LÓGICA EPISTÉMICA Y ACCIÓN COLECTIVA*

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

En este trabajo se ofrece un análisis formal de ciertas nociones, tales como el *conocimiento conjunto*, el *conocimiento común* y el *conocimiento implícito de un grupo*, que resultan particularmente provechosas para el análisis filosófico del conocimiento de grupos de agentes. Así, este artículo pretende mostrar algunas relaciones de interés entre la lógica epistémica y la teoría de la acción colectiva. En nuestra opinión, el tratamiento lógico de este tipo de nociones puede mejorar nuestra comprensión de aquellas situaciones en las que intervienen múltiples agentes. Esto puede resultar muy relevante en diversos campos de estudio, y especialmente en el análisis de la acción colectiva, en tanto que, con frecuencia, el conocimiento de los agentes determina sus diferentes cursos de acción. Finalmente, consideraremos de forma breve algunas ideas que apuntan a investigaciones futuras.

Palabras clave: lógica epistémica y doxástica, acción colectiva, conocimiento conjunto, conocimiento común, conocimiento implícito.

In this paper it is offered a formal analysis of some notions, such as "*everyone knows*", *common knowledge* and *implicit knowledge of a group*, which are particularly helpful for the philosophical analysis of group knowledge. Therefore, this paper is intended to show some interesting relationships between epistemic logic and theory of collective action. In our opinion, the logical treatment of these kinds of notions can improve our understanding of those situations in which multiple agents are concerned. This can be very important in several fields of research, and specially in the study of collective action, because the agents' knowledge often determines their different courses of action. Finally, we'll briefly discuss some perspectives for future research.

Key-words: epistemic and doxastic logic, collective action, "everyone knows", common knowledge, implicit knowledge.

1. Introducción

El estudio formal de las nociones epistémicas y doxásticas, tal y como éste se lleva a cabo en la lógica de las modalidades epistémicas, ha encontrado aplicación en ámbitos tan diversos como el modelado en

*Este trabajo es parte del proyecto de investigación "Experiencia y reflexión" (PI-2001/068), financiado por la Dirección General de Universidades de la Consejería de Educación, Cultura y Deportes del Gobierno Autónomo de Canarias.

economía y teoría de juegos, la lingüística, la inteligencia artificial, la informática, la teoría de la acción, entre otros. Precisamente será este último campo, el de la teoría de la acción, y, en concreto, de la acción colectiva, el que nos ocupará en este trabajo, pues pretendemos presentar algunas de las aportaciones fundamentales de la lógica epistémica al análisis formal de nociones de conocimiento que son centrales para el estudio de la acción colectiva.

En nuestro análisis el Sistema epistémico S5 nos servirá de base para la consideración de otros desarrollos formales relacionados con contextos en los que intervienen múltiples agentes. A partir de ahí, podremos presentar el tratamiento formal que desde la lógica epistémica se lleva a cabo respecto a las nociones de *conocimiento conjunto*, *conocimiento común* y *conocimiento implícito*. Estos tres conceptos epistémicos resultan de enorme interés para el análisis de acciones en las que está involucrado un grupo de agentes. En este trabajo concebiremos el conocimiento como un concepto modal, por lo cual nos ceñiremos al marco de la lógica modal epistémica. Y es que tal concepción de las nociones epistémicas ha acarreado enormes ventajas teóricas y prácticas desde los trabajos pioneros de Von Wright (1951) y Hintikka (1962) hasta nuestros días.

2. El Sistema S5 de la lógica epistémica

El lenguaje de la lógica epistémica proposicional (en adelante LEP) para un conjunto de h agentes consta de:

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

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

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

d) φ, ψ, etc. como variables metalingüísticas.

e) K (“*saber que*”) y M (“*considerar posible que*”) como operadores epistémicos[1].

[1] Los elementos recogidos en a), b), c), y d) son los propios de un lenguaje para la lógica clásica de proposiciones (en adelante LCP). Al ser la LEP una extensión de la LCP, para configurar su lenguaje lo que hacemos es añadir a un lenguaje para la LCP los operadores epistémicos monarios K y M del punto e).

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

1) Toda variable proposicional es una fbf.
2) $\neg\varphi$, donde φ es una fbf.
3) $(\varphi\rightarrow\psi)$, $(\varphi\wedge\psi)$, $(\varphi\vee\psi)$ y $(\varphi\leftrightarrow\psi)$, donde φ y ψ son fbfs.
4) $K_i\varphi$ ("*el agente i sabe que* φ"), donde φ es una fbf.
5) $M_i\varphi$ ("*el agente i considera posible que* φ"), donde φ es una fbf.

Nada que no se obtenga de la repetida aplicación de las reglas 1-5 es una fbf.

Los paréntesis externos de las fórmulas pueden ser eliminados.

Además, se establecen las siguientes equivalencias entre las conectivas lógicas:

$$\varphi\rightarrow\psi \Leftrightarrow \neg(\varphi\wedge\neg\psi)$$
$$\varphi\rightarrow\psi \Leftrightarrow \neg\varphi\vee\psi$$
$$\varphi\wedge\psi \Leftrightarrow \neg(\neg\varphi\vee\neg\psi)$$
$$\varphi\leftrightarrow\psi \Leftrightarrow (\varphi\rightarrow\psi)\wedge(\psi\rightarrow\varphi)$$

Este lenguaje formal nos va a permitir, por un lado, expresar los conocimientos (o la ignorancia) de un agente sobre hechos básicos del mundo, como haríamos, por ejemplo, en expresiones como K_ip, $K_i\neg q$, $\neg K_i(p\wedge q)$, $\neg K_i\neg r$, etc. Y, por otro lado, también nos va a permitir reflejar lo que sabe (o desconoce) un agente acerca de sus propios conocimientos o de los conocimientos de otros agentes, como ocurre, por ejemplo, con K_iK_ip, $K_i\neg K_iq$, $K_iK_j\neg p$, $\neg K_jK_iK_j(p\vee q)$, $\neg K_i\neg K_jK_i(p\wedge\neg q)$, etc.

Por otra parte, puesto que permaneceremos en el marco de la lógica modal epistémica, para la semántica de nuestra LEP tomamos como base la *semántica kripkeana de mundos posibles*. Así pues, definimos un modelo de la lógica epistémica como una estructura $\mathcal{M}=<M, R_1,...,R_h, v>$, donde

1. $M \neq \varnothing$, es un conjunto de estados o mundos epistémicamente posibles.

2. $R_i \subseteq M^2$ (para $i= 1, ..., h$), son relaciones de accesibilidad entre mundos o estados epistémicos, una para cada agente. Las propiedades que deben cumplir las relaciones R_i variarán en función del sistema axiomático de la LEP de que se trate[2].

3. v es una función de evaluación que asigna el valor verdadero o falso a cada fbf en un mundo o estado epistémico (siendo F el conjunto de fbfs, $v: F \text{x} M \rightarrow \{1,0\}$). La función de evaluación cumple las siguientes condiciones, para cualesquiera $m,n \in M$, $\varphi, \psi \in F$ y variable proposicional p:

(i) $v(p,m)=1$ o $v(p,m)=0$
(ii) $v(\neg\varphi,m)=1$ syss $v(\varphi,m)=0$
(iii) $v(\varphi\rightarrow\psi,m)=1$ syss $v(\varphi,m)=0$ o $v(\psi,m)=1$
(iv) $v(K_i\varphi,m)=1$ syss $\forall n \in M$: mR_in, $v(\varphi,n)=1$

La cláusula 3.(iv) expresa la interpretación modal usual de la noción de conocimiento, esto es, que un agente i sabe que φ ($K_i\varphi$), si φ es verdadero en todos los mundos que son compatibles con lo que i sabe.

Además, decimos que una fórmula φ es *válida* (y lo representamos como $\models\varphi$) si, para todo modelo y todo $m \in M$, $v(\varphi,m)=1$.

El *Sistema S5* de la lógica epistémica consta de los siguientes axiomas:

Ax1. Un conjunto suficiente de axiomas para derivar todas las tautologías de la LCP.
Ax2. $[K_ip \wedge K_i(p\rightarrow q)] \rightarrow K_iq$ *(Para i=1, ..., h)*
Ax3. $K_ip\rightarrow p$ *(Para i=1, ..., h)*
Ax4. $K_ip\rightarrow K_iK_ip$ *(Para i=1, ..., h)*
Ax5. $\neg K_ip\rightarrow K_i\neg K_ip$ *(Para i=1, ..., h)*

---

[2] A lo largo de este trabajo vamos a indicar que entre dos estados m y n se establece una relación de accesibilidad R_i de la forma siguiente: mR_in. Por otra parte, en el *Sistema S5* de la LEP, que será nuestro sistema axiomático de partida, las relaciones R_i son *relaciones de equivalencia*, esto es, *reflexivas*, *simétricas* y *transitivas*.

Como reglas de derivación tenemos el *Modus Ponens* (MP), la *Regla de Sustitución* y la *Generalización de K* (*Si* $\vdash\varphi$, *entonces* $\vdash K_i\varphi$, *para* $i=1, \ldots, h$).

Disponemos también de la siguiente definición metalingüística: $M_i\varphi =_{df} \neg K_i\neg\varphi$.

Por último, el *Sistema S5* es consistente y completo en relación a la semántica definida anteriormente[3].

3. Conocimiento conjunto y conocimiento común

Decimos que un grupo de *h* agentes tiene *conocimiento conjunto* de un hecho φ, si todos y cada uno de los agentes del grupo saben que φ. A su vez, φ es *conocimiento común* en dicho grupo si todos los miembros del mismo saben que φ, todos saben que todos lo saben, todos saben que todos saben que todos lo saben ... y así ad infinitum.

Las nociones de conocimiento conjunto y conocimiento común resultan extraordinariamente útiles para analizar el conocimiento de un grupo de agentes. Quien primero estudió formalmente estos conceptos fue D. Lewis en el contexto del análisis de las convenciones. Así, Lewis (1969) señala que para que algo funcione como una convención en un grupo de agentes, no basta con que ese algo forme parte del conocimiento conjunto de dicho grupo, sino que además ha de ser conocimiento común. Un sencillo ejemplo puede aclararnos esto: si el agente *i* cruza el paso de cebra cuando el semáforo está en verde, mientras que *j* detiene su coche cuando ve el semáforo en rojo, no es sólo porque *i* y *j* conozcan la convención que rige el funcionamiento de los semáforos, sino que ambos saben que ambos lo saben, y también saben que ambos saben que ambos lo saben, y así indefinidamente. De no ser así, ¿se atrevería *i* a cruzar la calle alguna vez?[4]. Estos conceptos también han tenido aplicación en estudios realizados en el ámbito de la economía, y asimismo intervienen en el entendimiento mismo

[3] Cf. Meyer y van der Hoek (1995, pp. 26-7).

[4] A nivel teórico y formal esta concatenación de conocimiento conjunto, en función de la cual se define el conocimiento común, es infinita. En la práctica está claro que el proceso no es infinito, pues actuamos sin necesidad de tener una seguridad total en que una convención de este tipo sea conocimiento común; en caso contrario, nadie habría cruzado un paso de cebra jamás, por mucho que un semáforo indicara que podía hacerlo.

de discursos donde hay suposiciones implícitas. Lewis (1969) sostiene que para que un oyente interprete correctamente lo que se propone comunicarle un hablante, es condición necesaria que toda la información contextual que se precisa para la interpretación sea no sólo conocimiento conjunto entre ambos agentes, sino que debe ser, además, conocimiento común. Respecto a este punto, Halpern (1986, p. 9) plantea el siguiente ejemplo: Supongamos que un agente *i* le pregunta a otro agente *j*: "¿Has visto la película que ponen en el Cine Roxy esta noche?". Pues bien, para que esta pregunta sea apropiadamente interpretada, no basta con que *i* y *j* sepan qué película ponen esta noche en el Cine Roxy, sino que *i* debe saber que *j* lo sabe, *j* debe saber que *i* sabe que *j* lo sabe, etc. Sobre estos importantes aspectos volveremos con más detalle un poco más adelante, cuando analicemos algunos problemas teóricos muy conocidos, como el enigma de los chicos embarrados o el problema del ataque coordinado, en función de las nociones epistémicas aquí presentadas. Ahora vamos a ocuparnos de los aspectos más formales de la lógica del conocimiento conjunto y del conocimiento común.

Necesitamos extender nuestro lenguaje para la LEP, presentado en el apartado 2, con los operadores monarios E (*"conocimiento conjunto"*) y C (*"conocimiento común"*), de tal modo que si φ es una fbf, también lo son $E\varphi$ (*"todos saben que* φ*"*) y $C\varphi$ (*"*φ *es conocimiento común"*).

De acuerdo con lo que hemos venido diciendo, tendremos, pues, las siguientes definiciones de nuestros nuevos operadores epistémicos:

$E\varphi =_{df} K_1\varphi \wedge \ldots \wedge K_h\varphi$ (siendo h el número de agentes del grupo).

$C\varphi =_{df} \varphi \wedge E\varphi \wedge EE\varphi \wedge EEE\varphi \wedge \ldots = \forall i \geq 0\ E^i\varphi$ (aquí $E^i\varphi$ es definido como $EEE\ldots\varphi$ -i operadores E-. Esta expresión sólo puede entenderse intuitivamente, en tanto que no podemos considerar infinitas conjunciones).

Para la definición semántica de E y C tomamos como base un modelo M como el presentado en el apartado 2, al que sólo necesitamos añadir las dos cláusulas siguientes para la función de evaluación v:

$v(E\varphi,m)=1$ *syss* $\forall n \in M$: mR_1n, $v(\varphi,n)=1$ *y ... y* $\forall n \in M$: mR_hn, $v(\varphi,n)=1$[5]
$v(C\varphi,m)=1$ *syss* $\forall n \in M$: mR^*n, $v(\varphi,n)=1$ *(también* mR^*m*)*

Si consideramos un único agente (es decir, si $h=1$), el conocimiento conjunto coincide con el conocimiento de dicho agente ($E\varphi \equiv K\varphi$).

Respecto a los nuevos operadores epistémicos *E* y *C*, podemos considerar los siguientes axiomas [Meyer y van der Hoek (1995, p. 46)]:

Ax6. $Ep \leftrightarrow K_1p \wedge ... \wedge K_hp$
Ax7. $Cp \rightarrow p$
Ax8. $Cp \rightarrow ECp$
Ax9. $[Cp \wedge C(p \rightarrow q)] \rightarrow Cq$
Ax10. $C(p \rightarrow Ep) \rightarrow (p \rightarrow Cp)$

El *Ax6* expresa la definición del operador *E* que hemos visto anteriormente. El *Ax7* nos indica que el conocimiento común establecido en un grupo de agentes versa siempre sobre hechos verdaderos. A su vez, los axiomas *Ax8*, *Ax9* (cierre bajo la implicación del operador *C*) y *Ax10* recogen la interacción formal entre los operadores *E* y *C*.

Además, introducimos la *Regla de Generalización de C*: *Si* $\vdash \varphi$*, entonces* $\vdash C\varphi$.

Como sistema axiomático para las nuevas nociones epistémicas merece destacar el *Sistema S5EC*, que consta del *Sistema S5* (recogido en el apartado 2), más los axiomas 6, 7, 8, 9 y 10 anteriores y la *Regla de Generalización de C*.

Este nuevo sistema es, asimismo, consistente y completo[6].

Podría pensarse que la noción de conocimiento común es baladí si contamos con la noción de conocimiento conjunto, pues cabría suponer que la primera no añade nada sustancial respecto a la segunda, a no ser una

[5] También podemos expresar la definición semántica de *E* de las dos formas siguientes:
$v(E\varphi,m)=1$ *syss* $\forall n \in M$: $(m,n) \in R_1 \cup ... \cup R_h$, $v(\varphi,n)=1$
$v(E\varphi,m)=1$ *syss* $v(K_1\varphi \wedge ... \wedge K_h\varphi,m)=1$

[6] Cf. Meyer y van der Hoek (1995, pp. 47-8).

especie de superconciencia colectiva respecto al conocimiento conjunto de un hecho, que nada importante aportaría en la práctica. Pero ya hemos visto cómo Lewis señala que para que algo funcione efectivamente como una convención para un conjunto de agentes, no basta con que dicho grupo tenga conocimiento conjunto de ese hecho, sino que se requiere, además, que sea conocimiento común. El ejemplo del semáforo que comentamos más arriba pretendía ilustrar esta idea. También podemos comprobar esto mediante el análisis de algunos conocidos problemas teóricos, valiéndonos de las nociones epistémicas que acabamos de presentar. Uno de estos problemas es el conocido como *enigma de los chicos embarrados*, que podemos encontrar recogido en muchos trabajos que estudian el conocimiento presente en un grupo de agentes.

El enigma de los chicos embarrados puede plantearse del siguiente modo[7]. Supongamos un número h de chicos que forman un círculo alrededor de su padre. Hay un número κ (con $1\leq\kappa\leq h$) de chicos que tienen la cabeza manchada de barro. Además, cada uno de ellos puede ver a todos los demás, pero no su propia cabeza, por lo cual cada chico sabe quiénes de los otros tienen barro en la cabeza, pero no si él mismo lo tiene. Se supone que los chicos no pueden comunicarse entre ellos, son completamente honestos (al igual que el padre) y bastante inteligentes (al menos por lo que al manejo de nociones epistémicas se refiere). Así las cosas, el padre dice en voz alta: “Hay al menos un chico que tiene barro en la cabeza. ¿Pueden dar un paso al frente todos los que sepan que su cabeza está manchada de barro?”

Si $\kappa>1$ ningún chico da un paso al frente. El padre formula la pregunta nuevamente (diciendo ahora “hay al menos dos chicos ...”). Si $\kappa>2$ tampoco hay respuesta por parte de los chicos. Este proceso se repite hasta que, cuando el padre ha formulado su pregunta κ veces, sorprendentemente todos los chicos que tienen la cabeza embarrada dan un paso al frente.

Las nociones de conocimiento conjunto y conocimiento común nos permiten analizar este problema y comprobar el porqué de su misterioso desenlace. Para ello vamos a describir cada situación o estado mediante una tupla de unos y ceros de la forma $<x_1, ..., x_h>$, donde $x_i=1$ si el chico i tiene

[7] En esta parte de nuestra exposición seguimos, en su mayor parte, el análisis de Fagin et al. (1995, pp. 3-7 y 24-30) y de Meyer y van der Hoek (1995, pp. 56-9).

la cabeza embarrada y $x_i=0$ si no la tiene. Así, por ejemplo, si $h=3$, una tupla de la forma $<1,0,1>$ expresaría que los chicos 1 y 3 tienen la cabeza embarrada, mientras que el 2 no la tiene. Supongamos que la tupla anterior describe la situación real. Entonces, antes de que el padre hable, el chico 1 consideraría posibles dos situaciones, descritas por las tuplas $<1,0,1>$ (la situación real) y $<0,0,1>$, puesto que no puede saber si su propia cabeza está o no manchada de barro. De igual modo, el chico 2 considera posibles dos situaciones: $<1,0,1>$ y $<1,1,1>$, y el chico 3 otras dos: $<1,0,1>$ y $<1,0,0>$.

Para nuestro análisis del problema vamos a necesitar un lenguaje formal que nos permita expresar si un chico del grupo tiene o no la cabeza manchada de barro. Por ello, tomamos un conjunto de variables proposicionales $\Phi=\{p_1,...,p_h, q_z, r_z\}$, donde, intuitivamente, p_i expresa que "el chico i tiene la cabeza embarrada", q_z, con $1\leq z\leq \kappa$, expresa que "al menos z chicos tienen la cabeza embarrada" y r_z que "hay exactamente z chicos con la cabeza manchada de barro". Vamos también a considerar una estructura $\mathcal{M}=<M, R_i, v>$, donde

1. M es un conjunto de estados o situaciones (cada una de las cuales está descrita por una tupla de la forma que hemos visto), siendo $|M|=2^h$.
2. $R_i\subseteq M^2$, son relaciones de accesibilidad entre situaciones, una para cada chico. Si consideramos dos situaciones $m,n \in M$, tales que el chico i considera posible el estado n desde el estado m, es decir, tales que mR_in, significa que las tuplas que representan ambas situaciones concuerdan en todo excepto posiblemente en el componente x_i. Esto supone que las relaciones R_i son relaciones de equivalencia.
3. v: $\Phi \text{x} M\rightarrow\{1,0\}$, es una función de asignación de valores de verdad a las variables proposicionales en cada situación. En nuestro caso, v cumple las tres condiciones siguientes:

(i) $v(p_i,<x_1,...,x_h>)=1$ *syss* $x_i=1$
(ii) $v(q_z,<x_1,...,x_h>)=1$ *syss* $z\leq\kappa$
(iii) $v(r_z,<x_1,...,x_h>)=1$ *syss* $z=\kappa$

Una vez que el padre enuncia en voz alta q_z, tenemos que Cq_z (y, por tanto Eq_z). Como cada chico embarrado sólo puede ver κ-1 chicos embarrados, si $z<\kappa$ ningún chico con la cabeza manchada de barro será capaz de concluir que él mismo es uno de los que tiene la cabeza manchada y, por consiguiente, ninguno dará un paso al frente; con lo cual tendremos para esa ronda: $C\neg r_z$ *(y* $\neg Er_z$*)*, y pasaríamos a la ronda siguiente, en la cual el padre enunciaría q_{z+1}. Por lo mismo, si $z=\kappa$ todo chico embarrado sabrá que él lo está y, por tanto, dará un paso al frente.

Así pues, por inducción sobre *z* (con $1\leq z\leq\kappa$) podemos probar que, tras la ronda número *z* de enunciados del padre, se obtiene Cq_z (y, por tanto, Eq_z). Entonces, tras la ronda número κ, se da Eq_κ, y así todos los chicos saben que al menos κ chicos tienen la cabeza embarrada. Y como cada chico embarrado sólo puede ver κ*-1* chicos embarrados, todo el que tenga la cabeza manchada sabrá que la tiene y, así, dará un paso al frente.

Si, como hemos venido suponiendo, la situación real estuviera descrita por la tupla *<1,0,1>*, tendríamos, antes de que el padre hablara, que $v(K_1\neg p_2,<1,0,1>)=1$, puesto que en esta situación el chico 2 no tiene la cabeza manchada de barro en ninguna de las dos situaciones que el chico 1 considera posibles. También se daría que $v(K_1p_3,<1,0,1>)=1$ y $v(\neg K_1p_1,<1,0,1>)=1$. De hecho, es conocimiento común que cada chico sabe si la cabeza de los demás está o no manchada de barro; serían válidas, pues, las dos expresiones siguientes: $C(p_i\rightarrow K_jp_i)$ y $C(\neg p_i\rightarrow K_j\neg p_i)$. Una vez que el padre dice en voz alta q_1, se da Cq_1 y Eq_1, y ninguno de los chicos es capaz de saber si su cabeza está manchada o no, pues 1 y 3 ven un chico manchado y 2 ve dos. En la siguiente ronda, cuando el padre dice q_2, se da Cq_2 y Eq_2. El chico 2 no da un paso al frente, pues ya sabía q_2, mientras que 1 y 3 sí lo dan, pues cada uno de ellos sólo ve un chico manchado de barro y, por tanto, deducen que ellos mismos han de estar manchados. En suma, cuando el padre enuncia q_κ se cumple que K_ir_κ para todo chico *i* que tiene la cabeza manchada de barro, con lo cual dará un paso al frente.

En las situaciones en las que hay dos o más chicos que tienen la cabeza manchada de barro, como ocurre con *<1,0,1>*, se cumple que *Ep* antes incluso de que el padre hable por primera vez. No obstante, después de que el padre diga que q_1, el estado del conocimiento del grupo varía, aun cuando todos los chicos supieran ya que q_1. Así, en la situación *<1,0,1>* el chico 1 considera posible la situación <*0,0,1*>, y en esta última el chico tres considera posible <*0,0,0*>. Por tanto, en el estado <*1,0,1*>, antes de que el padre hable, aunque todos saben que al menos uno de los chicos tiene barro en la cabeza, el chico 1 considera posible que el chico 3 considera posible que ninguno de ellos tiene la cabeza manchada de barro, lo cual deja de ocurrir una vez que el padre dice q_1.

Así pues, para que en el enigma de los chicos embarrados se dé el "misterioso" desenlace que hemos comentado, las aserciones del padre han de ser conocimiento común entre los miembros del grupo, no bastando con que sean simplemente conocimiento conjunto. Esto significa que las cosas no funcionarían, por ejemplo, si el padre se limitara a decir sus aserciones en el oído de cada chico sin que los demás supieran qué les decía, pues de esta forma nunca se alcanzaría tal conocimiento común.

Por otra parte, aunque la definición formal del conocimiento común encierra una conjunción infinita de operadores *E*, hemos de incidir en que esto no significa que en la práctica se requiera un tiempo infinito para que los chicos adquieran tal conocimiento común, sino que lo adquieren de una sola vez en cada ocasión que el padre hace su anuncio público. Como dijimos más arriba, la concatenación de operadores *E* es infinita a nivel teórico y formal, no así en el funcionamiento práctico del conocimiento común de un grupo de agentes.

Las nociones epistémicas que venimos presentando, en particular la de conocimiento común, también pueden ser aplicadas al estudio de sistemas distriduidos. El comportamiento de este tipo de sistemas puede ilustrarse mediante el muy discutido *problema del ataque coordinado*. Además, el análisis de este problema nos permite sacar a relucir los aspectos formales-epistémicos involucrados en la coordinación de ciertas acciones simultáneas.

Consideramos ahora una situación en la que hay dos regimientos de un ejército en dos colinas situadas a ambos lados de un valle que ocupa el ejército enemigo[8]. Si ambos regimientos atacan simultáneamente, vencerán al enemigo, si no es así, será el enemigo quien venza. Por ello, ninguno de los generales de los regimientos (llamémosles A y B) se decidirá a atacar a

---

[8] Para el planteamiento del problema seguimos, en líneas generales, la exposición del mismo que encontramos en Meyer y van der Hoek (1995, pp. 64-65).

menos que esté completamente seguro de que el otro regimiento atacará simultáneamente. El general A pretende coordinar un ataque con el general B. Para conseguir esto, sólo puede ponerse en contacto con él a través de un mensajero que lleve el mensaje p ("atacaremos el día D a la hora H"). Pero el mensajero ha de atravesar las líneas enemigas para cumplir su cometido, y el éxito de su empresa no está en absoluto garantizado. Por ello, para estar seguros de que el mensaje p ha sido recibido por el general B, éste último ha de enviar un acuse de recibo (K_Bp) al general A. Como este nuevo mensaje ha de ser enviado también por un mensajero, cuyo éxito tampoco es seguro, para que el ataque coordinado sea posible es necesario que A envíe a B un nuevo acuse de recibo expresando K_AK_Bp. Pero A necesita recibir ahora una confirmación de que B recibió este último mensaje, es decir, que $K_BK_AK_Bp$, pues de lo contrario no habría una garantía total de que los dos regimientos fueran a atacar simultáneamente. Este proceso tendría que repetirse infinitamente, pues siempre cabría la duda de si el último acuse de recibo fue o no recibido. Es decir, nos encontramos con que, a pesar de que es necesario el conocimiento común de p por parte de los generales A y B para estar seguros de que ambos atacarán simultáneamente, no es posible obtener tal conocimiento común en un tiempo finito.

Fagin et al. (1995) muestran que el conocimiento común de p (Cp) es necesario para solventar el problema del ataque coordinado, y sin embargo Cp no es obtenible en sistemas distribuidos en los que el acceso a la información pertinente por parte de los distintos componentes del sistema no está garantizado. Es decir, que el problema del ataque coordinado es irresoluble[9].

Ahora bien, como señala Halpern (1986, p. 11), para coordinar acciones en la práctica es suficiente con variantes del conocimiento común menos exigentes que la que venimos presentando, para cuyo establecimiento no sería necesaria una certeza absoluta respecto a que el flujo de información sea exitoso, sino que basta con que se tenga una confianza suficiente en que tal cosa ocurre. Algo parecido acontecería en el ejemplo del semáforo que comentábamos más arriba.

---

[9] Cf. Fagin et al. (1995, p. 140, teorema 4.5.4; y pp. 177-8).

4. Conocimiento implícito o distribuido en un grupo de agentes

Entre las nociones epistémicas relativas a grupos de agentes destaca la de *conocimiento implícito o distribuido*. Se dice que un grupo de agentes tiene conocimiento implícito o distribuido de un hecho φ, si el conocimiento "combinado" de los miembros de dicho grupo implica que φ [Fagin et al. (1995, p. 24)]. En general, para combinar el conocimiento de los agentes de un grupo eliminamos todos los mundos considerados epistémicamente imposibles por algún miembro del mismo. Esto lo hacemos formalmente como sigue.

Extendemos nuestro lenguaje formal para la LEP con el operador monario I ("*conocimiento implícito*"), resultando entonces que si φ es una fbf, también lo es $I\varphi$ ("*φ es conocimiento implícito*"). También extendemos el modelo semántico del apartado 2 mediante la siguiente definición de nuestro nuevo operador I:

$$v(I\varphi,m)=1 \text{ syss } \forall n \in M: (m,n) \in R_1 \cap \ldots \cap R_h,\ v(\varphi,n)=1$$

Meyer y van der Hoek (1995, p. 66) recogen el siguiente axioma relativo al conocimiento implícito:

$$Ax11.\ K_i\varphi \to I\varphi \quad (Para\ i=1,\ldots, h)$$

Claro está que este axioma sólo tiene sentido respecto a aquellos grupos de agentes en los que efectivamente se dé un conocimiento implícito, o, dicho de otro modo, respecto a aquellos grupos de agentes cuyos miembros comparten al menos una alternativa epistémica. Además, el *Ax11* se corresponde con la propiedad $R_I \subseteq R_i$, para $1 \leq i \leq h$, donde R_I representa la relación de accesibilidad asociada con el operador modal I.

En relación al conocimiento implícito cabe destacar el sistema axiomático *S5I*, que consta del *Sistema S5* más el *Ax11* y el *Sistema* $S5_I$[10]. Este *Sistema S5I* es consistente y completo[11].

---

[10] El *Sistema* $S5_I$ resulta de sustituir el operador K por el operador I en cada axioma del *Sistema S5*.

[11] Cf. Meyer y van der Hoek (1995, p. 67).

La comparación de las definiciones semánticas de los operadores *I* ("*conocimiento implícito*") y *E* ("*conocimiento conjunto*") nos muestra la diferencia que existe entre ambas nociones epistémicas. Si afirmamos, respecto a un conjunto de *h* agentes, que se da $E\varphi$, significa que todos y cada uno de los agentes del grupo, desde *1* hasta *h*, saben que φ. En cambio, $I\varphi$ indica que tal conocimiento está implícito o distribuido en el grupo de agentes, pudiendo darse que ninguno de los agentes del grupo sepa por sí mismo que φ (es decir: $\neg K_i\varphi$, para $i= 1, ..., h$).

Como señalan Meyer y van der Hoek (1995, p. 65), la intuición última tras la noción de conocimiento implícito o distribuido es que en un sistema cerrado un grupo de agentes cooperativos únicamente puede adquirir los conocimientos que están ya implícitos en dicho grupo. Y si los agentes de un grupo tal pudieran mancomunar su conocimiento, sólo podrían considerar como posibles aquellos mundos que fueran alternativas epistémicas para todos y cada uno de los agentes por separado.

Se suele expresar la intuición subyacente a $I\varphi$ como "*el hombre sabio conoce que* φ". Esta idea puede parecer, a primera vista, antiintuitiva, pues de acuerdo con ella cuanto menor sea el número de alternativas epistémicas que considere posibles un agente, más sabio será éste. Y así ha de ser, en efecto, si tenemos en cuenta la interpretación modal del conocimiento que se mantiene en la LEP, según la cual si un agente *i* sabe que φ, entonces φ es verdadero en todos los mundos que *i* considera epistémicamente posibles. Así pues, si un agente conoce un hecho φ, entonces sólo considera una posibilidad respecto a tal hecho, esto es, aquella en la que se da φ. En cambio, si el agente alberga dudas acerca de tal hecho, contemplará al respecto dos alternativas epistémicas, una en la que se da φ y otra en la que se da $\neg\varphi$, por lo que el número de sus alternativas epistémicas aumenta.

Para ilustrar algunas de las ideas que venimos exponiendo, consideremos un grupo formado por los agentes 1, 2 y 3. Tendremos que contemplar entonces tres relaciones de accesibilidad, una por agente, que denotamos por R_1, R_2 y R_3. Como las R_i son relaciones de equivalencia, y, por tanto, reflexivas, el mundo real (o actual) será siempre una de las alternativas epistémicas que consideren posibles los agentes 1, 2 y 3. Se da, pues, que $\models(p\wedge q)\rightarrow M_i(p\wedge q)$ (i= 1, 2, 3); aunque muy bien puede ocurrir que cada agente considere también posible la alternativa $\neg p$ y, por tanto se daría que $\neg K_1(p\wedge q)\wedge\neg K_2(p\wedge q)\wedge\neg K_3(p\wedge q)$. Veamos qué ocurre en cada caso:

(a) Supongamos que los agentes comparten alguna otra alternativa epistémica además de la correspondiente al mundo real. Por ejemplo, el agente 1 considera posible $p\wedge q$, $p\wedge\neg q$ y $\neg p\wedge\neg q$; el agente 2 tiene como alternativas epistémicas $p\wedge q$ y $p\wedge\neg q$; y el agente 3 considera posible $p\wedge q$, $p\wedge\neg q$ y $\neg p\wedge q$; dándose en el mundo real que $p\wedge q$. En función de esto tenemos que: $I[(p\wedge q)\vee(p\wedge\neg q)]$. Y, de acuerdo con lo dicho, podríamos afirmar que "el hombre sabio sabe que $(p\wedge q)\vee(p\wedge\neg q)$". Aquí vemos claramente que "el hombre sabio" (que en este caso coincide con el agente 2) es sabio no porque conozca lo que de hecho ocurre en el mundo real, sino porque sabe más que los demás miembros del grupo. Es decir, que tiene la suficiente información como para no admitir como posible ninguna alternativa epistémica que sea considerada imposible por alguno de los miembros del grupo. O también, tiene más información que los demás, y por eso puede eliminar alternativas no reales que, sin embargo, son consideradas posibles por otros miembros del grupo menos informados.

(b) El otro caso posible sería que la única alternativa epistémica que comparten los agentes del grupo sea la correspondiente al mundo real (en nuestro ejemplo $p\wedge q$). Entonces tendríamos que $I(p\wedge q)$. En este caso, "el hombre sabio" sabe más que los otros miembros del grupo por las mismas razones que antes, pero, además, es más sabio que "el hombre sabio" del apartado anterior, pues su información es tan completa que no considera posibles alternativas epistémicas distintas al mundo real en relación a los hechos p y q.

5. Consideraciones finales

Nociones como las de conocimiento común o conocimiento implícito, que acabamos de analizar, resultan muy útiles a la hora de estudiar los estados epistémicos de un grupo de agentes. De hecho, ambas nociones cuentan entre las más utilizadas para el estudio del conocimiento y la acción en contextos en los que intervienen múltiples agentes.

Al teorizar la acción colectiva, el conocimiento de los agentes constituye una de las variables fundamentales a tener en cuenta, pues dependiendo de tal conocimiento pueden seguirse unos u otros cursos de acción. En este trabajo hemos tratado de poner de manifiesto cómo las nociones epistémicas relativas a grupos de agentes pueden verse ampliamente clarificadas desde el paradigma de la lógica epistémica. Claro está que para el estudio de la acción colectiva, especialmente de sus vertientes formales, no basta con el desarrollo de una lógica del conocimiento, sino que también hay que contemplar la lógica de la creencia, la lógica temporal (que nos permita temporalizar estados mentales) y la lógica de los objetivos y las intenciones. En particular, resulta fácil desarrollar sistemas doxásticos en los que se formaliza las nociones de *creencia*, *creencia conjunta* y *creencia común* a partir de los elementos que hemos presentado para la lógica epistémica[12]. Asimismo, han sido propuestos desarrollos formales que tratan de configurar sistemas lógicos para los objetivos y las intenciones desde una perspectiva modal. También la lógica temporal nos proporciona interesantes elementos formales, articulándose sistemas que recogen distintas ontologías temporales (para tiempo lineal, ramificado hacia el futuro, etc.)[13]. Falta ahora desarrollar sistemas que combinen todos estos elementos en una lógica de la acción colectiva, de tal modo que pueda reflejarse el comportamiento formal de los mismos en situaciones que involucran a grupos de agentes.

BIBLIOGRAFÍA

Fagin, R. et al. (1995): *Reasoning about Knowledge*, MIT Press.

Halpern, J.Y. (1986): “Reasoning about Knowledge: An overview”, *Proceedings of the 1986 Conference Theoretical aspects of Reasoning about Knowledge*, Los Altos, California, Morgan Kaufmann Publishers, pp. 1-17.

Herrera González, J.R. & Vázquez Campos, M. (1999): “Combinaciones de lógica temporal y epistémica”, *Actas del Congreso La filosofía analítica en el cambio de milenio*, Servicio de Publicaciones de la Universidad de Santiago de Compostela, pp. 55-64.

[12] Cf., p.e., Kraus y Lehmann (1988).

[13] En Herrera & Vázquez (1999) se presentan algunas propuestas formales encaminadas a configurar sistemas lógicos temporales-epistémicos que permitan dar cuenta de cómo cambian los conocimientos y creencias de los sujetos a lo largo del tiempo.

Hintikka, J. (1962): *Knowledge and Belief*, Cornell, Cornell University Press.
Kraus, S. & Lehmann, D. (1988): "Knowledge, Belief and Time", *Theoretical Computer Science*, 58, 155-74.
Lewis, D.K. (1969): *Convention*, Cambridge, MA, Harvard University Press.
Meyer, J.-J. Ch. & van der Hoek (1995): *Epistemic Logic for AI and Computer Science*, Cambridge, Cambridge University Press.
Von Wright, G.H. (1951): *An Essay in Modal Logic*, Amsterdam, North-Holland.