CONOCIMIENTO E INCERTIDUMBRE. UNA APROXIMACIÓN DESDE LA LÓGICA DE LAS MODALIDADES EPISTÉMICAS[*]

JOSÉ RAFAEL HERRERA GONZÁLEZ
Dpto. de Historia y Filosofía de la Ciencia, la Educación y el Lenguaje
Facultad de Filosofía. Universidad de La Laguna
Camino de La Hornera s/n. Campus de Guajara
38071- La Laguna. Tenerife
rahego@ull.es

El objetivo primordial de este artículo es analizar diferentes formas de representar el conocimiento falible con ayuda de la semántica de mundos posibles. A modo de introducción, comenzaremos con la presentación de los sistemas epistémicos y doxásticos más importantes, incluyendo uno de los aspectos más destacables de este tema, a saber: el denominado problema de la omnisciencia lógica. A continuación, examinaremos algunos de los métodos más eficaces que se han propuesto para superar la omnisciencia lógica. A nuestro entender, de todo este análisis se desprende que los beneficios de la lógica modal epistémica y doxástica superan sus posibles inconvenientes.

Palabras clave: Conocimiento, creencia, incertidumbre, lógica epistémica y doxástica, problema de la omnisciencia lógica.

The main goal of this paper is to analyze different ways of representing fallible knowledge with the help of possible world semantics. By way of introduction, we'll start with a presentation of the more important epistemic and doxastic systems, including one of the most striking aspects of this issue, namely: the so-called logical omniscience problem. Then, we'll review some of the most efficient methods that have been suggested to overcome the logical omniscience. In our opinion, the lesson to be learned from this analysis is that the benefits of modal epistemic and doxastic logic outweigh its possible disadvantages.

Key words: Knowledge, belief, uncertainty, epistemic and doxastic logic, logical omniscience problem.

[*]Este artículo forma parte del trabajo llevado a cabo en el marco del Proyecto de Investigación *La realidad sin velos* (PI 2003/099), financiado por la Consejería de Educación del Gobierno Autónomo de Canarias, y del que Manuel Liz Gutiérrez es su investigador principal.

"Todos los hombres por naturaleza desean saber"
Aristóteles: *Metafísica*, Libro I

1. Introducción: La lógica de las modalidades epistémicas

La investigación acerca del conocimiento ha constituido, desde los albores mismos de la filosofía occidental, una de las piedras angulares de la reflexión filosófica. Esta preocupación por los temas relacionados con la epistemología se acentuó en la filosofía moderna, y ya para autores como Descartes, Hume o Kant (por citar sólo tres de los más significativos) la teoría del conocimiento, la indagación acerca del alcance y los límites del saber humano, se considera el punto de partida de toda filosofía que pretenda ser rigurosa y basarse en fundamentos seguros.

A este interés por el conocimiento no ha sido ajeno el desarrollo de la lógica, y ya el propio Aristóteles concibe la lógica misma como propedéutica para la ciencia (es decir, para el *saber verdadero*). Asimismo, los distintos métodos propuestos para alcanzar un conocimiento seguro y bien fundado, se han basado en presupuestos lógicos que pretenden garantizar la corrección de los razonamientos empleados.

No obstante, el estudio formal de las nociones de conocimiento y creencia no comienza a plantearse hasta mediados del siglo pasado, momento en que podemos situar los inicios de la lógica epistémica y doxástica, esto es, de la lógica del conocimiento y la creencia. Los primeros pasos de este tipo de lógicas los encontramos en Von Wright (1951, caps. IV y VI) y, sobre todo, en el ya clásico estudio de Hintikka (1962), donde las nociones de saber y creer son tratadas como *conceptos modales*, con lo que poco después su análisis lógico pudo basarse en la semántica kripkeana de mundos posibles[1].

Interpretar el conocimiento y la creencia como conceptos modales supone que, al analizarlos, hemos de concebir varios estados de cosas o mundos posibles de manera simultánea. Así, si afirmamos que "El sujeto i sabe que

[1] De hecho, aunque Hintikka (1962) hace uso de las ideas de *conjunto modelo* y *sistema modelo*, ambas nociones están muy próximas a las de *mundo posible* y *conjunto de mundos posibles*, propias de la semántica desarrollada poco tiempo después por S. A. Kripke.

p" o "El sujeto i cree que p", lo que estamos sosteniendo, desde un punto de vista modal, es que p es el caso en todos los posibles estados de cosas que son compatibles con lo que i sabe o cree (es decir, que p es verdadero en todas las alternativas epistémicas o doxásticas del agente i).

En cuanto a la interpretación intuitiva de los mundos posibles, Hintikka (1967, p.41) nos dice que:

> Consideremos un sistema de lenguaje dado [...] Llamemos una novela completa a todo conjunto de sentencias que sea máximamente consistente, es decir, a todo conjunto de sentencias consistente, pero que pierde esta propiedad tan pronto como se le añade una sentencia nueva. Entonces, un mundo posible es simplemente aquello que queda descrito por una novela completa.

Y, de acuerdo con lo que acabamos de decir, cuando consideremos sentencias que expresen lo que un sujeto *sabe* o *cree*, estas "novelas completas" describirán estados de cosas compatibles con los conocimientos o creencias de dicho sujeto.

En la lógica modal epistémica, por tanto, tomamos como base la semántica kripkeana de mundos posibles. Para ello los sistemas de la lógica modal se adaptan a contextos epistémicos, interpretando la *necesidad* como *conocimiento* o *creencia*, por lo que estas nociones heredarán las características de la necesidad lógica. Además, entre las distintas alternativas epistémicas (o doxásticas) se establece una *relación de accesibilidad* (*o de posibilidad*), por lo que podremos afirmar que, desde una determinada situación, un sujeto puede *considerar posible* (de acuerdo con sus conocimientos o creencias) que un determinado hecho sea o no el caso.

A continuación vamos a presentar los elementos formales de la lógica epistémica proposicional que recogen las ideas que hasta ahora hemos venido considerando de forma meramente intuitiva. Comenzaremos por configurar un lenguaje formal para la lógica epistémica; posteriormente, introduciremos una semántica para la interpretación de este lenguaje (la cual, como hemos adelantado, se enmarca dentro de la semántica kripkeana de mundos posibles); finalmente, expondremos los principales sistemas axiomáticos de la lógica epistémica proposicional, poniendo de relieve algunas de sus características más relevantes[2].

A) *El lenguaje formal*:

El lenguaje de la lógica epistémica proposicional para un conjunto de h agentes está formado por:

a) Un conjunto enumerable de variables proposicionales: p, q, r, etc.

b) $\rightarrow$, $\neg$, $\wedge$, $\vee$ y $\leftrightarrow$ como conectivas primitivas.

c) A, A', A'', etc. como variables metalingüísticas.

d) (,), [,] como signos auxiliares.

e) K ("saber que") y M ("considerar posible que") como operadores epistémicos.

Una *fórmula bien formada* (en adelante, fbf) de la lógica epistémica proposicional será una concatenación de los anteriores signos primitivos, de alguno de los siguientes tipos:

i) Toda variable proposicional es una fbf.

ii) $\neg$A , donde A es una fbf.

iii) (A$\rightarrow$A'), (A$\wedge$A'), (A$\vee$A') y (A$\leftrightarrow$A'), donde A y A' son fbfs.

iv) K_iA ("el agente i sabe que A"), donde A es una fbf.

v) M_iA ("el agente i considera posible que A"), donde A es una fbf.

Nada que no se obtenga de la repetida aplicación de las reglas i-v anteriores es una fbf. Podemos eliminar los paréntesis externos de las fórmulas, siendo posible, además, establecer las siguientes equivalencias entre las conectivas lógicas: A$\rightarrow$A' $\Leftrightarrow$ $\neg$(A$\wedge\neg$A'); A$\rightarrow$A' $\Leftrightarrow$ $\neg$A$\vee$A';

[2] Una presentación exhaustiva de los elementos lógico-formales que configuran la lógica epistémica proposicional, así como una extensa discusión de sus principales ventajas e inconvenientes, puede encontrarse en Meyer y Van der Hoek (1995[1]). Exposiciones de conjunto sobre el tema, breves e interesantes, se presentan en Freund (1995), Meyer (2001) y en Lenzen (2004).

$A \wedge A' \Leftrightarrow \neg(\neg A \vee \neg A')$; $A \leftrightarrow A' \Leftrightarrow (A \rightarrow A') \wedge (A' \rightarrow A)$. En la mayor parte de los casos, cuando se hace referencia a un único agente suele omitirse los subíndices de los operadores epistémicos, de tal modo que escribiríamos únicamente KA y MA.

B) *Semántica de la lógica epistémica proposicional*:

Definimos un modelo de la lógica epistémica como una estructura $M = \langle W, R_1, ..., R_h, v \rangle$, donde:

1) $W \neq \emptyset$, es un conjunto de *estados* o *mundos epistémicamente posibles* (también llamados *alternativas epistémicas*).

2) $R_i \subseteq W^2$ (para i= 1, ..., h), son *relaciones de accesibilidad* entre mundos o alternativas epistémicas, una para cada agente. Las propiedades de las relaciones R_i variarán en función del sistema axiomático de la lógica epistémica proposicional de que se trate. Asimismo, podemos indicar que entre dos alternativas epistémicas w_1 y w_2 se establece una relación de accesibilidad R_i de las dos formas siguientes: $w_1 R_i w_2$ o $(w_1, w_2) \in R_i$.

3) v es una *función de evaluación*, que asigna el valor verdadero (1) o falso (0) a cada fbf en un mundo o estado epistémico; así, siendo F el conjunto de fbfs, v: $F x W \rightarrow \{1,0\}$. Esta función de evaluación cumple las siguientes condiciones, para cualesquiera $w_1, w_2 \in W$, $A, A' \in F$ y variable proposicional p:

i) $v(p,w_1)=1$ o $v(p,w_1)=0$

ii) $v(\neg A,w_1)=1$ syss $v(A,w_1)=0$

iii) $v(A \rightarrow A',w_1)=1$ syss $v(A,w_1)=0$ o $v(A',w_1)=1$

iv) $v(K_i A,w_1)=1$ syss $\forall w_2 \in W / w_1 R_i w_2$, $v(A,w_2)=1$

Esta última cláusula expresa, de manera formal, la interpretación modal de la noción de conocimiento, a saber: *que, en un determinado estado w_1, un agente i sabe que A es el caso, si A es verdadero en todos los mundos* *que son posibles de acuerdo con el conocimiento de que dispone i en dicho estado w_1.*

Asimismo, decimos que una fbf A es *válida* (y lo representamos como $\vDash$A) si, para todo modelo y todo estado epistémico $w \in W$, $v(A,w)=1$.

C) *Sistemas axiomáticos de la lógica epistémica proposicional*:

Los siguientes *axiomas* (o *esquemas de axiomas*) ponen de manifiesto algunas propiedades fundamentales del conocimiento:

Ax1. Un conjunto suficiente de axiomas para derivar todas las tautologías de la lógica clásica de proposiciones.

Ax2. $[K_iA \wedge K_i(A \rightarrow A')] \rightarrow K_iA'$ (Para i=1, ..., h)

Ax3. $K_iA \rightarrow A$ (Para i=1, ..., h)

Ax4. $K_iA \rightarrow K_iK_iA$ (Para i=1, ..., h)

Ax5. $\neg K_iA \rightarrow K_i \neg K_iA$ (Para i=1, ..., h)

En función de esto, los *sistemas axiomáticos* básicos de la lógica epistémica proposicional serán los siguientes:

Sistema mínimo K: Formado por el Ax1 y el Ax2.

Sistema T: Constituido por el Sistema K más el Ax3.

Sistema S4: Constituido por el Sistema T más el Ax4.

Sistema S5: Consta del Sistema S4 más el Ax5[3].

Como *reglas de derivación* para todos estos sistemas axiomáticos contamos con el Modus Ponens (MP), la Regla de Sustitución y la Regla de Generalización de K (Si $\vdash$A, entonces $\vdash K_iA$, para i=1, ...,h). Además,

[3] Es interesante señalar que, en el plano semántico, las *relaciones de accesibilidad* en el Sistema S5 (el sistema más ampliamente aceptado de la lógica epistémica proposicional), son relaciones *reflexivas*, *simétricas* y *transitivas* (es decir, *relaciones de equivalencia*). Decimos que una relación R es *reflexiva* si $\forall w \in W$, wRw; una relación R es *simétrica* si $\forall w_1,w_2 \in W$ ($w_1Rw_2 \Rightarrow w_2Rw_1$); finalmente, una relación R es *transitiva* si $\forall w_1,w_2,w_3 \in W$ ($w_1Rw_2 \wedge w_2Rw_3 \Rightarrow w_1Rw_3$).

disponemos de la siguiente *definición metalingüística*: $M_iA =_{df} \neg K_i\neg A$. Así pues, la definición semántica del operador M será como sigue: $v(M_iA,w_1)=1$ syss $\exists w_2 \in W/w_1R_iw_2$ y $v(A,w_2)=1$.

Los anteriores sistemas axiomáticos K, T, S4 y S5 son consistentes y completos en relación con la semántica recogida en el apartado B anterior[4].

D) *La lógica doxástica. El Sistema axiomático KD45*:

Podemos configurar el lenguaje formal para la *lógica de la creencia* (*lógica doxástica proposicional*), referida a un conjunto de h agentes, reemplazando K (“saber que”) por B (“creer que”) en el lenguaje de la lógica epistémica presentado en el apartado A. Así pues, si A es una fbf, también lo es B_iA (“el sujeto i cree que A es el caso”).

El modelo semántico para la lógica doxástica proposicional se obtiene sustituyendo K_i por B_i en la semántica recogida en el apartado B anterior, con lo que ahora la cláusula iv de la función de evaluación v quedaría así:

$v(B_iA,w_1)=1$ syss $\forall w_2 \in W/w_1R_iw_2$, $v(A,w_2)=1$

Esta última cláusula expresa nuestra interpretación de la creencia como concepto modal, esto es: un sujeto i cree que A es el caso si A es verdadero en todos las *alternativas doxásticas* de i (es decir, en todas los posibles estados de cosas que son compatibles con las creencias de i). Y es que, al igual que ocurre con la lógica epistémica proposicional, en la lógica doxástica tomamos como base los sistemas de la lógica modal, interpretando la necesidad como creencia.

Uno de los sistemas axiomáticos de la lógica doxástica más ampliamente difundidos es el denominado Sistema KD45, que viene a ser una variante del Sistema S5 obtenido al reemplazar en este último K_i por B_i, y sustituyendo el Ax3 por una versión más débil del mismo, el denominado Axioma D:

---

[4] Cf., por ejemplo, las pruebas desarrolladas en Meyer y Van der Hoek (1995[1], pp. 18-9 y 26-8).

AxD. $B_iA \to \neg B_i \neg A$ (Para i=1, ...,h)

Este axioma también puede expresarse así: $\neg(B_iA \wedge B_i\neg A)$, o también: $\neg B_i(A\wedge\neg A)$[5].

El hecho de sustituir el Ax3 por el AxD en el Sistema KD45, se debe a que el primero de estos axiomas expresa que los conocimientos, una vez establecidos, son verdaderos; pero esta propiedad no es aceptable para la creencia, pues ya desde Platón se viene defendiendo que lo que distingue la creencia del conocimiento es precisamente que se pueden mantener creencias falsas, mientras que el conocimiento es siempre acerca de hechos verdaderos. Así, en *Gorgias*, 454d, podemos leer:

> SÓCRATES.- ¿Te parece que saber y creer son lo mismo o que son algo distinto el conocimiento y la creencia?
>
> GORGIAS.- Creo que son algo distinto, Sócrates.
>
> SÓCRATES.- Así es; lo comprobarás por lo siguiente. Si te preguntan: <<*¿Hay una creencia falsa y otra verdadera, Gorgias?>>, contestarías afirmativamente, creo yo.*
>
> GORGIAS.- Sí.
>
> SÓCRATES.- Pero *¿existe una ciencia falsa y otra verdadera?*
>
> GORGIAS.- *En modo alguno*.
>
> SÓCRATES.- *Luego es evidente que no son lo mismo*.
>
> GORGIAS.- Es cierto.

El AxD, por su parte, exige que *las creencias han de ser consistentes*, es decir, que un sujeto no puede mantener creencias que sean contradictorias entre sí. Y si bien no han faltado detractores de esta exigencia de consistencia para las creencias, no es menos cierto que el postular que las creencias de un mismo sujeto epistémico no pueden contradecirse se compadece muy bien con ciertas concepciones de la creencia, como aquella que la interpreta como *el estar dispuesto a actuar como si aquello en lo que*

[5] En el Sistema doxástico KD45 las *relaciones de accesibilidad* R_i son *seriales*, *transitivas* y *euclídeas*. Se dice que una relación R es *serial* si $\forall w_1 \in W, \exists w_2 \in W / w_1Rw_2$; asimismo, una relación R es *euclídea* si $\forall w_1,w_2,w_3 \in W (w_1Rw_2 \wedge w_1Rw_3 \Rightarrow w_2Rw_3)$.

se cree fuese verdad. Así, como apunta Lenzen (1978, p. 51), resulta difícil imaginar un curso de acción basado en la creencia simultánea de dos proposiciones contradictorias.

Respecto al Sistema doxástico KD45 también se obtienen resultados metalógicos estándar. En particular, este sistema es consistente y completo[6].

Los elementos formales que hemos venido presentando acarrean una serie de consecuencias para la teoría del conocimiento que es importante tomar en consideración. De entrada, el Ax1 pone de manifiesto que la lógica epistémica (y doxástica) proposicional es una extensión de la lógica proposicional clásica. El Ax2 (o Axioma K) indica que el conocimiento de un agente está cerrado bajo la implicación[7], lo que equivale a afirmar que un sujeto conoce (o, en su caso, cree) todas las consecuencias lógicas de sus conocimientos (o de sus creencias). El Ax3 (o Axioma T) establece que los hechos conocidos son verdaderos, poniendo de manifiesto la idea clásica según la cual se conocen sólo verdades. Por su parte, los axiomas Ax4 y Ax5 son axiomas de introspección. El Ax4 nos dice que si un sujeto sabe (o cree) algo, entonces sabe que lo sabe (o cree que lo cree, si nos referimos a la lógica de la creencia). A su vez, el Ax5 indica que si un agente no sabe algo (o no lo cree) entonces sabe que no lo sabe (o cree que no lo cree). Otro elemento de la lógica epistémica y doxástica proposicional a destacar es la Regla de Generalización de K (o de B), según la cual un sujeto conoce (o cree) todas las verdades lógicas.

Algunos de estos principios de la lógica epistémica y doxástica han sido objeto de crítica desde diversos frentes. Así, por ejemplo, Hintikka (1962, pp. 85-6, 92-3, y 138 y ss.) no acepta el Ax5 en su teoría formal del conocimiento y la creencia, y menos aún la interpretación introspectiva del Ax4. También se ha puesto en duda la pertinencia del AxD de la lógica doxástica, aduciéndose que las creencias de los agentes, especialmente si se trata de agentes humanos, pueden ser (y de hecho a menudo son) inconsistentes entre sí.

---

[6] Cf., por ejemplo, Meyer y Van der Hoek (1995[1], p. 70).

[7] Este axioma también puede ser expresado de la siguiente forma: $K_i(A \rightarrow A') \rightarrow (K_iA \rightarrow K_iA')$ (Para i=1, ..., h).

Pero son el Ax2 y la Regla de Generalización de K (o de B) los elementos de la lógica epistémica y doxástica proposicional que mayor controversia han generado, pues según muchos autores ambos principios conllevan una concepción excesivamente idealista del conocimiento y de la creencia. Así, un agente cuyo conocimiento o creencias se rigieran por estos postulados podría calificarse, a todas luces, de cognoscente ideal. Y es que, como ya hemos apuntado, el Ax2 (o Axioma K) implica que el agente en cuestión es capaz de conocer (o creer) todas las consecuencias lógicas de sus conocimientos (o de sus creencias), mientras que la Regla de Generalización de K (o B) conlleva que el sujeto sea capaz de conocer todas las fórmulas válidas[8].

Este tipo de dificultades que afectan a la lógica modal del conocimiento y la creencia constituyen lo que se ha dado en llamar el *problema de la omnisciencia lógica*, que ha llegado a convertirse en uno de los tópicos más extensamente tratados en los trabajos dedicados a esta rama de las lógicas no clásicas. La omnisciencia lógica reviste una enorme importancia para la lógica de las modalidades epistémicas y doxásticas, pues su origen se encuentra en el hecho mismo de concebir las nociones de conocimiento y creencia como conceptos modales.

De manera esquemática, podemos afirmar que el problema de la omnisciencia lógica tiene su fundamento último en las dos asunciones siguientes de la lógica modal epistémica y doxástica:

1) La *definición semántica del conocimiento y la creencia* que, tal y como hemos señalado más arriba, establecen, respectivamente, que: $v(K_iA,w_1)=1$ syss $\forall w_2 \in W/w_1R_i\, w_2$, $v(A,w_2)=1$ y $v(B_iA,w_1)=1$ syss $\forall w_2 \in W/w_1R_iw_2$, $v(A,w_2)=1$.

2) La noción de *validez lógica*, ya vista, según la cual una fórmula A es válida ($\vDash A$) si, para todo modelo y todo $w \in W$, $v(A,w)=1$.

---

[8] A este respecto, resulta interesante señalar que Hintikka (1962, p. 75) sostiene que condiciones como las expresadas por el Ax2 son difícilmente aplicables a lo que los sujetos humanos saben "activamente" (es decir, "realmente"), y sólo serían aceptables con relación a lo que dichos sujetos saben "virtualmente", esto es, en la medida en que sean capaces de extraer las consecuencias de lo que de hecho saben.

De acuerdo con el principio 2 anterior, toda fórmula lógicamente válida es verdadera en todo mundo posible, y esto junto con el postulado 1 trae consigo que todo sujeto epistémico haya de conocer (creer) todas las fórmulas lógicamente válidas (pues éstas son verdaderas en todas las alternativas epistémicas y doxásticas de cualquier agente). Es por esto que podemos afirmar que el problema de la omnisciencia lógica está en la base misma del tratamiento modal de las nociones epistémicas y doxásticas, pues depende nada menos que de dos de nuestros presupuestos lógicos de partida: la noción de validez y la definición modal del conocimiento y la creencia.

Como el problema que venimos comentando está intrínsecamente vinculado al enfoque modal basado en estructuras de Kripke, se han propuesto alternativas que tratan de modelar nociones epistémicas y doxásticas sin recurrir a la semántica kripkeana de mundos posibles[9]. Ahora bien, en vista de las enormes ventajas que acarrea el enfoque modal para la clarificación conceptual del conocimiento y la creencia, la mayor parte de los intentos de solución del problema de la omnisciencia lógica se mantienen dentro de los límites de la semántica kripkeana, si bien introduciendo modificaciones respecto de la aproximación modal estándar que hasta ahora hemos venido analizando.

Estos intentos por configurar sistemas lógicos epistémicos y doxásticos no omniscientes han dado lugar a buena parte de los desarrollos más interesantes dentro la lógica modal epistémica y doxástica. En las próximas tres secciones nos ocuparemos de algunas de estas propuestas, poniendo de relieve sus ventajas, y en algún caso sus inconvenientes. Estas propuestas tratan, en definitiva, de modelar nociones de conocimiento y creencia aplicables a agentes que, como los sujetos humanos, no son cognoscentes ideales, sino que mantienen un sistema de conocimientos y creencias en el que siempre cabe un margen más o menos amplio para la incertidumbre y la falibilidad.

---

[9] Entre este tipo de aportaciones cabe destacar el *modelo deductivo para la creencia* presentado por Konolige (1985).

2. La semántica no estándar de Hector Levesque y la lógica del razonamiento local de Fagin y Halpern

Hector Levesque (1984) presenta una *lógica de la creencia* que trata de mitigar el problema de la omnisciencia lógica al que nos hemos venido refiriendo. El sistema de Levesque se basa en la distinción entre *creencias implícitas* y *creencias explícitas*. Las primeras son un tipo ideal de creencia, de las que los sujetos no son conscientes, y para las que es admisible la omnisciencia lógica[10]. Con relación a esta forma de creencia implícita, Levesque (1984, p. 198) nos dice lo siguiente: si un sujeto considera que el mundo es de tal forma que en él p es verdadero, y si p implica lógicamente q, entonces dicho sujeto (aunque no se percate de ello) considera que el mundo es tal que en él q también es verdad. Dicho brevemente: si el mundo en el que el sujeto cree satisface p, entonces también debe satisfacer q. De acuerdo con esta interpretación, no tomamos en consideración lo que un sujeto cree directamente, sino *cómo sería el mundo si lo que tal sujeto cree fuera de hecho verdad*. Y es que en ocasiones nos interesa analizar las consecuencias de lo que un agente cree, aunque el agente en cuestión no tenga en cuenta dichas consecuencias de una manera consciente.

A diferencia de lo que ocurre con las creencias implícitas, las creencias denominadas explícitas se caracterizan porque los sujetos que las albergan son conscientes de las mismas, aunque no por ello han de creer explícitamente las consecuencias lógicas derivadas de las mismas. Respecto a este tipo de creencia, pues, no es adecuado admitir la omnisciencia lógica (cosa que sí ocurría con la creencia implícita). La distinción entre creencias implícitas y explícitas va a permitir a Levesque asumir que podemos creer todas las consecuencias lógicas de nuestras creencias tan sólo de una manera implícita, pero no explícita.

Para formalizar estas ideas Levesque configura un modelo semántico en el que, además de los mundos posibles estándar, encontramos *situaciones o mundos posibles parciales* (en relación con los cuales puede haber fbfs que no son ni verdaderas ni falsas), así como también *situaciones incoherentes*

---

[10] Las creencias implícitas de la lógica de Levesque se corresponderían con la noción clásica de creencia, esto es, con el tipo de creencia formalizada por sistemas como el KD45 presentado en la sección anterior.

(en las que alguna fbf es verdadera y falsa a la vez). Estos elementos formales van a permitir que un sujeto pueda mantener una imagen incoherente del mundo, sosteniendo creencias que son contradictorias entre sí, lo que abre las puertas de par en par a la incertidumbre; y todo ello conservando la consistencia y completud del sistema lógico desarrollado.

Fagin y Halpern (1988, pp. 46-8) llevan a cabo una serie de críticas, unas de carácter más técnico y otras de índole filosófica, a la lógica de Levesque que acabamos de presentar. Podemos sintetizar estas críticas en los siguientes puntos:

a) En primer lugar, para estos autores es difícil imaginar cómo podría un agente considerar explícitamente como posible una situación incoherente.

b) Levesque define la validez únicamente en función de las situaciones completas. Con ello se asegura de que todas las fórmulas válidas en la lógica clásica de proposiciones continúan siendo válidas en su lógica, y así su semántica puede dar lugar a sistemas consistentes y completos. Pero entonces resulta que habrá fórmulas (como, por ejemplo, $p\lor\neg p$) que no son verdaderas en algunas situaciones (concretamente, en las situaciones parciales y en las incoherentes) y que, empero, son válidas, pues como acabamos de apuntar su validez se determina sólo en función de las situaciones completas.

c) Aunque en el modelo de Levesque un agente no crea todas las consecuencias lógicas de sus creencias, resulta, no obstante, que en esta lógica los agentes se comportan como razonadores perfectos en lógica de la relevancia. Y, como señalan Fagin y Halpern (1988, p. 48), no parece que nuestra capacidad de razonar sea más perfecta en lógica de la relevancia de lo que puede serlo en la lógica clásica.

d) Finalmente, otro de los inconvenientes de la lógica de Levesque está en que no permite la anidación de creencias (es decir, fórmulas del tipo BBp, en las que hay ocurrencias de operadores doxásticos bajo el alcance de otros), además de que en ella sólo puede hacerse referencia a las creencias de un único agente. Y, de acuerdo con Fagin y Halpern (1988, p. 48), un tratamiento adecuado del comportamiento formal de la creencia ha de permitirnos tratar con creencias de segundo orden y también expresar las creencias que unos agentes mantienen acerca de las de otros.

Es por ello que Fagin y Halpern (1988) proponen una serie de desarrollos lógicos que tratan de hacer frente a la omnisciencia lógica y, al mismo tiempo, superar las anteriores dificultades. Una de las propuestas de estos autores es la denominada *lógica del conocimiento consciente*, que presenta un marcado carácter sintáctico[11]. El término "conocimiento consciente" hace referencia a aquellas fórmulas que el agente tiene a su disposición de una forma explícita, de tal modo que las consecuencias lógicas indeseables pueden ser desterradas al campo de lo no conscientemente conocido. En cualquier caso, la noción de conocimiento consciente aquí empleada tiene un sentido muy general, quizás próximo al de "percatarse de".

Por razones de espacio no entraremos aquí a considerar los aspectos formales de la lógica del conocimiento consciente de Fagin y Halpern; sólo diremos al respecto que, como principal ventaja, esta lógica logra eludir gran parte de las aserciones omniscientes propias de la lógica modal epistémica, presentando, además, la ventaja de permitir la anidación de creencias (implícitas y explícitas) y la referencia a múltiples agentes. No obstante, el punto débil de esta lógica reside en el hecho de que en ella no es posible expresar creencias inconsistentes, pues continúa siendo válido el Axioma D o de consistencia de las creencias, esto es: $BA \rightarrow \neg B \neg A$. De ahí que Fagin y Halpern (1988, pp. 58 y ss.) propongan una nueva lógica en la que puede conseguirse esto sin recurrir a estados incoherentes a la manera de Levesque.

La motivación última de esta nueva propuesta está en la idea de que los sujetos pueden sostener creencias inconsistentes por el hecho de que las creencias tienden a darse en "racimos o agregados de creencias" que no interactúan entre ellos, con lo que un mismo agente puede mantener diversos conjuntos (o racimos) de creencias que pueden ser contradictorios entre sí.

Postulamos ahora que desde un determinado estado un agente puede considerar posibles varios conjuntos de estados, cada uno de los cuales se corresponde con un agregado o racimo de creencias. Esto va a permitir que dicho agente pueda sostener creencias inconsistentes dependiendo del marco de referencia o contexto en que se encuentre.

---

[11] Traducimos aquí por "conocimiento consciente" el vocablo inglés "awareness".

El lenguaje formal para esta *lógica del razonamiento local de Fagin y Halpern* lo formamos añadiendo a un lenguaje para la lógica clásica de proposiciones los operadores monarios B ("creencia explícita") y L ("creencia implícita"). Pero, como se desprende de la definición semántica que de estos operadores daremos más abajo, estas nociones de creencia son muy distintas a las que emplea Levesque y a las de la lógica del conocimiento consciente que acabamos de presentar.

Asimismo, un modelo para el *razonamiento local* es una estructura $M=\langle W,Ç_1,...,Ç_h,v\rangle$, donde:

1) $W\neq\varnothing$, es un conjunto de estados.

2) $Ç_i$: $W\rightarrow P^x(P^x(W))$, siendo $P^x(W)=P(W)-\{\varnothing\}$, y $P^x(P^x(W))=P(P^x(W))-\{\varnothing\}$. Así, Çi es una función que nos da, para cada estado $w\in W$, el conjunto (no vacío) de conjuntos (no vacíos) de estados que considera posible el agente i en dicho estado w. A este conjunto de conjuntos lo llamamos $Ç_i(w)$, y a cada uno de los subconjuntos que lo forman lo simbolizamos por N (con un subíndice); así pues: $Ç_i(w)=\{N_1, ..., N_k\}$, siendo k el número total de racimos de estados que considera i en el estado w.

Para cada $w\in W$, $Ç_i(w)$ será entonces una colección no vacía de subconjuntos no vacíos (racimos) de W. De este modo, $Ç_i(w)$ nos indica qué marcos de referencia (o "marcos mentales") son considerados posibles por el agente i desde el estado w. Intuitivamente, si $Ç_i(w)=\{N_1, ..., N_k\}$, entonces en el estado w el agente i a veces considera que el conjunto de alternativas doxásticas es N_1, otras veces considera que es N_2, y así sucesivamente, dependiendo de su marco de referencia o del tipo de problemas que esté abordando en cada momento [Fagin y Halpern (1988, p.58)].

Si queremos formalizar la noción de conocimiento, en lugar de la de creencia, debemos añadir la condición de que $w \in N_1\cap...\cap N_k$, con lo cual w sería un elemento de todo subconjunto de $Ç_i(w)$[12].

---

[12] Esta condición se corresponde con la *propiedad reflexiva* de las relaciones de accesibilidad R_i (cf. supra, nota 3).

3) v: $FxW \rightarrow \{1,0\}$, es una función de asignación de valores de verdad a las fbfs en cada estado $w \in W$. Las cláusulas de la función de evaluación son las usuales[13], añadiendo únicamente una para B_iA y otra para L_iA:

$$v(B_iA,w_1)=1 \text{ syss } \exists\, N \in Ç_i(w_1)/\forall w_2 \in N,\ v(A,w_2)=1$$
$$v(L_iA,w_1)=1 \text{ syss } \forall w_2 \in N_1 \cap ... \cap N_k,\ v(A,w_2)=1$$

La validez se define aquí de la manera usual.

Interpretamos ahora B_iA como "el agente i cree que A en algún marco mental o respecto a un determinado contexto". Fagin y Halpern denominan este tipo de creencia explícita (B_iA) *creencia local*. A su vez, un agente i "cree implícitamente que A" (L_iA), si i creyera que A como resultado de considerar conjuntamente la información existente en sus diversos marcos mentales. De acuerdo con las definiciones semánticas anteriores, es fácil constatar que la creencia explícita (o local) implica la creencia implícita, es decir: $B_iA \rightarrow L_iA$.

En la lógica del razonamiento local la creencia explícita no está cerrada bajo la implicación, y así la fórmula $[B_ip \wedge B_i(p \rightarrow q)] \wedge \neg B_iq$ es satisfacible, porque en un marco mental el agente i puede creer que p, en otro puede creer que $p \rightarrow q$, pero podría no encontrarse nunca en un marco mental en que considerase esos dos hechos conjuntamente para concluir q. También un agente puede ahora mantener creencias inconsistentes, siendo satisfacible $B_ip \wedge B_i \neg p$, puesto que en un marco mental el agente puede creer que p y en otro que $\neg p$. Pero no puede darse $B_i(p \wedge \neg p)$, pues el agente no puede considerar posible la existencia de mundos incoherentes. Por otra parte, los sujetos se comportan como razonadores perfectos en el ámbito de cada marco mental, pues dentro de cada uno de tales marcos nos mantenemos en una lógica epistémica estándar como la representada por el Sistema S5.

Fagin y Halpern (1988, p. 60) ofrecen una serie de condiciones que podemos imponer al conjunto de conjuntos $Ç_i(w)$ para obtener así nociones de creencia que revistan diferentes propiedades:

---

[13] Cf. supra, sección 1.

a) El hecho de que $Ç_i(w)$ esté formado por conjuntos no vacíos asegura la validez de $\neg B_i(p\wedge\neg p)$. Si queremos que se cumpla que $\neg L_i(p\wedge\neg p)$, debemos postular que la intersección de los subconjuntos de $Ç_i(w)$ no sea vacía (es decir, que $N_1\cap...\cap N_k \neq \emptyset$).

b) Si queremos que en cada marco mental el agente i considere posible que efectivamente se encuentra en tal marco mental, debemos exigir que si $w_2 \in N \in Ç_i(w_1)$, entonces $N \in Ç_i(w_2)$. Esto asegura la validez de $B_ip\rightarrow B_iB_ip$ y de $L_ip\rightarrow L_iL_ip$.

c) Añadiendo la condición de que $\forall N\in Ç_i(w_1)$ y $\forall w_2\in N$, $Ç_i(w_2)\subseteq Ç_i(w_1)$, hacemos que se dé la introspección negativa respecto a la creencia explícita (o local) y a la creencia implícita, es decir, que hacemos válidas las expresiones $\neg B_ip\rightarrow B_i\neg B_ip$ y $\neg L_ip\rightarrow L_i\neg L_ip$.

d) También podemos modelar una situación en la que, en cada marco mental, un agente se negara a admitir que podría, coyunturalmente, encontrarse en otro marco mental distinto. Para ello basta con requerir, a nivel semántico, que si $w_2\in N\in Ç_i(w_1)$, entonces $Ç_i(w_2) =\{N\}$. Fagin y Halpern denominan "sujetos de mente estrecha" a aquellos para los que se cumple esta última restricción. Un agente tal creerá que es siempre consistente (incluso aunque no lo sea realmente) en tanto que en un marco mental dado se niega a reconocer que podría hallarse en otros marcos mentales diferentes. Así, en este caso es válido que $B_i(\neg(B_ip\wedge B_i\neg p))$, aunque $B_ip\wedge B_i\neg p$ sea consistente. Es más, puesto que en general un agente se comporta como un razonador perfecto dentro de un determinado marco mental, un sujeto de mente estrecha creerá siempre que él es un razonador perfecto. Así pues, para este tipo de agentes será válida la siguiente fórmula: $B_i[(B_ip\wedge B_i(p\rightarrow q)) \rightarrow B_iq]$.

Para concluir esta sección, cabe destacar como principal ventaja de los desarrollos de Fagin y Halpern, aparte del hecho de que logran atenuar en gran medida el problema de la omnisciencia lógica, el permitirnos modelar diversas nociones epistémicas y doxásticas que presentan características diferentes en función de nuestros intereses particulares a la hora de emplearlas. En todo caso, tanto la lógica del conocimiento consciente como la del razonamiento local se alejan del tipo ideal de conocimiento y creencia modelados por la lógica epistémica estándar, y permiten hacer referencia a sujetos epistémicos no omniscientes, cuyos razonamientos pueden estar, en mayor o menor medida, condicionados por la incertidumbre.

3. La semántica de mundos imposibles de V. Rantala

Para Hintikka (1975) el problema de la omnisciencia lógica surge de la presuposición de que las alternativas epistémicas que un agente considera posibles desde un estado dado, han de ser mundos lógicamente posibles. Por ello se propone justificar mundos que sean epistémicamente posibles y lógicamente imposibles. La motivación intuitiva de tales mundos se apoyaría precisamente en el hecho de que los agentes no son razonadores perfectos, es decir, que no son lógicamente omniscientes, sino que su conocimiento del mundo presenta frecuentemente dosis más o menos altas de incertidumbre[14].

Hintikka (1975, p. 475) analiza las causas del problema de la omnisciencia lógica en función de los siguientes supuestos:

De acuerdo con la interpretación modal del conocimiento[15], podemos afirmar que:

1) Una sentencia de la forma "i sabe que p" es verdadera en un mundo w_1 si y sólo si p es verdadera en todas las alternativas epistémicas que i considera posibles en w_1 (esto es, en todos los mundos epistémicamente posibles que son compatibles con lo que i sabe en el estado w_1).

[14] A lo largo de su exposición Hintikka hace referencia primordialmente a agentes humanos.

[15] Hintikka (1975) se centra en el análisis de la noción de conocimiento y en la omnisciencia lógica con ella asociada. No obstante, sus consideraciones son extensibles al tratamiento modal de cualquier actitud proposicional interpretada como necesidad y, por tanto, también a la noción de creencia tal y como se recoge, por ejemplo, en el Sistema KD45 (cf. supra, sección 1). A su vez, Rantala (1982), en cuyos desarrollos nos centraremos en esta sección, nos habla de actitudes proposicionales en general, si bien aquí referiremos sus consideraciones al caso del conocimiento para mantener así una mayor uniformidad en esta parte de nuestra exposición.

A su vez, la ausencia de omnisciencia lógica puede ser planteada así:

2) Se da que "i sabe que p", $p \rightarrow q$ es válido y, sin embargo, "i no sabe que q".

Por su parte, la noción de validez aquí empleada es la usual, esto es:

3) Una sentencia es válida si y sólo si es verdadera en todo mundo lógicamente posible.

Pues bien, aunque se ha argüido que la omnisciencia lógica se debe a la incompatibilidad entre los postulados 1, 2 y 3 anteriores, Hintikka sostiene que tal incompatibilidad sólo se da si se asume también que:

4) Todo mundo epistémicamente posible ha de ser también lógicamente posible (es decir, que toda alternativa epistémica a un determinado mundo w_1 ha de ser lógicamente posible).

La incompatibilidad entre 1, 2, 3 y 4 puede expresarse en los siguientes términos. Asumimos que en el estado w_1 se da 2. Entonces, de acuerdo con 1, que "i no sabe que q" significa que en w_1 el agente i considera posible una alternativa epistémica w_2 en la cual q es falsa. Asimismo, que "i sabe que p" supone, según 1, que p es verdadera en todas las alternativas epistémicas del agente i; en particular, p es verdadera en la alternativa epistémica w_2. En función de 4, tales alternativas epistémicas son también mundos lógicamente posibles (w_2 es, por tanto, un mundo lógicamente posible). Pues bien, la asunción de que $p \rightarrow q$ es válida trae consigo, de acuerdo con 3, que q es verdadera en todo mundo lógicamente posible en el cual p es verdadera. Como w_2 es uno de tales mundos, q debe ser verdadera en w_2. Y aquí tropezamos con una contradicción, pues anteriormente habíamos establecido que q es falsa en el mundo w_2.

Hintikka (1975, p. 476) señala que aunque diversos autores consideran que la interpretación modal del conocimiento (tal y como se lleva a cabo en 1) conduce inevitablemente a la omnisciencia lógica, no se han percatado de que la verdadera fuente de dificultades reside en la asunción, muchas veces implícita, del postulado 4, lo cual impide eliminar la omnisciencia lógica en función del presupuesto 2. Y es que asumir 4 supone admitir que un agente puede descartar todas las alternativas epistémicas que no sean lógicamente posibles, lo cual nos lleva automáticamente a convertirlo en un agente omnisciente. Pero, en realidad, un agente (al menos un agente humano) no puede seguir todas las consecuencias lógicas de su conocimiento ilimitadamente, y por ello usualmente ha de mantener alternativas epistémicas que aparentemente son posibles pero que contienen contradicciones ocultas.

Para solventar el problema de la omnisciencia lógica, concluye Hintikka, es preciso abandonar el supuesto 4, lo cual implica admitir la existencia de *mundos posibles imposibles* [Hintikka (1975, p. 477)], esto es, mundos que, por parecer posibles, pueden ser admitidos como alternativas epistémicas, pero que en realidad son lógicamente imposibles. Ya Kripke (1965, p. 210) distinguió entre "mundos posibles normales" y "mundos posibles no normales". Ahora bien, los "mundos no normales" de Kripke satisfacen las leyes de la lógica clásica (sólo que en ellos el comportamiento formal de la necesidad lógica difiere del usual), cosa que no ocurre con los "mundos posibles imposibles" de Hintikka.

Así pues, admitiendo la existencia de "mundos posibles imposibles" entre las alternativas epistémicas de un agente i, podemos sostener las condiciones 1, 2 y 3 juntas sin que se dé contradicción alguna entre ellas. En el estado w_1 un agente puede saber que p; la fórmula $p \rightarrow q$ puede ser lógicamente válida; y, sin embargo, en una alternativa epistémica w_2 del agente i en el estado w_1 (es decir, $w_1R_iw_2$) puede darse $\neg q$, pues w_2 puede ser uno de estos mundos epistémicamente posibles y lógicamente imposibles, con lo cual en dicho estado no tiene por qué darse $p \rightarrow q$. Ahora bien, Hintikka reconoce que la dificultad está en desarrollar un modelo semántico adecuado en el que intervengan este tipo de "mundos posibles imposibles".

Por su parte, Rantala (1982), basándose en las anteriores ideas de Hintikka y en los "mundos no normales" de Kripke, propone una semántica en la que los mundos posibles estándar son ampliados mediante un conjunto de *mundos imposibles*. En estos mundos imposibles no todas las fórmulas válidas (en el sentido clásico de validez) son verdaderas, mientras que fórmulas inconsistentes (en el sentido clásico de inconsistencia) sí pueden serlo.

Un modelo de Rantala para la *semántica de mundos imposibles* es una estructura del tipo $M=\langle W, W^{imp}, R_1,...,R_h, v, v^{imp}\rangle$, donde:

1) W y W^{imp} son conjuntos de mundos posibles e imposibles, respectivamente, con $W\neq\emptyset$.

2) $R_i \subseteq (W\cup W^{imp})x(W\cup W^{imp})$, son relaciones de accesibilidad, una para cada agente, que asocian con un mundo (posible o imposible) un conjunto de mundos (posibles o imposibles) accesibles.

3) Siendo F el conjunto de fbfs, $v: FxW\rightarrow\{1,0\}$, es una función de evaluación de fbfs en estados de W (es decir, en mundos posibles estándar); a su vez, $v^{imp}: FxW^{imp}\rightarrow\{1,0\}$ es una función de evaluación de fbfs en estados de W^{imp} (esto es, en mundos imposibles). Las cláusulas para la función de evaluación v son las usuales[16]. Para la función de evaluación v^{imp}, sin embargo, las condiciones de verdad son completamente libres.

Como es usual, decimos que una fórmula A es válida si, para todo modelo M de la forma anterior y todo mundo posible $w\in W$, $v(A,w)=1$.

Debido a la libertad a la hora de asignar valores de verdad a las fbfs en los mundos no normales o imposibles, la semántica de mundos imposibles presenta una gran capacidad para librarnos de la omnisciencia lógica. De hecho, en este marco un agente puede no mantener ningún principio lógico general, y así resulta que todas las formas de omnisciencia lógica pueden ser eliminadas.

En la semántica de mundos imposibles de Rantala, al igual que en la semántica no estándar de Levesque, la validez se establece únicamente en función de los mundos posibles estándar, y por ello nos encontramos con fórmulas que, como $p\vee\neg p$, pueden ser falsas en los mundos imposibles y que, sin embargo, son lógicamente válidas. Con esto se logra que aserciones no omniscientes sean verdaderas sin que por ello la lógica resultante deje de ser consistente y completa.

La semántica de mundos imposibles se considera, en general, un método excesivamente tosco para eliminar la omnisciencia lógica, pues eligiendo

[16] Cf. supra, sección 1.

adecuadamente el conjunto de mundos imposibles W^{imp} podemos deshacernos de la omnisciencia lógica de una manera trivial. El propio Rantala (1982, pp. 108 y 112) reconoce que la clave del éxito de esta semántica reside en el hecho de que respecto a los mundos no normales no se dan condiciones de verdad recursivas, y esto para muchos autores es motivo suficiente para calificar esta semántica de "poco lógica"[17]. Aun así, hay que reconocer que es éste el método más potente para solventar el problema de la omnisciencia lógica.

También Rantala (1982, p.113) considera importante encontrar interpretaciones intuitivas adecuadas de los mundos no normales o imposibles. Diversos autores se han ocupado, en efecto, de la motivación filosófica de los mundos imposibles[18]. A modo ilustrativo, Vander Laan (1997, pp. 598-600) sostiene que los mundos imposibles podrían interpretarse como aquellos modos en que las cosas no habrían podido suceder, considerando, además, que al igual que se defiende la existencia de conjuntos (Quine) o de mundos posibles (Lewis), también podría aducirse la existencia de un recurso lógico tan útil como pueden ser los mundos imposibles.

4. La lógica de las modalidades epistémicas graduadas de Meyer y Van der Hoek

La *lógica de las modalidades epistémicas graduadas* nos permite expresar la *confianza* que alberga un sujeto en que algo sea o no sea el caso. Este tipo de lógica nos capacita, además, para formalizar nociones que revisten tanto interés como pueden ser las de *incertidumbre* o *grados de certeza*, por lo que podemos afirmar que es un tipo de lógica para sujetos epistémicos no omniscientes.

El desarrollo de la lógica modal graduada comenzó en los años 70, siendo K. Fine uno de los primeros en estudiar este tipo de operadores modales. En los años 80 fueron redescubiertos por algunos lógicos italianos, como M. Fattorosi-Barnaba, F. de Caro y C. Cerrato. Y, ya en los 90, W. Van der Hoek y J.-J. Ch. Meyer comienzan a aplicar las modalidades graduadas a la

---

[17] Cf., por ejemplo, Meyer y Van der Hoek (1995^1, p. 88).

[18] Por ejemplo, la *Notre Dame Journal of Formal Logic*, volume 38, number 4, fall 1997, se encuentra enteramente dedicada a la temática de mundos imposibles.

lógica epistémica, poniendo de relieve cómo aumenta el poder expresivo de esta última[19].

La lógica epistémica graduada puede interpretarse como una *expresión cuantitativa de la confianza que un sujeto epistémico tiene respecto a que un hecho se dé o no*. Para ello añadimos al lenguaje formal de la lógica epistémica proposicional los operadores monarios M_n y K_n (donde *n es un número natural*). La interpretación intuitiva de estos nuevos operadores epistémicos es la siguiente:

M_nA indica que "hay más de n alternativas epistémicas en las que se da A".

K_nA indica que "a lo sumo n alternativas epistémicas refutan A".

Para la interpretación semántica de los operadores M_n y K_n, añadimos las siguientes cláusulas al modelo M=<W,R,v> de la lógica epistémica proposicional[20]:

$v(M_nA,w_1)=1$ syss $|\{w_2 \in W \ / \ w_1Rw_2$ y $v(A,w_2)=1\}| > n$

$v(K_nA,w_1)=1$ syss $|\{w_2 \in W \ / \ w_1Rw_2$ y $v(\neg A,w_2)=1\}| \leq n$

De acuerdo con estas definiciones, podemos establecer la siguiente equivalencia lógica:

$K_nA \equiv \neg M_n \neg A$

Ahora los operadores K y M de la lógica epistémica estándar (presentada en la sección 1) no son sino casos especiales de los operadores K_n y M_n de la lógica epistémica graduada; y así:

---

[19] Para profundizar en las referencias históricas sobre el desarrollo de la lógica epistémica graduada, cf. Meyer y Van der Hoek (1995[1], p. 104).

[20] Como hasta ahora, W representa un conjunto (no vacío) de mundos o estados posibles; w_1, w_2, etc., representan a cada uno de dichos mundos o estados. Es decir: $W=\{w_1, w_2, ...\}$.

$MA \equiv M_0A$

$KA \equiv K_0A$

Podemos también expresar el número exacto de alternativas epistémicas que satisfacen una fórmula A. Así, si $M!_nA$ expresa "exactamente n alternativas epistémicas satisfacen A", tendríamos que:

$M!_nA \equiv (M_{n-1}A \wedge \neg M_nA)$, si n>0

Siendo $M!_0A \equiv K_0\neg A \equiv K\neg A$

Meyer y Van der Hoek (1993[1], p. 4; 1993[2] , p. 11; y 1995[1], p. 105) presentan el Sistema axiomático Gr(S5) para la lógica epistémica graduada; este sistema consta de los siguientes axiomas (o esquemas de axiomas):

Ax0. Un conjunto suficiente de axiomas para derivar todas las tautologías de la lógica clásica de proposiciones.

Ax1. $K_0(A \rightarrow A') \rightarrow (K_nA \rightarrow K_nA')$

Ax2. $K_nA \rightarrow K_{n+1}A$

Ax3. $K_0\neg(A \wedge A') \rightarrow [(M!_nA \wedge M!_mA') \rightarrow M!_{n+m}(A \vee A')]$

Ax4. $\neg K_nA \rightarrow K_0\neg K_nA$

Ax5. $K_oA \rightarrow A$

Tanto n como m son números naturales. Como Reglas de derivación contamos con el Modus Ponens (MP), la Regla de Sustitución (RS) y la Generalización de K_0 (Si $\vdash A$, entonces $\vdash K_0A$).

El Sistema Gr(S5) es consistente y completo[21]. Por otra parte, en el plano semántico, las relaciones de accesibilidad R del Sistema Gr(S5), son reflexivas, simétricas y transitivas (relaciones de equivalencia) y euclídeas.

[21] Cf. Meyer y Van der Hoek (1995[1], p. 107).

El Axioma Ax0 y las Reglas de deducción MP y RS ponen de manifiesto que la lógica epistémica graduada es una extensión de la lógica clásica de proposiciones.

En virtud de la Regla de Generalización de K_0 el agente es capaz de conocer todas las verdades lógicas, tanto axiomas como teoremas, pero sólo en el caso de que estemos considerando un tipo de conocimiento absoluto como el representado por el operador epistémico K_0, que puede ser así interpretado como un tipo ideal de conocimiento[22]. En ningún caso esta regla de derivación es válida para K_n si $n \neq 0$.

El Axioma Ax1 pone de manifiesto que si un sujeto sabe que $A \rightarrow A'$ (es decir, $A \rightarrow A'$ en todas las alternativas epistémicas del agente de acuerdo con la información de que dispone), entonces, si imagina a lo sumo n excepciones para A, no podrá imaginar más de n excepciones para A', ya que toda excepción para A' también lo será para A.

El Axioma Ax2 indica que si un agente prevé a lo sumo n excepciones para A, también considerará, a lo sumo, n+1 excepciones para A. Este axioma es equivalente a $M_{n+1}A \rightarrow M_nA$, lo que nos permite aumentar grados en operador K_n y disminuirlos en M_n, lo cual equivale, de acuerdo con la interpretación de ambos operadores epistémicos, a *disminuir la confianza* que un agente tiene en que algo sea el caso[23].

El Axioma Ax3 expresa que si un sujeto sabe que A y A' son eventos mutuamente excluyentes, y concibe n situaciones en las que A es el caso y m situaciones en las que A' es verdad, en total cuenta con n+m situaciones en las que se da A o A' (o ambos). Este axioma permite aumentar grados en el operador M_n (es decir, *aumentar el grado de confianza* que un sujeto tiene en que un hecho sea el caso).

El Axioma Ax4 refleja un tipo de introspección negativa considerablemente atenuada si se compara con la de la lógica epistémica estándar presentada en la sección 1. En función de este axioma, si un agente no concibe más de n excepciones para A, entonces dicho agente sabe (con seguridad) que no concibe más de esas n excepciones.

---

[22] Tengamos en cuenta que K_0 equivale al operador epistémico K de la lógica epistémica no graduada.

[23] Así, al aumentar grados en el operador M_n, aumentamos la confianza o nivel de certidumbre del agente, y al disminuir grados, disminuimos esta confianza. Justo lo contrario ocurre con el operador K_n.

El Axioma Ax5 alude a la tradicional conexión entre conocimiento y verdad, estipulando que sólo se pueden conocer con seguridad hechos verdaderos. En el caso de la lógica epistémica graduada este axioma sólo es válido para el caso del conocimiento absoluto e ideal representado por el operador K_0. Hemos de tener en cuenta que la generalización de este axioma Ax5 (es decir, $K_nA \rightarrow A$) no es válida en ningún caso para n>0, pues si un sujeto no sabe con seguridad que A es el caso (es decir, si admite posibles excepciones para A), entonces no puede concluir con certeza que A es el caso. Esto pone de manifiesto que K_nA en realidad expresa una forma de "conocimiento incierto" o, incluso, de "creencia"[24]. Asimismo, muestra también que la lógica epistémica estándar (no graduada) puede concebirse como un caso especial de la lógica de las modalidades epistémicas graduadas.

La aplicación de las modalidades graduadas a la lógica epistémica incrementa la capacidad expresiva de esta última, al permitirnos hacer referencia al "grado de confianza" que tiene un agente en que algo sea o no el caso. Así, los operadores modales K_n, junto con sus duales M_n, forman un enorme espectro de operadores epistémicos de certeza decreciente: K_0 (que representa la certeza absoluta), K_1, K_2, ..., M_2, M_1, M_0 (que simplemente expresa que algo no se considera imposible). De acuerdo con esto, en la lógica de las modalidades epistémicas graduadas se formaliza una forma de conocimiento incierto, que refleja simplemente el grado de confianza o de justificación que tiene un agente a la hora de afirmar que algo es o no verdadero. Y con este tipo de desarrollos lógicos la noción de conocimiento puede quedar tan debilitada, que, en muchos casos, puede dejar de tener sentido la distinción clásica entre conocimiento y creencia. Este resultado está, en gran medida, en consonancia con la postura defendida por Williamson (2000), para quien frecuentemente la diferencia verdaderamente relevante es la que se establece entre conocer e ignorar, y no tanto la distinción clásica entre conocer y creer. Y es que, según Williamson, el concepto "conocer" no es susceptible de ser analizado en función de otras nociones más básicas, puesto que no hay nada más básico que él. Asimismo, este tipo de lógica epistémica nos permite formalizar un tipo de razonamiento que presenta cierto grado de incertidumbre, a

---

[24] Cf. Meyer y Van der Hoek (1995[1], p. 108).

diferencia de lo que ocurre, por ejemplo, con el tipo de razonamiento representado por el Sistema S5 de la lógica epistémica no graduada, en el cual, como hemos visto, el sujeto conoce la totalidad de las verdades lógicas y es capaz de determinar todas las consecuencias de sus conocimientos.

5. La lógica del conocimiento falible: balance final

Mitigar la omnisciencia lógica es el objetivo fundamental de los desarrollos que hemos venido analizando en las tres secciones precedentes. Se trata, en definitiva, de generar sistemas formales no omniscientes que conserven las correspondientes propiedades metalógicas (consistencia, completud, decidibilidad). Las propuestas que hemos comentado incluyen modelos que se mantienen más o menos próximos a la semántica estándar de mundos posibles, pues como vimos en su momento, aunque este tipo de sistemas lógicos epistémicos presentan algunos inconvenientes, no le faltan considerables virtudes a la hora de analizar distintas nociones de conocimiento y creencia. En cualquier caso, es de destacar que es precisamente a partir de este tipo de intentos por superar el problema de la omnisciencia lógica de donde han surgido muchos de las aportaciones más interesantes en el ámbito de la lógica epistémica y doxástica proposicional.

Aquí nos hemos centrado en una serie de sistemas lógicos epistémicos que tratan no sólo de disipar la omnisciencia lógica, sino también de dar cabida a la incertidumbre en el ámbito del conocimiento, y todo ello sin menoscabo de las propiedades metalógicas a las que aludíamos hace un momento. En este sentido, hemos presentado y discutido la semántica no estándar de Hector Levesque, la lógica del conocimiento consciente y el razonamiento local de Fagin y Halpern, la semántica de mundos imposibles de V. Rantala y la lógica de las modalidades epistémicas graduadas de Meyer y Van der Hoek. Esta lista no agota, ni mucho menos, los intentos formales de soslayar el problema de la omnisciencia lógica, pero sí da cuenta, en buena medida, de las líneas de investigación más sobresalientes que se han emprendido para alcanzar dicho objetivo[25].

---

[25] Una extensa panorámica de los desarrollos lógicos epistémicos que tratan de superar el problema de la omnisciencia lógica puede encontrarse en Meyer y Van der Hoek (1995[2]).

¿Cuál de los anteriores sistemas lógicos epistémicos es el que mejor modela las nociones epistémicas y/o doxásticas que en ellos se recogen? Esta pregunta no tiene una respuesta simple, como tampoco la tendría si nos interrogásemos acerca de cuál es la caracterización filosófica correcta del conocimiento o la creencia. Aquí defendemos que no hay una concepción única, completamente satisfactoria, de las nociones epistémicas y doxásticas, sino que puede haber varias definiciones aceptables de los conceptos de conocimiento y creencia, y el aceptar una u otra de estas concepciones va a depender de los objetivos que persigamos en cada momento, del contexto en que nos encontremos, del tipo de sujetos epistémicos al que nos refiramos, etc. Así, en muchos casos podremos conformarnos con la visión clásica del conocimiento, que lo contempla, ya desde Platón, como una forma de creencia verdadera y justificada[26]. Pero, como muestra Gettier (1963), en algunos casos esta definición de lo que significa conocer es excesivamente genérica y, por consiguiente, poco adecuada; y de ahí que a partir de la crítica de Gettier diversos autores hayan tratado de encontrar caracterizaciones del conocimiento más precisas, muchas veces añadiendo requisitos extra a la concepción clásica de Platón.

En consonancia con este punto de vista general sobre el conocimiento (que vale también para la creencia), tampoco creemos que haya un sistema lógico epistémico que sea "el definitivo" a la hora de formalizar las nociones epistémicas y/o doxásticas, sino que podemos emplear uno u otro sistema dependiendo de cada situación particular, pues las características adecuadas de los conceptos de conocimiento y creencia que intentemos modelar en cada ocasión dependerán estrechamente del tipo de agentes que estemos considerando, de los contextos epistémicos que pretendamos simular, del grado de racionalidad asumible o deseable, etc., por lo que se requerirá, en cada caso, de desarrollos formales que eliminen unos u otros aspectos de la omnisciencia lógica. Es por ello que no podemos decidir *a priori* qué lógica modal epistémica es la adecuada, y así, en lo relativo a este punto nos mostramos plenamente de acuerdo con la postura que se mantiene en trabajos como los de Fagin y Halpern (1988, p. 41), Fagin, Halpern, Moses y Vardi (1995, p. 313) o Meyer y Van der Hoek (1995[1],

---

[26] Así, en *Teeteto*, 202c, se nos dice que "una opinión verdadera acompañada de explicación es saber".

p.89). Lo que sí debemos tener en cuenta es que los métodos para superar la omnisciencia lógica que son excesivamente generales, como la semántica de mundos imposibles de Rantala, si bien logran soslayar el problema, resultan, quizá, excesivamente "ilógicos"; por su parte, los métodos que introducen desarrollos más parciales, como los de Levesque o de Fagin y Halpern, conservan, en mayor o menor medida, ciertos "restos" de omnisciencia lógica.

Como señalamos en la sección 1 de este artículo, la omnisciencia lógica está firmemente arraigada en el tratamiento modal de las nociones de conocimiento y creencia. Pero abandonar por ello la perspectiva modal para el análisis de las nociones epistémicas y doxásticas supondría también renunciar a las ventajas que aportan este tipo de perspectivas; sería tanto como, según la consabida y elocuente expresión, "tirar al niño con el agua de la bañera". A nuestro juicio, el empeño de la investigación futura deberá centrarse en el desarrollo de nuevos sistemas lógicos epistémicos que sean lo suficientemente realistas como para poder ser aplicados a sujetos humanos, esto es, a sujetos cuyo conocimiento y capacidad de razonamiento no son, ni mucho menos, infalibles. Se trataría, en suma, de profundizar en la configuración de lógicas modales epistémicas no omniscientes, o con el menor grado de omnisciencia posible, o, mejor aun, con el grado de omnisciencia lógica que sea adecuado para cada situación.

Si, como sostiene Aristóteles en la cita que encabeza este artículo, "todos los hombres por naturaleza desean saber", las distintas aportaciones de la lógica modal epistémica pueden contribuir notablemente al análisis y mejor comprensión de lo que significa ese anhelo por saber que nos distingue en tanto que seres humanos.

BIBLIOGRAFÍA

FAGIN, R. & HALPERN, J.Y. (1988): "Belief, Awareness and Limited Reasoning", *Artificial Intelligence* 34, pp. 39-76.

FAGIN, R., HALPERN, J.Y., MOSES, Y. & VARDI, M.Y. (1995): *Reasoning about Knowledge*, MIT Press.

FREUND, M. (1995): "Lógica epistémica". En: Alchourrón, C.E., Méndez, J.M. y Orayen, R. (eds.): *Enciclopedia Iberoamericana de Filosofía,* vol. 7, Madrid, Trotta, pp. 205-14.

GETTIER, E.L. (1963): "Is Justified True Belief Knowledge?", *Analysis* 23, pp. 121-23.

HINTIKKA, J. (1962): *Knowledge and Belief*, Cornell, Cornell University Press. Versión castellana de Acero, J.J. (1979): *Saber y creer*, Madrid, Tecnos.

_______ (1967): "Individuals, possible worlds and epistemic logic", *Noûs* 1, pp. 33-62.

_______ (1975): "Impossible possible worlds vindicated", *Journal of Philosophical Logic* 4, pp. 475-84.

KONOLIGE, K. (1985): "Belief and Incompleteness". In: Hobbs, J. & Moore, R. (comps.): *Formal Theories of the Commonsense world*, N.J, Norwood, pp. 259-404.

KRIPKE, S.A. (1965): "Semantical analysis of Modal Logic II. Non-normal modal propositional calculi". In: Addison, J.W., Henkin, L. & Tarski, A. (eds.): *The Theory of Models*, Amsterdam, North-Holland, pp. 206-20.

LENZEN, W. (1978): "Recent work in Epistemic Logic", *Acta Philosophica Fennica* 30.

_______ (2004): "Epistemic Logic". In: Niiniluoto, I., Sintonen, M. & Wolenski, J. (eds.): *Handbook of Epistemology*, Kluwer Academic Publishers, pp. 963-83.

LEVESQUE, H.J. (1984): "A Logic of Implicit and Explicit Belief", *Proceedings of the National Conference on Artificial Intelligence*, Texas, pp. 198-202.

MEYER, J.-J. Ch. (2001): "Epistemic Logic". In: Goble, L. (ed.): *Philosophical Logic*, Blackwell Publishers, pp. 183-202.

MEYER, J.-J. Ch. & VAN DER HOEK, W. (1993^1): "Graded Modalities in Epistemic Logic". In: *Graded Model and Epistemic Logic,* Technical Report RUU-CS-93-44, Utrecht University.

_______ (1993^2): "Modalities for Reasoning about Knowledge and Uncertainties". In: *Graded Model and Epistemic Logic*, Technical Report RUU-CS-93-44, Utrecht University.

_______ (1995^1): *Epistemic Logic for AI and Computer Science*, Cambridge, Cambridge University Press.

_______ (1995[2]): *Modal Logics for Representing Incoherent Knowledge*, Technical Report UU-CS-1995-19, Utrecht University.

PLATÓN: *Gorgias*. Versión castellana de J. Calonge Ruiz, Madrid, Gredos, 1983.

_______ *Teeteto*: Versión castellana de A. Vallejo Campos, Madrid, Gredos, 2000.

RANTALA, V. (1982): "Impossible worlds semantics and logical omniscience", *Acta Philosophica Fennica*, 35, pp. 106-15.

VANDER LAAN, D.A. (1997): "The ontology of impossible worlds", *Notre Dame Journal of Formal Logic* 38, nº 4, fall 1997, pp. 597-620.

VON WRIGHT, G.H. (1951): *An Essay in Modal Logic*, Amsterdam, North-Holland.

WILLIAMSON, T. (2000): *Knowledge and its Limits*, Oxford, Oxford University Press.