

# **REPRESENTACIÓN ARGUMENTATIVA DEL RAZONAMIENTO SOBRE INTERPRETACIONES DEL CONOCIMIENTO CIENTÍFICO**

**Gustavo Adrián BODANZA**

*Universidad Nacional del Sur  
CONICET*

---

## **Abstract**

Scientific explanations and predictions are made of arguments. The theories provide the premises that, under specific initial conditions, support the conclusions: statements describing the phenomena to be explained or predicted. These arguments are defeasible, since one theory might elicit better arguments than another. This lead us to the idea that the process by which scientific explanations and predictions are justified can be studied through models of defeasible argumentation. This work introduces a simple model in which arguments conformed by structures of the form  $\langle \text{theory}, \text{fact} \rangle$  interact through a defeat relation towards a final acceptance (or rejection) within the represented scientific field. Taking relations like explicative strength and simplicity as implicit in the defeat relation, the model can represent the usual circumstances in science where a weakly explicative theory survives because of methodological advantages like simplicity.

## **Keywords**

< Representation of scientific knowledge > < Defeasible argumentation >  
< Scientific explanation >

---

1.

La representación del conocimiento ha devenido en uno de los temas de investigación principales de la inteligencia artificial, y ha motivando el trabajo de lógicos y filósofos en búsqueda de los fundamentos del razonamiento de sentido común (cf. por ejemplo McCarthy y Hayes, 1969). La no-



Fecha de recepción: 12 Sep. 2005 - Fecha de aceptación: 11 Ago. 2006

*Representaciones*, Vol. II, N° 2 – Nov. 2006, pp 1-15

© SIRCA Publicaciones Académicas – [representaciones@sircaestudios.org](mailto:representaciones@sircaestudios.org)

monotonía, *i.e.* la propiedad de que ciertas conclusiones preestablecidas pueden ser revisadas en presencia de nueva información, se ha identificado como una de las características que diferencian las inferencias de sentido común de las inferencias deductivas. Así, sobre fines de los años '70 distintas líneas de investigación empiezan a desarrollarse a fin de hallar formalismos que dieran cuenta de ella. Entre éstas se han destacado las lógicas *default* (Reiter, 1980), las lógicas no-monótonas (McDermott y Doyle, 1980), las lógicas autoepistémicas (Moore, 1985), y las lógicas circunscriptivas (McCarthy, 1980).

Alternativamente a estos enfoques, todos ellos basados en lógica, a mediados de los años '80 comienza a desarrollarse otro tipo de sistemas en los que la modelación matemática cobra importancia: los llamados "sistemas argumentativos" o "sistemas de argumentación rebatible" (Poole, 1985; Loui, 1987, 1991; Simari y Loui, 1992; Prakken, 1993; Vreeswijk, 1993; etc.). Según este enfoque, el razonamiento de sentido común procede construyendo argumentos a favor y en contra de una determinada tesis. De un análisis de las relaciones de preferencia y ataque que puedan darse entre los argumentos surgiría un subconjunto de argumentos justificados, los que darían una respuesta tentativa acerca de la aceptabilidad o no de la tesis. El carácter no-monótono del razonamiento de sentido común se da, bajo esta interpretación, cuando nuevos argumentos considerados pueden provocar el rechazo de otros previamente justificados.

El razonamiento no-monótono es manifiesto en la resolución cotidiana de problemas. Pero también podemos verlo activamente en el marco de una discusión científica. Como ejemplo claro, estamos pensando en el razonamiento de los físicos a partir de las distintas interpretaciones ontológicas de los modelos matemáticos que dan cuenta del comportamiento de la luz: las interpretaciones que la ven como ondas, las que la ven como partículas y las que la ven como una entidad de naturaleza dual onda-partícula. De lo que nos ocuparemos es de plantear la posibilidad de representar el uso de tales interpretaciones a través de un modelo de argumentación rebatible. Procuraremos ver, por ejemplo, si el uso que se hace de la interpretación ondulatoria de la luz para la explicación y predicción de fenómenos tales como la difracción o la interferencia, halla racionalidad al ser reconstruido sobre la base de un sistema argumentativo que, a la vez, incluye representaciones de información contraria a la hipótesis ondulatoria, como propiedades cuánticas o evidencia sobre el efecto fotoeléctrico. La tolerancia en la ciencia de incongruencias tales sugiere que los límites de la racio-

nalidad se extienden más allá de la clásica noción de consistencia. Los sistemas de razonamiento no-monótono, y en especial los sistemas de argumentación rebatible, constituyen un intento de llevar a cabo tal extensión. Aunque estos sistemas hayan sido desarrollados inicialmente con el propósito de representar el razonamiento de sentido común dentro de la inteligencia artificial, su aplicación al estudio del pensamiento científico ha sido prontamente aprovechado (cf. *Journal of Applied Logic*, 2004).

Debemos advertir que nuestro propósito es meramente exploratorio, lo que se apreciará en el tratamiento sobresimplificado del problema. Por otra parte debe quedar claro que hablamos de *interpretaciones* sugeridas por teorías o hipótesis y no de teorías o hipótesis mismas, y de *razonamiento de sentido común* en torno al pensamiento científico en vez de razonamiento científico a secas. En consecuencia, nuestro aporte no llegará a ser una reconstrucción sistematizada del razonamiento científico ni tampoco lo pretende.

En la sección 2 nos explayamos sobre las motivaciones para la construcción del modelo y los supuestos sobre los que se construirá. En la sección 3 exponemos algunos conceptos básicos de argumentación rebatible a través de un sencillo sistema construido sobre el de Dung (Dung, 1995). A través de esta herramienta, en la sección 4 modelaremos tres situaciones distintas en torno al dilema onda-partícula de la interpretación de la naturaleza de la luz, lo que nos permitirá justificar la aceptación de las distintas interpretaciones según el contexto de información (e.g., según se quiera explicar un fenómeno de interferencia o de efecto fotoeléctrico). Sobre la base de nuestro análisis concluiremos que el enfoque argumentativo permitiría una representación ajustada del problema. En la sección 5 presentamos una breve discusión sobre la clasificación de los sistemas no-monótonos en escépticos o crédulos, según su comportamiento frente a situaciones de conflicto sin salida. Esto es relevante ya que un caso especial del problema es el de la aceptabilidad de dos hipótesis rivales entre las que no hay ventajas claras. En la literatura de sistemas no-monótonos se caracterizaría como escéptico al comportamiento de rechazar a ambas, y de crédulo al de optar indistintamente entre una y otra, y distintos sistemas se han propuesto para modelar un variado rango de actitudes.

## 2.

Nuestro interés está enfocado en el razonamiento de sentido común tendiente a decidir qué hipótesis resulta más apropiada desde cierto punto de vista pragmático, específicamente, cómo los fenómenos observados sugieren interpretaciones en términos de determinadas hipótesis. Por otra parte, no pretendemos buscar una fundamentación epistemológica de las hipótesis. La estrategia consiste en asumir que, *desde el sentido común*, los científicos asocian a cada hipótesis la creencia de que sus consecuencias observacionales verificadas la confirman y, por lo tanto, sugieren una interpretación de la realidad observada en los términos de la hipótesis. Por ejemplo, como es sabido, la hipótesis ondulatoria de las radiaciones electromagnéticas, entre ellas, la luz, permite explicar de un modo sencillo fenómenos tales como la difracción, la reflexión, la refracción o la interferencia. Así podemos decir, por un lado, que de la hipótesis ondulatoria *O* se deduce que, bajo ciertos supuestos auxiliares *A* y dadas determinadas condiciones iniciales *C*, se observará el fenómeno de interferencia *I*; y por otro lado tenemos que dado el fenómeno *I*, si se verifican *A* y *C* entonces *prima facie* se concluirá *O*. Esta última es la creencia de sentido común acerca de la confirmación de la hipótesis que sugiere interpretar a los fenómenos lumínicos como manifestaciones de una entidad de tipo ondulatorio<sup>1</sup>. Consideremos ahora la hipótesis corpuscular *Q*, rival de la ondulatoria *O*, que permite explicar fenómenos que *O* no puede explicar, tales como el efecto fotoeléctrico. Respecto de esta hipótesis tendremos, por un lado, que de *Q* se deduce conjuntamente con ciertos supuestos auxiliares *A'* y dadas determinadas condiciones iniciales *C'* un enunciado observacional acerca del efecto fotoeléctrico *F*; y por otro lado tenemos que dado el fenómeno *F*, si se verifican *A'* y *C'* entonces *prima facie* se confirmará *Q*. A partir de esto, ya podemos vislumbrar cómo será la construcción de argumentos que nos interesa.

En segundo lugar, será de utilidad tener en cuenta además ciertas consideraciones acerca de cómo se maneja la contradicción en el plano del sentido común aplicado a las interpretaciones científicas. El caso de la mecánica cuántica es paradigmático en este punto. Ahora bien, pensemos que, previo a la aceptación de la dualidad de la luz como onda-partícula, la creencia en la naturaleza ondulatoria de la luz tenía cierto estatus *ontológico*, esto es, se creía que *realmente* la luz estaba constituida por ondas. Pero la aceptación posterior de la dualidad se mantiene en un plano que pare-

ce no ir más allá de lo *metodológico*, al apelar, según el fenómeno a tratar, a propiedades o *bien* ondulatorias o *bien* corpusculares. El caso del efecto fotoeléctrico es esclarecedor. El fenómeno se produce al hacer incidir un haz de luz sobre una placa metálica; entonces ocurre una eyección de partículas cargadas provocada por la energía radiante de la luz. La famosa fórmula de Einstein que gobierna este efecto,  $E_k = h\nu - \omega$ , expresa que la energía cinética  $E_k$  del electrón eyectado de la superficie del metal es una función de la frecuencia del rayo de luz incidente  $\nu$  ( $h$  es la constante de Planck, y  $\omega$  es el umbral de energía que debe superar el electrón para vencer la resistencia del metal que lo contiene). Ahora bien, de la hipótesis ondulatoria se puede inferir la implicación contrastadora de que la energía cinética del electrón eyectado variará con la intensidad de la onda lumínica incidente sobre el metal del cual el electrón se desprende. Sin embargo, puede establecerse experimentalmente que la energía cinética no varía con la intensidad de la luz. En cambio, como muestra la fórmula, la energía cinética  $E_k$  del electrón sí es una función de la frecuencia  $\nu$  de la luz. Esto hace falsa a la implicación contrastadora, lo cual implica a su vez que la hipótesis ondulatoria de la luz es también falsa. Pero si la hipótesis ondulatoria es falsa y la energía del electrón varía con la frecuencia ¿de qué hablamos cuando hablamos de ‘frecuencia’, si ésta es a su vez una propiedad ondulatoria? La solución *ad usum* a la paradoja que se presenta es *interpretar* la magnitud  $\nu$  “como si” se tratara de la frecuencia de una onda (lumínica) mientras la energía de la luz  $y$ , por lo tanto, su intensidad, se da en cuantos discretos, “como si” fuera de naturaleza corpuscular. Esto implica que, en el mismo fenómeno, se aceptan propiedades corpusculares de la luz junto con algunas propiedades ondulatorias, pero no todas, puesto que el fenómeno mismo las contradice. Lo que vemos, entonces, es que lo que inicialmente aparece como una restricción aparentemente ontológica, a saber, que la luz no puede *ser en realidad* a la vez onda y partícula, se debilita pasando a ser una restricción meramente metodológica, aplicable sólo en algunos casos. Por ejemplo, para calcular ciertos fenómenos como la difracción de la luz se obvian las propiedades cuánticas y se asume que se trata simplemente de ondas. Esto se debe a que la hipótesis ondulatoria facilita los cálculos a la vez que da resultados de igual precisión a los de la teoría cuántica para ese tipo de fenómenos.

Nótese que aquí no es necesario considerar a las hipótesis como enunciados o leyes *nórmicas* en el sentido de Schurz (2004) –tomado a su vez de Scriven (1959)–, es decir, enunciados del tipo ‘los A normalmente

son B', o 'los haces de luz normalmente se comportan como ondas'. Es claro que en este caso cualquier inferencia a partir de hipótesis o leyes normicas serían de carácter no-monótono, lo cual ha dado lugar a un amplio campo de investigación, como ya mencionamos antes. Pero no es por esta razón que nosotros hablamos del carácter no-monótono del razonamiento científico. Antes bien, aún cuando las hipótesis fueran consideradas enunciados universales estrictos, cualquier inferencia que justificara *prima facie* su confirmación daría como resultado una conclusión no-monótona, conclusión que en este caso sería la hipótesis misma. Es decir, inferir que una hipótesis está confirmada a partir de la observación de un fenómeno favorable es un razonamiento no-monótono, más allá del carácter de la hipótesis misma. Nuestro interés estará en estas inferencias en tanto son las que dan lugar a una interpretación posible de la realidad.

Lo que haremos a continuación es intentar una representación del cambio en las interpretaciones de este tipo mediante un formalismo adecuado, a saber, mediante la construcción de un sistema de argumentación rebatible. Del análisis anterior recordemos que tenemos las pautas sobre cómo vamos a construir los argumentos, y las restricciones que completan el contexto en el que éstos van a interactuar.

### 3.

Un marco argumentativo está compuesto, en líneas generales, por un conjunto de entidades llamadas *argumentos* y una relación binaria arbitraria entre argumentos interpretada como una relación de ataque. Cuáles son los argumentos y cuáles de éstos atacan a otros en un marco determinado, es algo decidido por el constructor del modelo. Esto expresa el concepto de *marco argumentativo* tal como es introducido por Dung: formalmente, se trata de un par  $\langle AR, \Rightarrow \rangle$ , donde  $AR$  es un conjunto de argumentos y  $\Rightarrow$  es una relación binaria entre los argumentos de  $AR$ , de modo que ' $\sigma \Rightarrow \tau$ ' significa que el argumento  $\sigma$  ataca al argumento  $\tau$ . Luego podemos dar un paso más —apartándonos de Dung— para definir la composición interna de cada argumento. Para nuestros fines, bastará asumir que un argumento es una estructura  $\sigma = \langle \Gamma_\sigma, c_\sigma \rangle$ , donde  $\Gamma_\sigma$  es un conjunto de premisas y  $c_\sigma$  es una conclusión, siendo premisas y conclusión fórmulas de un lenguaje de primer orden. La *base de información* con la que se construyen los argumentos de un marco está compuesta por dos conjuntos de fórmulas del lenguaje, uno,

$K$ , llamado *base de conocimiento*, y otro,  $\delta$ , a cuyos elementos llamamos reglas *default*. Para todo argumento  $\sigma$  tenemos que  $\Gamma_\sigma \subseteq \delta$  y  $\Gamma_\sigma \cup K \vdash c_\sigma$  donde mediante ' $\vdash$ ' simbolizamos la relación de consecuencia lógica clásica. Como criterio de racionalidad mínimo vamos a postular que, para cada argumento  $\sigma$ , el conjunto  $\Gamma_\sigma \cup K$  debe ser consistente. La única condición necesaria que vamos a imponer sobre la relación de ataque entre argumentos es que para que un argumento  $\sigma$  ataque a otro argumento  $\sigma'$  es necesario que  $\Gamma_{\sigma'} \cap \delta \neq \emptyset$ , es decir, que el argumento atacado esté construido con al menos una regla *default*. Esto es porque un argumento sin reglas *default* no es rebatible, sino concluyente (un argumento es concluyente si el aceptar  $K$  implica que necesariamente debe aceptarse la conclusión del mismo). Las condiciones suficientes del ataque, en cambio, deben ser definidas por el usuario del sistema, quien decide cuando un argumento tiene más fuerza que otro con el cual entra en conflicto<sup>2</sup>.

El conjunto de argumentos "justificados" en un marco, *i.e.* aquellos que resultarán "ganadores", se puede definir siguiendo distintos criterios según los cuales podría seleccionarse un único conjunto de argumentos o bien varios alternativos, diferencia que usualmente se explota para representar alternativamente decisiones "escépticas" y "crédulas"<sup>3</sup>. Para nuestros fines basta con considerar por ahora un criterio que determine un único conjunto de argumentos justificados. Seguimos en esto también a Dung, para usar su noción de *extensión fundada* (*grounded extension*), aunque introduciremos en ella ciertas modificaciones para adaptarla a nuestro propósito. Sea  $AR$  el conjunto de argumentos de un marco argumentativo genérico, y sea  $S$  un subconjunto de  $AR$ ; decimos que un argumento  $\sigma \in AR$  es *aceptable* en  $S$  si y sólo si para todo  $\sigma' \in AR$ , si  $\sigma' \Rightarrow \sigma$  ó  $\{\Gamma_\sigma, \Gamma_{\sigma'}\} \cup K \vdash \perp$ , entonces existe un argumento  $\sigma'' \in S$  tal que  $\sigma''$  ataca a  $\sigma'$ . Sea ahora  $Acept(S) = \{\sigma \in AR: \sigma \text{ es aceptable en } S\}$ , entonces la *extensión fundada* del marco es el menor punto fijo del operador  $Acept(\cdot)$ , o sea, el menor subconjunto  $S \subseteq AR$  tal que  $Acept(S) = S$ . Así, los argumentos justificados de un marco argumentativo son aquellos que pertenecen a la extensión fundada del marco.

#### 4.

Los argumentos que sugieren las interpretaciones que queremos representar dependen de los fenómenos que quieran explicarse o predecirse. Para modelar esto, podemos pensar distintas bases de información en torno

a un mismo conjunto de reglas *default* que indican cómo se conectan los enunciados observacionales con las interpretaciones de los mismos. Para nuestra conveniencia, y siguiendo las pautas de la sección 2, vamos a considerar reglas *default* bidireccionales, cada una de las cuales expresa, de izquierda a derecha, una implicación contrastadora (e.g., ‘si  $x$  es una onda electromagnética entonces  $x$  puede manifestar interferencias’), y de derecha a izquierda, una interpretación de un fenómeno bajo una hipótesis determinada, (e.g., que  $x$  manifieste interferencias “sugiere” que  $x$  es una onda electromagnética). En nuestro ejemplo, las reglas serán:

$$\delta = \{ \begin{array}{l} (\text{luz}(x) \wedge \text{onda}(x)) \leftrightarrow \text{interferencia}(x), \\ (\text{luz}(x) \wedge \text{partícula}(x)) \leftrightarrow \text{fotoeléctrico}(x), \\ (\text{luz}(x) \wedge \text{onda}(x)) \leftrightarrow \neg \text{fotoeléctrico}(x) \end{array} \}.^4$$

(donde ‘ $\text{luz}(x)$ ’ significa ‘ $x$  es un haz de luz’; ‘ $\text{interferencia}(x)$ ’ significa ‘ $x$  puede interferirse con otras radiaciones’; ‘ $\text{fotoeléctrico}(x)$ ’ significa ‘ $x$  produce efecto fotoeléctrico’; los demás predicados tienen el significado obvio.)

(i) Supongamos un primer sistema que consta de  $\delta$  y de una base de conocimiento  $K$  donde tenemos solamente los enunciados observacionales ‘ $\text{luz}(a)$ ’ e ‘ $\text{interferencia}(a)$ ’, donde ‘ $a$ ’ es una constante del lenguaje. Entonces podemos construir un marco argumentativo cuyos argumentos serán contruidos a partir de la pregunta ‘¿el haz de luz  $a$  es de naturaleza ondulatoria o está compuesto por partículas?’. Es decir, buscaremos todos los argumentos permitidos que concluyan alguna de las siguientes fórmulas: ‘ $\text{onda}(a)$ ’, ‘ $\text{partícula}(a)$ ’, ‘ $\neg \text{onda}(a)$ ’, ‘ $\neg \text{partícula}(a)$ ’. Así obtenemos dos argumentos:

$$\begin{array}{l} \sigma = \langle \{ (\text{luz}(x) \wedge \text{onda}(x)) \leftrightarrow \text{interferencia}(x) \}, \text{onda}(a) \rangle, \\ \sigma' = \langle \{ (\text{luz}(x) \wedge \text{onda}(x)) \leftrightarrow \text{interferencia}(x), (\text{luz}(x) \wedge \text{onda}(x)) \leftrightarrow \\ \neg \text{fotoeléctrico}(x), (\text{luz}(x) \wedge \text{partícula}(x)) \leftrightarrow \text{fotoeléctrico}(x) \}, \\ \neg \text{partícula}(a) \rangle. \end{array}$$

(Nótese que como en esta base de conocimiento no consta que  $a$  produzca el efecto fotoeléctrico, no hay problemas con la construcción de  $\sigma'$ .) En este contexto, así como no hay incompatibilidades entre  $\sigma$  y  $\sigma'$ , podemos pensar que tampoco hay ataques entre ellos (antes bien, parecen apoyarse mutuamente). Así la relación de ataque en este marco será  $\Rightarrow = \emptyset$ . Luego, ambos argumentos  $\sigma$  y  $\sigma'$  formarán parte de la extensión fundada del marco.

(ii) Supongamos ahora un segundo sistema donde cambiamos la base de conocimiento  $K$ , de modo que ahora tenemos la información  $K = \{ \text{luz}(a), \text{fotoeléctrico}(a) \}$ . Entonces, guiados por la misma pregunta que antes, no podremos construir los argumentos  $\sigma$  y  $\sigma'$ , en el último caso porque introduciría una contradicción con 'fotoeléctrico( $a$ )'. En cambio, tendríamos los argumentos:

$$\sigma'' = \langle \{ (\text{luz}(x) \wedge \text{partícula}(x)) \leftrightarrow \text{fotoeléctrico}(x) \}, \text{partícula}(a) \rangle;$$

$$\sigma''' = \langle \{ (\text{luz}(x) \wedge \text{onda}(x)) \leftrightarrow \neg \text{fotoeléctrico}(x) \}, \neg \text{onda}(a) \rangle;$$

Nuevamente, si no podemos determinar ataques entre  $\sigma$  y  $\sigma''$ , ambos argumentos formarán parte de la extensión fundada, concluyendo esta vez que  $a$  es partícula pero no onda.

(iii) Veamos ahora dos escenarios más interesantes, similares al anterior salvo que en éstos conocemos otro experimento donde el haz de luz  $a$  produce el efecto fotoeléctrico, o sea, ahora es parte de las respectivas bases de conocimiento el subconjunto  $\{ \text{luz}(a), \text{interferencia}(a), \text{fotoeléctrico}(a) \}$ . En el primer sistema (iii.a) la base de conocimiento será precisamente  $K = \{ \text{luz}(a), \text{interferencia}(a), \text{fotoeléctrico}(a) \}$ . Entonces los argumentos del marco —siempre guiados por la misma pregunta— serán  $\sigma$ ,  $\sigma''$  y  $\sigma'''$ . Nótese que  $\sigma'$  no puede construirse tampoco en este marco porque entraría en contradicción con  $K$ , ya que implica ' $\neg \text{fotoeléctrico}(a)$ '. Es claro que  $\sigma$  y  $\sigma'''$  no pueden ser aceptados consistentemente a la vez. Ahora bien, aunque no ocurrieran ataques entre los argumentos, igualmente tenemos una incompatibilidad entre  $\sigma$  y  $\sigma'''$ . Entonces la extensión fundada es  $\{\sigma''\}$ , justificando la interpretación corpuscular del efecto fotoeléctrico, y manteniendo la incertidumbre acerca de la interpretación ondulatoria. Por otra parte, es claro que si en el sistema se da  $\sigma \Rightarrow \sigma'''$  entonces la extensión fundada será  $\{\sigma, \sigma''\}$ , y si  $\sigma''' \Rightarrow \sigma$  entonces la extensión fundada será  $\{\sigma''', \sigma''\}$ . El primer caso se correspondería con la actitud de darle más importancia al éxito de la hipótesis ondulatoria que a su refutación, poniendo a esta última entre paréntesis como una mera anomalía, al estilo kuhniano. El segundo caso, en cambio, se correspondería con la actitud contraria, más al estilo popperiano. Ahora bien, también podemos pensar que haya alguna forma de aceptar o bien  $\sigma$  o bien  $\sigma'''$  cuando no hay ataques entre ellos —aunque sean incompatibles—. Para esto podemos apelar a otra noción de extensión. Una alternativa es la semántica de extensiones preferidas (*preferred extensions*) de Dung. Una *extensión preferida* es un conjunto máximamente admisible de argumentos. Un conjunto de argumentos es *admisible* si y sólo si todos sus argumentos

son aceptables en él, y no hay ataques internos ni inconsistencias en unión con  $K$ . En nuestro ejemplo, tendríamos dos extensiones preferidas, ambas conteniendo a  $\sigma$  (concluyendo que  $a$  es partícula), pero una conteniendo a  $\sigma$  pero no a  $\sigma''$  ( $a$  es además onda), y la otra conteniendo a  $\sigma''$  pero no a  $\sigma$  (no se concluye nada sobre si  $a$  es onda o no). Así, al sancionar distintas extensiones esta semántica permite representar la decisión de escoger *razonablemente* cualquiera de las extensiones, aunque no ambas a la vez. Nótese que en la primera se aceptaría una dualidad partícula-onda, mientras en la segunda no se descarta la hipótesis ondulatoria. La elección entre ambas alternativas coincidiría con la actitud corriente en la ciencia de apelar a una u otra interpretación según conveniencias prácticas, como por ejemplo, la facilidad de los cálculos o la ayuda intuitiva para la comprensión de los fenómenos.

En el sistema anterior, la validación de la dualidad partícula-onda en una de las extensiones preferidas se da gracias a que no existe en la base de conocimiento ninguna restricción que impida que una misma entidad se comporte de una y otra manera a la vez. Observemos qué ocurre en el segundo sistema (iii.b) donde vamos a suponer que en nuestra ontología ningún ente puede ser a la vez de naturaleza corpuscular y ondulatoria, tal como parecía indudable para la comunidad científica previa al desarrollo de la teoría cuántica. Aquí la base de conocimiento será  $K = \{ \text{luz}(a), \text{interferencia}(a), \text{fotoeléctrico}(a), \forall x (\text{partícula}(x) \rightarrow \neg \text{onda}(x)) \}$ , y los argumentos que pueden construirse son nuevamente  $\sigma$ ,  $\sigma''$  y  $\sigma'''$ . Supongamos que no se dan ataques entre ellos. Igualmente, tenemos incompatibilidad entre  $\sigma$  y  $\sigma'''$  y entre  $\sigma$  y  $\sigma''$ . Notemos que la extensión fundada de este sistema es  $\emptyset$ , es decir, desde un punto de vista más bien escéptico ningún argumento resulta convincente y, por ende, ninguna de las hipótesis. Así, la extensión fundada bien podría representar el estado de perplejidad producido por el choque entre el supuesto ontológico y los datos experimentales. Un punto de vista menos escéptico queda representado por las extensiones preferidas. El sistema sanciona dos:  $\{\sigma\}$ , que concluye que  $a$  es un ente ondulatorio y no concluye nada acerca de si es corpuscular, y  $\{\sigma'', \sigma'''\}$ , que concluye que  $a$  es corpuscular pero no ondulatorio. Lo que este ejemplo nos muestra es que, si son correctas las representaciones de los puntos de vista “escéptico” y “crédulo” a través de las extensiones fundadas y preferidas, respectivamente, la ciencia parece haberse inclinado por la credulidad antes que por el escepticismo, al menos en lo que respecta a la teoría de la luz.

## 5.

La última observación merece que nos detengamos un momento a analizar en qué medida estos conceptos de escepticismo y credulidad son correctamente representados por los distintos tipos de extensiones mencionados. En general, cuando se habla de escepticismo y credulidad en el campo del razonamiento no-monótono se hace referencia a la capacidad de los sistemas de generar menos o más conclusiones en los casos en que se presentan choques de inferencias irresolubles unívocamente, tal como ocurría en nuestros sistemas del punto (iii). En el caso de los sistemas argumentativos, no hablamos de menos o más *conclusiones*, lo cual es típico de las lógicas *default* o no-monótonas, sino de menos o más *argumentos*. Previo a la caracterización “escéptico-crédulo” otros autores también han hablado de soluciones “cautas” en lugar de “escépticas” (cf. Simari y Loui (1992)), o de soluciones “liberales” en lugar de “crédulas” (cf. Makinson y Schlechta (1991)). Para evitar ambigüedades o vaguedades indeseables, remitámonos a caracterizaciones más precisas en términos más formales que descriptivos o metafóricos.

En general, hay dos modos formales posibles de definir las extensiones, independientemente de cómo se usen para representar soluciones crédulas o escépticas. Usando los términos de Prakken y Vreeswijk (2001), estos modos se pueden clasificar como *asignaciones de estado único* (*unique-status assignment*) y *asignaciones de estado múltiple* (*multiple-status assignment*). La asignaciones de estado único se obtienen definiendo un mecanismo que lleva directamente a una única extensión. Entre los investigadores que utilizan esta técnica en sistemas argumentativos podemos mencionar principalmente a Pollock (1990), quien construye una definición inductiva donde los argumentos pueden estar *in* o *out* en una serie infinita de niveles. Más precisamente, todos los argumentos están *in* en un nivel-0, y un argumento  $\sigma$  está *in* en un nivel- $(n+1)$  si y sólo si no existe ningún argumento que está *in* en el nivel- $n$  que ataque a  $\sigma$ ; luego, todos los argumentos  $\sigma$  para los cuales existe un nivel  $m$  tal que  $\sigma$  está *in* en todos los niveles- $n$ , para  $n > m$ , son los argumentos *justificados* del sistema. Simari y Loui (1992) usan también una variante de esta idea para su sistema. Además, las extensiones fundadas de Dung, como vimos, son otra alternativa dentro de las asignaciones de estado único. Por su parte, las asignaciones de estado múltiple pueden determinar varias extensiones para un mismo sistema. Las extensiones preferidas de Dung son un ejemplo dentro de esta estrategia. Otro ejemplo puede ser el de las *extensiones estables*: son conjuntos

de argumentos libres de conflictos que atacan a todo argumento que no pertenece a ellos. Esta es una idea interesante, aunque puede presentar inconvenientes por el hecho de no estar bien definida para todo marco argumentativo. Por ejemplo, el marco  $AF = \langle \{\sigma\}, \{(\sigma, \sigma)\} \rangle$  no tiene extensión estable, puesto que el único subconjunto libre de conflictos es  $\emptyset$ , el cual no ataca al argumento externo  $\sigma$ , como es claro.

Ahora bien, mientras las asignaciones de estado único permiten representar sólo actitudes escépticas, las asignaciones de estado múltiple permiten representar tanto actitudes crédulas como escépticas. Las actitudes crédulas se logran permitiendo elegir indistintamente una cualquiera entre las múltiples extensiones sancionadas; una actitud escéptica, en cambio, usualmente se logra definiéndola como la intersección de todas las extensiones obtenidas. La teoría de Dung, como vimos, es una excepción a esto, ya que para la actitud crédula establece las extensiones preferidas mientras para la actitud escéptica elige una asignación de estado único como la de la extensión fundada. Con esto logra un comportamiento más escéptico que tomando la intersección de las preferidas, ya que la extensión fundada siempre es un subconjunto de esa intersección, y en algunos es un subconjunto propio. En resumen, si bien hay distintas técnicas para definir las extensiones de un sistema argumentativo, no hay una clasificación formal y materialmente adecuada de éstas que permita discernir cuál es el tipo apropiado para representar cada tipo de actitud epistémica respecto de la justificación de argumentos. Antes bien, todo depende de cuestiones prácticas acerca de cómo se manejan tales definiciones para lograr los comportamientos esperados. Concluyendo, la distinción escéptico-crédulo en este contexto, sigue siendo una cuestión filosófica subyacente a los formalismos.

## Conclusión

Tal como han sido planteados estos sistemas, si se aceptan tanto la representación de la información como los criterios según los cuales se establecen los ataques entre argumentos, entonces se aceptará que, en principio, hay una buena base para pensar que los sistemas de argumentación rebatible permiten modelar adecuadamente algunas decisiones pragmáticas del pensamiento científico que involucran sentido común. Al reflejar, aún cuando sea de un modo grosero, los cambios en la interpretación de la teoría de la luz, la revisión de argumentos producida con los cambios de exten-

siones fundadas según varía la base de información en nuestro ejemplo sugieren una aprehensión de la racionalidad de las decisiones de adoptar una u otra interpretación.

Nuestro propósito ha sido presentar algunas ideas acerca de la dialéctica que envuelve al pensamiento científico y de cómo representarla, y esto creemos haberlo cumplido. Se ha tratado de mostrar que ciertos aspectos pragmáticos, como las interpretaciones de las teorías tendientes a dar explicaciones más simples o más abarcadoras de ciertos fenómenos, podrían ser capturados por un modelo de argumentación rebatible. Las pautas para la formalización del modelo están basadas en la idea de ataque entre argumentos, la noción de aceptabilidad y distintos tipos de extensiones. Por supuesto, hay que tener en cuenta que lo visto es apenas un esbozo, dado que los sistemas hasta el momento se basan en sobresimplificaciones extremas, y que para lograr resultados que realmente conduzcan a una reconstrucción racional del pensamiento científico hay mucho trabajo por hacer.

## Agradecimientos

Este trabajo ha recibido el apoyo de FONCyT - Agencia Nacional de Promociones Científicas y Técnicas (Argentina), mediante el subsidio PICT 2003, N° 13995, y de la Secretaría de Ciencia y Técnica de la Universidad Nacional del Sur, a través del PGI 24/ZI13.

## Notas

<sup>1</sup> Esto puede hacer pensar que nuestro modelo basado en razonamiento no-monótono va a enfocar el problema de la abducción. Es conveniente notar, entonces, que esta implicación confirmatoria es más bien de tipo inductivo, y no pretendemos mostrar que nuestro modelo permite justificar tal inducción, sino el de usar esa implicación *como premisa de un argumento de sentido común* que propone interpretar el fenómeno en los términos de la hipótesis. El lector interesado en la abducción y sus diferencias con la inducción desde un punto de vista del razonamiento no-monótono puede consultar, por ejemplo, Lachiche (2000).

<sup>2</sup> Para una concepción similar de los argumentos rebatibles cf. Simari y Loui (1992), donde además se da un criterio de comparación entre argumentos según la noción de especificidad, y a partir de éste se establece una noción de derrota asimilable a la de ataque que estamos usando aquí.

<sup>3</sup> En el parágrafo 5 se dará una explicación más detallada de esto.

<sup>4</sup> El lector familiarizado con la lógica *default* de Reiter puede interpretar cada *default* bidireccional nuestro  $A \leftrightarrow B$  como dos *default* de Reiter:  $A:B/B$  y  $B:A/A$ . A las variables libres vamos a interpretarlas como universalmente cuantificadas a los efectos de aplicar las reglas de inferencia.

## Referencias

- Dung, Phan Mihn (1995) "On the acceptability of arguments and its fundamental role in non-monotonic reasoning, logic programming and n-person games", *Artificial Intelligence* 77, 321-357.
- Journal of Applied Logic*, número especial de 2004 conteniendo una selección de trabajos del *Third International Workshop on Computational Models of Scientific Reasoning and Applications (III CMSRA)*.
- Lachiche, Nicolas. (2000) "Abduction and induction from a non-monotonic reasoning perspective", en P. A. Flach y A. C. Kakas (eds.) *Abductive and inductive reasoning: essays on their relation and integration*, Kluwer Academic Publishers, 107-116.
- Loui, Ron (1987) "Defeat Among Arguments"; *Computational Intelligence* 2, 100-106.
- Loui, Ron (1991) "Ampliative Inferences, Dialectic and Computation", en R. Cummins y J. Pollock (eds.) *Philosophy and AI. Essays at the interface*, MIT Press, 141-155.
- Makinson, David y Karl Schlechta (1991) "Floating conclusions and zombie paths: two deep difficulties in the 'directly skeptical' approach to inheritance nets", *Artificial Intelligence* 48, 199-209.
- McCarthy, John (1980) "Circumscription: a form of non-monotonic reasoning", *Artificial Intelligence* 13, 27-39.
- McCarthy, John y Patrick Hayes (1969) "Some philosophical problems from de standpoint of AI", *Machine Intelligence* 4, 463-502.
- McDermott, Drew y John Doyle (1980) "Non-Monotonic Logic I", *Artificial Intelligence* 13, 41-72.
- Moore, Richard (1985) "Semantical considerations on nonmonotonic logic", *Artificial Intelligence* 25, 75-94.
- Pollock, John (1990) *Nomic probability and the foundations of induction*, New York: Oxford University Press.
- Poole, David (1985) "On the comparison of theories: preferring the most specific explanation", *Proc. of the Ninth IJCAI*, Los Altos, 144-147.
- Prakken, Hendrik (1993) *Logical Tools for Modeling Legal Arguments*, Tesis Doctoral, Amsterdam: Vrije Universiteit.

- Prakken, Hendrik y Gerhard Vreeswijk (2001) "Logics for defeasible argumentation". En D. Gabbay y F. Guenther (Eds.), *Handbook of Philosophical Logic*, 2da ed., Dordrecht: Kluwer Academic Publishers.
- Reiter, Raymond (1980) "A logic for default reasoning", *Artificial Intelligence* 13, 81-132.
- Schurz, Gerhard (2004) "Normic laws, non-monotonic reasoning, and the unity of science". En S.Rahman, J. Symon, D. Gabbay y J. P. van Bendegem (eds.) *Logic, Epistemology and the Unity of Science*, Dordrecht: Kluwer Academic Publishers, 181-212.
- Scriven, M. (1959) "Truisms as ground for historical explanation". En P. Gardiner (ed.) *Theories of history*, New York: The Free Press.
- Simari, Guillermo y Ron Loui (1992) "A mathematical treatment of defeasible reasoning" *Artificial Intelligence* 53, 125-157.
- Vreeswijk, Gerhard (1993) *Studies in Defeasible Argumentation*; Tesis Doctoral, Amsterdam: Vrije Universiteit.
- 

**Gustavo A. BODANZA**

**ccbodanz@criba.edu.ar**

PhD. is Assistant Professor in the Universidad Nacional del Sur, Argentina, member of the Center of Research on Logic and Philosophy of Science (CILF) and of the Artificial Intelligence Laboratory (LIDIA) of that university. He is also Assistant Researcher of the Argentinian National Council of Science and Technology (CONICET). He is currently undertaking research on logic and defeasible argumentation, artificial intelligence and reasoning under uncertainty.