El giro dinámico en la epistemología formal: el caso del razonamiento explicativo*

(The Dynamic Turn in Formal Epistemology: the Case of Explanatory Reasoning)

Fernando SOLER TOSCANO

Recibido: 14.6.2012
Versión Final: 20.9.2012

BIBLID [0495-4548 (2014) 29: 80; pp. 181-199]
DOI: 10.1387/theoria.6347

ABSTRACT: We explore the possibilities that dynamic epistemic logic offers to model abductive reasoning. We show that many of the problems with formal approaches to abduction based on classical logic can be solved when considering an epistemic agent that reasons and acts.

Keywords: Abduction; dynamic epistemic logic; formal epistemology; plausibility models; belief revision

RESUMEN: Exploramos las posibilidades que ofrece la lógica epistémica dinámica para modelar el razonamiento abductivo. Mostramos que muchos de los problemas que encuentran los tratamientos formales de la abducción basados en lógica clásica pueden ser resueltos al considerar un agente epistémico que razona y actúa.

Descriptores: Abducción; lógica epistémica dinámica; epistemología formal; modelos de plausibilidad; revisión de creencias

1. *Introducción*

El razonamiento explicativo, o abductivo, ha despertado en los últimos años un gran interés en áreas tan diversas como la inteligencia artificial, la lingüística o la filosofía de la ciencia. Hintikka (1998) llegó a calificar la abducción como el problema fundamental de la epistemología contemporánea.

Para describir la abducción como forma de inferencia, es frecuente recurrir al esquema que proporciona C.S. Peirce[1]:

> El hecho sorprendente, φ, es observado.
> Pero si ψ fuera verdad, entonces φ sería aceptado como algo evidente.
> Por lo tanto, hay razón para sospechar que ψ es verdad (CP 5.189, 1903).

Este esquema, si bien es sugerente, resulta insuficiente para modelar una forma de inferencia que se pretende aplicar en una variedad tan grande de disciplinas como las que hemos citado más arriba. Por tanto, parece conveniente una sistematización, para la

* El presente artículo se ha realizado en el marco de los proyectos *Interpretaciones Alternativas de Lógicas no Clásicas* (Junta de Andalucía, HUM-5844) y *Conciencia, Lógica y Computación* (Ministerio de Ciencia e Innovación, FFI2011-29609-C02-01). El autor agradece a los dos revisores anónimos sus valiosos informes, que han contribuido notablemente a mejorar este trabajo.

[1] Introducimos las variables φ y ψ que serán usadas más adelante al hacer referencia a este esquema de Peirce.

cual se considera que la lógica puede resultar de utilidad. A pesar de que el tratamiento formal no llegue a agotar toda la variedad de prácticas inferenciales que en ocasiones se agrupan bajo el esquema de Peirce, el estudio lógico del razonamiento abductivo abre la puerta a numerosas aplicaciones (Kakas, Kowalski y Toni 1998). En (Aliseda 2006) aparecen las definiciones que se usan en lógica clásica para caracterizar los problemas abductivos y sus soluciones. Tomamos L como un lenguaje formal y $\models$ la relación de consecuencia lógica clásica.

Definición 1 (Problema abductivo). *Sean* $\Theta \subset L$ *y* $\varphi \in L$. *Decimos que* (Θ, φ) *es un* problema abductivo *sii (si y solo si):*

$$\Theta \not\models \varphi$$

Aliseda (2006) introduce, además, la siguiente clasificación,

- *Decimos que* φ *es una* novedad *sii* $\Theta \not\models \neg\varphi$
- *Decimos que* φ *es una* anomalía *sii* $\Theta \models \neg\varphi$

Definición 2 (Solución abductiva). *Sea* (Θ, φ) *un problema abductivo novedoso. Decimos que* $\psi \in L$ *es una* solución abductiva *al mismo sii*

$$\Theta \cup \{\psi\} \models \varphi$$

Llamamos a ψ *solución* plana. *Pero además,*

- *Decimos que* ψ *es una solución* consistente *sii* $\Theta \cup \{\psi\} \not\models \bot$
- *Decimos que* ψ *es una solución* explicativa *sii* $\psi \not\models \varphi$

Si (Θ, φ) es un problema abductivo anómalo, lo que usualmente se propone es comenzar haciendo la contracción $\Theta \ominus \neg\varphi$ para obtener el problema novedoso $(\Theta \ominus \neg\varphi, \varphi)$ y resolverlo tal como indica la definición 2. En cuanto a la contracción, se suele usar el modelo AGM (Alchourrón, Gärdenfors y Makinson 1985) como referencia.

Una de las discusiones habituales sobre la abducción es si esta se debe entender como un producto o un proceso. Es decir, si lo importante es que las conclusiones que se generan (productos) satisfagan la definición 2, o bien que el proceso seguido hasta obtenerlas satisfaga ciertas condiciones epistémicas sugeridas por el esquema de Peirce.

En cuanto a los productos, los procesos abductivos que han aparecido dentro de la lógica clásica (Reyes-Cabello, Aliseda y Nepomuceno-Fernández 2006, Soler-Toscano, Nepomuceno-Fernández y Aliseda 2009) ofrecen soluciones abductivas que satisfacen todos los requisitos de la definición 2. Sin embargo, las críticas que habitualmente reciben los métodos lógicos se dirigen hacia sus procesos. Hintikka (1998) toma cuatro tesis de Kapitan (1997) sobre razonamiento abductivo que se han convertido en requisitos que la epistemología exige a la lógica si queremos hablar de un proceso abductivo. Estas tesis son:

> **Tesis inferencial**. La abducción es, o incluye, un *proceso* inferencial.
> **Tesis de objetivo**. El propósito de la abducción científica es doble: en primer lugar generar nuevas hipótesis y posteriormente seleccionar las mejores para su análisis.

Tesis de comprensión. La abducción científica incluye todas las operaciones por las que se engendran las teorías.
Tesis de autonomía. La abducción es un tipo de razonamiento irreductible tanto a la deducción como a la inducción.

Los modelos que tradicionalmente se han formulado desde la lógica clásica (Soler-Toscano 2012) encuentran grandes dificultades para satisfacer, al menos, las tesis de *objetivo* y de *autonomía*. En cuanto a la primera, si bien se generan posibles soluciones abductivas, la selección de la mejor explicación es un problema al que los métodos lógicos difícilmente dan respuesta. Respecto de la tesis de *autonomía*, estos métodos suelen adolecer de una gran dependencia de la deducción[2] como consecuencia de la propia caracterización del razonamiento abductivo por medio de las definiciones 1 y 2.

Si bien existen tratamientos formales del razonamiento abductivo en lógicas no clásicas, siguen casi siempre los mismos patrones que los modelos clásicos. Se define igualmente el problema abductivo en términos de una teoría y una fórmula que no se infiere de ella y se busca completar la teoría. De este modo, estas propuestas de abducción en lógicas no clásicas caen en los mismos problemas.

Lo que proponemos en este trabajo es una reinterpretación del razonamiento abductivo dentro de la lógica epistémica dinámica (en adelante, LED). No ofrecemos ningún cálculo concreto, sino algo previo, pues a diferencia de lo que hemos comentado de otros acercamientos a la abducción en lógicas no clásicas, nos enfrentamos a la tarea de reinterpretar las propias nociones de problema y solución abductiva en el contexto de un agente epistémico que tiene información sobre su entorno y es capaz de actuar en función de dicha información.

El desarrollo de LED ha supuesto un cambio de orientación en la lógica, un *giro dinámico*, como lo llama van Benthem (2003), quien destaca que en LED se unen los productos con los procesos. Por ello, pensamos que LED ofrece un marco mucho más adecuado que la lógica clásica para modelar la abducción.

La sección 2 introduce la semántica de los *modelos de plausibilidad*, que servirá de referencia durante todo el trabajo. Las principales ideas del artículo se encuentran en la sección 3, donde se muestran las ventajas del uso de LED (mejor dicho, diversas variantes de LED) al estudio del razonamiento abductivo. Mostramos cómo algunos de los problemas difíciles de resolver en lógica clásica pueden ser elegantemente abordados con las nuevas herramientas.

2. *Modelos para el conocimiento y la creencia*

Una de las formas posibles de modelar los conocimientos y creencias de un agente la ofrecen los *modelos de plausibilidad* (Baltag y Smets 2008). En esta sección presentamos

[2] A lo largo de este trabajo, subrayamos en diversas ocasiones la dependencia del modelo clásico de razonamiento abductivo respecto de la deducción. Bien entendidas, las definiciones 1 y 2 no dependen de la deducción, sino de la relación de consecuencia lógica. Ahora bien, cuando en el modelo clásico se buscan cálculos para producir soluciones abductivas se recurre siempre, de una u otra forma, a cálculos deductivos. Por ello, la dependencia de la que hablamos no está en las definiciones clásicas en sí mismas, sino en los cálculos a los que dichas definiciones conducen.

los rudimentos de esta semántica, ya que más adelante estudiaremos las posibilidades de aplicación de estos modelos al razonamiento abductivo. No pretendemos que esta sea la única ni la mejor semántica para capturar todas las características de la abducción. De hecho, en algunos casos plantearemos extensiones y alternativas.

Definición 3 (Lenguaje $\mathcal{L}$)**.** *Dado un conjunto* P *de proposiciones, las fórmulas* φ *del lenguaje* $\mathcal{L}$ *están dadas por*

$$\varphi ::= p \mid \neg\varphi \mid \varphi \vee \varphi \mid \langle\leq\rangle\,\varphi \mid \langle\sim\rangle\,\varphi$$

donde $p \in \mathtt{P}$. *Las fórmulas de tipo* $\langle\leq\rangle\,\varphi$ *se leen "hay un mundo al menos tan plausible como el actual donde se verifica* φ*", y las fórmulas como* $\langle\sim\rangle\,\varphi$ *se leen "hay un mundo epistémicamente indistinguible del actual donde se verifica* φ*". Se pueden definir otras conectivas (*$\wedge$*,* $\rightarrow$*,* $\leftrightarrow$*), así como las modalidades universales* $[\leq]$ *y* $[\sim]$ *del modo habitual (*$[\leq]\,\varphi := \neg\langle\leq\rangle\,\neg\varphi$ *y* $[\sim]\,\varphi := \neg\langle\sim\rangle\,\neg\varphi$*).*

Como vemos, el lenguaje $\mathcal{L}$ es como el de la lógica proposicional al que incorporamos dos modalidades, $\langle\leq\rangle$ y $\langle\sim\rangle$, que nos van a permitir definir las nociones de creencia y conocimiento. Como se ha apuntado, estas nociones dependerán de un *orden de plausibilidad* que el agente establece entre los mundos del modelo. Veamos cómo se construyen estos modelos.

Definición 4 (Modelo de plausibilidad)**.** *Sea* P *un conjunto de proposiciones atómicas. Un* modelo de plausibilidad *es una estructura* $M = \langle W, \leq, V\rangle$ *donde*

- W *es un conjunto no vacío de* mundos posibles.
- $\leq\, \subseteq (W \times W)$ *es un preorden localmente conexo, cuya inversa es bien fundada*[3]*, que se llama* relación de plausibilidad*, y representa el orden de plausibilidad que el agente establece entre los mundos (leemos* $w \leq u$ *como "*u *es al menos tan plausible como* w*").*
- $V : W \rightarrow \wp(\mathtt{P})$ *es la* función de evaluación atómica *que indica qué proposiciones son verdaderas en cada mundo.*

Representamos mediante (M, w) *un modelo de plausibilidad* M *con un* estado distinguido $w \in W$.

A partir de la relación de plausibilidad $\leq$ podemos definir las creencias del agente como aquellas fórmulas que son verdaderas en los mundos más plausibles. Para definir el conocimiento del agente, necesitamos una relación de accesibilidad $\sim$ que divide el conjunto W en clases de equivalencia de mundos *indistinguibles*, tal como es habitual en LED (van Ditmarsch, van der Hoek y Kooi 2007). Así, el agente conocerá φ en el mundo

[3] Un preorden es una relación reflexiva y transitiva. Para ser localmente conexo se requiere que cada vez que dos elementos sean comparables a un tercero, esos dos elementos sean comparables entre sí, es decir: para todo $w, w_1, w_2 \in W$, si se dan Rww_1 o Rw_1w y también Rww_2 o Rw_2w, entonces tienen que darse Rw_1w_2 o Rw_2w_1. Finalmente, una relación $R \subseteq (W \times W)$ tiene su inversa bien fundada si no existe ninguna cadena $\overline{R}$-ascendente de longitud infinita, siendo $\overline{R}$ la versión *estricta* de R, es decir $\overline{R}wu$ sii Rwu y no Ruw.

w sii φ es verdadera en todos los mundos indistinguibles de w. Dadas las propiedades de $\leq$, podemos definir la *relación epistémica de indistinguibilidad* $\sim$ como la unión de $\leq$ y su inversa, es decir, $\sim := \leq \cup \geq$. Así, aunque un mundo u sea más plausible que w, $w \leq u$, ambos son epistémicamente posibles para el agente, y por ello la preferencia de u no excluye la posibilidad de w, por lo que a efectos de conocimiento debemos considerar por igual w y u.

Ahora podemos ver cómo se evalúa una fórmula en un modelo de plausibilidad. Las dos modalidades $\langle\leq\rangle$ y $\langle\sim\rangle$ se interpretan de forma estándar, con ayuda de sus respectivas relaciones.

Definición 5 (Evaluación semántica). *Sea* (M, w) *un modelo* $M = \langle W, \leq, V\rangle$ *con el estado distinguido* $w \in W$. *Indicamos mediante* $(M, w) \Vdash \psi$ *que la fórmula* $\psi \in \mathcal{L}$ *es verdadera en el estado* w *de* M. *Según la forma lógica de* ψ,

$$\begin{array}{lll}
(M,w) \Vdash p & \textit{sii} & p \in V(w), \textit{ para cualquier } p \in \mathtt{P} \\
(M,w) \Vdash \neg\varphi & \textit{sii} & (M,w) \nVdash \varphi \\
(M,w) \Vdash \varphi \wedge \psi & \textit{sii} & (M,w) \Vdash \varphi \textit{ y } (M,w) \Vdash \psi \\
(M,w) \Vdash \langle\leq\rangle\,\varphi & \textit{sii} & \textit{existe un } u \in W \textit{ tal que } w \leq u \textit{ y } (M,u) \Vdash \varphi \\
(M,w) \Vdash \langle\sim\rangle\,\varphi & \textit{sii} & \textit{existe un } u \in W \textit{ tal que } w \sim u \textit{ y } (M,u) \Vdash \varphi
\end{array}$$

Los operadores $[\leq]$ *y* $[\sim]$ *se definen como los duales de* $\langle\leq\rangle$ *y* $\langle\sim\rangle$, *respectivamente. Esto es,* $[\leq]\varphi := \neg\langle\leq\rangle\neg\varphi$ *y* $[\sim]\varphi := \neg\langle\sim\rangle\neg\varphi$.

3. *Caracterización del razonamiento explicativo en LED*

En esta sección presentamos las posibilidades que ofrecen las herramientas de LED para caracterizar el razonamiento abductivo. Vamos a referirnos principalmente a los modelos de plausibilidad, aunque en ocasiones mencionaremos otras lógicas. Comenzaremos por mostrar cómo LED permite una mejor caracterización del agente epistémico, sujeto del razonamiento abductivo. A continuación, pasaremos a analizar cómo se interpretan en LED las nociones de problema abductivo y solución abductiva.

3.1. *El agente epistémico*

El principal mérito de LED ha sido desplazar el centro de gravedad de la lógica desde la relación de consecuencia hasta un agente que tiene información sobre el mundo, sobre sí mismo y sobre otros agentes. Un agente capaz de actuar en función de la información que tiene. Esto nos permite igualmente reinterpretar el razonamiento abductivo. A continuación, presentamos diversas características del agente epistémico y las posibilidades que ofrecen para esta labor de reinterpretación.

3.1.1. Tipos de información

Al abordar el razonamiento abductivo desde la lógica clásica, la noción principal es la de teoría, de la cual se deducen ciertas fórmulas. Tenemos, pues, un solo tipo de información. Sin embargo, al tratar con agentes epistémicos se distinguen típicamente al menos dos tipos de información, conocimiento y creencia. Esta distinción resulta muy útil, como veremos, para modelar el razonamiento abductivo, ya que mientras que parte de la información del agente se puede interpretar como conocimiento, la solución abductiva solo podrá ser asimilada como creencia.

El conocimiento, en los modelos de plausibilidad, se define a partir de la relación de indistinguibilidad. Así, el agente conoce φ en cierto mundo w si y solo si φ es verdad en todos los mundos que el agente no puede distinguir de w, es decir, los que considera epistémicamente posibles. Pero, sin embargo, entre esos mundos existe un orden de plausibilidad, no todos son igualmente plausibles para el agente. Aquí entra la noción de creencia. El agente cree φ en cierto mundo w si y solo si φ es verdad en los mundos más plausibles a los que se puede llegar desde w. Debido a las propiedades de la relación de plausibilidad, φ es verdad en los mundos más plausibles desde w si y solo si siguiendo el orden de plausibilidad, a partir de cierto momento solo alcanzamos φ-mundos (Baltag y Smets 2008). Podemos expresar esta idea con las modalidades $\langle\leq\rangle$ y $[\leq]$. Formalmente,

El agente *conoce* φ	$K\varphi := [\sim]\,\varphi$
El agente *cree* φ	$B\varphi := \langle\leq\rangle\,[\leq]\,\varphi$

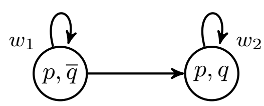

Figura 1: Ejemplo de modelo de plausibilidad

La figura 1 muestra un ejemplo de modelo de plausibilidad. Llamemos M a este modelo. La relación de plausibilidad $\leq$ está representada por las flechas entre mundos. Vemos que para el agente, el mundo w_2 es más plausible que w_1. En este caso, tenemos que p es verdadera en ambos mundos, pero q solo lo es en w_2. Así pues, el agente conoce p en w_1 pero no conoce q, es decir $(M, w_1) \Vdash Kp \wedge \neg Kq$. Sin embargo, el agente sí cree q, $(M, w_1) \Vdash Bq$. Por supuesto, también cree p, $(M, w_1) \Vdash Bp$.

Cuando una fórmula φ es verdadera en todos los estados de cierto modelo M decimos que es válida en M, y escribimos $M \Vdash \varphi$. En nuestro ejemplo, $M \Vdash Kp \wedge \neg Kq \wedge Bq$.

3.1.2. Agentes que actúan

En el modelo clásico de abducción, así como en AGM (Alchourrón, Gärdenfors y Makinson 1985), disponemos de operaciones para agregar o quitar información de la teoría. Se trata de operaciones sintácticas que hacemos “desde fuera”, por decirlo de alguna forma. Ahora tenemos un agente que puede actuar. Las acciones epistémicas modifican la información del agente. En este apartado vamos a ver las dos acciones principales que pueden realizar los agentes en los modelos de plausibilidad, una modifica el conocimiento y otra la creencia. Para más detalles sobre las propiedades de estas acciones ver (Baltag y Smets 2008).

La primera operación que vamos a presentar, la *observación*, modifica el conocimiento del agente. Se define de modo muy natural sobre los modelos de plausibilidad, pues consiste en eliminar todos los mundos donde la fórmula observada no se satisface, de modo que se reduce el dominio del modelo.

Definición 6 (Observación). *Sea* $M = \langle W, \leq, V\rangle$ *un modelo de plausibilidad. La operación de* observación *de* ψ *produce el modelo* $M_{\psi !} = \langle W', \leq', V'\rangle$ *donde* $W' := \{w \in W \mid (M, w) \Vdash \psi\}$*,* $\leq' := \leq \cap (W' \times W')$ *y, para cada* $w \in W'$*,* $V'(w) := V(w)$.

Esta operación elimina mundos de W, dejando solo aquellos que satisfacen la ψ observada. La relación de plausibilidad queda restringida a los mundos que sobreviven.

Otra operación que podemos hacer es modificar la relación de plausibilidad. Esto se puede hacer de diversas formas. La operación que llamamos *conjetura* es conocida como *radical upgrade* en la literatura.

Definición 7 (Conjetura). *Sea* $M = \langle W, \leq, V\rangle$ *un modelo de plausibilidad y* ψ *una fórmula. La* conjetura *de* ψ *produce el modelo* $M_{\psi\Uparrow} = \langle W, \leq', V\rangle$*, que solo difiere de* M *en la relación de plausibilidad, que ahora es:*

$$\begin{aligned}\leq' := &\big\{(w,u) \mid w \leq u \ y \ (M,u) \Vdash \psi\big\} \cup \\ &\big\{(w,u) \mid w \leq u \ y \ (M,w) \Vdash \neg\psi\big\} \cup \\ &\big\{(w,u) \mid w \sim u \ y \ (M,w) \Vdash \neg\psi \ y \ (M,u) \Vdash \psi\big\}\end{aligned}$$

La nueva relación de plausibilidad indica que después de conjeturar ψ, todos los ψ-mundos devienen más plausibles que todos los $\neg\psi$-mundos. El orden que hubiera previamente dentro de los ψ-mundos o dentro de los $\neg\psi$-mundos no cambia (van Benthem 2007). Esta operación preserva las propiedades de la relación de plausibilidad, tal como se muestra en (Velázquez-Quesada 2010).

Más adelante usaremos estas operaciones para redefinir las nociones de problema abductivo y solución abductiva en el marco de LED.

3.1.3. Agentes no omniscientes

Los agentes con los que trabajamos en lógica epistémica son habitualmente *omniscientes*, lo que significa que verifican las dos propiedades siguientes:

$$\begin{gathered}\Vdash \varphi \quad \text{implica} \quad \Vdash K\varphi \\ \Vdash K(\varphi \to \psi) \to (K\varphi \to K\psi)\end{gathered}$$

La primera propiedad implica que el agente conoce todas las tautologías, y la segunda que el agente conoce todas las consecuencias de su conocimiento. Es decir, que si el agente conoce $\varphi \to \psi$ y también conoce φ entonces necesariamente conoce ψ. Las dos propiedades juntas hacen que el conocimiento del agente sea cerrado bajo consecuencia lógica.

Las definiciones de conocimiento y creencia que hemos introducido en la sección 3.1.1 producen agentes omniscientes tanto en su conocimiento como en sus creencias.

La omnisciencia lógica, cuando se trata de modelar agentes reales, suele verse como un problema. En el momento que damos a un agente omnisciente la información φ le estamos dando todas las consecuencias lógicas de φ. Aunque en el tratamiento clásico de la abducción no se considera el agente epistémico, subyace un agente omnisciente. Como hemos probado en (Nepomuceno-Fernández, Soler-Toscano y Velázquez-Quesada 2012), también el modelo AGM presupone un agente omnisciente. De hecho, es frecuente considerar las teorías cerradas bajo consecuencia lógica. Incluso aunque no se consideren cerradas, las definiciones 1 y 2 trabajan de forma equivalente: algo es un problema abductivo o una solución en virtud de que se cumpla o no la relación de consecuencia lógica entre ciertos conjuntos de fórmulas.

Si queremos modelar el razonamiento de agentes reales no podemos usar una noción de conocimiento omnisciente. Es imposible que un científico disponga, al buscar una explicación, de todas las consecuencias lógicas de su conocimiento. Sin embargo, al trabajar desde la lógica clásica despreciamos este hecho. Afortunadamente, existen modificaciones de la lógica epistémica que permiten definir nociones de conocimiento y creencia no omniscientes (van Benthem y Velázquez-Quesada 2010).

Una de las las formas de modelar agentes no omniscientes es distinguir en la semántica entre información (conocimiento o creencia) *implícita* y *explícita*. Para ilustrarlo, podemos pensar en un agente que tiene la información p y la información $p \rightarrow q$, pero no ha aplicado la regla del modus ponens, por lo que no tiene la información q. Es información explícita la que el agente maneja conscientemente (usamos "conscientemente" en el sentido del término inglés "awareness", tal como se usa, por ejemplo, en (van Benthem y Velázquez-Quesada 2010), en nuestro caso p y $p \rightarrow q$) e implícita la que eventualmente podría obtener a partir de la información explícita (para nuestro agente, q). Como vemos, el conocimiento implícito puede, eventualmente, hacerse explícito si el agente *razona*. Así pues, el razonamiento se convierte en una operación fundamental para agentes no omniscientes.

Como sugerimos en (Nepomuceno-Fernández, Soler-Toscano y Velázquez-Quesada 2012), no tenemos por qué considerar que el conocimiento implícito contenga todas las consecuencias lógicas del conocimiento explícito, sino que podemos considerar agentes con capacidades deductivas incompletas. Para estos agentes, un problema abductivo no aparece cuando φ *no se sigue* de su información (como hace la definición clásica 1), sino cuando el agente *no puede derivar* φ (Soler-Toscano y Velázquez-Quesada 2012).

Para tratar desde la semántica la distinción entre conocimiento implícito y explícito, una posibilidad explorada en (Velázquez-Quesada 2009) consiste en dotar a cada mundo del modelo de plausibilidad de un conjunto A de fórmulas que el agente reconoce (explícitamente) como verdaderas en dicho mundo. Así, la noción estándar de conocimiento ($(M, w) \Vdash K\varphi$ sii $(M, u) \Vdash \varphi$ para todo u indistinguible de w) se convierte en el conocimiento implícito (φ es verdadera pero el agente no necesariamente lo reconoce). Para que el agente conozca explícitamente φ se requerirá no solo que φ sea verdadera en todos los mundos indistinguibles, sino que sea reconocida como verdadera (pertenezca al conjunto A) en todos ellos.

3.1.4. Sistemas multiagente

Los procesos de explicación se dan, en ocasiones, en el seno de una comunidad de agentes que razonan e intercambian información. La información que tiene cada uno de los agentes del grupo no tiene por qué ser la misma. Si los tratamientos clásicos de la abducción olvidan el papel del agente, aún más olvidan el lugar de la comunidad, ya sea la comunidad científica o simplemente dos agentes que discuten para buscar la explicación de algo que acaban de observar. Tradicionalmente se habla, como hacen las definiciones de la sección 1, de una teoría Θ, despreciando las distintas actitudes que cada uno de los agentes puede tener respecto de las fórmulas de Θ.

En LED disponemos de operadores de conocimiento de grupos, cuya aplicación resulta de sumo interés para el razonamiento explicativo. Dado un conjunto de agentes B, los operadores más habituales son los siguientes:

Es *conocimiento común* φ	$C_B\varphi$
Todos saben (es *conocimiento general*) φ	$E_B\varphi$
Es *conocimiento distribuido* φ	$D_B\varphi$

Para mostrar la semántica de estos operadores, debemos considerar una relación de indistinguibilidad $\sim_a$ para cada agente $a \in B$ (recordemos que en la definición 4 disponíamos de una única relación $\leq$ que usamos para definir $\sim$). Definimos también

$$\sim_E \;:=\; \bigcup_{a\in B} \sim_a \qquad \sim_D \;:=\; \bigcap_{a\in B} \sim_a$$

Comencemos por el operador más sencillo, el conocimiento general $E_B\varphi$, que requiere que todos los agentes de B conozcan φ. El conocimiento se evalúa según la relación de indistinguibilidad (ver las definiciones que dimos en el apartado 3.1.1). Así, el conocimiento general requiere que por cada uno de los agentes de B se cumpla que en todos los mundos indistinguibles para él se verifique φ. Formalmente,

$$(M, w) \Vdash E_B\varphi \;\text{ sii }\; \forall u \in W\,(\, w \sim_E u \;\Rightarrow (M, u) \Vdash \varphi)$$

El conocimiento distribuido es algo más débil. Como ejemplo, pensemos que cierto agente a conoce la fórmula p mientras que otro agente b conoce q. Ninguno de ellos conoce $p \wedge q$, sin embargo decimos que esta fórmula es conocimiento distribuido para ambos, $D_{\{a,b\}}(p \wedge q)$. Así pues, la idea intuitiva es unir el conocimiento de los agentes de B. Formalmente,

$$(M, w) \Vdash D_B\varphi \;\text{ sii }\; \forall u \in W\,(\, w \sim_D u \;\Rightarrow (M, u) \Vdash \varphi)$$

Finalmente, el conocimiento común se puede considerar como el punto fijo del operador de conocimiento general. Para que se dé $C_B\varphi$ se requiere que todos sepan φ ($E_B\varphi$), pero también que todos sepan que todos saben φ ($E_BE_B\varphi$), etc. Formalmente, si $\sim_C$ es el cierre de equivalencia de $\sim_E$,

$$(M, w) \Vdash C_B\varphi \;\text{ sii }\; \forall u \in W\,(\, w \sim_C u \;\Rightarrow (M, u) \Vdash \varphi)$$

Resulta interesante emplear estos operadores en la búsqueda de soluciones abductivas, especialmente el de conocimiento distribuido, dado que podemos pensar en soluciones abductivas distribuidas, a las que los agentes solo podrán llegar mediante acciones epistémicas por las que intercambian información. Así, para explicar el hecho sorprendente r, el agente a puede conocer $p \rightarrow q$, y b puede conocer $q \rightarrow r$. Entonces, b puede sugerir q como posible explicación, a lo que a sugerirá que p puede explicar q y por tanto r. De esta manera, tenemos una comunidad de agentes que se enfrentan de forma conjunta a la búsqueda de una explicación. Como vemos, LED nos ofrece herramientas para modelarlo.

3.2. El problema abductivo

La definición clásica de problema abductivo resulta muy pobre si la comparamos con la descripción de Peirce. Como vimos, se trata solo de comprobar que cierta fórmula no es consecuencia de la teoría, mientras que Peirce nos habla de la observación de un hecho sorprendente. A continuación mostramos cómo LED, sin pretender agotar toda la gama de fenómenos que en ciertas ocasiones se consideran incluidos en el razonamiento abductivo, permite modelar la noción de problema abductivo recogiendo ciertas ideas de Peirce sobre la observación y la sorpresa.

3.2.1. Acciones que producen un problema abductivo

En la definición clásica de problema abductivo, la fórmula φ constituye un problema abductivo porque no se deriva de la teoría Θ. ¿Pero de dónde viene φ? Para Peirce, φ es una observación, es decir, proviene de una acción epistémica del agente. Como hemos visto en la definición 6, podemos modelar en LED la acción de observar φ. ¿Qué significa, entonces, que φ sea un problema abductivo? Tras observar φ, si se trata de una fórmula proposicional[4], el agente conoce φ, por lo que no podemos decir que se produzca un problema abductivo si el agente no conoce φ. Sin embargo, podemos retroceder al momento anterior de observar φ; si el agente no conocía entonces φ, podemos decir que tras observarla se convierte en un problema abductivo. Formalmente,

$$\varphi \text{ es un problema abductivo en } (M_{\varphi!}, w) \text{ sii } (M, w) \nVdash K\varphi \tag{1}$$

De este modo, recogemos la idea de Peirce de que el problema abductivo aparece cuando el agente observa φ. En el siguiente apartado discutimos el carácter sorprendente de esta observación.

Hemos optado por definir la noción de problema abductivo en (1) en términos de conocimiento. Podríamos hacerlo también en términos de creencia y considerar que φ es un problema abductivo en $(M_{\varphi!}, w)$ sii $(M, w) \nVdash B\varphi$. Así, la condición para que φ sea un problema abductivo se hace más fuerte que en (1), ya que $\neg B\varphi$ implica $\neg K\varphi$.

[4] En el caso de que φ no sea una fórmula proposicional puede ser que después de que el agente la observa no la conozca. El caso más típico es $p \wedge \neg Kp$. Esta fórmula puede ser verdadera (si p es verdad pero el agente no lo sabe) pero sin embargo, si el agente la observa se vuelve falsa, por tanto, no la puede conocer. Por el contrario, en el caso de información proposicional, se cumple que siempre que se observa se vuelve conocida.

También podemos pensar en otro tipo de problemas abductivos que no surgen tras una observación sorprendente sino al preguntarnos el porqué de algo que ya conocíamos. Muchos autores ven como un problema abductivo, por ejemplo, la pregunta de si el quinto postulado de la geometría euclídea es independiente o no de los cuatro anteriores. ¿Qué es lo que hace alguien que se cuestione esto? Intentará, sin usar el quinto postulado, demostrarlo a partir de los anteriores. Gracias a la distinción entre conocimiento explícito e implícito que hemos presentado en el apartado 3.1.3 podemos modelar este tipo de problemas. En (Nepomuceno-Fernández, Soler-Toscano y Velázquez-Quesada 2012) definimos una acción $\langle Dis_\varphi \rangle$ que sirve para descartar φ del conocimiento del agente. Entonces, un agente no omnisciente puede realizar esta acción y preguntarse si con sus recursos deductivos puede derivar φ para hacerlo de nuevo explícito. En caso contrario, tratará φ como un problema abductivo.

En el siguiente apartado veremos que es posible, también, dar cuenta de anomalías.

3.2.2. Grados de sorpresa

Peirce nos dice que aparece un problema abductivo cuando se observa cierto hecho sorprendente φ. Hemos visto que LED nos permite modelar el agente que ejerce de sujeto de tal observación, así como la propia observación, pero nos falta ver cómo podemos modelar el carácter de *sorpresa*. Esta cuestión ha sido estudiada con detalle en (Lorini y Castelfranchi 2007). Ahora presentamos, con los recursos que hemos mostrado en el apartado 3.1.1, una de las posibilidades.

En (1) entendimos que φ es un problema abductivo tras ser observado si y solo si el agente no conocía φ antes de su observación. ¿Pero creía o no φ? Caben tres posibilidades:

- El agente lo creía, $B\varphi$. En este caso, no existe sorpresa, dado que el agente esperaba aquello que ha observado. Hay, más bien, la confirmación de una creencia. Aún así, el agente puede preguntarse por qué se ha dado φ y tendríamos una forma débil de problema abductivo. Tratará de justificar, en términos de conocimiento, una información que ya tenía como creencia, una vez que los hechos la han confirmado.

- El agente no creía φ ni su negación, $\neg B\varphi \wedge \neg B\neg\varphi$. En este caso sí hay sorpresa, porque el agente no esperaba encontrar φ. Se parece a la novedad abductiva de Aliseda (definición 1).

- El agente no creía φ pero sí creía su negación, $\neg B\varphi \wedge B\neg\varphi$. Aquí la sorpresa es mucho mayor, porque el agente observa algo contrario a sus creencias. De alguna forma, se corresponde con la anomalía de Aliseda (definición 1).

Obsérvese que no es posible modelar el caso en que el agente tenía conocimiento de $\neg\varphi$ antes de observar φ, ya que el conocimiento, en LED, se supone verdadero. Esto parece dejar fuera ciertos casos, bien estudiados en filosofía de la ciencia, de descubrimiento de inconsistencias en teorías científicas. Pero podemos modelar tales casos con nociones más débiles de creencias. Tal como hemos definido el operador $B\psi$, se requiere que ψ sea verdadera en los mundos más plausibles, y debido a las propiedades de la relación de plausibilidad, no son posibles creencias contradictorias. Pero existen nociones más débiles de creencia. Por ejemplo, podemos requerir que ψ sea verdadera en algún mundo más plausible que el actual. En ese caso, el agente tendría distintos grados de creencia para distintas fórmulas $\psi_1, \psi_2, \ldots$, en función de la plausibilidad de los mundos en que tales fórmulas sean verdaderas. Entonces, es posible tener creencias contradictorias si hay un mundo más plausible que el actual donde es verdad ψ y otro donde es verdad $\neg\psi$, con lo que podemos modelar las situaciones de teorías inconsistentes.

En agentes no omniscientes con información implícita y explícita, podemos considerar aún más grados de sorpresa, según haya o no creencia implícita o explícita de φ o su negación. Por ejemplo, si el agente no creía explícitamente φ antes de observarla pero la creía implícitamente, hay una sorpresa pero no muy fuerte, ya que el agente podría haber empleado sus recursos inferenciales antes de observar φ para hacer explícita su creencia, y estaría en el primer caso de los que enumeramos anteriormente, en el que como vimos no hay sorpresa.

3.2.3. Problemas abductivos de diferente grado

En este apartado vamos a presentar la posibilidad de estudiar desde LED ciertos tipos de problemas abductivos que no están en la formulación original de Peirce. Se supone que cuando el agente observa φ, se trata de un hecho que tiene que ver con el mundo, no con sus actitudes proposicionales, ni con las de otros agentes. Pero sin embargo, la formulación de problema abductivo dada en (1) es válida también cuando φ es una fórmula epistémica, es decir, aparte de variables proposicionales y conectivas booleanas contiene algún operador de conocimiento.

Un agente puede sorprenderse al observar su propia ignorancia (o conocimiento) sobre cierto hecho φ. Este es un problema abductivo distinto del que surge al observar φ. ¿Cómo puede ser que φ sea verdadero y yo no lo haya sabido? En este sentido hablamos de distintos grados de problemas abductivos, dado que podemos definir el grado epistémico de una fórmula como el mayor número de operadores epistémicos (de conocimiento o creencia) anidados que contenga. En la semántica, el grado epistémico de una fórmula determina el conjunto de mundos donde debe evaluarse. Así, para comprobar $(M, w) \Vdash p$ solo debemos atender a w, mientras que para evaluar $(M, w) \Vdash Kp$ acudimos a todos los mundos accesibles desde w, y para evaluar $(M, w) \Vdash KKp$ a todos los mundos accesibles desde mundos accesibles desde w y así sucesivamente. El grado epistémico cobra una importancia especial al considerar el conocimiento de varios agentes.

El estudio de estos problemas puede resultar interesante, especialmente por su relación con paradojas como la de Fitch (Salerno 2009). Dado que si un agente ignora cierta fórmula φ (de grado epistémico n) existen verdades incognoscibles como $\varphi \wedge \neg K\varphi$ (de grado $n + 1$), las fórmulas de este tipo pueden dar lugar a problemas abductivos irresolubles, ya que no existe ninguna acción epistémica que el agente pueda hacer para conocer $\varphi \wedge \neg K\varphi$.

3.3. La solución abductiva

Respecto de la solución abductiva, LED nos permite también una caracterización más rica que la definición 2, que se limitaba a comprobar ciertas relaciones de consecuencia lógica.

3.3.1. Acciones abductivas

Según el esquema de Peirce, el agente sabe que si ψ fuera verdadero entonces la verdad del hecho sorprendente φ sería algo obvio. En LED expresamos esto diciendo que el agente conoce $\psi \to \varphi$, es decir, $K(\psi \to \varphi)$. Entonces, cuando el agente se enfrenta a un problema abductivo φ y sabe $\psi \to \varphi$, ¿cómo lo soluciona? De nuevo, Peirce nos dice que *hay razones para sospechar* ψ. En este apartado discutimos qué significa *sospechar* ψ en LED y cómo podemos ver esta sospecha como una acción epistémica.

Algo que suele quedar fuera de los tratamientos lógicos del razonamiento abductivo es la integración de la solución. Precisamente porque en lógica clásica no hay manera de *sospechar* una fórmula. Pero en lógica epistémica disponemos de las creencias. Es una forma de información más débil que el conocimiento, como hemos visto. Por tanto, la creencia parece la candidata más idónea para ocupar el lugar de la sospecha. Así se distingue entre el conocimiento que el agente tiene de $\psi \to \varphi$ y la creencia de ψ, según la idea de Peirce.

Pero para que esa sospecha sea *razonable*, como nos dice Peirce, es necesario que el agente conozca $\psi \to \varphi$. Juntando todo lo expuesto, podemos decir que, dado el problema abductivo φ en $(M_{\varphi!}, w)$ (ver (1)), la fórmula ψ es una solución al mismo si y solo si

$$(M, w) \Vdash K(\psi \to \varphi) \tag{2}$$

Obsérvese que no podemos exigir $(M_{\varphi!}, w) \Vdash K(\psi \to \varphi)$ porque esto se verifica trivialmente en todos los casos en que φ es una fórmula proposicional, ya que tras observarla el agente la conoce, y por tanto sabe también $\psi \to \varphi$ para cualquier ψ.

¿Cómo hace ahora el agente para sospechar ψ? La acción abductiva más adecuada para integrar ψ en la información del agente sería la de *conjeturar* ψ (definición 7). De este modo, ψ se integra en la información del agente como una creencia.

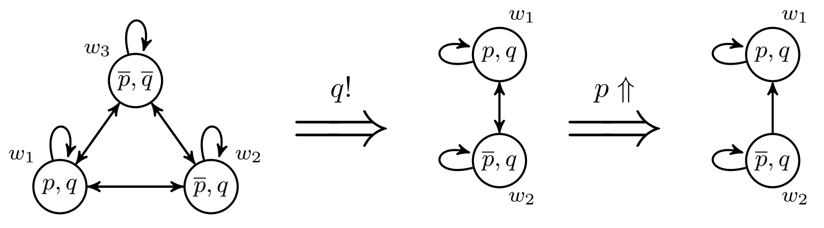

Figura 2: Solución de un problema abductivo

En la figura 2 encontramos ilustrado todo el proceso que hemos explicado. En el modelo de la izquierda se verifica $K(p \to q)$, sin embargo $\neg Kq$. En el modelo central, tras observar q, el agente sabe q, y por supuesto sigue sabiendo $p \to q$, es decir, $Kq \wedge K(p \to q)$. Entonces, q es un problema abductivo, dado que el agente lo conoce tras la observación pero no lo conocía antes. Igualmente, p es una solución posible. Tras realizar la acción abductiva de conjeturar p, a la derecha vemos que el agente cree p, Bp, dado que p es verdadero en el mundo más plausible.

3.3.2. Tipos de solución abductiva

En el apartado anterior hemos presentado los recursos que nos ofrece LED para determinar cuándo una fórmula se puede considerar una solución abductiva. Pero la definición 2 introduce dos tipos (no mutuamente excluyentes) de soluciones abductivas: consistentes y explicativas. Esta clasificación adolece del problema que hemos comentado varias veces: se definen los términos abductivos (en este caso soluciones) dependiendo de términos deductivos. Así, cierta solución será o no consistente o explicativa según se deduzcan de ella ciertas cosas.

Una de las aportaciones más importantes de LED al tratamiento del razonamiento abductivo consiste precisamente en poder definir estos tipos de solución sin referirlos a condiciones deductivas. Obviamente, no obtendremos las mismas propiedades de la definición 2, sino que tendremos que interpretar, para un agente epistémico, qué significa que una solución sea consistente o explicativa. Veamos cómo es posible hacerlo.

En primer lugar, la idea del criterio de consistencia es descartar aquellas explicaciones que son incompatibles con la información del agente. En lógica epistémica, una solución inconsistente ψ cumple que $K\neg\psi$, por lo que para que ψ sea consistente podemos exigir $\neg K\neg\psi$, lo que habitualmente se escribe $\widehat{K}\psi$. En la semántica de modelos de plausibilidad (definición 5) esto se interpreta como que existe algún mundo accesible (no importa si es más o menos plausible que el actual) donde ψ es verdadera.

Es más interesante aún el tratamiento de la condición explicativa en LED. Este requisito, que como vimos exige $\psi \not\models \varphi$ en su formulación clásica, trata de evitar soluciones triviales, que explicarían φ por sí mismas, sin el concurso del resto de la información que tiene el agente. Una de las posibles soluciones no explicativas es la propia φ así como cualquier fórmula de la que φ sea consecuencia. ¿Cómo podemos interpretar el requisito explicativo? Una posibilidad es considerar que una solución ψ es explicativa cuando su incorporación (que como hemos visto se puede hacer por medio de la acción de conjeturar ψ) produce un verdadero cambio epistémico en el agente. Para comprender qué significa esto, recordemos que la operación de conjeturar (definición 7) modifica la relación de plausibilidad entre mundos. Así pues, una conjetura que no produzca ningún cambio en la relación de plausibilidad será una conjetura trivial, y por tanto una solución no explicativa, porque no cambia nada en la información del agente. Al conjeturar la propia φ que suscitó el problema abductivo ocurre esto. A medida que se producen más cambios en la relación de plausibilidad, las conjeturas son más relevantes epistémicamente.

Además de la noción de creencia presentada en el apartado 3.1.1, existe la noción de *creencia fuerte*, que se define (Baltag y Smets 2009):

El agente *cree fuertemente* φ $\quad SB\varphi := [\sim](\varphi \to [\leq]\varphi)$

El agente cree fuertemente φ si y solo si todos los mundos donde φ es verdadero son más plausibles que todos los mundos donde φ es falso. Se trata, por tanto, de una forma de creencia más fuerte que $B\varphi$, que solo requería que φ fuera verdad en los mundos más plausibles. Con este nuevo tipo de creencia, podemos distinguir tres grados de relevancia explicativa de una posible solución abductiva ψ, en función de la situación epistémica del agente tras la observación del problema abductivo φ:

- Si $SB\psi$, entonces ψ *no es relevante*, ya que conjeturar ψ no aporta ningún cambio epistémico en el agente (la relación de plausibilidad se verá inalterada).
- Si $\neg SB\psi \wedge B\psi$, entonces ψ es *débilmente relevante*, ya que al conjeturar ψ se pueden producir cambios en las creencias del agente, pero ya antes creía ψ.
- Si $\neg B\psi$, entonces ψ es *fuertemente relevante*, ya que al conjeturar ψ se producirán grandes cambios en las creencias del agente que, además, pasará a creer ψ.

3.3.3. Selección de la mejor explicación

Uno de los mayores problemas con los que se enfrentan los tratamientos lógicos del razonamiento abductivo es el de la selección de la mejor explicación. Si bien desde la epistemología es usual relacionar la abducción con la inferencia de la mejor explicación (Lipton 1991), en la lógica clásica la selección queda como un criterio extralógico, pues la relación de consecuencia lógica solo nos dice qué se sigue y qué no, pero no tiene sentido preguntarse qué se sigue mejor. Incluso en la deducción, tampoco tiene sentido preguntar por la mejor conclusión de α y $\alpha \to \beta$, dado que hay infinitas posibles y todas son consecuencia lógica.

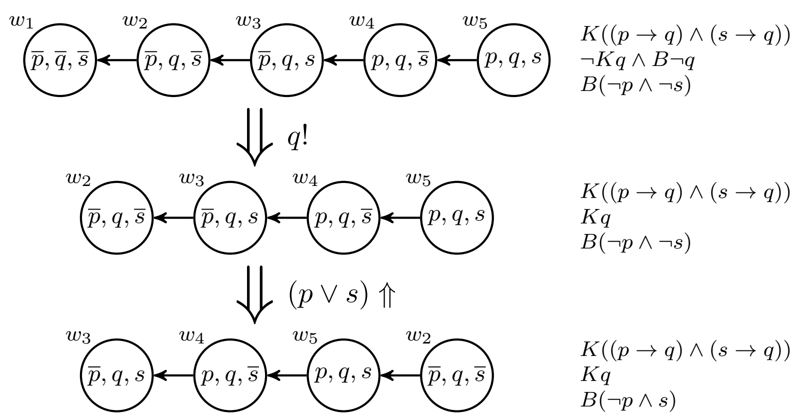

Figura 3: El agente selecciona la mejor explicación según sus preferencias previas

Sin embargo, en LED tiene sentido preguntarse por la mejor explicación. Al disponer en la semántica de una relación de plausibilidad entre mundos, la pregunta por la mejor explicación ya no resulta extralógica.

La figura 3 muestra una posibilidad para integrar dentro de la búsqueda de explicaciones la selección de la mejor. Hemos omitido algunas flechas en la relación de plausibilidad (aquellas debidas a la reflexividad y transitividad) para hacer más legibles los modelos. En la figura aparecen tres modelos, el primero de ellos corresponde a la información inicial del agente, donde conoce $p \to q$ y $s \to q$, pero no cree que sean verdaderas ni p ni q ni s. En la imagen central, aparece el modelo tras observar q. Vemos que el agente pasa de creer $\neg q$ a conocer q. Se produce así una sorpresa abductiva que se corresponde con una anomalía. El agente se pregunta por qué q, busca una explicación.

Dado que el agente conoce $p \to q$ y $s \to q$, dispone de dos explicaciones atómicas[5] posibles para q, a saber, p y s. Tal como explicamos en el apartado 3.3.1, el agente podría realizar la acción de conjeturar una cualquiera de las dos fórmulas y explicaría q. Pero, ¿cuál es la mejor explicación?

Aquí es donde podemos aprovechar las creencias previas del agente. Por ello encontramos sumamente elegante la solución que LED nos ofrece para seleccionar la mejor explicación, ya que descansa en las preferencias que el agente tenía antes de enfrentarse al problema abductivo. En este sentido, no existe "la mejor explicación", sino la mejor explicación para cierto agente, que puede ser distinta a la de otro agente cuyas creencias previas fueran distintas. Por ello encontramos tan natural esta solución.

Para aprovechar las creencias previas del agente, una posibilidad es conjeturar, no una explicación u otra (p o s), sino la disyunción de todas las explicaciones posibles, en este caso, $p \vee s$. El tercer modelo de la figura 3 muestra esta situación tras aplicar la acción $(p \vee s) \Uparrow$. El agente conjetura la disyunción de las dos explicaciones posibles e inmediatamente pasa a creer s, debido a que según sus preferencias previas el mundo w_3 era el más plausible de los que satisfacen $p \vee s$.

3.3.4. El carácter no monótono del razonamiento abductivo

Una de las propiedades más interesantes y mejor estudiadas del razonamiento abductivo es su no monotonía (Nepomuceno-Fernández y Soler-Toscano 2007). Es decir, aunque el agente proponga una solución abductiva, ésta puede quedar anulada en presencia de más información. La no monotonía es debida al carácter conjetural del razonamiento abductivo. A diferencia de la deducción, sus conclusiones son plausibles, no necesarias, y pueden ser refutadas en presencia de más información.

En LED tenemos también la posibilidad de modelar elegantemente la no monotonía del razonamiento abductivo. La figura 4 muestra qué es lo que ocurre si, tras explicar q en la figura 3 conjeturando $p \vee s$, el agente observa $\neg s$. La solución abductiva que el agente había obtenido, s, se demuestra falsa. Por tanto, tiene que revisar sus creencias de nuevo. Pero gracias a que en la figura 3 había realizado la acción abductiva $(p \vee s) \Uparrow$,

[5] Siguiendo la terminología de Aliseda tan solo existen las dos explicaciones *atómicas* p y s. Existen explicaciones *conjuntivas* como $p \wedge s$, pero no es *minimal*. A partir de las explicaciones atómicas se pueden obtener las *disyuntivas* como $p \vee s$. Nos centramos en las explicaciones atómicas.

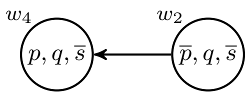

Figura 4: Selección de una nueva explicación

el agente ha pasado a creer la nueva explicación, p. Así, si inicialmente s se presentaba como la mejor explicación según las creencias previas del agente, la nueva información ha hecho que p sea ahora seleccionada.

El lector puede comprobar que si en la figura 3 el agente hubiera conjeturado s en vez de $p \vee s$, tras la observación de $\neg s$, la observación q volvería a quedarse sin explicación.

4. *Consideraciones finales*

La idea principal que hemos querido mostrar en este trabajo es que LED permite el estudio del razonamiento abductivo sin depender de un sistema deductivo previo. En el enfoque de la lógica clásica, parece como si lo primero fuera la deducción y en función de ella definimos qué es el razonamiento abductivo y cuáles son los criterios de calidad de las posibles soluciones. La deducción se convierte así en juez de la abducción. El problema con esto no es solo dar a la deducción una prioridad de difícil justificación epistémica (para muchos autores, la abducción está mucho más presente que la deducción) sino que algunas cuestiones fundamentales del razonamiento abductivo (como la selección de la mejor explicación) se convierten en sumamente complejas, y solo se pueden tratar de modo artificial.

Sin embargo, la lógica epistémica introduce una semántica que permite trabajar con conocimientos y creencias de los agentes. Frente al esquema tradicional de premisas y conclusión, tenemos un agente que maneja cierta información (al menos conocimiento y creencia, aunque existen otros tipos de información, como hemos visto) y que puede realizar acciones que la modifican. Hemos visto la posibilidad de definir acciones abductivas que cambian las creencias del agente. Así pues, el mismo agente puede aplicar en unos contextos razonamiento deductivo y en otros abductivo. Como hemos ido mostrando, este esquema resuelve de forma natural muchos de los problemas que típicamente encuentran los acercamientos formales al razonamiento abductivo basados en la lógica clásica.

REFERENCIAS

Alchourrón, C., P. Gärdenfors, y D. Makinson. 1985. On the logic of theory change: partial meet contraction and revision functions. *Journal of Symbolic Logic* 50: 510–530.

Aliseda, A. 2006. *Abductive reasoning: Logical investigations into discovery and explanation*, Volumen 330 de *Synthese Library*. Springer.

Baltag, A., y S. Smets. 2008. A qualitative theory of dynamic interactive belief revision. En *Logic and the Foundations of Game and Decision Theory (LOFT7)*, editado por G. Bonanno, W. van der Hoek, y M. Wooldridge, Volumen 3 de *Texts in Logic and Games*, 13–60. Amsterdam University Press.

Baltag, A., y S. Smets. 2009. Learning by questions and answers: From belief-revision cycles to doxastic fixed points. En *WoLLIC*, editado por H. Ono, M. Kanazawa, y R. J. G. B. de Queiroz, Volumen 5514 de *Lecture Notes in Computer Science*, 124–139. Springer.

Hintikka, J. 1998. What is abduction? The fundamental problem of contemporary epistemology. *Transactions of the Charles S. Peirce Society* 34: 503–533.

Kakas, A., R. Kowalski, y F. Toni. 1998. The role of abduction in logic programming. En *Handbook of logic in Artificial Intelligence and Logic Programming*, 235–324. Oxford University Press.

Kapitan, T. 1997. Peirce and the structure of abductive inference. En *Studies in the Logic of Charles Sanders Peirce*, editado por N. Houser, D. Roberts, y J. van Evra, 477–496. Indiana University Press.

Lipton, P. 1991. *Inference to the best explanation*. New York: Routledge.

Lorini, E., y C. Castelfranchi. 2007. The cognitive structure of surprise: looking for basic principles. *Topoi* 26: 133–149.

Nepomuceno-Fernández, A., y F. Soler-Toscano. 2007. Metamodelling abduction. *Theoria* 22: 285–293.

Nepomuceno-Fernández, A., F. Soler-Toscano, y F. R. Velázquez-Quesada. 2012. Dinámica de la información en agentes no omniscientes. En *Ensayos sobre lógica, lenguaje, mente y ciencia*. Alfar.

Reyes-Cabello, A. L., A. Aliseda, y Á. Nepomuceno-Fernández. 2006. Towards abductive reasoning in first-order logic. *Logic Journal of the IGPL* 14: 287–304.

Salerno, J, ed. 2009. *New essays on the knowability paradox*. Oxford University Press.

Soler-Toscano, F. 2012. *Razonamiento abductivo en lógica clásica*. College Publications.

Soler-Toscano, F., A. Nepomuceno-Fernández, y A. Aliseda. 2009. Abduction via c-tableaux and δ-resolution. *Journal of Applied Non-Classical Logics* 19: 211–225.

Soler-Toscano, F., y F. R. Velázquez-Quesada. 2012. A dynamic-epistemic approach to abductive reasoning. En *Logic of Knowledge. Theory and Applications*, editado por S. Magnier, F. J. Salguero, y C. Barés, Volumen 3 de *Dialogues and the Games of Logic. A Philosophical Perspective*. London, UK: College Publication (London).

van Benthem, J. 2003. Logic and the dynamics of information. *Minds and Machines* 13: 503–519.

van Benthem, J. 2007. Dynamic logic for belief revision. *Journal of Applied Non-Classical Logics* 17: 129–155.

van Benthem, J., y F. R. Velázquez-Quesada. 2010. The dynamics of awareness. *Synthese (Knowledge, Rationality and Action)* 177: 5–27.

van Ditmarsch, H., W. van der Hoek, y B. Kooi. 2007. *Dynamic epistemic logic*, Volumen 337 de *Synthese Library Series*. Springer.

Velázquez-Quesada, F. R. 2009. Inference and update. *Synthese (Knowledge, Rationality and Action)* 169: 283–300.

Velázquez-Quesada, F. R. 2010. Dynamic epistemic logic for implicit and explicit beliefs. En *MALLOW 2010*, editado por O. Boissier, A. E. F. Seghrouchni, S. Hassas, y N. Maudet, Volumen 627. CEUR Workshop Proceedings.

Fernando Soler Toscano pertenece al Grupo de Lógica, Lenguaje e Información de la Universidad de Sevilla. Trabaja en el desarrollo de modelos formales del razonamiento abductivo.

Dirección: Dpto. Filosofía, Lógica y Filosofía de la Ciencia. Universidad de Sevilla. C/ Camilo José Cela, s/n. Sevilla, 41018, España. E-mail: fsoler@us.es