

CONOCIMIENTO E INFORMACIÓN: ALGUNAS CUESTIONES REPRESENTACIONALES*

JESÚS M. LARRAZABAL Y FERNANDO MIGURA

ILCLI. Dpto. Lógica y Filosofía de la Ciencia
Universidad del País Vasco. Apdo. 220
20080 San Sebastián

El presente trabajo viene motivado por cierta insatisfacción producida por la ausencia de abordajes conjuntos de las nociones de conocimiento e información. Nuestra tarea consistirá en considerar las relaciones entre el conocimiento y la información desde una doble perspectiva: la que trata de dar cuenta del conocimiento partiendo de una idea objetiva de información y la que trata de dar cuenta de la información desde las formalizaciones modales clásicas del conocimiento. En un primer momento la introducción de una serie de distinciones conceptuales servirá al análisis de ambas nociones y sus relaciones, para más adelante contrastarlas con el concepto de conocimiento formalizado en las axiomatizaciones modales usuales.

Palabras clave: Información, Cognición, Representación, Semántica, Inferencia.

1. La concepción heredada

Desde la filosofía se ha definido clásicamente el conocimiento como la *creencia verdadera* que está *justificada*, vinculándolo a una empresa encaminada a la búsqueda de certezas. Sin embargo, no es desde la filosofía sino desde la ciencia, al afrontar algunas aplicaciones tecnológicas, de donde proviene la también clásica noción matemática de información entendida como "función de reducción de la incertidumbre".

La presente exposición, en cuanto adopta un punto de vista lógico e íntimamente ligado a los desarrollos de la Inteligencia Artificial (en adelante

* Este artículo ha sido escrito en el marco del Proyecto de Investigación del Gobierno Vasco "PGV9225". Nuestro agradecimiento a los miembros del Seminario de Lógica y del Seminario de Lenguaje y Comunicación de la Universidad del País Vasco por las críticas y sugerencias recibidas.

IA), se desmarca en el análisis de dichas nociones tanto de las cargas filosóficas como de las reducciones puramente cuantitativas.

1.1. Significado e información

Antes de seguir adelante sería conveniente deshacer una serie de malentendidos conceptuales relativos al conocimiento y la información que están fuertemente arraigados en ámbitos ligados a la IA. Adoptando una manera informal de hablar, en ocasiones se identifica la información que una señal nos proporciona con su significado, al decir que una señal es más "significativa" cuanto menos probable es su mensaje. Aunque la información que porta una señal está relacionada con su significado, veremos que sólo lo está de una manera accidental. A nuestro entender dicha confusión viene motivada, entre otras, por estas tres razones:

- Por los usos del término "significado" asociados a expresiones del tipo "el humo significa fuego" correspondientes a la distinción original de Grice (1957) entre "significado natural" y "significado no natural". Es únicamente el significado natural de una señal el que se puede identificar con la información que transmite.

- Porque por extensión se toma como identidad lo que no es sino una relación accidental entre el significado convencional de los signos lingüísticos y la información proporcionada por ellos, ya que de hecho al comunicarnos usualmente explotamos dicho significado para transmitir información. Expresamos información usando signos que tienen un significado que corresponde a la información que deseamos expresar.

- Porque en la Empresa de Representación del Conocimiento en IA ha sido dominante la perspectiva de la "interpretación de estructuras simbólicas" en una base de conocimiento. En este contexto la asignación de contenido informativo a las representaciones consiste simplemente en la especificación recursiva de sus condiciones de verdad por parte del diseñador del sistema, lo que usualmente se identifica con su "significado".

Ahora bien, si algo es claro es que las señales informativas han de ser distinguibles de las puramente significativas. Es condición necesaria para las primeras, aunque no para las segundas, que cuenten con la "propiedad

de veracidad". La información que nos proporciona una señal es lo que nos puede decir "con verdad" sobre un estado de cosas. Asimismo, consideramos necesario el carácter verídico de la información para que nos pueda proporcionar conocimiento.

Algo en lo que coinciden todos los análisis del concepto de conocimiento es que si bien un agente puede creer y desear lo que quiera, sólo puede conocer lo que es el caso. Por otra parte, cualquiera que sea la naturaleza de los "estados mentales" a los que denominamos "creencias", el sentido común nos dice que del hecho de que alguien sepa algo se sigue que lo crea.

Atendiendo a lo dicho podemos afirmar que el que un agente A sepa que p implica que A cree que p e implica, a su vez, que la proposición p es verdadera. Ahora bien, no basta que A crea que p y p sea el caso para que podamos calificar la creencia de A de conocimiento. Ni incluso si la creencia verdadera de A está justificada (piénsese en los clásicos contraejemplos de Gettier (1963)), a no ser que la idea de justificación incorporada excluya la posibilidad de que podamos estar justificados al creer algo falso.

1.2. El problema

Constatamos que en ámbitos como el de la Representación del Conocimiento se hace caso omiso del papel central jugado por la noción de verdad, perdiéndose distinciones conceptuales esenciales para una teoría rigurosa de la información. Como ya hemos dicho, la atribución de contenido informativo a las representaciones por parte del diseñador o usuario del sistema coincide con su significado. Puesto que la creencia entendida simplemente como "estado interno" del sistema contará con el significado asignado por la interpretación correspondiente, contará a su vez con la información correspondiente. En consecuencia dejará de tener sentido la distinción entre conocimiento y simple creencia. Si el enunciado p figura en la base de datos del sistema, éste operará con el dato p como si fuera verdadero independientemente de que lo sea, al margen de si proporciona o no información. Desde el punto de vista del sistema, todo es-

tado calificado como creencia constituye conocimiento. Todo lo que signifique que p es información de que p .

2. La perspectiva informacional

La precariedad de los análisis dedicados a una determinación formal del concepto de conocimiento es debida en gran medida a la ausencia de una caracterización apropiada de la idea de información. Usualmente la noción de información se ha dado como sobreentendida o como definida implícitamente en la interpretación de los estados del sistema o agente cognitivo modelizado.

Por nuestra parte, recogeremos selectivamente algunos de los rasgos de la idea semántica de información propuesta en Dretske (1981) y desarrollada en versiones (Barwise & Perry (1983), Barwise (1989), Israel & Perry (1990), Devlin (1991)) de la Teoría de Situaciones (en adelante TS).

La concepción de la información a la que nos referimos es aquella que asume un mundo "inundado" por información, entendida como un producto objetivo cuya generación y transmisión no requiere de su interpretación por parte de un agente cognitivo. Desde esta perspectiva la información se define en términos de una red de relaciones nómicas entre eventos, hechos, estados de cosas o situaciones. La naturaleza de la información es esencialmente relacional y se puede hacer manifiesta del modo siguiente: "x lleva la información de que y"

- El portador (x) de la información es un hecho (en la versión de la TS: lleva información al ser "soportado" por una situación-parte-del-mundo).

- La información (y) portada por un hecho es una proposición verdadera.

- Son las leyes, relaciones legaliformes, regularidades nómicas (modelizadas por las restricciones entre "tipos de situaciones" en la versión de la TS) las que permiten el flujo de la información, es decir, que un estado de cosas (una parte del mundo) nos proporcione información sobre otro estado de cosas (parte del mundo).

Estos rasgos generales de la idea de información son recogidos por la definición del contenido informativo propuesta en Dretske (1981):

Una señal r lleva la información de que p si y sólo si la probabilidad condicional de que p dada r (y K) es 1 (pero dado sólo K es menor que 1)¹.

Es claro que el contenido informativo de una señal r así entendido hace el hecho de que r lleve la información de que p dependiente de qué otros estados de cosas además de p pueden producir r .

En Dretske (1981) no sólo se propone una teoría semántica de la información sino que se pretende además erigir una epistemología sobre la misma. En este sentido es central el concepto de conocimiento defendido en el libro: un agente A sabe que p si y sólo si la creencia de A de que p está causada (o causalmente sostenida) por la información de que p .

En virtud de esta concepción del conocimiento vemos que éste no consiste simplemente en un estado mental en el sentido en que entendemos la creencia. Se nos hace manifiesta su naturaleza esencialmente relacional. Si bien el estado mental involucrado en el conocimiento es el de la creencia, lo que hace que una creencia sea calificable de conocimiento es su conexión con la información pertinente.

2.1. Semántica informacional y conocimiento

El hecho de que sean ciertas creencias las que constituyan conocimiento por razón de su origen informacional no nos hace concebir la creencia como elemento primitivo sobre el que se defina el conocimiento. En lo que sigue simplemente nos serviremos de la creencia y la información para destacar algunas propiedades estructurales satisfechas por el concepto de conocimiento desde la perspectiva informacional.

Asumida esta visión que parte de una idea objetiva de información es claro que el que una creencia se califique de conocimiento depende de su

¹ K hace referencia al conocimiento de que dispone el agente respecto a las alternativas relevantes en la fuente. Dicho brevemente, la consideración del parámetro K hace que el contenido de una señal sea relativo al conocimiento de que dispone el que la recibe.

origen o apoyo en la información. En consecuencia, no se exige del agente cognoscente que tenga evidencia de ese origen o soporte en la información. Desde su punto de vista todo cuenta como creencia. Desde la perspectiva del agente no hay manera de distinguir el conocimiento de la creencia, al menos cuando la creencia es adquirida de una fuente tomada por fiable.

Decir que el agente A conoce que p supone atribuirle conocimiento sobre la fiabilidad del origen de su creencia, esto es, supone concebir a un agente con acceso a la información sobre dicha fiabilidad.

Parece ser que a lo máximo que podemos aspirar desde el punto de vista del agente es a que crea que conoce. Pero incluso en este caso, lo que crea que conoce no tiene nada que ver con lo que realmente conoce, sino concierne únicamente a su estado de creencia sobre sus creencias particulares.

Así, desde la perspectiva informacional el concepto de conocimiento atribuido es "externo" al agente cognitivo. En última instancia, se trata de un concepto que sólo tiene sentido desde el punto de vista del teórico que modeliza los estados cognitivos del agente.

El que el agente conozca que conoce es ciertamente más plausible para una concepción "internalista" del conocimiento en la que el sujeto tiene acceso a la distinción entre simple creencia verdadera y conocimiento. Por ejemplo, cuando el agente razona sobre sus propios estados cognitivos, se tiende a pensar que si el sujeto posee justificación para su creencia de que p, entonces está justificado en creer que está así justificado al creer que p. Pero por contra desde un enfoque puramente externalista lo que cuenta como conocimiento depende de si existe una transmisión causal apropiada de información al sujeto cuando considera el estado de cosas objeto de su creencia. Así, desde el punto de vista del teórico que modeliza sus estados cognitivos, el agente conoce que p cuando está adecuadamente conectado a la "restricción" informacional correspondiente, pudiendo el agente no saber que conoce al no tener acceso al hecho de que se satisface una relación propiamente informativa. La propiedad intros-

pectiva discutida parece más apropiada para caracterizar estados de creencia en los que se atribuye al agente creencias sobre sus creencias.

Pero debe quedar claro que dicho externalismo no es exclusivo de la perspectiva informacional. La aproximación clásica a la Representación del Conocimiento basada en la interpretación de estructuras simbólicas es también externalista aunque por razones distintas. Desde el punto de vista de un Sistema Basado en Conocimiento toda representación en su base de datos constituye conocimiento "en relación al modelo propuesto por quien diseñó el sistema". Sin embargo nos encontramos ante "sistemas basados en creencias" desde un punto de vista filosófico absoluto, manifestado en la perspectiva del teórico.

3. Lógica y cognición

Adaptando un ejemplo que figura en Barwise & Perry (1983), supongamos que A sabe que el café de su taza es negro puesto que al servírselo la información adquirida mediante percepción, al verlo, hizo que se formara la creencia de que era negro. Asumamos, a su vez, que en el paso desde la percepción a la creencia se conserva la información. Atendiendo a la fiabilidad de las restricciones impuestas por nuestro aparato perceptivo y a nuestro conocimiento sobre la persistencia de los colores, podemos decir que la creencia de A de que el café es negro lleva la información de que el café es negro. Ahora bien, el estado cognitivo de A puede llevar esa información sin que A tenga una creencia al respecto, si p. e., no se plantea ese extremo al estar absorto en su trabajo. Aunque A estaría dispuesto a confirmar que el café de su taza es negro en caso de que se le interrogara al respecto.

Cualquier situación o estado de cosas lleva mucha información en la que ni creemos ni dejamos de creer. Así, el estado del estómago de A puede llevar información de que A ha consumido recientemente una taza de café, a pesar de que A lo haya olvidado. Es más, en ese caso nada impide decir que el estado de su estómago lleva la información de que A tomó café o menta, incluso si A no tiene noticia de que existan infusiones de menta.

La cuestión reside en que un agente cognitivo sólo puede saber o creer una cantidad finita de información, mientras que puede llevar (o codificar) una cantidad infinita de información.

Dretske (1981) expresa esta misma característica de la estructura de la información, al mantener que si la información de que q está analíticamente incluida en la información de que p , la señal, situación o estado que lleva la información de que p también lleva la información de que q . Así, por ejemplo, cuando en circunstancias normales el teléfono de nuestra casa suena no sólo lleva la información de que alguien nos llama, sino que además lleva la información de que nuestra línea funciona, de que alguien ha marcado nuestro número, de que algún teléfono de nuestra calle está sonando, etc.

En definitiva, la estructura de la información es "anidada" y una señal situación o estado que lleva la información de que p lleva, entre otras, toda la información lógicamente implicada por p y toda la incluida en p en virtud de las leyes de la naturaleza.

Esta característica de la información sirve para distinguirla tanto del concepto de significado no natural como de la idea de conocimiento entendido simplemente como "tener información".

Si alguien profiere el enunciado "Elisabeth vive en París" quiere decir que Elisabeth vive en París (sea lo que sea lo que la persona que lo emite quiera decir). Pero por supuesto no significa que Elisabeth vive en París o en Biarritz. Por otro lado, es claro que el enunciado implica que Elisabeth vive en París o en Biarritz, a pesar de que no signifique eso. Si el enunciado lleva la información de que Elisabeth vive en París, lleva por esa razón la información de que Elisabeth vive en París o en Biarritz al estar una información analíticamente incluida en la otra (en concreto, al ser lógicamente implicada por la otra).

De esta manera nos encontramos con que las propiedades del conocimiento se diferencian de la estructura de la información. Del hecho de que alguien crea p no se sigue necesariamente que crea q siempre que p implique q . Análogamente para el conocimiento. De esto se han dado perfecta cuenta los lógicos al tener que hacer frente al conocido problema

de la "omnisciencia lógica" que se deriva de las axiomatizaciones tradicionales del concepto de conocimiento.

Vemos que la estructura de la información así concebida la hace "cerrada bajo consecuencia lógica". Por contra no podemos decir lo mismo del conocimiento.

El hecho de que el conocimiento participe en un complejo entramado intencional de creencias y deseos distingue los agentes cognitivos de los simples procesadores de información como son los termostatos, voltímetros, etc.

Entendido así el conocimiento, es claro que el concepto de conocimiento es más fuerte que el de información. Alguien puede conocer que p , siguiéndose q de p (lógica o nómicamente), pero no creer que q por no saber que q se sigue de p . Pero incluso sabiéndolo, puede simplemente no "considerar" que q en virtud de sus disposiciones, creencias, deseos y propósitos inferenciales (fenómenos genuinamente intencionales). En este sentido alguien puede llevar la información de que q sin saber que q .

Alguien que crea que p y q , desde luego, cree que p y cree que q . Pero un agente puede, por un lado, creer que p y, por otro, que q sin tener necesariamente que creer que p y q , si no conecta mentalmente p y q , no teniendo ninguna disposición respecto a su conjunción lógica. Análogamente alguien puede creer que pasará la noche en Hendaya o en Irún sin creer que pasará la noche en Hendaya ni creer que pasará la noche en Irún.

Puesto en palabras de Barwise (1988):

La información viaja a la velocidad de la lógica, el conocimiento genuino únicamente viaja a la velocidad de la cognición y la inferencia.

Aunque en virtud de las precisiones anteriores parece ser que la lógica del conocimiento no es la lógica de la información, creemos que no estará de más la integración de ambas nociones en un marco común que nos

permita hablar con simplicidad de sus relaciones. En el siguiente apartado nos proponemos hacer esto mismo.

4. Lógica epistémica e información

Ya que la axiomatización de un concepto pretende integrar en un marco formal las propiedades destacadas en el análisis de dicho concepto, sin duda nos encontraremos con que las consideraciones previas tienen repercusiones en los intentos de plasmar formalmente un concepto de conocimiento basado en información.

Un buen punto de partida pueden ser las lógicas epistémicas propuestas para axiomatizar las nociones de creencia y conocimiento construídas sobre la semántica de los mundos posibles.

Como sabemos la idea intuitiva tras dicha semántica es que junto a los estados de cosas que se dan en la realidad, hay también otros estados de cosas posibles o mundos posibles. Algunos de esos mundos posibles pueden ser indistinguibles del mundo real para el agente. Así, decimos que un agente A conoce el hecho p si p es verdadero en todos los mundos que concibe posibles. Por ejemplo, si un agente piensa como posibles dos estados del mundo de modo que en uno el tiempo es soleado en P.A. mientras que en el otro llueve en P.A. y, sin embargo, en ambos llueve en S.S., entonces diremos que el agente sabe que llueve en S.S. pero no sabe qué tiempo hace en P.A.

Hasta aquí la historia es perfecta salvo en que tan sólo es una parte de la historia. En dicha caracterización del conocimiento se hace abstracción de un aspecto que estimamos ha de tenerse en cuenta. Nuestra intuición nos dice que en circunstancias normales el agente conoce que llueve en S.S. precisamente porque está en S.S. (o está debidamente conectado a una fuente que transmite de manera fiable la información relativa al tiempo en S.S.: radio,...) y no conoce el tiempo en P.A. porque se trata de una situación distante de la que no recibe la información pertinente, aunque podría recibirla. En el caso más simple la idea es que la información que el agente recibe al percibir visualmente el estado del tiempo en S.S. hace que sepa que llueve en S.S. y la ausencia de esta información, transmitida

por este u otro medio, hace que desconozca el tiempo en P.A. En última instancia, se trata de "situar" al agente y considerar sus conexiones con el flujo de información relevante. Dado que el agente tiene acceso a la información de que llueve en S.S., los estados de cosas compatibles con la información disponible en su situación actual son aquellos en que llueve en S.S. y llueve o hace sol en P.A.

Quizás se alegue que las axiomatizaciones usuales del conocimiento no pretenden dar cuenta de su origen o modos de adquisición. En todo caso, creemos que puede resultar esclarecedor considerar cómo la información y el conocimiento interactúan en un marco formal.

Como sabemos en las aproximaciones axiomáticas al concepto de conocimiento, éste se ha definido tradicionalmente como un operador (modal) epistémico (Hintikka (1962)).

En este orden, gran parte del trabajo en IA y en ciencias de la computación puede ser visto como un intento de axiomatización de la noción de información basado más o menos directamente en los desarrollos previos de las lógicas modales epistémicas. Por ejemplo, la semántica de los mundos posibles interpretada "parcialmente" ha servido para dar cuenta de las estructuras informacionales más simples definidas en términos de "extensiones o inclusiones entre estados informacionales": $\langle W, \leq \rangle$, donde " \leq " es al menos un orden parcial (van Benthem (1990)).

La delimitación intensional más simple para estas estructuras es la dada por una lógica proposicional modal, con una función de evaluación V en modelos $M = \langle W, \leq, V \rangle$ que caracterizan S4.

Por otra parte, como es sabido, viene siendo usual formalizar la noción de conocimiento mediante un operador modal fijado axiomáticamente por S5. Es bien sabido que dichas axiomatizaciones han sido ampliamente debatidas por ser poco realistas, debido a las consecuencias contraintuitivas que generan en relación al problema de la omnisciencia lógica.

En todo caso, hemos de admitir que el sistema S5 y otras propuestas cercanas (entre otras, KD45 para creencias) han contado con gran aceptación en ámbitos asociados a la IA. Piénsese en los sistemas distribuidos (Halpern & Moses (1984)) como intentos de caracterizar la estructura y el

flujo de la información entre los componentes de sistemas complejos en términos de la adscripción de conocimiento a esos componentes. Dicha noción de conocimiento corresponde a la que, definida en las estructuras de Kripke, satisface los axiomas de la lógica clásica del conocimiento S5.

5. Una propuesta de análisis formal

Nuestra propuesta sigue en cierta medida el camino inverso. Partiremos de un intento de plasmación formal de la idea objetiva de información que nos permita hablar del conocimiento basado en información.

En este acercamiento evitaremos detenernos en los detalles y asumiremos familiaridad con las semánticas kripkeanas y las axiomatizaciones modales más usuales (T, S4, S5, KD45), así como con la "Awareness Logic"² (Fagin & Halpern (1988).)

Frente a la insatisfacción producida por las soluciones sintácticas y mixtas para el problema de la omnisciencia lógica, optaremos por un abordaje semántico-informacional del mismo.

Trataremos de mostrar que las axiomatizaciones modales pueden servir para determinar las propiedades de un concepto de información inspirado en la "creencia implícita", definida en Levesque (1984), que permite un acercamiento a una noción de conocimiento realista fundada en el orden natural. Esta versión del conocimiento incorpora "rasgos psicológicos" al introducir el concepto de "awareness".

Sin más preámbulos pasamos a presentar ciertos conceptos necesarios para definir nuestra propuesta. Por simplicidad nos limitaremos a considerar el lenguaje de la lógica proposicional con sus conectivas usuales definido sobre un conjunto de proposiciones primitivas Φ , extendido con los operadores modales I y K (restringiendo su composición de manera que sus alcances no se interfieran) para componer fórmulas de la forma $I\phi$ y $K_i\phi$ interpretadas (con respecto a "situaciones") como "llevar la información de que ϕ " e "i conoce que ϕ ". Puesto que consideraremos el caso

² Traduiremos "to be aware" como "tener constancia de".

más simple con un único agente, en adelante prescindiremos de la referencia al mismo.

Sea $M = \langle S, B, V, F \rangle$ una estructura para nuestro lenguaje, donde S es un conjunto de situaciones primitivas (que son en esencia las propias de la TS). B es un subconjunto de S . Pensamos en B como en el conjunto de situaciones que podrían ser reales atendiendo a lo que es conocido y V y F son funciones que asignan a las proposiciones primitivas de Φ subconjuntos de S . Intuitivamente, $V(p)$ consta de todas las situaciones que soportan la verdad de la proposición p , mientras que $F(p)$ de todas las situaciones que soportan la falsedad de p . En esta versión de la aproximación modal junto a los mundos posibles clásicos que por definición describen configuraciones completas de estados de cosas relativas al conjunto Φ de proposiciones primitivas, contamos con situaciones que son parciales en cuanto que pueden no decidir sobre toda proposición primitiva. Las estructuras así introducidas no nos garantizan la verdad o falsedad de toda proposición primitiva p en una situación dada (si s es parcial, para alguna p , $s \notin V(p) \cup F(p)$), aunque, por otro lado, exijamos para ellas el requisito de coherencia por lo que ninguna situación soportará a un tiempo la verdad y la falsedad de una proposición primitiva (para ninguna situación s y proposición primitiva q se cumple que $s \in V(q) \cap F(q)$.)

Siguiendo la terminología clásica diremos que una situación completa o mundo posible es aquélla que soporta la verdad o falsedad de toda proposición primitiva: para toda situación completa s' y para todo p tenemos que $s' \in V(p) \cup F(p)$.

Tomando una versión de la relación de compatibilidad entre situaciones introducida en Levesque (1984) definimos el conjunto B^* de todos los mundos posibles o situaciones completas compatibles con las situaciones s en B como sigue:

Siendo $s \in B$, $B^*(s) = \{s' \in S \mid \text{para todo } p \in \Phi, s' \in V(p) \text{ o bien } s' \in F(p) \text{ y si } s \in V(p), \text{ entonces } s' \in V(p) \text{ y si } s \in F(p), \text{ entonces } s' \in F(p)\}$ Así tenemos que $B^* = \cup B^*(s)$ para toda s en B .

Puesto que esta semántica incorpora situaciones parciales la definición inductiva del concepto de verdad es sustituida por las relaciones de soporte \models_V y \models_F entre situaciones y fórmulas con las cláusulas siguientes:

1. a) $M, s \models_V p$, donde p es una proposición primitiva, syss $s \in V(p)$,
 b) $M, s \models_F p$, donde p es una proposición primitiva, syss $s \in F(p)$,
2. a) $M, s \models_V \neg\phi$ syss $M, s \models_F \phi$,
 b) $M, s \models_F \neg\phi$ syss $M, s \models_V \phi$,
3. a) $M, s \models_V \phi \wedge \psi$ syss $M, s \models_V \phi$ y $M, s \models_V \psi$,
 b) $M, s \models_F \phi \wedge \psi$ syss $M, s \models_F \phi$ o $M, s \models_F \psi$,
4. a) $M, s \models_V \phi \vee \psi$ syss $M, s \models_V \phi$ o $M, s \models_V \psi$,
 b) $M, s \models_F \phi \vee \psi$ syss $M, s \models_F \phi$ y $M, s \models_F \psi$,
5. a) $M, s \models_V I\phi$ syss $M, s' \models_V \phi$, para toda $s' \in B^*$.
 b) $M, s \models_F I\phi$ syss $M, s \models_V \neg I\phi$,
6. a) $M, s \models_V K\phi$ syss $M, t \models_V \phi$, para todo $t \in B$.
 b) $M, s \models_F K\phi$ syss $M, s \models_V \neg K\phi$.

Como es usual si $M, s' \models_V \phi$ y s' es un mundo posible (esto es, una situación completa) diremos que la fórmula ϕ es verdadera en s' . A su vez, diremos que ϕ es válida (esto es, $\models \phi$) si es verdadera en todas las situaciones completas para cualquier estructura M .

Una vez establecida la semántica para \models_v y \models_f se sigue sin dificultad que si un agente conoce que ϕ en s , entonces s lleva la información de que ϕ (y además soporta la verdad de ϕ).

Proposición:

$$\models K\phi \rightarrow I\phi^3$$

De lo expuesto también se sigue que es la información la que es cerrada bajo implicación:

Si una situación s lleva la información de que ϕ y ϕ implica ψ , entonces s lleva la información de que ψ .

En el caso de que ϕ sea una fórmula válida ($\models \phi$) tenemos que toda situación s lleva la información de que ϕ (lo que podría ser considerado como un caso degenerado de transmisión de información) aunque curiosamente no se sigue que toda situación s soporte la verdad de ϕ , con lo que esta presentación hace justicia a la intuición de que no toda situación es relevante para fijar el valor de ϕ a pesar de que ϕ sea lógicamente válida.

La exclusión de situaciones incoherentes en nuestra presentación hace que en nuestra propuesta se dé el cierre bajo el condicional: si $(K\phi \wedge K(\phi \rightarrow \psi)) \rightarrow K\psi$; aunque, eso sí, evite el cierre bajo la implicación y no exija el conocimiento de todas las fórmulas válidas (si bien $\phi \rightarrow (\phi \wedge (\psi \vee \neg\psi))$ es válida, $K\phi \wedge \neg K(\phi \wedge (\psi \vee \neg\psi))$ es satisficible así como $\neg K(\phi \vee \neg\phi)$), debido a la parcialidad de ciertas situaciones que hace posible que el agente cognoscente ni siquiera considere ciertas proposiciones.

En Levesque (1984) las situaciones incoherentes permiten al agente creer enunciados simultáneamente insatisficibles, p.e., $\phi \wedge \neg\phi$. Cuando tratamos con el concepto de conocimiento, tal posibilidad no parece adecua-

³ En esencia nos encontramos ante una interpretación alternativa de la proposición que establece que la creencia explícita implica a la creencia implícita tal como son definidas en Levesque (1984).

damente motivada. Además, opinamos, que en todo caso las intuiciones que empujan a tomar en cuenta tales creencias inconsistentes quedan mejor recogidas en el modelo de "razonamiento local" en términos de "marcos mentales" propuesto por Fagin & Halpern (1988).

En la versión del conocimiento que desarrollamos no se exige al agente que conozca todas las fórmulas lógicamente válidas ni las consecuencias de toda implicación lógica, debido a la parcialidad asociada a las situaciones. Sólo si el agente "tiene constancia de" una proposición ϕ (en símbolos, $A\phi$), puede ser conocido lo expresado por ella.

Si asociamos la siguiente semántica formal al operador "A" (tener constancia de):

$$\begin{aligned} M, s \models_v Ap \text{ syss } M, s \models_v K(p \vee \neg p), \\ M, s \models_f Ap \text{ syss } M, s \models_v Ap \end{aligned}$$

tenemos que un agente considera una proposición primitiva p syss conoce $p \vee \neg p$. De aquí se sigue que basta que el agente conozca p o conozca $\neg p$ para que tenga constancia de p .

En virtud de la semántica atribuida al operador que representa "tener constancia de" una proposición primitiva por parte del agente sostenemos que la fórmula Ap es verdadera sólo en esas situaciones que soportan la verdad o la falsedad de p . Así, $M, s \models Ap$ viene a expresar que p es de alguna forma relevante en s . Si tomamos prestado de Fagin & Halpern (1988) el conjunto $\text{Prim}(\phi)$ de proposiciones primitivas que aparecen en ϕ , en adelante, $A\phi$ será una abreviación para la conjunción de las fórmulas Ap , para todo $p \in \text{Prim}(\phi)$.

Es fácil ver que en virtud de las consideraciones anteriores se cumple el siguiente resultado:

Proposición:

Si ϕ es una fórmula booleana válida entonces

$$\models A\phi \rightarrow K\phi.$$

Fagin & Halpern (1988) introduce sucesivamente una versión sintáctica para el operador "A", una "lógica general" para "A" y una lógica del "razonamiento local" que suponen mejoras sobre la formulación original de Levesque (1984) en la medida en que permiten analizar las distintas propiedades de los operadores modales y sus relaciones en un marco más flexible.

En nuestra propuesta el conocimiento se define como la información que además es tomada en cuenta por el agente. En nuestra semántica se establece que un agente conoce ϕ si y sólo si se da la información de que ϕ y el agente tiene constancia de ϕ , esto es, $\models K\phi \leftrightarrow (I\phi \wedge A\phi)$.

No es nuestra intención llevar estas comparaciones hasta el límite sino estudiar las relaciones entre los conceptos de información y conocimiento en el presente marco.

La caracterización de la información que hemos presentado incorpora la estructura "anidada" de la información a la que aludíamos más arriba (sección 3). Toda situación lleva una cantidad infinita de información (incluidas todas las fórmulas proposicionales válidas) lo que podría ser considerado, como hemos dicho, un caso "degenerado" de transmisión de información. A pesar de la apariencia contraintuitiva de este resultado, puede ser plausible si reparamos en la distinción establecida en la TS entre la información "soportada" y la "llevada" por una situación. Una situación puede llevar cierta información sin hacerla ella misma verdadera. Por ejemplo, una situación-parte-del-mundo s en la que percibimos humo o en la que alguien dice que hay un incendio puede llevar la información de que en otra situación s' hay fuego en función de "restricciones" apropiadas, sin que en s mismo haya fuego. Podemos comprobar además que de la presente caracterización de la información se sigue su veracidad: si una situación s lleva cierta información, aunque ella misma no la haga verdadera, ésta habrá de ser verdadera en la situación completa correspondiente al mundo real. Por otro lado, por razones dadas más arriba, del hecho de que toda situación lleve la información de que ϕ , si ϕ es válida, no se sigue que toda situación soporte la verdad de ϕ .

Asimismo vemos que es la información llevada (o portada) por una situación la que es cerrada bajo consecuencia lógica.

Debe quedar claro que la interpretación anterior no constituye un todo acabado; al contrario, se trata de una primera aproximación a la estructura de la información inspirada en un marco, si no clásico, si más familiar que el de la perspectiva informacional. Sus insuficiencias son manifiestas. Hemos presentado la idea de información "total" o "absoluta" llevada por una situación haciendo abstracción de la referencia explícita a las restricciones que rigen las relaciones informacionales (Israel (1985)). Toda información es relativa a una o varias restricciones y los mismos hechos y situaciones pueden portar informaciones muy diferentes en virtud de las distintas restricciones consideradas. Sin duda el carácter esencialmente relacional de la información no se recoge en dicho marco. Nuestra próxima tarea será integrar en tal marco estas restricciones operantes en el flujo de la información.

Para terminar haremos algunas consideraciones sobre el concepto de conocimiento caracterizado en este marco.

Aquí el conocimiento no sufre en toda su gravedad del problema de la omnisciencia lógica al no ser cerrado bajo la implicación ni bajo la validez de las fórmulas, aunque sea cerrado bajo el condicional como lo manifiesta la validez del esquema $(K\phi \wedge K(\phi \rightarrow \psi)) \rightarrow K\psi$. Como sabemos la razón de su validez hay que buscarla en la ausencia en nuestra propuesta de situaciones inconsistentes.

Por otra parte, tampoco debemos olvidar que vista desde la información la validez de tal esquema dependerá en última instancia de las restricciones que imponamos sobre el operador A.

Entendiendo el conocimiento como la información disponible de la que además tiene constancia el agente $(K\phi \leftrightarrow (I\phi \wedge A\phi))$ se sigue el principio de veracidad para el conocimiento $K\phi \rightarrow \phi$.

Puesto que nos encontramos ante un concepto de conocimiento claramente externalista basado en información no nos debe extrañar que tanto el axioma de "introspección positiva" $(K\phi \rightarrow KK\phi)$ como el de "introspección negativa" $(\neg K\phi \rightarrow K\neg K\phi)$ dejen de ser válidos. La razón

de que el axioma de introspección positiva no se satisfaga en todos los casos reside en que el agente puede recibir la información de que ϕ ($I\phi$) y tener constancia de ϕ ($A\phi$) sin tener constancia de la información de que ϕ ($A(I\phi)$), esto es, que ϕ es información más que "pseudoinformación". Esto parece plausible desde una perspectiva puramente externalista como ésta en la que el agente no tiene acceso a la distinción entre simple creencia y conocimiento basado realmente en información. Como ya dijimos estos axiomas parecen más apropiados para una caracterización de la creencia en la que el agente razona sobre sus propias creencias.

Análogas consideraciones se extienden al axioma de introspección negativa.

Conclusiones

Entre las lecciones que podemos extraer de cuanto precede destacaremos en primer lugar la que sugiere que mientras un abordaje de los fenómenos informacionales desde la lógica puede ser apropiado, éste no es el caso sin más para los fenómenos asociados a la cognición. Aunque contemos con una noción de conocimiento basada en información, sus peculiaridades "psicológicas" escapan a una caracterización puramente lógica en términos de conceptos definidos "ad hoc" como el de "tener constancia de".

La plausibilidad de la idea de conocimiento que defendamos dependerá de hasta qué punto da cuenta de estos aspectos psicológicos derivados de su inserción en el correspondiente "entramado intencional". Asimismo nos parece que el sistema S5 no es el adecuado para formalizar un concepto externalista realista de conocimiento basado en información. La asimilación de las capacidades epistémicas del agente a los procesos inferenciales para un sistema formal correcto y completo respecto a un concepto absoluto de conocimiento desemboca en una propuesta poco realista. Otra cosa bien distinta es que para el análisis de sistemas artificiales se haya mostrado fructífera una atribución de propiedades epistémicas a los mismos a partir de una noción de conocimiento entendido como operador modal S5.

Para concluir diremos que aunque nos hayamos servido de una idea de información representada explícitamente mediante el operador modal I, no estamos defendiendo la utilidad de hacer explícita en el lenguaje tal noción más allá de un contexto analítico como el presente. Nuestra propuesta pretendía integrar en un mismo marco intensional una idea objetiva de información (con las limitaciones apuntadas más arriba) y un concepto de conocimiento definido en un ámbito que da entrada a la parcialidad.

REFERENCIAS BIBLIOGRÁFICAS

- Barwise, J. (1988): "On the Model Theory of the Common Knowledge", Stanford, CA: CSLI Report n° 122.
- Barwise, J. (1989): *The Situation in Logic*, Stanford, CA: CSLI Lecture Notes I, n° 17.
- Barwise, J. & J. Perry (1983): *Situation and Attitudes*, Cambridge, Mass.: MIT Press, 1983.
- Devlin, K. (1991): *Logic and Information*, Cambridge: Cambridge University Press, 1991.
- Dretske, F. (1981): *Knowledge and the Flow of Information*, Cambridge, Mass.: MIT Press, 1981.
- Fagin, R. & J. Y. Halpern (1988): "Belief, Awareness and Limited Reasoning", *Artificial Intelligence* 34, 77-96.
- Gettier, E. (1963): "Is Justified True Belief Knowledge?", *Analysis* 23, 121-123.
- Grice, H. (1957): "Meaning", *The Philosophical Review*, July 1957, 66, 377-88.
- Halpern, J. Y. & Y. O. Moses (1984): "Knowledge and Common Knowledge in a Distributed Environment". In: *Proceedings of 3rd ACM Conference on Principles of Distributed Computing*, 50-61.
- Hintikka, J. (1962): *Knowledge and Belief*, N.Y.: Cornell University Press. (V. cast. de J. J. Acero, *Saber y Creer. Una Introducción a la lógica de las Dos Nociones*. Madrid: Tecnos, 1979.)
- Israel, D. (1985): "Concepts of Information: Comparative Axiomatics". In: R. H. Thomason (ed.), *Philosophical Logic and Artificial Intelligence*, 35-72.

- Israel, D. & J. Perry (1990): "What is Information?", Stanford, CA: CSLI Report n° 145.
- Konolige, K. (1984): "Belief and Incompleteness", SRI Artificial Intelligence Note 319. Menlo Park, CA: SRI International.
- Levesque, H. (1984): "A Logic of Implicit and Explicit Belief". In: *Proceedings AAAI-84*, Austin, TX, 198-202.
- van Benthem, J. (1990): "Modal Logic as a Theory of Information", Stanford, CA: CSLI Report n° 144.