. For example, the extension of a kind changes from one time to another.

The epistemological argument is based on Quine's suggestion that appealing to kinds could help dealing with the new riddle of induction. This lead some philosophers, like Gärdenfors [42], to take kinds to be concepts and to represent them as regions in conceptual spaces (moreover, this would make hypotheses about the structure of these domains empirically testable). The issues concerning proyectability and induction were then interpreted in terms of the concept learning processes of a given epistemic subject. The idea was that conceptual categorization in terms of prototypical instances could help dealing with these epistemological problems. A very similar approach to induction was made by Carnap, in terms of his attribute spaces [131]. Although some connections between the standard conceptual spaces approach and the models introduced in this thesis were made, it would be interesting to see these relations in more detail. Furthermore, ideally such an approach should be compatible with current standard answers to the problems of confirmation and induction in formal epistemology (say, Bayesian approaches).

The semantic argument is based on Kripke's and Putnam's idea that natural kind terms behave similarly to proper names, since they are rigid designators (expression whose reference remains invariant across possible worlds). This lead to discussions on natural kind essentialism. Adding modal features to the model of kinds introduced in this thesis would help with discussing such matters. Again, such an approach should be compatible with standard frameworks in formal semantics.

The naturalistic argument suggests that kinds are described by scientific classifications. I have not focused on scientifically accepted classifications, such as the ones given by the different taxonomic programs in biology or the periodic system in chemistry. However, philosophers of science have argued that traditional approaches to the topic of natural kinds are at odds with some of these classifications, for example with the biological ones (see [37]). I find this diagnosis somewhat surprising. The mathematical structures that were used in this thesis to model kinds (lattices, concept lattices, trees, similarity relations, and so on) are among the most used mathematical models of classifications in science (see [100]). Furthermore, the ontological assumptions made by these models fit very nicely many aspects of the more traditional conception of kinds (they remain silent on some issues such as essentialism). Given the minimal assumptions made in this thesis, I find it hard to believe that the structure of scientific classifications is really in conflict with such an approach to natural kinds. Therefore, it would be interesting to check this issue in more detail.

Although I think that this approach can be used to say a great deal about each of these topics, that still requires further investigation. For the time being, I hope that the approach to natural kinds taken in this thesis has shown to be fruitful enough to be developed further in the future.

Bibliography

[1] Amilburu, A. M. (2015): *La Naturaleza de los Géneros Naturales. Un Estudio Crítico sobre la Contribución de esta Noción a la Comprensión de las Prácticas Clasificatorias en Ciencia*. PhD Thesis, UPV-EHU.

[2] Armstrong, D. M. (1978): *Nominalism & Realism: Universals & Scientific Realism, Volumes 1-2*. Cambridge: Cambridge University Press.

[3] Armstrong, D.M. (1997): *A World of States of Affairs*. Cambridge: Cambridge University Press.

[4] Armstrong, D. M. (1989): *Universals: an opinionated introduction*. Boulder: Westview Press.

[5] Bealer, G. (1987): The Philosophical Limits of Scientific Essentialism. *Philosophical Perspectives*, **1**, Metaphysics, 289–365.

[6] Béziau, JY. (2012): The Power of the Hexagon. *Logica Universalis*, **6**, 1–43.

[7] Béziau, JY.; Basti, G. (ed.) (2017): *The Square of Opposition: A Cornerstone of Thought*. Birkhäuser.

[8] Bird, A. (2007): *Nature's Metaphysics: Laws and Properties*. Oxford: Clarendon.

[9] Bird, A. (2009): Essences and Natural Kinds. In Le Poidevin, R.; Simons, P.; McGonigal, A.; P. Cameron, R. (ed.) *The Routledge Companion to Metaphysics*. London: Routledge.

[10] Bird, A. (2015): The Metaphysics of Natural Kinds. *Synthese*, 1–30.

[11] Bird, A.; Tobin, E. (2017): Natural Kinds. In *The Stanford Encyclopedia of Philosophy* (Spring 2017 Edition), Edward N. Zalta (ed.), forthcoming URL = https://plato.stanford.edu/archives/spr2017/entries/natural-kinds/.

[12] Boyd, R. (1999): Homeostasis, Species and Higher Taxa. In R. Wilson (ed.), *Species: New Interdisciplinary Essays*. Cambridge, Massachusetts: MIT Press.

[13] Breysse, O.; De Glas M. (2007): A New Approach to the Concepts of Boundary and Contact: Towards an Alternative to Mereotopology. *Fundamenta Informaticae*, 78, 2, 217–238.

[14] Brockhaus, K., (1963): *Untersuchungen zu Carnaps Logischem Aufbau der Welt*. Dissertation, Universität Münster.

[15] Busse, R. (2016): The Adequacy of Resemblance Nominalism about Perfect Naturalness. *Philosophy and Phenomenological Research*, XCVI, **2**, 443–469.

[16] Carnap, R. (1923): Die Quasizerlegung – Ein Verfahren zur Ordnung nichthomogener Mengen mit den Mitteln der Beziehungslehre. Unpublished manuscript RC-081-04-01, University of Pittsburgh.

[17] Carnap, R. (1967 [1928]): *The Logical Structure of the World & Pseudoproblems in Philosophy*. Translated by Rolf A. George. Berkeley, Los Angeles: University of California Press.

[18] Carnap, R. (1963): *Logical Foundations of Probability*. Chicago: The University of Chicago Press.

[19] Carnap, R. (1980): *A Basic System of Inductive Logic, part II*. In Richard C. Jeffrey (ed.) *Studies in Inductive Logic and Probability Vol.II* (pp. 7-156). Berkeley, Los Angeles: University of California Press.

[20] Caspard, N.; Leclerc, B.; Monjardet, B. (2012): *Finite Ordered Sets, Concepts, Results and Uses*. Cambridge: Cambridge University Press.

[21] Chakravartty, A. (2007): *A Metaphysics for Scientific Realism: Knowing the Unobservable*. Cambridge: Cambridge University Press.

[22] Cohnitz, D.; Rossberg, M. (2016): Nelson Goodman. In *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), Edward N. Zalta (ed.), URL = ¡https://plato.stanford.edu/archives/win2016/entries/goodman/¿.

[23] Cooper, R. (2005): *Classifying Madness: A Philosophical Examination of the Diagnostic and Statistical Manual of Mental Disorders*. Dordrecht: Springer.

[24] Corcoran, J. (1972): Completeness of an Ancient Logic. *The Journal of Symbolic Logic*, **37**, 4, 696–702.

[25] Craver, C. F. (2009): Mechanisms and natural kinds. *Philosophical Psychology*, **22** (5), 575-594.

[26] Davey, B.A.; Priestley, H.A. (2012): *Introduction to Lattices and Order*. Cambridge: Cambridge University Press.

[27] Decock. L.; Douven, I. (2009): Two Accounts of Similarity Compared. In H. Alexander and H. Leitgeb (eds.)(2009), *Reduction, Abstraction, Analysis*. Ontos Verlag.

[28] Devitt, M. (1980): Ostrich Nominalism or Mirage Realism? *Pacific Philosophical Quarterly*, **61**, 433–439.

[29] Devitt, M. (2011): Natural Kinds and Biological Realisms. In J. K. Campbell; M. O' Rourke; M. H. Slater (ed.) *Carving Nature at its Joints: Natural Kinds in Metaphysics and Science*. Cambridge, London: The MIT Press.

[30] Douven, I.; Gärdenfors, P. (2019): What are Natural Concepts? A Design Perspective. *Mind Lang.*, 1–22.

[31] Douven, I. (2019): Putting Prototypes in Place. *Cognition*, **193**.

[32] Dupré, J. (1993): *The Disorder of Things: Metaphysical Foundations of the Disunity of Science*. Harvard: Harvard University Press.

[33] Eddon, M. (2007): Armstrong on Quantities and Resemblance. *Philosophical Studies*, **136**, 385–404.

[34] Ellis, B. (2001): *Scientific Essentialism*. Cambridge: Cambridge University Press.

[35] Ellis, B. (2002): *The Philosophy of Nature: A Guide to the New Essentialism*. Chesham: Acumen.

[36] Ereshefsky, M. (2004): *The Poverty of the Linnaean Hierarchy: A Philosophical Study of Biological Taxonomy*. Cambridge: Cambridge University Press.

[37] Ereshefsky, M.; Reydon, T.A.C. (2015): Scientific Kinds. *Philosophical Studies*, **172**, 969–986.

[38] Evans, G. (1978): Can There be Vague Objects? *Analysis*, **38**, 208.

[39] Fine, K. (1994): Essence and Modality. In J. Tomberlin (ed.) *Philosophical perspectives 8: Logic and Language*, 1–16. Atascadero, CA: Ridgeview.

[40] Ganter, B.; Wille, R. (1999): *Formal Concept Analysis: Mathematical Foundations*. Springer-Verlag: Berlin and Heidelberg.

[41] García, R. K. (2015): Is Trope Theory a Divided House? In M. Loux; G. Galluzo (ed.) *The Problem of Universals in Contemporary Philosophy*. Cambridge University Press.

[42] Gärdenfors, P. (2000): *Conceptual Spaces: The Geometry of Thought*. Cambridge: MIT Press.

[43] Gärdenfors, P. (2014): *The Geometry of Meaning: Semantics based on Conceptual Spaces*. London: MIT Press.

[44] Gentner, D. (1983): Structure-Mapping: A Theoretical Framework for Analogy. *Cognitive Science*, **7**, 155–170.

[45] Goldstone, R. L.; Medin, D. L.; Gentner, D. (1991): Relational Similarity and the Nonindependence of Features in Similarity Judgements. *Cognitive Psychology*, **23**, 222–262.

[46] Goldstone, R. L.; Son J. Y. (2005): Similarity. In K. J. Holyoak, R. G. Morrison (ed.) *The Oxford Handbook of Thinking and Reasoning*. Oxford: Oxford University Press.

[47] Goodman, N. (1983 [1954]): *Fact, Fiction and Forecast*. Cambridge, London: Harvard University Press.

[48] Goodman, N. (1966 [1953]): *The Structure of Appearance*. Dordrecht: Boston Studies in the Philosophy and History of Science, Synthese Library.

[49] Goodman, N. (1972): *Problems and Projects*. Indianapolis: Bobbs-Merrill.

[50] Guigon, G. (2009): *The Metaphysics of Resemblance*. PhD Thesis. Université de Genève.

[51] Hacking, I. (1991): A Tradition of Natural Kinds. *Philosophical Studies*, **61**, 109-126.

[52] Hacking, I. (2007a): Natural Kinds: Rosy Dawn, Scholastic Twilight. *Royal Institute of Philosophy Suplemment*, **61**, 203–239.

[53] Hacking, I. (2007b): Trees of Logic, Trees of Porphyry. In J. Heilbron (ed.), *Advancements of Learning: Essays in Honour of Paolo Rossi*. Firenze: Olschki.

[54] Hahn, U.; Chater, N (1997): Concepts and Similarity. In Koen Lamberts, David Shanks (ed.) *Knowledge, Concepts and Categories* (pp. 43–92). UK: Psychology Press.

[55] Hahn, U.; Chater, N.; Richardson, L. B. (2003): Similarity as Transformation. *Cognition*, **87**, 1–32.

[56] Hausman (2010): *Logic and Philosophy. A Modern Introduction.* Wadsworth: Cengage Learning.

[57] Hawley, K.; Bird, A. (2011): What are Natural Kinds. *Philosophical Perspectives*, **25**, 205–221.

[58] Hazen, A. P.; Humberstone, L. (2004): Similarity Relations and the Preservation of Solidity. *Journal of Logic, Language and Information*, **13**, 25–46.

[59] Hempel, C. G. (1970): Fundamentals of Concept Formation in Empirical Science. In O. Neurath, R. Carnap, C. Morris (ed.), *Foundations of the Unity of Science, Vol. 2* (pp. 651–746). Chicago, London: The University of Chicago Press.

[60] Hendry, R. F. (2006): Elements, Compounds and Other Chemical Kinds. *Philosophy of Science*, **73**, 864–75.

[61] Hendry, R.F. (2015): Are Chemical Kinds Natural Kinds? In U. Mäki, I. Votsis, S. Ruphy, G. Schurz (eds.), *Recent Developments in the Philosophy of Science: EPSA13 Helsinki*, 251–261. Switzerland: G. Springer International Publishing 1.

[62] Inwagen, V. I. (2016): The neo-Carnapians. *Synthese*, 1–26.

[63] Kant, I. (1988; 1800): *Logic*. Translated by R. S. Hartmann; W. Schwarz from *Immanuel Kants Logik: Ein Handbuch zu Vorlesungen*. New York: Dover Publications.

[64] Khalidi, M.A. (2013): *Natural Categories and Human Kinds. Classification in the Natural and Social Sciences*. Cambridge: Cambridge University Press.

[65] Khalidi, M.A. (2015): Natural Kinds as Nodes in Causal Networks. *Synthese*.

[66] Kripke, S. (1971): Identity and Necessity. In M. K. Munitz (ed.) *Identity and Individuation*. New York University Press, 135–164.

[67] Kripke, S. (1981 [1972]): *Naming and Necessity*. Malden, Oxford, Carlton: Blackwell Publishing.

[68] Kuhn, T. (2000): Afterwords. In T. Kuhn, J. Conant, J. Haugeland (eds.), *The Road since Structure* (pp. 224–252). Chicago: Chicago University Press.

[69] Lange, M. (2009): *Laws and Lawmakers: Science, Metaphysics and the Laws of Nature*. Oxford: Oxford University Press.

[70] Leitgeb, H. (2007): A New Analysis of Quasianalysis. *Journal of Philosophical Logic*, **36**, 181–236.

[71] Leitgeb, H. (2011): Logic in General Philosophy of Science: Old Things and New Things. *Synthese*, **179**, 339–350.

[72] Lewis, D. (1973): *Counterfactuals*. London, Oxford: Basil Blackwell.

[73] Lewis, D. (1983): New Work for a Theory of Universals. *Australasian Journal of Philosophy*, **61**, 343–377.

[74] Lewis, D. (1986): *On the Plurality of Worlds*. New York: Basil Blackwell.

[75] Loux, M. J. (2015): An Exercise in Constituent Ontology. In M. Loux; G. Galluzo (ed.) *The Problem of Universals in Contemporary Philosophy*. Cambridge: Cambridge University Press.

[76] Lowe, E. J. (1982): On the Alleged Necessity of True Identity Statements. *Mind*, **91**, 364, 579–584.

[77] Lowe, J. (2006): *The Four-Category Ontology*. Oxford: Oxford University Press.

[78] Lowe, E. J. (2007): A Problem for A Posteriori Essentialism Concerning Natural Kinds. *Analysis*, **67**, 4, 286–292.

[79] Mackie, P. (2006): *How Things Might Have Been. Individuals, Kinds, and Essential Properties*. Oxford: Oxford University Press.

[80] Magnus, P. D. (2012): *Scientific Enquiry and Natural Kinds. From Planets to Mallards*. London: Palgrave Macmillan.

[81] Markman, A. B.; Gentner, D. (1993): Splitting the Differences: A Structural Alignment view of Similarity. *Journal of Memory and Language*, **32**, 517–535.

[82] Martin, J. N. (1997): Aristotle's Natural Deduction Reconsidered. *History and Philosophy of Logic*, **18**, 1, 1–15.

[83] Mayr, E.; Bock, W.J. (2002): Classifications and other Ordering Systems. *Journal of Zoological Systematics and Evolutionary Research*, **40**, 169–194.

[84] McLaughlin, P.; Schlaudt, O. (2020): Kant's Antinomies of Pure Reason and the 'Hexagon of Predicate Negation'. *Logical Universalis*, **14**, 51–67.

[85] Medin, D. L.; Goldstone, R. L.; Gentner, D. (1993): Respects for Similarity. *Psychological Review*, **100**, 254–278.

[86] Mill, J. S. (1882): *A System of Logic*. In J. M. Robson (ed.) (1973) *The Collected Works of John Stuart Mill*. London: Routledge and Kegan Paul.

[87] Mormann, T. (1994a): A Representational Reconstruction of Carnap's Quasi-Analysis. In M. Forbes, (ed.), *PSA 1994*, vol. 1, 96–104.

[88] Mormann, T. (1994b): Accesibility, Kinds and Laws: A Structural Explication. *Philosophy of Science*, **61**, 3, 389–406.

[89] Mormann, T. (2009): New Work for Carnap's Quasi-Analysis. *Journal of Philosophical Logic*, **38**, 249–282.

[90] Mormann, T. (2020): Topological Models of Columnar Vagueness. *Erkenntnis*. Published Online.

[91] Mormann, T. (forthcoming): Prototypes, Poles and Tessellations. Towards a Topological Theory of Conceptual Spaces. *Synthese*. Published Online.

[92] Mosterín, J. (1984): Taxonomía Formal. *Enrahonar: Quaderns de Filosofia*, **7**, 8, 109–119.

[93] Mumford, S. (2005): Kinds, Essences, Powers. *Ratio*, XVIII.

[94] Niiniluoto, I. (1987): *Truthlikeness*. Dordrecht: D. Reidel Publishing Company.

[95] Nosofsky, R. M. (1991): Stimulus bias, Assymmetric Similarity and Classification. *Cognitive Psychology*, **23**, 94–140.

[96] Nosofsky, R. M. (2011): The Generalized Context Model: an Exemplar Model of Classification. In Emmanuel Pothos (ed.) *Formal Approaches in Categorization* (pp. 18–39). Cambridge: Cambridge University Press.

[97] Oliver, A. (1996): The Metaphysics of Properties. *Mind*, **105**, 1–80.

[98] Orilia, F.; Swoyer, C. (2017): Properties. In *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition), E. N. Zalta (ed.), URL = ¡https://plato.stanford.edu/archives/win2017/entries/properties/¿.

[99] Parsons, T. (2017): The Traditional Square of Opposition. In *The Stanford Encyclopedia of Philosophy* (Summer 2017 Edition), E. N. Zalta (ed.), URL = ¡https://plato.stanford.edu/archives/sum2017/entries/square/¿.

[100] Parrochia, D.; Neuville, P. (2013): *Towards a General Theory of Classifications*. Basel: Birkhäuser.

[101] Paseau, A. (2012): Resemblance Theories of Properties. *Philosophical Studies*, **157**, 361–382, Published Online 2010.

[102] Price, H. H. (1953): Universals and Resemblances. In *Thinking and Experience*. London: Hutchinson's University Library.

[103] Putnam, H. (1975): The Meaning of Meaning. In H. Putnam *Mind, Language and Reality (Philosophical Papers) (Vol. II)*. Cambridge University Press.

[104] Putnam, H. (1990): Is Water Necessarily H_2O? In J. Conant (ed.) *Realism with a Human Face*. Harvard University Press.

[105] Quine, W. V. O. (1953): On What There Is. In *From a Logical Point of View*. Cambridge: Harvard University Press.

[106] Quine, W. V. O. (1964): Ontological Reduction and the World of Numbers. *The Journal of Philosophy*, **61(7)**, 209–216.

[107] Quine, W. V. O. (1970): Natural Kinds. In N. Rescher et al (eds.) *Essays in Honor of Carl G. Hempel*. Dordrecht: Reidel.

[108] Roca-Royes, S. (2011): Essential Properties and Individual Essences. *Philosophy Compass*, **6/1**, 65–77.

[109] Rodriguez-Pereyra, G. (2002): *Resemblance Nominalism*. Oxford: Oxford University Press.

[110] Rodriguez-Pereyra, G. (2004): Paradigms and Russell's Resemblance Regress. *Australasian Journal of Philosophy*, **82**, 4, 644–651.

[111] Rooij, R. V.; Schulz, K. (2014): A Question of Priority. In E. McCready et al (eds.) *Formal Approaches to Semantics and Pragmatics*, Studies in Linguistics and Philosophy 273. Dordrecht: Springer

[112] Rosch, E. (1975): Cognitive Representations of Semantic Categories. *Journal of Experimental Psychology*, **104**, 3, 192–233.

[113] Rosch, E. (1988): Principles of Categorization. In Allan Collins, Edward Smith (ed.) *Readings in Cognitive Science: A Perspective from Psychology and Artificial Intelligence* (pp. 312–322). California: Morgan Kaufmann Publishers.

[114] Rumfitt, I. (2015): *Boundary Stones of Thought*. Oxford: Oxford University Press.

[115] Ruphy, S. (2010): Are Stellar Kinds Natural Kinds? A Challenging Newcomer in the Monism/Pluralism and Realism/Antirealism Debates. *Philosophy of Science*, **77**, 5, 1109–1120.

[116] Russell, B. (2009 [1948]): *Human Knowledge: Its Scope and Limits*. London, New York: Taylor & Francis Routledge Classics.

[117] Salmon, N. (1979): How Not to Derive Essentialism from the Theory of Reference. *The Journal of Philosophy*, **76**, 12, 703–725.

[118] Salmon, N. (1982): *Reference and Essence*. Oxford: Basil Blackwell.

[119] Santana, C. (2019): Mineral Misbehavior: Why Mineralogists don't deal in Natural Kinds. *Foundations of Chemistry*, 1–11.

[120] Schaffer, J. (2004). Two Conceptions of Sparse Properties. *Pacific Philosophical Quarterly*, **85**, 92–102.

[121] Schreider, J. A. (1975): *Equality, resemblance and order*. Translated from Russian by M. Greendlinger. Moscow: Mir Publishers.

[122] Shepard, R. N. (1987): Toward a Universal Law of Generalization for Psychological Science. *Science*, **237**, 1317–1323.

[123] Slater, M. (2014): Natural Kindness. *The British Journal for the Philosophy of Science*, **0**, 1–37.

[124] Slater, M.; Borghini, A. (2011): Lessons from the Scientific Butchery. In J. K. Campbell, M. O' Rourke y M. H. Slater (ed.) *Carving Nature at its Joints: Natural Kinds in Metaphysics and Science*. Cambridge, London: The MIT Press.

[125] Smiley, T. (1973): What is a Syllogism? *Journal of Philosophical Logic*, **2**, 136–154.

[126] Soames, S. (2005): Kripke on Epistemic and Metaphysical Possibility: Two Routes to the Necessary A posteriori. In A. Berger (ed.) *Saul Kripke*. Cambridge University Press.

[127] Stevens, S. S. (1946): On The Theory of Scales of Measurement. *Science*, **103**, 2684, 677–680.

[128] Suppes, P.; Krantz, D.; Luce, D.; Tversky, A. (1989): *Foundations of measurement, Volumes I-III*. New York: Academic.

[129] Swoyer, C. (1991): Structural Representations and Surrogative Reasoning. *Synthese*, **87**, 449-–508.

[130] Swoyer, C. (1995): Leibniz on Intension and Extension. *Nous*, **29**, 1, 96–114.

[131] Sznajder, M. (2016): What Conceptual Spaces can do for Carnap's late inductive logic. *Studies in History and Philosophy of Science*, Series A, **56**, 62–71.

[132] Sznajder, M. (2017): *Inductive Logic on Conceptual Spaces*. PhD Thesis.

[133] Tahko, T. (2012): Boundaries in Reality. *Ratio*, XXV.

[134] Tahko, T. (2015): Natural Kind Essentialism Revisited. *Mind*, **124**, 495.

[135] Thomason, R. (1969): Species, Determinates and Natural Kinds. *Nous*, **3**, 1, 95–101.

[136] Tobin, E. (2010): Crosscutting Natural Kinds and the Hierarchy Thesis. In Helen Beebee, Nigel Sabbarton-Leary (eds.) *The Semantics and Metaphysics of Natural Kinds*. London, New York: Routledge Studies in Metaphysics.

[137] Tversky, A. (2004 [1977]): Features of Similarity. In Eldar Shafir (ed.) *Preference, Belief and Similarity, Selected writings of Amos Tversky* (pp. 7–46). Cambridge, London: MIT press.

[138] Varzi, A. (2014): Realism in the Desert. In M. Dell'Utri, F. Bacchini, S. Caputo (eds.) *Realism and Ontology without Myths*. Newcastle: Cambridge Scholars Publishing.

[139] Wiggins, D. (1980): *Sameness and Substance*. Oxford: Basil Blackwell.

[140] Wiley, E.O.; Lieberman, B.S. (2011): *Phylogenetics: Theory and Practice of Phylogenetic Systematics*. New Jersey: Wiley-Blackwell.

[141] Willard, S. (2004): *General Topology*. Mineola, New York: Dover Publications.

[142] Wille, R. (2000): Boolean Concept Logic. In B. Ganter, G. W. Mineau (eds.) *International Conference on Conceptual Structures 2000*. Berlin, Heidelberg: Spinger-Verlag.

[143] Williamson, T. (1988): First Order Logics for Comparative Similarity. *Notre Dame Journal of Formal Logic*, **29**, 4.

[144] Williamson, T. (2013): *Modal Logic as Metaphysics*. Oxford: Oxford University Press.

[145] Williamson, T. (2017): Model-Building in Philosophy. In R. Blackford and D. Broderick (eds.) *Philosophy's Future: The Problem of Philosophical Progress*. Wiley-Blackwell.

[146] Zenker, F.; Gärdenfors, P. (2015) (eds.): *Applications of Conceptual Spaces, The Case for Geometric Knowledge Representation*. Springer.