KNOWLEDGE REPRESENTATION, REFLEXIVE REASONING AND DISCOURSE PROCESSING†

Jesús EZQUERRO
Mauricio IZA*

Manuscript submitted: 1994.11.30.
Final version: 1996.1.30.

* Department of Logic and Philosophy of Science, Apt. 1.249, 20080 San Sebastian, Spain. E-mail: ylpezmaj@sf.ehu.es (Jesús Ezquerro), ylbizmim@sf.ehu.es (Mauricio Iza).

BIBLID [ISSN 0495-4548 (1996) Vol. 11: No 26; p. 125-145]

ABSTRACT: Classical approaches such as *frames*, *scripts*,... have been unable to deal with the kind of inferences necessary in natural language processing situations such as text comprehension. Shastri & Ajjanagadde (1993), proposed a local connectionist model for the sort of reasoning requiring such a fast inference. The problem with this system is that it controls only the adequacy of argument-fillers, leaving untouched the activation control issue, namely, why we perform certain inferences, and not others, in a given situation. The aim of this paper is to examine how a focus, in the sense of Grosz (1981). could operate as a theoretic constraint with the above reasoner system to handle two aspects: inference control and anaphora resolution (i.e., antecedent activation), during text understanding.

Keywords: Anaphora resolution, Connectionist systems, Focus Theory, Knowledge representation, Reflexive reasoning, Text processing.

1. Introduction

Human agents, despite operating with a large knowledge base, take a few hundred milliseconds to perform a broad range of cognitive tasks, such as object recognition, spoken and written language understanding, and inference performance such as: "Tweety is a bird, so it flies". Representation of conceptual information and the cognitive processes that access it are such that not only are relevant facts automatically 'retrieved', but certain kinds of 'inferences' also exhibit 'extreme efficiency'.

In text understanding readers perform inferences in order to achieve a comprehension. For example, in anaphora processing, its interpretation usually depends upon some arbitrary part of general knowledge not easily accessible. Therefore, anaphora resolution requires inference processes based on world knowledge (Garnham, 1985) or on context, that have to be performed quickly.

Two general kinds of inference are of particular interest: *Necessary* and *elaborative* inferences. In the first case, readers make necessary inferences such as the ones required for preserving referential coherence or for establishing causal relations[1]. In the second, readers perform elaborative inferences in order to predict anticipated consequences or information[2]. For example, perhaps as soon as people reads the sentence "The actress fell from the 14th floor", they predict that the actress died. In doing so, they generate a predictive or forward inference. Experimental data show that this kind of inferences becomes only partially encoded inside the mental representation of a text (McKoon & Ratcliff, 1986). For instance, in Sanford & Garrod's (1981) account, it is almost misleading to affirm that given "Jane unlocked the door", the conclusion "She used a key" will be drawn. For Sanford (1990) such *knowledge* becomes *accessible*; although, in the case of instantiation, context may be restrictive enough to define the nature of the role. Therefore, the two crucial determinants of inference making are the establishment of local coherence and ready availability of knowledge, that can underlie elaboration.

Instantiation is a type of elaborative inference that depends on knowledge about empirical regularities. It occurs when a general term, such as *fish*, takes on a more specific interpretation from its context. Garnham (1981) shows psychological evidence for on-line instantiation. He found that:

"The shark swam rapidly through the water"

was read slowly after:

"The fish avoided the swimmer"

since there is nothing in that sentence indicating that the fish is a shark. However, with a context that cued an instantiation:

"The fish attacked the swimmer"

the use of *shark* in the following sentence did not produce any difficulty, suggesting that, when people read that a fish attacked a swimmer, they represent the fish as a shark or something like one.

Therefore, context can be used to elaborate a representation of the object. Such elaboration is based on world knowledge, not on lexical semantics. This is consistent with the claim that the content representation of a sentence, namely, its mental model, is not linguistic in nature. Its components would represent objects, not word meanings[3].

In Cognitive Psychology and Artificial Intelligence, inference has been usually considered as an 'establisher' of connections between fragments of the surface structure, recurring in context and world knowledge[4]. This approach to inference is based in text and relies on the notion of inference process that searches significant relationships between different propositions in the text. This idea is not new, it underlies semantic networks (Quillian, 1968), frames (Minsky, 1975), scripts (Schank, 1972; Schank & Abelson, 1977), schemes, or other ways of describing a 'chunk' of knowledge (Anderson, 1983; Kintsch, 1988). However, there are more computational relations of cognition that the ones captured by this approach[5].

If we consider these computations as operations governed by systematic rules in symbolic structures (that is, inferences) we confront the following challenge: some generalised notion of inference is intractable, we are however able to perform fast reasoning processes in real time[6]. One example posed by Schubert (1989), from the Little Red Riding Hood story (LRRH) ["The wolf heard the wood cutters nearby and decided to wait"], shows how quickly one catches the implicit meaning of a sentence as a natural reflex of reading. Understanding this sentence requires a chain of reasoning that Shastri & Ajjanagadde (1990) described as 'reflexive reasoning'.

In understanding the story of Red Riding Hood, we rapidly access inferences about the nature of wolves, little girls, grandmothers and wood-cutters, and about their likely intentions towards each other. We must do this *on-line*, because otherwise we would not be able to retrieve, for example, the necessary references for pronouns or choices of lexical meaning. This sort of reasoning is characterised by the fact that it involves accessing knowledge about words chosen unpredictably from an extremely large vocabulary. The 'data base' and rules relating them are enormous and contain many relational concepts (Stenning & Oaksford 1992, p. 15).

Could a connectionist system account for this reflexive reasoning? Shastri & Ajjanagadde (1990; henceforth, S & A) claim that this sort of reasoning involves systematic inference which a system performs on structured knowledge in the long-term memory. The reason is that it requires high level functions such as those executed by knowledge-based systems. In particular, systematic inference on structured knowledge requires a representation of dynamic variable bindings and their propagation in an inference chain.

The dynamic variable binding problem in connectionist systems arises when the system needs a structured knowledge representation whose composite structures are continuously changing through time. A connectionist system that supports reflexive reasoning should be able to represent structured knowledge such as n-ary predicates. At the same time, the system should be able to create and destroy dynamically the relationships between arguments and constants of predicates, without disrupting the structure of the network. Thus, such a dynamic variable binding does need a very fast propagation of the variable bindings from one dynamic fact to the other. This is not plausible in static variable bindings given that it cannot change the links between argument nodes and constants (see Park, 1992, pp. 10-11).

S & A (1990) propose a local connectionist model for one reasoning type requiring such a fast inference. They describe in detail a technique for using 'neuron-like' elements to encode facts and rules involving n-ary predicates with variables. They adopt temporal synchrony to achieve dynamic binding of variables. This *dynamic binding* problem has persisted as one of the weaknesses of connectionist systems. In Shastri & Ajjanagadde's case, a key part of their solution is to match rhythmic patterns of activity, i.e. temporal synchrony, within the network.

However, as we will see below, this model needs "something more" to explain this sort of reasoning. In story understanding, each inference (and some can be very complex) must be chosen from a large knowledge base about objects, people and things. The practical deployment of inferential capabilities for any task requires control: knowing what to infer when, and knowing when to stop (Sidner, 1983). Likewise, for solving anaphoric reference, some computation similar to the above chain of reasoning (LRRH example) might occur in some situations, when the sentence in question is processed.

Anaphora resolution involves, at least, two cognitive processes: antecedent search and selection, and, anaphora meaning assignment (Garnham, 1985). Both processes make it possible to establish semantic coherence between text utterances and its subsequent integration because of the coreference established between antecedent and anaphora (in the LRRH example, the actor of the two actions).

However, the 'search + selection' hypothesis requires some complementary idea to explain the effects of the semantic and representational features. Specifically, the search process is conditioned by the degree of activation of the antecedent [in working memory]. While the antecedent remains activated, the utterance containing the anaphora is rapidly integrated in the building representation about text content. The presence of an activated candidate allows the exclusion of other potential antecedents, thus facilitating the selection process. Finally, when the antecedent is not activated in memory, both meaning assignment and anaphora integration require a "backward" search in text representation (O'Brien, 1987; O'Brien, Duffy & Myers, 1986), or even building an inferential bridge (that could reduce the search process) to previous knowledge, requiring additional cognitive means.

Focus theory has tried to account for this antecedent activation. Focus and focusing mechanisms, by means of attentional states, explain how some anaphors can be automatically interpreted during discourse processing (dialogue). Also, they can control when an inference is necessary and when this inference must end [finish].

The relationship between language and focusing is twofold: what is said influences focusing; what is focused on influences what is said. The speaker provides cues for the hearer both to what s/he is currently focused on, and to what s/he wants to focus on next. These cues may be linguistic or may come from shared linguistic or non-linguistic knowledge. The hearer depends on shared beliefs about what entities are highlighted to interpret such things as the appropriate sense of a particular word, and the object or event corresponding to a definite description. The link between the entities discussed in an utterance and the entities focused on when the utterance is spelled out is thus an important aspect in producing and understanding such utterance. In the case of definite descriptions, the entities on which a speaker/hearer is focused influences both the kinds of descriptions used and how their descriptions are interpreted (Grosz, 1981).

In the remainder of this paper, we describe the above mentioned S & A's connectionist model for reflexive reasoning. Next, we show the purposes of Focus theory, specifically, concerning anaphora resolution (antecedent activation) and inference control through attentional states [Why we infer what we do?]. Finally, in the last section, we propose and discuss how focus theory could operate together with the reasoner mechanism (their possible compatibility) and its outcome for explaining both the sorts of inferences and anaphora interpretation during text understanding.

2. A connectionist system for reflexive reasoning

We usually understand typical sentences in one to two seconds (or less). Given that reflexive reasoning takes place during language understanding, it follows that 'episodes of reflexive reasoning may take as little as a few hundred msec' (S & A 1992, p. 3). This kind of reasoning needs to establish dynamic and temporal bindings in any situation that involves reasoning with representations using variables (that is, anaphora resolution). However, to perform general and flexible reasoning, a representation system must be expressive enough to encode 'abstract' and 'systematic' knowledge about structured and composite entities, that is, encoding rules with variables and polyadic predicates.

Regarding how humans perform certain inferences with extreme efficiency, the structured connectionist approach offers an appropriate framework that explains these symbolic relationships and meets the computational effectivity challenge mentioned above. The relative simplicity of a neuron's processing ability as regards to the needs of symbolic computation, and the complexity of messages exchanged by neurons, impose strong constraints on the nature of neural representations and processes (Feldman & Ballard, 1982; Feldman, 1989; Shastri, 1990). A reasoning system must be able to encode systematic and abstract knowledge and to 'instantiate' it in specific situations to draw appropriate inferences. The system must solve a complex version of the 'variable binding problem' (Feldman, 1982; von der Malsburg, 1986).

In this respect, systematic inference can be considered as the dynamic outcome of a systematic propagation of structures, within structured knowledge, from one knowledge structure to another. The system for systematic inference must be capable of dynamic creation and destruction of the facts as well as the propagation of the fillers from one fact to another according to the encoded rules in the knowledge base (Park 1992, pp. 7-9). This problem of chaining such inference steps and the accurate propagation of bindings along the inference chain has recently received considerable attention (see, Touretzky & Hinton, 1988; Smolensky, 1987; Dolan & Dyer, 1988; S & A, 1990, 1992).

The system proposed by S & A (1990) can represent knowledge expressed in the form of 'facts' and 'rules', and is able to determine whether a given 'question' can be answered or not by facts and rules encoded in the system. These capacities are possible because the system maintains and propagates several variable bindings along inference chains using a synchronized temporal activation.

The fundamental features of this temporal synchrony solution lie in a separation of one period of oscillation cycle into several phases, and in the use of phase-sensitivity nodes. On the basis of these properties, dynamic variable binding is represented by argument nodes and their constant, sharing the same phase in any particular period of oscillation during a reasoning process. Consequently, the system is able to represent composite structures dynamically and systematically to propagate them in such a way that other composite structures are generated. In short, S & A propose a neurally plausible solution to the variable binding problem and describe a system that can represent a large body of systematic knowledge and perform some class of inferences rapidly.[7]

In order to compute the solutions in only one flow of spreading activation, the dependencies between knowledge parts must be 'acyclic' in this connectionist network. That is, the information flow has to pass through a directed acyclic graph (DAG) where each node in the graph is a processing element and each connection is a well-connected link ('hardwired')[8]. This constraint, however strong, allows the capture of a broad range of common-sense reasoning situations. Among other things, it allows restricted types of 'causal reasoning', for instance, reasoning about actions and events where there is no circularity. It also allows reasoning with definitional knowledge of concepts/terms (Ajjanagadde, 1990). This is due to the encoding of predicates, individual concepts and rules such as:

V x,y,z [*give*(x,y,z) ==> *own*(y,z)],

V x,y [*own*(x,y) ==> *can-sell*(x,y)],

V x,y [*buy*(x,y) ==> *own*(x,y)],

where links between arguments reflect the correspondence between the arguments on the antecedent and consequent sides of the rules.

Therefore, S & A (1990, 1992) describe how a connectionist system with simple elements can encode millions of facts and rules involving variables and polyadic predicates. This system can perform several kinds of inferences (forward and backward reasoning) in 'hundred milliseconds', even, interacting with a type hierarchy (IS-A; Mani & Shastri, 1991).

In the system proposed, the time spent in an inference chain (the propagation of the rhythmic patterns of activity) depends only on the length of an inferential chain. But it is independent from the number of rules and facts encoded in a knowledge base (though the space requirement is linear in the size of the knowledge base).

Nevertheless, there are three kinds of fundamental constraints acknowledged by S & A (1993) which come out as a natural consequence of adopting a temporal solution:

- The number of different entities permitted to participate simultaneously in dynamic variable binding is bounded by biologically motivated values (cf., Stenning, 1992).

- The problem of dealing with multiple instantiations (two dynamic facts) to the same predicate simultaneously (see, Mani & Shastri, 1991).

- Some forms of rules are not supported in the proposed system for backward and forward reasoning[9].

The first limitation is a natural outcome of using temporal synchrony to solve the dynamic variable binding problem since each entity of a predicate should occupy one phase in a period of oscillation in order to avoid cross-talk problem. So, the maximum number of different entities allowed at one time in the proposed mechanism cannot exceed the number of phases (psychologically, S & A justify this value in relation with 7 +/- 2, a commonly accepted measure of short term-memory capacity). S & A (1993) justify that this basic constraint is a reasonable assumption in reflexive reasoning:

> if we restrict ourselves to a single episode of reasoning, understanding a few sentences, or observing a simple scene. But, what happens when the agent is participating in a dialogue, reading a newspaper, or scanning a complex scene where the total number of significant entities exceeds the number of distinct phases that can coexist. In such situation the set of entities in 'focus' must keep changing constantly with entities shifting in and out of focus in a dynamic manner. Identifying the mechanisms that underlie such internal shifts of attention and cause the system's oscillatory activity to evolve smoothly so that new entities start firing in a phase while entities presently firing in a phase gradually 'release' their phase, remains a challenging open problem (p. 47).

In addition, for S & A (1993), there must exist constraints on the class of reasoning that may be performed in a reflexive manner. Not surprisingly, cognitive agents can perform only a limited class of inferences with extreme efficiency. Nevertheless, the proposed system controls only the argument-filler's adequacy, but leaves unexplained the activation control issue, that is, the question about why we perform just some particular inferences, and not others, in a given situation.

In the last section, we will try to show how this open problem could be solved by incorporating some proposals of Focus Theory. First, we will introduce briefly the Focus Theory. In our opinion, this theory can help to explain the two problems just mentioned: focus shifting and inference control during anaphora resolution.

3. A representation of discourse focus

Computational theories and models for use of referring expressions try to determine the entity (or entities) which a pronoun or definite description refers to. Both the problem of specifying the range of possible referents a description makes available, and the problem of choosing between different possibilities have been worked out from two approaches. In one of them, questions about how referring expressions interact with attentional state are considered primary (Grosz, 1981; Sidner, 1983; Grosz, Pollack & Sidner, 1989; Grosz & Sidner, 1990). In the other, (Hobbs, 1986), referent identification is subsumed by more general processes of inference.

In the first approach, the concepts of focus and the process of focusing (Grosz, 1978; Grosz, 1981; Grosz & Sidner, 1986) have played central roles in the treatment of definite descriptions, used both for first reference and as anaphoric noun phrases. Accordingly, focusing is defined as the movement of the focus of attention of the discourse participants as the discourse progresses (task-dialogue). Two levels of focusing, global and local, have been identified (Grosz & Sidner, 1986). Global focusing is modelled by a stack of focus spaces and it affects the use and interpretation of definite descriptions. Focusing at the local level is modelled with centers and centering. The center of a given segment is an element of the attentional state, and at the beginning of each new segment a new center is introduced. Centering affects the use and interpretation of pronouns.

Each individual space on the global focus stack contains representations of the entities that the participants focus on during a corresponding discourse segment, as well as the discourse-segment purpose. The entities currently in focus (that is, in some space on the focus stack) are the primary candidates for referents of definite descriptions; they are also the source of implicitly focused entities (a phrase may refer to an item related to something in a current focus space; for example, "the cover" may be used to refer to the cover of a book when the book is in focus). The set of entities in the global focus also provides constraints on the content of subsequent definite descriptions; for example, a speaker must include enough descriptions to distinguish the entity to which he wishes to refer from other entities in focus (Grosz, Pollack & Sidner, 1989, pp. 445-446).

Thus, focusing is an active process. As a dialogue progresses, the participants shift their focus to new entities or to new perspectives on entities previously highlighted by the dialogue. Focus can be understood in terms of the ease of accessibility of the potential antecedents and the inference control patterns[10]. In this sense, it should be easier to retrieve the representations currently in focus than to retrieve the ones stored in long-term memory. This fact has clear consequences concerning the facility for reference resolution[11].

Furthermore, focus plays an important role during inference performance as well. Murray, Klin & Myers (1993) show that the consequences of events are inferred when they are extremely predictable and strongly in focus. They suspect that propositions which are both in the text and in the knowledge base are activated in parallel, and yield either a backward or a forward inference, depending on whether the cause of the focal event has to be found in something that has already happened in the text, or in some future action or event that is highly predictable given the text and the reader's general knowledge. For instance, Duffy (1986) contains some passages where the protagonist had just been served soup when the train "screeched to a stop". In this example, subjects were tested with the probe word "spill", and no evidence was found indicating that the inference (that the soup spilled) was active. However, in Duffy's material, the sentence focusing on the soup and the sentence stating that the train suddenly stopped were separated by a sentence in which the protagonist reached for the salt. That is to say, the soup was no longer in focus when the train halted and the failure to activate the targeted inference is not surprising (Murray et al., 1993, p. 471). Given this psychological evidence, the system needs a mean of highlighting those elements in its knowledge base that correspond to the entities currently being focused on and must be able both to use this highlighting (e.g., to interpret and generate descriptions) and to change it appropriately as the dialogue progresses. In addition, during retrieval and deduction operations, this highlighting enables the system to access firstly to the more relevant information.

The encoding of task information also plays a role in shifting focus. In addition to highlighting those items 'explicitly focused on' by the dialogue participants, by placing them on focus spaces, the focusing mechanisms access differentially to certain information associated with these items. Concretely, the sub-actions and objects involved in a task are 'implicitly focused' on whenever that task is focused on (Grosz, 1981, p. 95).

This way, the procedures for interpreting anaphoric definite noun-phrases and implicitly focused references choose items between those in global focus. Focusing techniques predict the anaphoric referent of a definite noun-phrase. Also, they reduce the necessary search to find the representation of a previously mentioned entity -those entities that aren't in focus aren't available, and those in focus and associated with the more immediate subtask are searched before less salient ones.

During anaphora interpretation, confirming the expected focus often requires some computation and may be very complex. Focusing also simplifies the inference processes offering a restriction to how inferences are performed: Focus techniques "bind" the pronoun to the specification of the focus and then look for an inference chain that supports the resulting sentence. A contradiction may be reached, which indicates that the expected focus must be rejected (and it is necessary to see whether some of the items in the focus-stack comes close to matching the anaphora and, if so, which is the closest). In this respect, focusing simplifies the inference process because it specifies the initial and the final propositions that the inferring processes use, and it governs which inference can be taken back if some contradiction results (Sidner, 1983, p. 289).

To sum up, the importance of focusing, both on the interpretation and on the generation of definite descriptions comes from the highlighting function it serves. By separating those items currently highlighted from those that are not, focusing provides a boundary for interpretation in providing a small set of items from which to choose. For generation purposes, the boundary circumscribes those items from which the entity being described must be distinguished, and thus provides some means for determining when a description is sufficiently complete. It should be noted, however, that discourse focus theory has some shortcomings related to computational implementation ['architecture'] on the one hand, and to the fact of not having been experimentally tested, on the other[12]. In the next section, we discuss how focus could be integrated in the above 'reflexive' reasoner system and examine our main proposals concerning the problems mentioned in the previous section. In our opinion, some insights can be obtained integrating focus theory in S & A's reflexive reasoning system.

4. Discussion

Our limited knowledge of how common-sense reasoning -deduction, plausible reasoning, planning and plan recognition- can be dealt with in a computational system, has been a source of problems for the attempts to model language processing (Grosz, 1980). As we have seen in the second section, a different representation of each argument and predicate together with an explicit encoding of inferential chain seems essential if a system is applied to the large number of dynamic binding, between argument predicates, that results from these rule applications.

S & A (1990, 1992) propose a solution to the variable binding problem (Lange & Dyer, 1989; Smolensky, 1987) and show that it leads to the design of a connectionist reasoning system that can represent systematic knowledge involving n-ary predicates and 'variables', and perform a broad class of reasoning with extreme efficiency. This solution to the variable binding problem maintains and propagates a large number of bindings 'simultaneously', as long as the number of different entities participating in the bindings during any given episode of reasoning remains bounded. Reasoning in the proposed system is the transient but systematic flow of 'rhythmic' patterns of activation, where each 'phase' in the variable binding is represented as the synchronous firing of the appropriate argument and constant nodes. A fact behaves as a temporal pattern matcher that becomes 'active' when it detects that the bindings corresponding to it are present in the system's pattern of activity. Finally, rules consist of interconnection patterns that propagate and transform rhythmic patterns of activity. Reflexive reasoning behaves in this system like a commonsense reasoning. It needs millions of rules and facts encoded and its 'length' of inference chain is not very large. The time taken for a chain of inference in the proposed system is linear [O(d)] to the diameter of the inferential dependency network that contains the query predicate. As S & A (1990) affirm, it is independent of the entire number of rules and facts encoded in the system: all relevant rules form an inferential dependency network and facts are associated with predicates in the rules. As a result, the long-term knowledge base of the proposed system is considered as a set of disconnected inferential sub-networks.

In spite of its high efficient inference, this system needs "something more" to account for the sort of automatic inferences people usually executes. This system achieves most of his efficacy due to the type of pre-representation used. But, is it possible to establish a connection between different inferential sub-networks? One of the strengths of human reasoning, in understanding 'fictional' texts, relies on the ability to represent richly specified scenarios about a small number of individuals (such as the LRRH story). When humans do not have pre-existent representations of the individuals involved, they depend on their representations of the predicates and their logical relations. Human exhibit the extreme facility of reflexive reasoning when reading texts about concepts whose interrelations they understand, such as little girls and wolves. Indeed, coherent text understanding depends upon an enormous number of inferences from the suppressed premises of general knowledge.

Some authors have suggested the strategy of extending S & A's system so that it can exploit its mechanism for inferencing about Long Term Memory-represented constants in reasoning about stories (Stenning, 1992). Similarly, we analyse how this system could be complemented with the above described Focus theory (like a separate mechanism based on medium term memory structures) to deal with inference control and anaphora interpretation.

The target of applying a "meta-reasoning" tool like focus (context) to inference control is to perform inferences that have some 'justification' (or, 'utility' for a correct understanding), and to avoid fallacious inferences in a given process of reasoning. This idea is supported by the fact that humans guess that fitness in a model ('prediction'), and the executed inferences seem to be confirmed in this model. Likewise, in Mental Models Theory, the representation is a source of highly accessible information that supports both the judgement of salience and the process of drawing automatic inferences.

In the S & A's system, the only activation control is the specified argument correspondence depending upon the features of the argument-fillers involved in a given situation. Thus, it captures type restrictions that argument-fillers must satisfy to execute a rule. It may also indicate type preferences and provide a graded measure of a rule's applicability in such a situation (S & A, 1992, pp. 7-8). However, given a case where there are two different predicates that could be inferred from an activated one (even if the arguments are shared, or are the same for all predicates), this activation control is unable to decide which one of the two predicates should be chosen. How could focus help us to solve this problem?

Considering this sort of reasoning as the computation of truth-in-a-model, focus could constrain the inference boundaries to avoid unnecessary inferences in a particular process of reasoning. A possible solution in the above system could be to establish "focus-inhibitory-links" between different predicates to control the activation flow. But, this solution could be a bit complicated, because we would have to connect each inhibitory link with a different focus due to the local representation used by the reasoner system.

On the other hand, one of the first demonstrations that referential processes have a direct impact on sentence comprehension, came from Haviland & Clark (1974). They showed that reading time depends, at least in part, on the reader's ability to recover antecedents for the definite anaphors in the sentence. They demonstrated that readers spent more time on the sentence "The beer was warm" when it followed "Mary unpacked some picnic supplies", rather than "Mary unpacked some beer". To explain the difference, they argued that in the former case the reader has to draw a time-consuming *bridging inference* to establish a putative antecedent for the referential expression "The beer" (e.g., some "beer" associate with the "picnic" supplies) before comprehension is achieved.

This referential process involves searching back through a representation of potential antecedents to discover one that matches. If the representation has no matching antecedent, then one must be inferred. Nevertheless such an account raises questions about the matching process itself and the form of the antecedent representation. One problem concerns the range of the search domain and how it can be limited to recover the immediate antecedent. Here, attentional factors may play an important role to explain in a general way the importance of recency as well as how it can be overridden by other factors such as topicalization (Garrod & Sanford, 1990, p. 467).

As we have seen in Focus Theory, the attentional state of discourse structure[13] reflects the focus of attention of discourse participants as the discourse progresses. It is modelled as a (pushdown) stack of focus spaces, one for each segment of the discourse. Focus spaces contain representations of the discourse-segment purpose and the entities referred to in the segment. The stack grows when discourse-segments are introduced and shrinks when their intentions are satisfied. A new space is 'pushed' onto the stack when a new segment is introduced into the discourse, and that a space 'pops out' when its purpose is satisfied. The focus space model of attentional state constrains processing as the discourse unfolds. Those entities and purposes represented in the stack are the most salient at this point of the discourse[14].

Therefore, attentional state explains anaphora resolution because an antecedent can remain activated in working memory by any of the following causes:

1. intermediate utterances between anaphora and its antecedent continue appealing to it;

2. the antecedent is not mentioned in intermediate utterances but, it remains associated to the primary theme in the mental model that the reader dynamically constructs;

3. the antecedent corresponds to a central or accessible concept in the schematic knowledge available to the reader, so that this can activate such antecedent through a thematic inference[15].

In the present case, this is also difficult for the above mentioned system. However, this antecedent activation for subsequent anaphora could be achieved by providing determined entities remain with some activation degree through time. S & A (1993) discuss a possibility where the system is composed of several modules (such as the perceptual, linguistic, and reasoning modules), because representing all distinct entities of these modules that should be active at a given time requires a number of phases. This way, each module has its own phase distribution and frequency of oscillation, so that each module may maintain binding involving about ten entities.

Given that, how do modules communicate with each other in order to exchange variable bindings information consistently? For communication between modules, they appeal to Aaronson's (1991) work on a connectionist interface. This interface allows two phase-based modules, each with its own phase structure, to exchange binding information from a source module to a target module quickly (S & A, 1992, p. 48).

We think about the possibility of building a composite system containing a connectionist interface between modules to perform a wide range of inference with high efficiency. To account for the control of inference, this solution sets some constraints in discourse interpretation, and similarly for anaphora resolution. Shastri[16] thinks that it is possible to incorporate the rules of focus shifting as an additional module to his reflexive reasoner (see also Aaronson, 1991). This suggestion involves the employment of one stack (not too deep) for global focus[17], so that a new constraint [*in-focus*(x)] is added for a given reasoning process. This way, the entity in focus helps to solve some anaphora cases and to control inference activation. Considering it as an additional module, it could encode the focusing rules seen above. Grosz's (1986) attentional state contains information about objects, properties, relations, and discourse intentions that are most salient at any given point. It essentially summarizes information from previous utterances that becomes crucial for processing subsequent utterances, obviating the need for keeping a complete history of the discourse. It is inherently dynamic, recording objects, properties, and relations that are salient at each point in the discourse[18].

An alternative solution is to incorporate focus within the same reasoner system, without appealing to an additional module. For this purpose, biologically motivated constraints on the Focus system as a whole could be applied to the number of tokens held at any time and the number of role or type mappings that they can sustain. As units are binary, we could differentiate the activation of focused entities by means of collecting the amount of arguments that an entity has instantiated during the process of understanding[19]. So, entities recently mentioned and entities mentioned several times along the text would constitute the focused elements within this conception of the problem.

Consequently, it is possible, with the help of some additional mechanism such as focus, to design a system that combines forward as well as backward reasoning. Such a system would be capable of (i) representing incoming information and making predictions based on this information by using its long-term knowledge, and (ii) generating explanations for, and testing the consistency of, incoming information by referring to its long-term knowledge. Nevertheless, if reasoning presupposes the existence of some semantic justification, like discourse focus, that controls which operations constitute a correct inference, it should be required for a connectionist reasoner system, not to only be able to represent structured knowledge, but also able to operate upon such representations systematically, so that it would dynamically 'generate' new representations of structured objects [learning]. Humans probably perform inferences jumping steps in the inference chain (due to previous experiences) thus shortening the inferential chain length and the required time.

During text processing, readers have some domain knowledge to be able to build event representations. For example, knowledge about properties of entities that produce the event, the kind of relationships in which these entities can enter, and how this relationship can be represented. In highly familiar domains, this knowledge remains easily available, providing the construction of mental models that help the execution of automatic inference and anaphoric reference interpretation.

To sum up, the system proposed by Shastri & Ajjanagadde (1990, 1992) could be extended to deal with some open problems and to account for the sort of inferences executed in a given situation. However, its implementation is complicated to some extent due to its local representation. Some insight could be gained using distributed systems, such as DISCERN (Miikkulainen, 1990). This system show some interesting properties to account for the sort of inferences and anaphora resolution (definite descriptions; see, Garrod & Sanford, 1990) in scripts. Finally, it would be interesting try to explain our two problems discussed here in a distributed system, mainly because of the role that learning issues could play.

Notes

† *Acknowledgements.* A first draft of this paper was delivered at the International Colloquium on Cognitive Science (ICCS), San Sebastian, May, 1993, and then published as the technical report 'Reflexive Reasoning, Focus Theory and Anaphora Resolution', Logic Seminar Report LPHS-EHU-02.3, Dpto. de Lógica y Filosofía de la Ciencia, UPV/EHU. The present paper is an extended and modified version of the latter one. We would like to thank K. Stenning, Nam Seog Park and the friends of the Human Communication Research Centre of the University of Edinburgh for their kind comments to earlier versions of this paper. We also like to thank specially L. Shastri, M. Dascal, Tim Smithers, Victor Sánchez de Zavala and the anonymous referees of THEORIA for their helpful comments. Agustín Vicente and Stuart Silvers helped us to reduce our linguistic bugs, many thanks. The support of the Research Project Number PS92-0041 of the Ministerio de Educación y Ciencia (DGICYT) is gratefully acknowledged.

1 See, among others, McKoon & Ratcliff (1986) and O'Brien, Duffy & Myers (1986) for coherence, and Myers, Shinjo & Duffy (1987) for causal relations.

2 See, for instance, McKoon & Ratcliff (1986), O'Brien, Shank, Myers & Rayner (1988) and Garrod, O'Brien, Morris & Rayner (1990).

3 Garnham (1979) showed that verbs as well as nouns can be instantiated. *Fried* is a better recall cue than *cooked* for: "The housewife cooked the chips".

4 As we will see below, an important insight has been to increase the syntax of the representation language so that the 'form' (i.e., syntactic structure) of the representation 'reflects' directly the 'inferential structure' of knowledge, utilized also by some connectionist models.

5 This approach has left to pragmatics the explanation of some problems no satisfactorily solved, such as speaker/writer intentions. However, it is necessary to take into account the 'symbiotic' relationship between the nature of representation, inference effectivity and the computational architecture in which the computations are placed.

6 See, Shastri & Ajjanagadde (1990).

7 See S & A (1993) for a more refined version of this system.

8 See S & A (1990).

9 See also, Park (1992, pp. 59-60); for some difficulties in dealing with variable arguments and existentially quantified variables.

10 See, for example, Grosz (1977); Hudson, Tannenhaus & Dell (1986); Sanford & Garrod (1981).

11 Pronoun reference is demonstrably sensitive to focus. Anderson, Garrod & Sanford (1983) showed that pronominal reference to characters which depend upon being situated in a scene are more difficult to process when that scene is cued as completed. (An example of dependence is that a "waiter" is dependent on a restaurant scene.) However, main characters, not dependent upon a particular

scene, can easily be referred to by a pronoun after a change from the scene in which they have just appeared (see, also, Garrod & Sanford, 1990).

12 It seems that other processes also take part in anaphor interpretation (see Nelson, Stevenson & Stenning, 1992); also, there are some directions in which focus mechanisms must be extended to treat general problems posed by focusing and definite descriptions in dialogue (see, Grosz, 1978b, 1981): for instance, the interaction between different cue types (i.e., linguistic) and, the inexact matching issue (i.e., metaphoric and metonymic references (see, Gibbs, 1990).

13 Also, there are a linguistic structure and an intentional structure (Grosz & Sidner, 1986).

14 Constraints may use this fact to stipulate when various linguistic expressions can be us ed and to help to determine when a given discourse-segment purpose (intention) may dominate or satisfaction-precede another (Grosz, Pollack & Sidner, 1989, pp. 442-443).

15 This issue of maintaining possible antecedents and how they bind with a possible subsequent anaphor has been also mentioned in script distributed models (role-binding cases) (Miikkulainen, 1990)

16 Personal communication.

17 This limitation depends on the number of instantiations of a same predicate within the system. Concerning this issue, Mani & Shastri (1992) propose a solution for several instantiations (e.g., three) of a same predicate.

18 Focus theory has tried to explain the two above mentioned problems: focus shifting and inference control during anaphora resolution. In this theory, the global component of attentional state is modelled by a set of *focus spaces*; changes in attentional state are modelled by a set of transition rules that specify the conditions for adding and deleting spaces.

19 That is to say, entities with more activation through time or more referred entities along the text.

BIBLIOGRAPHY

Aaronson, J.: 1991, 'Dynamic fact communication mechanism: A connectionist interface', in *Proceedings of the Thirteenth Conference of the Cognitive Science Society*, N. J., Lawrence Erlbaum.

Ajjanagadde, V.: 1990, 'Reasoning with function symbols in a connectionist system', in *Proceedings of the Twelfth Conference of the Cognitive Science*, N. J., Lawrence Erlbaum.

Anderson, A., Garrod, S., & Sanford, A.J.: 1983, 'The accessibility of pronominal antecedents as a function of episode shifts in narrative texts', *Quarterly Journal of Experimental Psychology* 35A, 427-440.

Anderson, J.R.: 1983, 'A Spreading Activation Theory of Memory', *Journal of Verbal Learning and Verbal Behavior* 22, 261-295.

Brow, G. & Yule, G.: 1983, *Discourse analysis*, Cambridge University Press.

Dolan, C.P., &, Dyer, M.G.: 1988, 'Parallel Retrieval and Application of Conceptual Knowledge', CSD-880010, UCLA, Computer Science Department.

Duffy, S.A.: 1986, 'Role of expectations in sentence integration', *JEP: Learning, Memory and Cognition* 12, 208-219.

Feldman, J.A.: 1982, 'Dynamic connections in neural networks', *Bio-Cybernetics* 46, 27-39.

Feldman, J.A.: 1989, 'Neural Representation of Conceptual Knowledge', in L. Nadel, L.A. Cooper, P. Culicover and R.M. Harnish (eds.): *Neural Connections, Mental Computation*, MIT Press.

Feldman, J.A. & Ballard, D.H.: 1982, 'Connectionist models and their properties', *Cognitive Science* 6 (3), 205-254.

Garnham, A.: 1979, 'Instantiation of verbs', *Quarterly Journal of Experimental Psychology* 31, 207-214.

Garnham, A.: 1981, 'Anaphoric reference to instances, instantiated and non-instantiated categories: A reading-time study', *British Journal of Psychology* 72, 377-384.

Garnhan, A.: 1985, *Psycholinguistics. Central topics*, Cambridge, Cambridge University Press.

Garrod, S., Freudenthal, D. & Boyle, E.: 1992, 'The Role of different types of anaphor in the on-line resolution of sentences in a discourse', Submitted to *Journal of Memory and Language.*

Garrod, S., O'Brien, E.J., Morris, R.K., & Rayner, K.: 1990, 'Elaborative Inferencing as an Active or Passive Process', *J.E.P.: Learning, Memory and Cognition* 16, No. 2, 250-257.

Garrod, S.C. & Sanford, A.J.: 1990, 'Referential processes in reading: Focusing on roles and individuals', in D.A. Balota, G.B.F. d'Arcais and K. Rayner (eds.): *Comprehension Processes in Reading*, Hillsdale, N.J., Lawrence Erlbaum Associates, pp. 465-486.

Gibbs, R.W.: 1990, Comprehending Figurative Referential Descriptions, *J.E.P.: Learning, Memory and Cognition* Vol. 16, No. 1.

Grosz, B.: 1977, 'The representation and use of focus in dialogue understanding', *Technical note* 15, SRI International Artificial Intelligence Center.

Grosz, B.J.: 1978, 'Focusing in Dialogue', in Waltz (ed.): *Theoretical issues in Natural Language, Processing-2,* Urbana, IL, 96-103.

Grosz, B.J.: 1980, 'Utterance and Objective: Issues in Natural Language Communication', *AI Magazine* 1, 11-20.

Grosz, B.J.: 1981, 'Focusing and description in natural languages dialogues', in Aravind Joshi, Bonnie Webber and Ivan Sag (eds.): *Elements of discourse understanding*, Cambridge, Cambridge University Press.

Grosz, B.: 1986, 'A theory of discourse structure', in *Proceedings of the Eighth Annual Conference of the Cognitive Science Society*, pp. 432-441.

Grosz, B.J., Joshi, A.K. & Weinstein, S.: 1983, 'Providing a unified account of definite noun phrases in discourse', in *Proceedings of the 9th International Joint Conference on Artificial Intelligence*, Cambridge, M.A.

Grosz, B.J., Pollack, M.E. & Sidner, C.L.: 1989, 'Discourse', in Posner, M.I. (ed.): *Foundations of Cognitive Science*, Bradford, MIT Press.

Grosz, B.J. & Sidner, C.L.: 1986, Attention, Intentions, and the Structure of Discourse, *Computational Linguistics*, 12 (3), 175-204.

Grosz, B.J. & Sidner, C.L.: 1990, 'Plans for Discourse', in Philip R. Cohen, Jerry Morgan and Martha E. Pollack (eds.): *Intentions in Communication*, Cambridge, Mass, MIT Press.

Haviland, S.E. & Clark, H.H.: 1974, 'What's new? Acquiring new information as a process in comprehension', *Journal of Verbal Learning and Verbal Behaviour* 13, 512-21.

Hobbs, J.R.: 1986, 'Resolving pronoun references', in Grosz, B.J., Sparck Jones, K. and Lynn Webber, B. (eds.): *Readings in natural language processing*, Morgan Kaufmann Publishers, Inc.

Hudson, S.B., Tanenhaus, M.K., & Dell, G.S.: 1986, 'The effect of the discourse center on the local coherence of a discourse', in *Proceedings Eighth Annual Conference of the Cognitive Science Society*, Amherst, MA, LEA, pp. 96-101.

Kintsch, W.: 1988, 'The role of knowledge in discourse comprehension: A construction integration model', *Psychological Review* 95 (2), 163-182.

Lange, T.E. & Dyer, M.G.: 1989, 'Frame Selection in a Connectionist Model of High-Level Inferencing', *Technical Report UCLA-AI-89-05*, UCLA, Computer Science Department.

von der Malsburg, C.: 1986, 'Am I thinking assemblies?', in G. Palm and A. Aertsen (eds.): *Brain Theory*, Springer-Verlag.

Mani, D.R. & Shastri, L.: 1991, 'Combining a Connectionist type Hierarchy with a Connectionist Rule-Based Reasoner', in *Proceedings of the Thirteenth Conference of the Cognitive Science Society*, N.J., Lawrence Erlbaum.

Mani, D.R. & Shastri, L.: 1992, 'A connectionist solution to the multiple instantiation problem using temporal synchrony', in *Proceedings of the Fourteenth Conference of the Cognitive Science Society*, Erlbaum.

McKoon, G., & Ratcliff, R.: 1986, 'Inferences about predictable events', *Journal of Experimental Psychology: Learning, Memory and Cognition* 12, 82-91.

Miikkulainen, R.: 1990, 'A Neural Network Model of Script Processing and Memory', *Technical Report UCLA-AI-90-0*, UCLA, Artificial Intelligence Laboratory, Computer Science Department.

Minsky: 1975, 'A frame for representing knowledge', in Winston, P.H. (ed.): *The Psychology of Computer Vision*, N. York, McGraw-Hill, pp. 211-277.

Murray, J.D., Klin, C.M., & Myers, J.L.: 1993, 'Forward inferences in narrative text,' *Journal of Memory and Language* 32, 464-473.

Myers, J.L., Shinjo, M., & Duffy, S.A.: 1987, 'Degree of causal relatedness and memory', *Journal of Memory and Language* 26, 453-465.

Nelson, A.W.R., Stevenson, R.J. & Stenning, K.: 1992, 'Parallelism in pronoun comprehension', *HCRC/92*, University of Edinburgh.

O'Brien, E.J.: 1987, 'Antecedent Search Processes and the Structure of Text', *J.E.P.: Learning, Memory and Cognition* 13 (2), 278-290.

O'Brien, E.J., Duffy, S.A. & Myers, J.L.: 1986, 'Anaphoric Inference During Reading, *J.E.P.: Learning, Memory and Cognition* 12 (3), 346-352.

O'Brien, E.J., Shank, D.M., Myers, J.L., & Rayner, K.: 1988, 'Elaborative inferences during reading: do they occur online?' *JEP: Learning, Memory and Cognition* 14, 410-420.

Park, N.S.: 1992, 'A Connectionist Representation of Rules, Facts, and Dynamic Bindings of Variables' *MSc, Department of Artificial Intelligence*, University of Edinburgh.

Quillian, M.R.: 1968, 'Semantic memory', in M. Minsky (ed.): *Semantic information processing*, Cambridge, Mass., MIT Press.

Sanford, A.J.: 1990, 'On the nature of text driven inference', in D.A. Balota, G.B. Flores d'Arcais and K. Rayner (eds.): *Comprehension Processes and Reading*, Hillsdale, NJ, LEA, pp. 515-533.

Sanford, A.J., & Garrod, S.C.: 1981, *Understanding Written Language: Explorations in Comprehension Beyond the Sentence*, Chichester, John Wiley.

Sanford, A.J. & Moxey, L.M.: 1991, 'Notes on Plural Reference and the Scenario-Mapping Principle in Comprehension', *HCRC/RP-23.*

Schank, R.: 1972, 'Conceptual Dependency: A Theory of Natural Language Understanding', *Cognitive Psychology* 3, 55-63.

Schank, R. & Abelson, R.P.: 1977, *Scripts, Plans, Goals and Understanding*, N. Jersey, Lawrence Erlbaum.

Schubert, L.K.: 1989, 'An Episodic Knowledge Representation for Narrative Texts', in *Proceedings of KR-89*, Toronto.

Shastri, L.: 1990, 'Connectionism and the computational effectiveness of reasoning', *Theoretical Linguistics* 16 (1), 65-87.

Shastri, L. & Ajjanagadde, V.: 1990, 'From simple associations to systematic reasoning: A connectionist representation of rules, variables and dynamic bindings', *Technical Report: MS-CIS-90-05*, Computer and Information Science Department, University of Pennsylvania.

Shastri, L. & Ajjanagadde, V.: 1993, 'From simple associations to systematic reasoning: A connectionist representation of rules, variables and dynamic bindings using temporal synchrony', *Behavioral and Brain Sciences* 16 (4).

Sidner, C.L.: 1983, 'Focusing in the Comprehension of Definite Anaphora', in Brady, M. and Berwick, R.C. (eds.): *Computational Models of Discourse*, MIT Press, Cambridge, Massachusetts.

Smolensky, P.: 1987, 'On variable binding and the representation of symbolic structures in connectionist systems', *Technical Report CU-CS-355-87*, Department of Computer Science, University of Colorado at Boulder.

Stenning, K.: 1992, 'Measuring Shastri's phase-locked reflexive reasoning against human performance', Forthcoming in *Behavioral and Brain Sciences*, September, 1993.

Stenning, K. & Oaksford, M.: 1992, 'Rational Reasoning and Human Implementations of Logic', *Research Paper HCRC/RP-28.*

Stenning, K., Shepherd, M. & Levy, J.: 1988, 'On the Construction of Representations for Individuals from Descriptions in Text', *Language and Cognitive Processes* 3 (2), 129-164.

Touretzky, D.S. & Hinton, G.E.: 1988, 'A Distributed Connectionist Production System', *Cognitive Science* 12 (3), 423-466.

Jesus Ezquerro is Ph.D. in Philosophy and Lecturer of Logic and Philosophy of Science at the Universidad del Pais Vasco/EHU. Its research interests are Philosophy of Science, Language and Mind, and has worked and conducted research on the foundations of Cognitive Science and on Cognitive Architectures. He has published articles in several journals and books, and has edited *Cognition, Semantics and Philosophy* (1992), and *Philosophy and Cognitive Science: Categories, Consciousness and Reasoning* (1996).

Mauricio Iza is Ph.D. in Psychology and has currently a Post-Doc Position (MEC) for doing research on computational models of language processing in the *Institut de Recherche en Informatique de Toulouse* (France). He has published articles in several journals on text processing, inferential models and Anaphora resolution, being the most recent 'A hybrid architecture for text comprehension' in *Pragmatics and Cognition* (1995).