

# UN ANALISIS DEL CONCEPTO DE COGNOSCIBILIDAD DESDE LA SEMANTICA DE MUNDOS POSIBLES†

*(An Analysis of the Notion of Knowability in the Field of  
Possible Worlds Semantics)*

Javier VILANOVA\*

Manuscrito recibido: 1998.4.3.

Versión final: 1999.5.3.

\* Departamento de Lógica y Filosofía de la Ciencia, Facultad de Filosofía, Universidad Complutense de Madrid, 28040 Madrid. E-mail: vilanova@eucmax.sim.ucm.es

BIBLID [0495-4548 (1999) 14: 36; p. 413-429]

RESUMEN: Las nociones epistémicas modales se definen como aquellos conceptos epistémicos que, como el de cognoscibilidad o el de indudabilidad, incluyen una nota modal. Según se defiende en este trabajo, la semántica de mundos posibles y algunas de sus extensiones (especialmente las llevadas a cabo para lógica temporal, lógica epistémica y lógica condicional) son instrumentos adecuados para deshacer el nudo de las intensionalidades superpuestas en estas nociones especialmente esquivas al análisis. Para mostrarlo, se proporcionan una serie de análisis sucesivos de la noción de cognoscibilidad que a partir de una interpretación naïf van salvando una serie de presuposiciones, problemas y paradojas hasta dar con una análisis que se presume satisfactorio.

Descriptores: epistémico, modal, cognoscibilidad, mundos posibles.

ABSTRACT: *"Modal epistemic notions" are those epistemic concepts which in some way or another has a modal element. These modal epistemic notions, although they could appear intuitively clear, they turn out to be particularly obscure, slippery, when one subjects them to a formal analysis. In this paper we will try to show that possible world semantics (and its extensions for epistemic, temporal and conditional logic) is an appropriate instrument for the explanation of these notions. Four successive analysis of the notion of "knowability" are given, ranging from a naïve account to an analysis that gets to the bottom of the problem.*

Keywords: *modality, possible worlds semantics, epistemic, knowability.*

## SUMARIO

1. Introducción
  2. Análisis I
  3. Análisis II
  4. Análisis III
  5. Análisis IV
  6. Conclusiones
- Bibliografía

### 1. Introducción<sup>1</sup>

El tema de este artículo son aquella clase de nociones epistémicas que de un modo u otro incluyen un elemento modal. Conceptos como "cognoscibilidad" o "inverificabilidad", cuyo ámbito temático es el de la epistemología o la teoría del conocimiento, poseen una indudable nota modal; es por ello que por lo que utilizaré a partir de ahora el término genérico "*nociones epistémicas modales*" para referirme a ellos. Son conceptos especialmente esquivos al análisis, ya que en las oraciones donde aparecen se entrecruzan dos tipos de contextos intensionales: contextos epistémicos y contextos modales. Confluyen de este modo dos problemas tradicionales de la filosofía analítica, como son los de las actitudes epistémicas y las nociones modales.

Desde Frege es ya conocida la dificultad que entraña la representación y el análisis formal de verbos que expresan actitudes epistémicas tales como "saber", "creer", "ignorar" o "sospechar",... El mismo Frege advirtió que los enunciados donde aparecen predicados como los anteriores violan en muchas oraciones el principio de extensionalidad (también denominado "principio de substitutividad") según el cual dos expresiones con la misma referencia pueden intercambiarse en una oración compleja sin que cambie el valor de verdad de la oración. Compárese "Arquímedes *sabía* que el agua podía convertirse en gas" con "Arquímedes *sabía* que el H<sub>2</sub>O podía convertirse en gas". El mismo fenómeno se produce en aquellos contextos afectados por verbos o expresiones modales. Compárese "Es necesario que la raíz cuadrada de nueve sea igual a tres" con "Es necesario que el número de mosqueteros de la novela de Dumas sea igual a tres".

En las nociones que vamos a discutir en este artículo se cruzan los problemas producidos por los dos tipos de contextos intensionales citados: modal y epistémico. El principal problema a este respecto, como mostraremos más tarde, es que el modo en que las dos aspectos intensionales se combinan en, por ejemplo, la palabra "cognoscibilidad" no es el de una simple superposición. El análisis descubre que no basta, por ejemplo, con disponer una función de sapiencia, "se sabe que", y una función de posibilidad, "es posible que", y superponer una sobre la otra de tal modo que "es cognoscible que p" signifique lo mismo que:

"Es posible que (se sabe que (p))"

Estamos, al revés, ante una noción compleja en la que, por otra parte, resulta fácil contrabandear premisas filosóficas o teorías materiales como

si formaran parte del significado de la palabra. Tales hechos exigen que afilemos al máximo nuestros instrumentos de análisis, no sólo para desenredar el ovillo de las intensionalidades superpuestas, sino también para producir una definición filosóficamente neutral respecto a doctrinas epistemológicas como el realismo, intuicionismo, idealismo, escepticismo, etc.

El propósito de este artículo es mostrar que la semántica de mundos posibles es un instrumento adecuado para el análisis formal de los conceptos modales epistémicos. En efecto, es bien conocida la inmensa capacidad expresiva de esta semántica, que rangua desde el nivel de abstracción de la semántica universal de Montague hasta el grado de concreción y complejidad de los desarrollos de la misma efectuados para la lógica condicional, la lógica monomótona o la lógica borrosa. Por otra parte, desde la semántica de mundos posibles se han efectuado satisfactorios acercamientos tanto a las modalidades (especialmente a partir de Carnap (1947) y Kripke (1963)), como a las actitudes epistémicas (a partir del trabajo pionero de Hintikka (1962)), lo que induce a confiar en la posibilidad de la combinación de ambos enfoques que los conceptos epistémicos modales requiere. Este carácter "integrador" de la semántica de mundos posibles será aprovechado al máximo en las siguientes páginas; así, introduciremos definiciones y conceptos no sólo de la semántica mundiposibilística de la Lógica Modal y la Lógica Epistémica, sino además de semánticas para Lógica Temporal y Lógica Condicional.

En este artículo concentraremos nuestra atención en una noción especialmente importante y problemática: la noción de cognoscibilidad. Para ello, partiremos del análisis naif que presentamos en el párrafo anterior, mostraremos cuáles son sus defectos, y a continuación iremos proporcionando una serie de análisis sucesivos, cada uno de los cuales resolverá problemas que presentaban los anteriores e introducirá un nivel de claridad más alto. Como se verá, este plus de claridad requerirá a su vez un mayor nivel de complejidad, ya que los últimos análisis incorporan un aparato formal muy sofisticado. Dado que el objetivo de este artículo es más de explicación conceptual que de invención matemática, las definiciones formales serán dadas a un nivel general y más intuitivo que técnico.

## 2. *Análisis I*

En adelante hablaremos de un lenguaje objeto  $L$  formado por enunciados  $p, q$ , etc. Hablaremos también de un conjunto de índices o mundos

$U = \{u, v, x, y, z, \dots\}$ , una función de interpretación  $\# : L \rightarrow \text{Pot}(U)$  que asigna a cada enunciado  $p$  del lenguaje un subconjunto de  $U$  formado por todos y sólo aquellos mundos donde  $p$  es verdadera. Simbolizemos con la letra  $S$  el predicado "se sabe que..." o "...es sabido", y simbolizemos con la letra  $N$  el predicado "...es cognoscible". Nuestro primer análisis corresponde a la aproximación *prima facie* que presentamos en la introducción: " $C_p \equiv \diamond S p$ ". En semántica de mundos posibles " $\diamond p$ " es verdadero si y sólo si  $p$  es verdadero en alguno de los elementos de  $U$ . Desembocamos así en nuestro primer análisis:

$$(A.1) C_p \equiv \exists v \in U (v \in \#(S p))$$

Enseguida surge una cuestión con respecto al análisis 1. ¿Qué representan los índices que componen  $U$ ? Siguiendo una práctica habitual en semántica de mundos posibles, podemos interpretar el universo  $U$  como el conjunto de las alternativas al mundo actual que son compatibles con las leyes de la lógica. En otras palabras,  $U$  estaría compuesto por *todos* los mundos *lógicamente* posibles. No es necesario más que reflexionar unos instantes para darse cuenta de que la noción de mundo lógicamente posible es demasiado débil. Cuando pensamos en situaciones alternativas en las cuales podría ser conocido algo que desconocemos actualmente, con seguridad pretendemos que en tales situaciones se conserven otros aspectos importantes de la realidad además de la validez de las verdades lógicas. Aspectos relativos a nuestro sistema cognitivo o la estructura de nuestro conocimiento deben, sin duda alguna, ser preservados (por ejemplo, un mundo donde seamos omniscientes no es un mundo admisible). Lo mismo ocurre respecto a la constitución ontológica del mundo actual (un mundo sin tiempo tampoco debería ser tenido en cuenta) y, al menos si no queremos trivializar nuestra noción de cognoscibilidad, rasgos estructurales como las leyes físicas o la persistencia de entidades señaladas. Consideremos entonces la noción de *mundo epistémicamente posible*, entendiendo por tal un mundo que preserva todos los rasgos que acabamos de enunciar, y limitemos  $U$  al conjunto de mundos epistémicamente posibles. Con esta modificación el análisis resulta más plausible.

Adolece, sin embargo, de un importante defecto. Para exponerlo apelaremos a la conocida Paradoja de la Cognoscibilidad, que constituye una muestra perfecta de este problema<sup>2</sup>. Esta paradoja ha sido discutida en el contexto de la polémica entre el realismo y el antirrealismo. Según la postura intuicionista, algo es el caso si y sólo si es posible llegar a saberlo;

es decir,  $Cp \leftrightarrow p$ . Si aceptamos la existencia de un hecho no conocido actualmente ( $p \ \& \ \neg Sp$ ), entonces llegamos a una contradicción<sup>3</sup>:

- 1	$\forall p (Sp \rightarrow p)$	
- 2	$\forall p \forall q (S(p \wedge q) \rightarrow Sp \wedge Sq)$	
- 3	$\forall p (p \rightarrow \diamond Sp)$	
- 4	$\exists p (p \wedge \neg Sp)$	
[ 5	$(r \wedge \neg Sr)$	
6	$\diamond S(r \wedge \neg Sr)$	M.P. 3, 5
7	$\diamond (Sr \wedge S \neg Sr)$	M.P. 2, 6
8	$\diamond (Sr \wedge \neg Sr)$	M.P. 1, 7
9	$\perp$	Contradicción
10	$\perp$	$\exists \exists$ 4, 5-9

¿Qué es lo que va mal con la paradoja de la cognoscibilidad? Aunque la demostración anterior ha sido utilizada por los realistas como arma arrojada contra el principio de que toda verdad es cognoscible, pienso que este uso es indebido. Lo que la paradoja muestra es no que vaya algo mal con el principio de que toda verdad es cognoscible, sino otra cosa muy distinta: que la definición de cognoscibilidad que se propone no es neutral respecto a doctrinas epistemológicas materiales, ya que es incompatible lógicamente con algunos principios epistemológicos cuya validez no es una cuestión *a priori* (y que no son, por lo tanto, analíticos). En efecto, si se observa el argumento con ecuanimidad, incluso para el realista constituye una paradoja que " $p \leftrightarrow Cp$ " y " $\exists p (p \ \& \ \neg Sp)$ " sean contradictorios. De hecho, la premisa 4 constituye ya un enunciado incómodo para ambas partes. Aunque tengamos razones para pensar que existe un enunciado que es verdadero pero que no conocemos, sin embargo no podemos señalar un enunciado en concreto  $p$  tal que  $p$  es verdadero pero no lo es  $Sp$ . Se trata, de hecho, de la conocida paradoja de Moore, que enuncia la imposibilidad de encontrar una proposición verdadera que no es sabida. Y esto es imposible porque para exhibir una proposición tal debemos saber que es verdadera, lo cual contradice el supuesto de que no lo sabemos.

Ahora bien, lo que la Paradoja de Moore prohíbe es que podamos exhibir en un momento determinado una proposición verdadera que no conocemos. Sin embargo, las personas acostumbramos a exhibir ejemplos de enunciados, que si bien ahora sabemos su valor de verdad, *en algún momento pasado* éramos ignorantes al respecto. Por ejemplo, hasta hoy lo ignoraba todo sobre las gardenias pero hoy ha venido a parar a mis manos un manual

de jardinería y me he enterado de que son plantas que aman el calor, así que puedo decir a reguardo de toda contradicción: las gardenias son plantas de clima cálido pero ayer lo ignoraba.

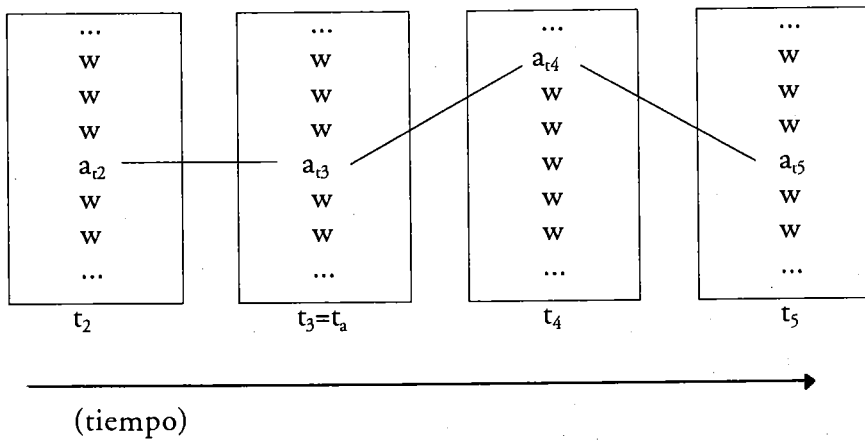
Lo que intento sugerir es que, para dar cuenta del concepto de cognoscibilidad, debemos desligar el predicado de sapiencia del momento actual y adoptar una perspectiva temporal. Para ello, hemos de relativizar *S* a momentos temporales, y hablar no de conocimiento a secas sino de conocimiento en un momento determinado. Ello supone convertir *S* en un predicado diádico, "...es sabido en el instante...", para lo cual necesitamos introducir en nuestro modelo instantes temporales. Lo que es cognoscible incluirá, además de lo que es sabido en el momento actual, lo que será sabido en el futuro. Además, en la medida en que necesitamos desmarcar claramente el hecho de que se sepa que *p* en el momento determinado del hecho de que *p* sea verdadera en el momento actual, precisamos introducir en nuestra semántica nuevos elementos que permitan interpretar de modo realista el predicado "*S*". Esto es lo que haremos en el análisis siguiente, adoptando para ello instrumentos de la Lógica Epistémica de Hintikka.

### 3. Análisis II

En la ya clásica obra *Knowledge and Belief* Hintikka propone representar un corpus de conocimiento (el conjunto de proposiciones que una persona o comunidad conoce) mediante un conjunto de mundos posibles, a saber, el conjunto de mundos lógicamente compatibles con lo que el individuo que detenta ese corpus de conocimientos sabe. Denominemos a ese conjunto de mundos posibles "*situación epistémica*" (el término es mío). Una vez que diseñamos la situación epistémica correspondiente a un corpus de conocimiento o un estado de sapiencia, es fácil proporcionar una interpretación para los predicados epistémicos. Una proposición *p* es sabida si es el caso que *p* es verdadera en todo mundo perteneciente a la situación epistémica. Una proposición *p* es compatible con nuestro conocimiento si es el caso que *p* es verdadera en algún mundo a la situación epistémica. Hintikka utiliza parámetros para individuos cognoscentes, convirtiendo el predicado de sapiencia en un predicado diádico con dos argumentos: la proposición sabida y el sujeto cognoscente. En nuestro análisis prescindimos de introducir el parámetro para el individuo cognoscente, que aparece implícitamente recogido en el predicado *S*, y cuyo uso sólo resulta necesario a la hora de analizar algunos casos específicos de actitudes epistémicas que

se escapan de la perspectiva general adoptada aquí (aquella situación en que una persona sabe que otra persona sabe determinada cosa y otras similares).

En nuestra semántica precisamos introducir una situación epistémica para cada instante temporal; la situación epistémica compuesta por los mundos epistémicamente posibles que son compatibles con nuestro conocimiento en ese instante. Para ello, definamos un *marco epistémico temporal* como una estructura  $\langle U, T, <, \{\equiv_t\}_{t \in T}, \{a_t\}_{t \in T}, t_a \rangle$  tal que  $U$  es un conjunto no vacío de mundos,  $T$  es un conjunto no vacío de índices numéricos (los instantes temporales),  $<$  es un orden lineal definido sobre  $T$ ,  $\{\equiv_t\}_{t \in T}$  es una familia de relaciones de equivalencia en  $U$ ,  $\{a_t\}_{t \in T}$  es un conjunto de elementos designados de  $U$ , y  $t_a$  es un elemento señalado de  $T$ . Dado un índice temporal determinado  $t$ ,  $a_t$  es el mundo actual en  $t$  el instante  $t$ , y el conjunto de mundos  $u$  tales que  $a_t \equiv_t u$  constituye la situación epistémica en el momento  $t_a$ . Podemos representar el marco epistémico temporal a través de un diagrama en el que cada situación epistémica que determina el copus de conocimiento en el mundo actual en el instante  $t$  es simbolizada como una "rebanada de mundos", y las rebanadas aparecen una detrás de otra conectadas por la flecha del tiempo:



El predicado  $S$  posee, ahora, dos argumentos: la proposición  $p$  que es conocida, y el momento  $t$  en que se sabe que  $p$ . Su definición es la siguiente:

$$(1) Spt \equiv \forall u \in U (a_t \varepsilon_t u \rightarrow u \in \#(p))$$

Nuestro segundo análisis de la noción de cognoscibilidad surge de forma inmediata. Decimos que p es cognoscible en el momento actual si existe algún momento en el futuro en el que p será conocido<sup>4</sup>. Escribamos "t≥t'" si y sólo si no es el caso que t'<t. Entonces:

$$(A. II) C_p \equiv \exists t \geq t_a (Spt)^5$$

Es fácil ver que la paradoja de la cognoscibilidad ya no se produce. Reescribamos ahora la demostración utilizando los conceptos recién definidos. La premisa 3 de la deducción dice que si p es verdadera entonces existe un momento en el futuro en que p será sabida. La premisa 4 dice que existe una proposición y un instante temporal en el que dicha proposición no es sabida.

- 1  $\forall p \forall t (Spt \rightarrow p)$
- 2  $\forall p \forall q \forall t (S(p \wedge q)t \rightarrow Spt \wedge Sqt)$
- 3  $\forall p (p \rightarrow \exists t > t_a (Spt))$
- 4  $\exists p \exists t (p \wedge \neg Spt_a)$
- 5  $(r \wedge \neg Srt_a)$
- 6  $\exists t > t_a (S(r \wedge \neg Srt_a)t)$  M.P. 3, 5
- 7  $\exists t > t_a (Srt \wedge S(\neg Krt_a)t)$  M.P. 2, 5
- 8  $\exists t > t_a (Srt \wedge \neg Srt_a)$  M.P. 1, 7
- ...

Un modelo de las premisas en nuestra semántica sería cualquier interpretación en la que una proposición p es verdadera en el mundo actual, falsa en algún mundo de la situación epistémica actual, y verdadera en todo mundo perteneciente a alguna situación epistémica futura. Además, es fácil ver por qué la deducción de la contradicción no se produce. Consideremos la línea 7 de la anterior argumentación. El primer miembro de la conjunción dice que hay un momento en el futuro en el que sabemos q. El segundo miembro de la conjunción dice que aunque en el presente no sabemos q, existe un momento futuro en el que sabemos que entonces (en el pasado), no era sabida. Como se ve, el paso de 7 a 8, en el que se hace uso de la premisa 1, no conduce a ninguna contradicción: el enunciado 8 afirma que r se sabe en un momento t pero no se sabe en el momento t<sub>a</sub>, lo cual no resulta contradictorio. Adviértase que las líneas 6 y 7 son verdaderas en virtud del



principio de cognoscibilidad y el supuesto 4. En efecto, según la premisa 3, si  $p$  es verdadera en  $a_{ta}$ , entonces hay un  $t' \geq t_a$  tal que  $Cp_{t'}$ ; de donde se sigue que existe un momento  $t'$  en el futuro en el que  $p$  es verdadera en todo  $u$  tal que  $a_t \approx_{t'} u$ , lo cual hace verdadero el primer miembro de la conjunción en 7. Teniendo en cuenta que  $\neg Spt_a$  es verdadero si existe un  $u'$  donde  $p$  es falso y tal que  $a_{ta} \approx_{ta} u'$ , el supuesto 5 hace verdadero al segundo miembro de la conjunción en 7.

#### 4. Análisis III

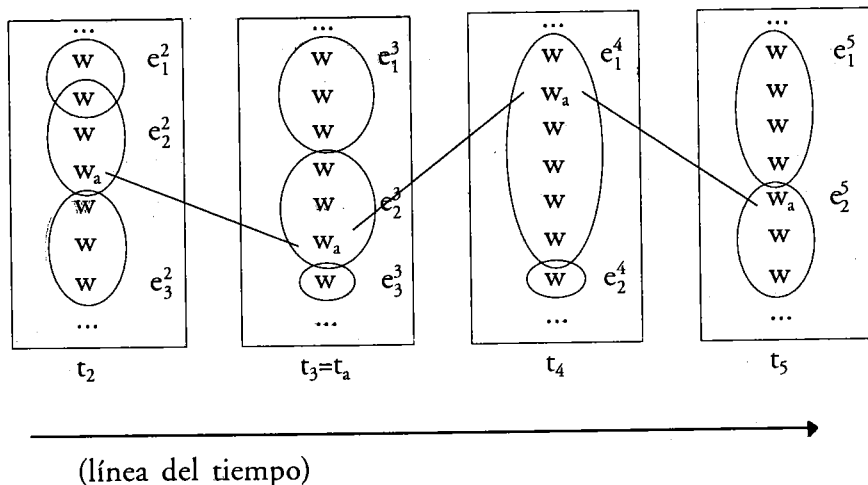
El análisis anterior constituye una explicación bastante plausible de la noción de cognoscibilidad. Sin embargo, todavía esconde en su interior una presuposición no justificada. En efecto, el análisis II depende del supuesto de que todo lo que de un modo u otro podemos llegar a saber "será" efectivamente conocido en algún momento del futuro. Llamemos a este principio "*principio del optimismo gnoseológico*", y examinemos su validez. Si bien en el contexto de la investigación matemática y de la teoría del sujeto ideal, el principio del optimismo gnoseológico resulta completamente plausible, cuando pasamos a considerar teorías empíricas (físicas, biológicas, historiográficas, etc.), el principio del optimismo gnoseológico empieza a causar problemas. Por un lado, si consideramos comunidades científicas humanas y fácticas en vez del sujeto ideal (incluso aunque consideremos una hipotética comunidad científica universal, que acoja a todos los seres humanos de todos los tiempos pasados, presente, y futuros) la idea que todo lo cognoscible llegará a ser sabido corresponde antes al nivel deóntico que al nivel óntico. Por poner un caso extremo, si de aquí a unos años tiene lugar la tan temida tercera guerra mundial y la humanidad llega a su trágico fin, resulta claro que el principio del optimismo no se cumple. Y no hay, que sepamos, ninguna razón de peso que omita la posibilidad del cataclismo.

Además, hay muchas cuestiones empíricas que casi con toda seguridad nunca podremos llegar a resolver. Por ejemplo, tomemos la debatida cuestión de si Santiago Apóstol estuvo o no en Galicia. La ausencia de evidencias tanto en pro como en contra nos hace pensar que nunca podremos establecer los hechos. Por supuesto, uno puede aducir que siempre pueden aparecer en algún momento del futuro nuevas evidencias que resuelvan la cuestión, pero todavía quedarían infinidad de cuestiones históricas por establecer, y el principio del optimismo gnoseológico no tiene ninguna base sobre la que asegurar que aparecerán evidencias para absolutamente todas las

cuestiones. Otro ejemplo, más cerca de la vida cotidiana, puede ayudar a explicar en qué falla el principio del optimismo gnoseológico. No cabe duda que a las doce en punto del mediodía del 3 de Junio de diez años atrás yo sabía perfectamente si estaba sentado o de pié. Sin embargo, en este momento ignoro completamente ese dato, y (salvo una muy fortuíta concatenación de casualidades), nunca llegaré a resolver mi duda.

Hasta cierto punto, problemas como el anterior se resuelven si en vez de definir la cognoscibilidad como conocimiento en el futuro, la definimos como conocimiento en cualquier momento pasado, presente o futuro (es decir,  $\exists t$  (Cpt)). En el caso del apostol Santiago, aunque nosotros y nuestro herederos no sabremos nunca si visitó Galicia, en su época si fue sabido, al menos fue sabido por el propio Santiago y sus allegados. De cualquier manera, tampoco esta modificación en la definición de cognoscibilidad basta para cubrir todos los casos. Tomemos el ejemplo clásico de la moneda: saco una moneda del bolsillo, la lanzo al aire, la recojo con la mano y la vuelvo a meter en el bolsillo sin haber mirado si salió cara o cruz. Dado que la moneda o bien salió cara o bien salió cruz, parece claro que estamos ante una proposición verdadera (ya sea "la moneda salió cara", o "la moneda salió cruz") que nunca será sabida o fue sabida. Sin embargo, la cuestión es decidible, al menos "fue" decidible en su momento, cuando tenía la moneda en la mano antes de llevarla al bolsillo.

Ejemplos como el de la moneda hacen preciso el recurso a una modalidad no temporal. Es decir, hemos de tener en cuenta no sólo situaciones epistémicas que efectivamente se han dado en el futuro, sino también situaciones epistémicas que pudieron darse en algún momento del pasado o podrán darse en algún momento del futuro. En nuestro ejemplo anterior, debemos tener en cuenta la situación en que abro la mano y descubro que es cara, la situación en que desconozco de qué valor es la moneda, e incluso la situación en que no me he fijado en el objeto que he arrojado al aire e ignoro si es una moneda o un botón. Para ello, introduciremos en nuestro marco semántico no una sino una variedad de situaciones epistémicas en cada rebanada temporal, correspondientes a los posibles corpus de conocimiento en el momento  $t$  al que corresponde la rebanada. Gráficamente, el marco modal que nos proponemos definir se parece a la siguiente figura:



Para ello, definimos una función  $\Delta: T \rightarrow \text{Pot}(\text{Pot}(U))$ , que asigna a cada instante temporal  $t$  el conjunto de situaciones epistémicas posibles en  $t$ ,  $\{e_a^t, e_1^t, e_2^t, \dots\}$  (siendo  $e_a^t$  la situación epistémica actual: aquella que corresponde al conocimiento que de facto se posee en el instante  $t$ ), y donde cada  $e_i^t$  es un conjunto de mundos posibles. Según el análisis III, un enunciado  $p$  es cognoscible si hay una situación epistémica posible donde se conoce  $p$ , es decir, un  $e_i^t$  tal que todos los mundos que lo componen son mundos donde  $p$  es verdadero:

$$(A. III) C_p \equiv \exists t \in T \exists e_i^t \in \Delta(t) [\forall v \in e_i^t (v \in \#(p))]$$

### 5. Análisis IV

El análisis anterior capta muy acertadamente lo que intuitivamente entendemos como la posibilidad de conocer una determinada verdad. Lamentablemente, tampoco es inmune a paradojas. Tomemos el ejemplo de la moneda. Supongamos que la moneda salió cara en el mundo actual, y que desconocemos que la moneda fue cara, pero que hay una situación epistémica distinta a la actual donde la moneda salió cruz, y donde es sabido que la moneda fue cruz. En tal caso llegamos al extraño resultado de que en el mundo actual es cognoscible que la moneda fue cruz.

El ejemplo anterior nos advierte de una carencia del marco semántico del análisis III. Hemos dicho que los  $\{e_a^t, e_1^t, e_2^t, \dots\}$  correspondientes a un momento  $t$  son las situaciones epistémicas alternativas en el momento  $t$ , pero no hemos especificado suficientemente las condiciones que debe cumplir un subconjunto de  $U$  para ser considerado una alternativa plausible a nuestro estado de conocimiento actual. En otras palabras, precisamos explicitar con más detalle cómo surgen las situaciones epistémicas en tanto que alternativas plausibles a un estado de conocimiento determinado. Consideremos otra vez el ejemplo de la moneda. A la hora de concebir situaciones epistémicas alternativas a la actual permitimos modificar algunos de los hechos que se dan en el mundo actual, por ejemplo, consideramos una alternativa apta aquella en la que abrimos la mano que esconde la moneda (algo que no se da en la situación real). Sin embargo, sólo consideramos aquellas modificaciones que tienen directa relación con las acciones que está en nuestra mano llevar a cabo en orden a adquirir evidencias acerca de hechos que desconocemos. No es una alternativa aceptable, por ejemplo, aquella en que mis pantalones están hechos de una tela invisible, porque no es el resultado de ninguna acción dirigida a la producción de evidencias. Como tampoco es admisible una situación epistémica en la que monto en la maquina de Wells y retrocedo en el tiempo para observar qué salió en la moneda, ya que no es algo que esté en mi mano hacer.

Consideremos, pues, un conjunto  $\mathcal{R}_t$  que representa el repertorio de acciones dirigidas a la producción de evidencias que están disponibles en el momento  $t$ .  $\mathcal{R}_t$  incluirá acciones que suponen una transformación directa del entorno físico, como puede ser desenterrar un yacimiento arqueológico o construir un acelerador de electrones. También estarán en  $\mathcal{R}$  acciones que no suponen cambios drásticos en el medio material sino tan sólo procesos perceptivos o cognitivos, tales como llevar a cabo una determinada observación o concentrar la atención en un aspecto del fenómeno. Cada una de esas acciones puede ser representado mediante un enunciado del tipo "A hace X", siendo A el agente epistémico, y X el acto que se lleva a cabo. Consecuentemente, el repertorio de acciones,  $\mathcal{R}$ , será introducido en nuestro modelo como un conjunto de enunciados.

¿Cómo se generan las nuevas situaciones epistémicas a partir del repertorio de acciones epistémicas? La idea es que debemos considerar aquellas (y sólo aquellas) situaciones epistémicas que surgen como resultado de la realización de una o más constituyentes de  $\mathcal{R}_t$ . Para introducir esta idea en nuestro marco semántico haremos uso de una noción definida originalmente por Stalnaker mientras desarrollaba una semántica de mundos posibles

para interpretar enunciados condicionales. Esta es la noción de "mínima revisión de un mundo". La tesis de Stalnaker es que, dado un enunciado condicional  $p \rightarrow q$  cuyo antecedente es falso en el mundo actual, para conocer su valor de verdad recurrimos a los mundos máximamente similares al mundo actual donde  $p$  es verdadero. Así, comprobamos si en esos mundos  $q$  es verdadero, y de este modo atribuimos un valor de verdad al condicional. Independientemente del uso que hace Stalnaker de ella, la noción de revisión mínima parece estar involucrada en muchos procesos semánticos. En nuestro caso, la conexión entre las situaciones epistémicas alternativas y la situación epistémica actual en un momento determinado puede ser aprehendida a través del concepto de revisión mínima o de máxima similitud. En efecto, para cada situación epistémica posible en un momento determinado  $e_t^i$  sea  $a_t^i$  el mundo que corresponde a la situación que es el caso en  $e_t^i$  (el mundo actual en  $e_t^i$ ). Lo que necesitamos es considerar sólo aquellas situaciones  $e_t^i$  tales que  $a_t^i$  es uno de los mundos máximamente similares a  $a_t$  de aquellas en los que se ha aplicado alguna de las acciones de  $\mathfrak{R}_t$ .

Sea  $\Phi$  una función que asigna a cada mundo  $v$  y cada repertorio  $\mathfrak{R}_t$  el conjunto de mundos  $u$  tales que  $u$  es una revisión mínima de  $v$  para uno o varios de los enunciados de  $\mathfrak{R}_t$ . Nuestro análisis final es el siguiente:

$$(A. IV) C_p \equiv \exists t \in T \exists e_t^i \in \Delta(t) [a_t^i \in \Phi(a_t, \mathfrak{R}_t) \& \forall v \in e_t^i [v \in \#(p)]]$$

El lector observará que, mediante el recurso a la noción de revisión mínima, salvamos la paradoja de la moneda que afectaba al análisis anterior. En efecto, la noción de revisión mínima nos permite expresar que cada " $a_t^i$ " es precisamente el mundo que hubiera tenido lugar si se hubiera llevado a cabo alguna de las acciones epistémicas de  $\mathfrak{R}_t$  en  $a_t$ . En el caso de la moneda, esto nos asegura que no cambiamos en ningún  $a_t^i$  respecto a  $a_t$ , más que aquellas consecuencias de la acción de abrir la mano donde tenemos la moneda, y entonces, si la moneda es cara en  $a_t$  también es cara en  $a_t^i$  (es completamente implausible que nuestra acción epistémica pueda cambiar el hecho de que la moneda es cara).

Remarquemos que el análisis IV es neutral con respecto a concepciones epistemológicas materiales tales como el realismo o el intuicionismo. Esto quiere decir que las posturas intuicionista, realista, escéptica, idealista, etc. pueden ser expresadas en la semántica simplemente introduciendo restricciones adicionales que reflejen los principios de cada teoría episte-

mológica. Incidentalmente: en el marco modal que hemos proporcionado la paradoja de la cognoscibilidad no rige ni siquiera en ejemplos extremos como el de la moneda. En efecto, el principio que dice que todo enunciado verdadero es cognoscible no conduce a ninguna paradoja. " $p \rightarrow Cp$ " dice que, si  $p$  es verdadero en el mundo actual en el momento presente, entonces existe un momento  $t$  y una situación epistémica  $e_t^i$  tales que  $p$  es verdadero en el instante  $t$  en todo mundo perteneciente a  $e_t^i$ . Este principio resulta coherente con la hipótesis de que existe una proposición verdadera que no es conocida en el momento actual  $e$ , incluso, con la hipótesis más fuerte de que existe una proposición verdadera que no es ni será nunca conocida. En el caso de la moneda, sabemos que hay una situación epistémica en la que se sabe una proposición no sabida en la situación epistémica actual; a saber, o la proposición que dice que la moneda salió cara, o la proposición que dice que la moneda salió cruz. El hecho de que no sepamos cuál de las dos es la proposición en cuestión no contradice el hecho de que existe una proposición verdadera no sabida, como tampoco interfiere con el hecho de que dicha proposición resulta cognoscible.

## 6. Conclusiones

Nuestro análisis final resulta ser, me temo, mucho más complicado de lo que a primera vista el lector podría esperar. El *definiens* es una expresión compleja que incluye tres cuantificadores (dos existenciales y uno universal) y hace uso de varias nociones que a primera vista no parecen estar presupuestas en la simple noción de cognoscibilidad: mundo posible, situación epistémica, índice temporal, repertorio de acciones epistémicas, revisión mínima. Una objeción posible, ante esta alta complejidad, surgiría de la sospecha de que durante los sucesivos análisis se haya producido una inflación de teoría, una introducción indiscriminada de conceptos y formalismos innecesarios. A mi modo de ver, esto no se ha producido en absoluto. La complejidad está presente ya en la misma noción de cognoscibilidad, cuya aparente cogencia es sólo eso, apariencia, y cuyo significado es más oscuro y resbaladizo de lo que la aproximación *prima facie* sugiere. Según mi firme creencia, ninguno de los conceptos mencionados en la lista anterior sobra. De hecho, el análisis IV es un correlato teórico de la siguiente definición informal de cognoscibilidad, que a la postre es la que aquí se ha mostrado como adecuada:

*(A.4) Una proposición p es cognoscible si en algún momento pasado, presente o futuro existía (existe, existirá) un procedimiento experimental u observacional que de ser seguido hubiera dado lugar al conocimiento de p.*

A su vez, la definición anterior no es más que una explicación más completa y precisa de la definición del análisis I, que, como ya hemos dicho, puede parafrasearse informalmente de este modo:

*(A.1) Una proposición p es cognoscible si es posible saber que p.*

Los análisis II y III pueden verse como pasos intermedios entre el primer definiens, vago y equívoco, y el último, claro y determinado. Informalmente, el análisis II puede parafrasearse así:

*(A.2) Una proposición p es cognoscible si en algún instante en el futuro se sabrá que p.*

Mientras que el análisis III puede enunciarse, de manera informal, como:

*(A.3) Una proposición p es cognoscible si existe un momento pasado, presente, futuro en el que se podía (puede, podrá) saber que p.*

A mi modo de ver, el hecho de que todas estas definiciones puedan ser producidas dentro de la semántica de mundos posibles (y lo que es más importante, el hecho de que cada una de ellas haya sido inspirada desde la propia semántica de mundos posibles y los defectos que ésta ha detectado en la precedente) es un resultado alentador que muestra la potencia que este aparato formal posee en orden a dar cuenta no sólo de otras nociones epistémicas modales, sino en general para dar cuenta de contextos intensionales en los que se juntan dos tipos de intensionalidades diferentes.

*Notas*

† Este trabajo ha sido elaborado al amparo de los proyectos de investigación XUGA 20506B96 y DG PB95-0863.

- <sup>1</sup> Este artículo representa la continuación de una investigación comenzada conjuntamente con los profesores Concepción Martínez Vidal y José Miguel Sagüillo Fernández-Vega de la Universidad de Santiago de Compostela, y cuyos resultados aparecen recogidos en Martínez, Sagüillo y Vilanova (1997). La autoría de muchas de las ideas recogidas aquí corresponde, por lo tanto, a los Profesores Martínez y Sagüillo, sin cuya ayuda nunca hubiera existido este trabajo, y a quienes agradezco efusivamente su amable permiso para incluir sus ideas aquí. También quiero agradecer al anónimo informante de esta revista, por las sugerencias y correcciones efectuadas.
- <sup>2</sup> La Paradoja de la Cognoscibilidad aparece formulada por primera vez en Hart (1979), quien a su vez hace uso de un principio formulado originalmente en Fitch (1963), y que aquí denominamos "Teorema de Fitch".
- <sup>3</sup> Adviértase que en las líneas 7 y 8 hacemos uso de la regla de lógica modal que permite aplicar el *modus ponens* en contextos modales:  $\forall p \forall q ((p \rightarrow q \wedge \Diamond p) \rightarrow \Diamond q)$ .
- <sup>4</sup> El lector observará que esta definición asume una concepción acumulativa del conocimiento, según la cual todo lo que sabemos ahora se seguirá sabiendo en el futuro. A quien esta concepción le parezca una simplificación excesiva puede sustituir en la definición A. II el " $\exists t \geq t_0$ " por " $\exists t \in T$ ", sin que por ello varíe substancialmente la validez de los argumentos que siguen. En cualquier caso, esta concepción acumulativa será también puesta en entredicho en el siguiente apartado, y sólo se acepta aquí "provisionalmente".
- <sup>5</sup> Adviértase que S y C, dada esta formulación, pertenecen al metalenguaje. Podrían introducirse operadores en el lenguaje objeto L a partir de los cuales es posible definir C y S directamente en L (de un modo similar a como se introducen los operadores "alguna vez en el pasado", "siempre en el pasado" o "alguna vez en el futuro" en la lógica de los tiempos verbales). Sin embargo, la complejidad técnica que ello produce desaconseja seguir este camino al menos el contexto de esta exposición.

*BIBLIOGRAFIA*

- Carnap, R.: 1947, *Meaning and Necessity: a Study in Semantics and Modal Logic*, Chicago, Chicago University Press.
- Cozzo, C.: 1994, 'What can we learn from the paradox of knowability', *Topoi* 13, 71-78.
- Edgington, D.: 1985, 'The Paradox of Knowability', *Mind* 1985, 557-68.
- Fitch, F.B.: 1963, 'A Logical Analysis of Some Value Concepts', *The Journal of Symbolic Logic* 28/2, 135-142.
- Hintikka, J.: 1962, *Knowledge and Belief*, New York, Cornell University Press.
- Kripke, S.: 1963, 'Semantical Considerations on Modal Logic', *Acta Philosophica Fennica* 16, 83-94.
- Martínez, C., Sagüillo, J.M., Vilanova, J.: 1997, 'Fitch's Problem and the Knowability Paradox: Logical and Philosophical Remarks', *Logica Trianguli* 1, 1-22.



- Sagüillo, J.M.: 1994, 'Paradoxical Argumentations', in Van Eemeren et al. (eds.): *Analysis and Evaluation*, Amsterdam, 132.
- Stalnaker, R.C.: 1968, 'A Theory of Conditionals', *American Philosophical Quarterly*, 98-112.
- Thomason, R.H.: 1984, 'Tense and modality', in D. Gabbay and F. Guentner (eds): *Handbook of Philosophical Logic*, vol. II, 136-166.
- Williamson, T.: 1982, 'On the paradox of knowability', *Mind* 96, 256-261.

*Javier Vilanova* se doctoró en Filosofía en 1997 en la Universidad de Santiago de Compostela con una tesis titulada "Similaridad Intermundana y Lógica Condicional". Desde 1997 es profesor asociado en el Departamento de Lógica y Filosofía de la Ciencia de la Universidad Complutense de Madrid. Ha publicado artículos en temas de Lógica Condicional, Epistémica, Filosofía del Lenguaje y de Lógica Filosófica.