

Una Formalización del Razonamiento Hipotético e Inductivo en Teoría de la Ciencia

Claudio Delrieux

Grupo de Investigación en Programación Declarativa e Inteligencia Artificial
Universidad Nacional del Sur
Alem 1253 - (8000) - Bahía Blanca - ARGENTINA
e-mail:usdelrie@criba.edu.ar

Palabras Clave: Razonamiento Revisable - Razonamiento Hipotético -
Razonamiento Inductivo - Teoría de la Ciencia

Resumen

Se presenta una formalización del razonamiento hipotético adecuada para integrarse dentro del contexto del razonamiento científico. Dicha formalización está fundamentada en la importancia epistémica que se asigna a la información factual. Esto permite caracterizar los distintos tipos de conocimiento que se utilizan, es decir, el conocimiento lógico-deductivo, conocimiento tentativo proveniente de reglas *default* y el conocimiento proveniente de opiniones o hipótesis. Cada tipo de conocimiento tiene una importancia epistémica diferente.

En [2] se presentó una correspondencia entre los aspectos formales de la teoría de la ciencia y un modelo de razonamiento revisable. El comportamiento de una teoría científica dentro del contexto de explicación es visto como un caso particular de razonamiento no monotónico con importancia epistémica. El modelo de razonamiento hipotético e inductivo presentado aquí extiende dicha correspondencia al contexto de descubrimiento, dentro del cual una teoría es utilizada para generar y testear hipótesis experimentales.

Una Formalización del Razonamiento Hipotético e Inductivo en Teoría de la Ciencia

1 Introducción

El razonamiento hipotético es el razonamiento que se efectúa a partir de premisas (hipótesis) cuyo valor de verdad se desconoce o es falso. Este tipo de razonamiento es muy útil en varios contextos. Por ejemplo, en el razonamiento matemático es muy común la demostración por el absurdo, en la cual se introduce una premisa falsa con el objeto de demostrar que ello lleva a contradicción.

La problemática del razonamiento hipotético comienza a partir de su representación por medio de sentencias condicionales. Si el antecedente del condicional es verdadero, estamos en el caso de un *condicional factual*, el cual es representado dentro del lenguaje con una implicación material y es sometido a la interpretación veritativo-funcional correspondiente. Sin embargo, cuando el valor de verdad del antecedente del condicional se desconoce o es falso, estamos en el caso del *condicional problemático* y del *condicional contrafáctico*, respectivamente. En estos casos se introduce una hipótesis (el antecedente del condicional) que se sabe que no es verdadera, y se busca conocer las consecuencias que ésto acarrea. En ambos casos, la representación del condicional *dentro* del lenguaje es inadecuada porque no es posible efectuar una interpretación veritativo-funcional de la sentencia resultante.

El problema más importante que introduce el razonamiento hipotético en un marco lógico-deductivo proviene de las contradicciones que produce introducir o aceptar el antecedente del condicional. Es decir, la hipótesis produce una *sobredeterminación* en el valor de verdad de una o más sentencias. Aceptar el antecedente del condicional requiere abandonar otras creencias que son contradictorias con el mismo. En algunos casos puede existir un criterio que determine cuáles creencias deben abandonarse y cuáles son retenidas. Este procedimiento de inferencia se denomina *exducción* [9]. En otros casos, sin embargo, se produce una ambigüedad contextual, es decir, el contexto no permite decidir cuáles creencias deben ser abandonadas.

Se propone en este trabajo una formalización del razonamiento hipotético, fundamentado en un análisis de la importancia epistémica de los distintos tipos de conocimiento que se utiliza. Esto da origen a un sistema de razonamiento ampliativo, es decir, que aumenta la capacidad de inferencia de un sistema deductivo. Dicho sistema es integrable al razonamiento no monotónico en general. De esa manera es posible encontrar una correspondencia entre los aspectos formales de la teoría de la ciencia y el sistema presentado. La formalización del razonamiento hipotético propuesta aquí permite extender dicha correspondencia al comportamiento de las teorías dentro del contexto de descubrimiento.

En la sección siguiente se plantean algunos conceptos introductorios de la teoría de la ciencia y de sus problemas característicos. En la sección 3 se consideran los distintos tipos de conocimiento que deben tenerse en cuenta dentro de las teorías científicas. En la siguiente sección se repasan brevemente los conceptos del razonamiento *default* de Reiter [8] y del sistema \mathcal{P} de razonamiento plausible [13]. En la sección 5 se muestra cómo el sistema propuesto permite representar adecuadamente el comportamiento de una teoría de la ciencia dentro del contexto de explicación [2]. Luego se presenta el modelo de razonamiento hipotético e inductivo aquí propuesto. Se muestran ejemplos de razonamiento hipotético, generación de hipótesis y la dinámica de teorías frente a nuevo conocimiento experimental. En la sección 7, por último, se discuten las conclusiones y el trabajo futuro.

2 Teoría de la ciencia

A lo largo su historia, dentro de la comunidad científica se estableció un conjunto de procedimientos que permiten llevar adelante en forma adecuada las tareas científicas que incluyen, entre otras, la explicación y predicción de fenómenos, la generación de teorías, y la verificación y falsación de las mismas. Las primeras teorías de la ciencia se basaron fundamentalmente en la lógica deductiva clásica, por lo que se encontraron dificultades insalvables (especialmente en las ciencias empíricas) originadas en los diversos tipos de conocimiento científico que era necesario representar. Esto llevó a ampliar los sistemas propuestos de modo tal que, sin apartarse de un lenguaje lógico, utilizaran procedimientos de inferencia ampliativos además de la deducción.

La descripción que efectuaremos a continuación tiende a ser normativa y no descriptiva, es decir, refleja una norma subyacente a los casos históricos sin describir ninguno de ellos. Las teorías y leyes científicas tienen la función de establecer conexiones sistemáticas dentro del conjunto de lo que resulta inmediatamente evidente. De ese modo es posible la deducción de determinados hechos a partir de otros. Según sea que la deducción se haya realizado antes o después que los hechos deducidos se hayan comprobado, la deducción se denomina *predicción* o *explicación* [4].

Uno de los objetivos principales de la actividad científica es la abstracción de un conjunto de casos particulares en lo que se denomina una *ley*. Una ley científica es una pieza de conocimiento hipotético, dado que al no ser observable su condición de verdad, no puede verificarse. En Hempel [4], la lógica de la explicación procede según un esquema $L, C \vdash e$, donde L, C , el *explanans* es un conjunto L de leyes generales (que se deducen de los principios básicos —que constituyen el punto de partida de la ciencia en cuestión— y del conocimiento formal) y un conjunto C de condiciones particulares, y e , el *explanandum* es el fenómeno o hecho a explicar. Este esquema llevó a Popper [7] a plantear que la ciencia empírica es *falsacionista*. Es decir, un fenómeno mal explicado donde $L \vdash e$, pero $\neg e$ es del caso, es razón para abandonar una ley.

Un programa de investigación tiende a organizar el conocimiento de su área en un conjunto E de evidencia (los fenómenos a explicar) y en un conjunto de leyes L que producen un *cubrimiento* de E . En la tarea de investigación, además de la búsqueda de evidencia, destacamos dos procedimientos heurísticos. La heurística positiva busca, dado un nuevo fenómeno e a explicar, un conjunto de condiciones C particulares a este caso tales que se cumpla $L, C \vdash e$. La heurística negativa busca, para un fenómeno e mal explicado, un conjunto de condiciones C particulares a este caso tales que $L, C \not\vdash e$, de modo tal de proteger al núcleo L de ser falsado. El poder sistematizador de L es, entonces, la relación entre los casos de E que son cubiertos con dichas heurísticas y el total de los casos. De esto se sigue que el carácter del núcleo L de una teoría es siempre tentativo, y está sujeto a cambios, revisiones y excepciones.

Según su propia terminología, en la actividad científica es corriente trabajar con *postulados hipotéticos*, y con *generalizaciones accidentales* [3]. Los primeros son postulados que se introducen deliberadamente al conocimiento cierto, porque al hacerlo se sigue lógicamente un conjunto de evidencia, o porque se simplifica tal derivación. Es decir, muchas veces las condiciones particulares C aludidas más arriba tienen un carácter hipotético. Encontrar estas condiciones particulares para explicar el éxito o fracaso de un núcleo L que sistematiza un conjunto de evidencia E es la actividad principal de las heurísticas positiva y negativa.

Las generalizaciones accidentales, por su parte, actúan sintácticamente como *leyes* sin llegar a serlo. Una ley es una generalización de un conjunto de observables, la cual una comunidad científica acepta gradualmente, y que una vez establecida, está comprometida a

utilizar y retener aún al gran costo de rechazar evidencia [11]. Por ello es que las leyes constituyen el núcleo de una teoría. La diferencia entre una ley y una generalización accidental es que la primera es verdadera (dentro de un contexto más amplio que apela a principios fundamentales), mientras que la segunda no. Pero, como vimos, la determinación de la falsedad de una ley es *a posteriori*.

3 Tipos de conocimiento

Puede observarse que un punto importante quedó insatisfactoriamente atendido en el anterior análisis. Nos referimos al manejo de la importancia epistémica de los diversos elementos que constituyen el conocimiento científico. Es decir, no se ha propuesto una forma adecuada para representar la mayor o menor importancia que tienen la evidencia, los principios, las leyes, los postulados hipotéticos y otros tipos de sentencias. En este trabajo se propone introducir diferentes tipos de conocimiento no deductivo en el marco Hempeliano. Esto obedece a la necesidad de utilizar dichos tipos de conocimiento dentro de algún marco de razonamiento, de modo de poder aprovechar el valor heurístico que dicho conocimiento encierra.

El acercamiento formal en la lógica consistió tradicionalmente en incorporar un operador modal \Box a un lenguaje proposicional o de primer orden, de modo que $\Box p$ es una fórmula modal. Si p representa una sentencia arbitraria, estas fórmulas modales se encuadran dentro de lo que la tradición escolástica denominaba “modalidad *de dicto*”, en tanto que el operador *modifica la proposición*, es decir, al enunciado en sí. Si expresamos la predicación de la propiedad p sobre el individuo X como $X \in p$, entonces afirmar (*de dicto*) que necesariamente el individuo X tiene la propiedad p se efectúa con la sentencia modal $\Box(X \in p)$. Otra forma de entender la modalidad *de dicto* es decir que se aplica a *estados de cosas*, es decir, un agente tiene una creencia *de dicto* cuando *afirma* que un determinado estado de cosas es del caso, y tiene una creencia *de dicto* verdadera si dicho estado de cosas existe [1].

Puede ocurrir también que se califique *la predicación* de una propiedad sobre un individuo. Estas expresiones modales se denominan *de re* (de la cosa en sí). En este caso afirmar (*de re*) que el individuo X necesariamente tiene la propiedad p se efectúa con la sentencia (no modal) $X \Box \in p$. Un agente tiene una creencia *de re* cuando atribuye a un individuo una determinada propiedad, y tiene una creencia *de re* verdadera si atribuye una propiedad a un individuo que tiene dicha propiedad [1].

Las diferencias entre modalidades *de dicto* y *de re* son inexpresables en los lenguajes de primer orden. Una sentencia modalizada *de dicto* es esencialmente una sentencia modal, es decir, está fuera de la lógica clásica, mientras que la sentencia modalizada *de re* puede pensarse como una sentencia *dentro* del lenguaje. La modalidad *de dicto*, al referirse a estados de cosas, atribuye propiedades a individuos concretos (sentencias sin variables libres) y como su nombre sugiere, se puede aplicar sensatamente a conocimiento extensional o particular, es decir, a elementos del lenguaje que reflejen *casos particulares*. Dicho de otra forma, carece de sentido referirse *de dicto* a individuos o estados de cosas que no existen. La modalidad *de re*, en cambio, modifica la inherencia de un concepto en otro. Es decir, se refiere a un conocimiento intensional o general, sea o no del caso que existan individuos con las propiedades referidas.

En el aspecto pragmático, un agente que razone con conocimiento no deductivo debe dar un tratamiento diferente a las premisas *de dicto* y a las premisas *de re*. El primer tipo de premisas se agrega a su conocimiento lógico acerca de un estado de cosas actual. De ese modo es posible extender el contexto del conocimiento cierto que tiene el agente con un conjunto de premisas tentativas que expresan verdades *prima facie* acerca de individuos

que existen en dicho contexto. Como el conocimiento tentativo *de dicto* no está dentro del lenguaje, para incorporarlo al mismo es necesario que tome la forma de una sentencia afirmativa. Un agente cuando utiliza este tipo de creencias expresará “*Afirmo* que tal estado de cosas es del caso”.

El segundo tipo de premisas representa las tendencias que el agente ha elaborado (por ejemplo, como abstracción de un gran conjunto de estados de cosas) con respecto a las propiedades de determinadas clases de individuos. De ese modo es posible extender el conocimiento general que tiene el agente con un conjunto de premisas tentativas que expresan implicaciones *prima facie*, es decir, para individuos que poseen ciertas propiedades es posible inferir tentativamente otras propiedades. Podemos ejemplificar estas diferencias del siguiente modo:

	<i>de dicto</i>	<i>de re</i>
Particular	(afirmo que) Opus es un pingüino	(normalmente) Opus es un pingüino
General	(afirmo que) Las aves vuelan	(normalmente) Las aves vuelan

Es notable observar la coincidencia que existe entre las premisas *de dicto* y los postulados hipotéticos y entre las premisas *de re* y las generalizaciones accidentales. En este trabajo nos referiremos además a la relación existente entre las premisas *de re* y las *reglas default* [8], y a las premisas *de dicto* con un nuevo tipo de premisas denominadas *premisas plausibles*. De esa manera es posible formalizar la teoría Hempeliana de la ciencia dentro del marco del razonamiento no monotónico aumentado con el razonamiento a partir de premisas plausibles.

La clasificación mostrada más arriba permite clarificar la relación mencionada. Cuando se razona con una regla *default*, se está incorporando una observación acerca de las características que un conjunto de individuos normalmente posee. Por ello es que las reglas *default* son una forma tentativa *de re*. Cuando se acepta conocimiento plausible, se está incorporando como información afirmaciones que son supuestas o informadas por otro agente. Por ello es que las premisas plausibles son una forma tentativa *de dicto*. También es importante observar que naturalmente tiene sentido aceptar reportes de conocimiento particular y que las reglas tentativas tienen sentido para conocimiento general. En el primer caso, no es adecuado aceptar de otro agente la afirmación de que *Las aves vuelan*, puesto que no parece posible que dicho agente haya podido observar a todo el conjunto de las aves. Es decir, si se acepta una afirmación general como conocimiento tentativo, el tratamiento a dar a la misma es el de conocimiento tentativo *de re*. En el segundo caso, aceptar que *Normalmente Opus es un pingüino* no parece tener sentido alguno. Es decir, si se acepta una afirmación particular como conocimiento tentativo, el tratamiento a dar a la misma es el de conocimiento tentativo *de dicto*.

Un análisis epistemológico caracteriza a los tipos conocimiento según su importancia. El más importante es el conocimiento lógico-deductivo general \mathcal{K} , que expresa el conjunto de verdades de razón. Luego sigue el conocimiento particular que representa lo que es inmediatamente evidente o es informado por otros agentes. Por último tenemos el conocimiento tentativo Δ expresado en reglas *default*, que representa un conjunto de presuposiciones tales que su contenido intensional no es seguro, *i.e.*, pueden existir excepciones. Para presentar la formalización de la teoría Hempeliana de la ciencia dentro del marco del razonamiento no monotónico, resumiremos brevemente en la sección siguiente una introducción a la lógica *Default* de Reiter y al sistema \mathcal{P} de razonamiento plausible.

4 Razonamiento revisable

El objetivo de esta sección es proveer el contexto formal necesario para el resto del trabajo. Podemos pensar que se desea construir un modelo del razonamiento de un agente idealmente racional AR , el cuál tiene conocimiento lógico (*i.e.*, deductivamente válido) \mathcal{K} , y además tiene conocimiento *revisable*, es decir, conocimiento que expresa opiniones o tendencias generales, pero que pueden tener excepciones. Estas excepciones no son necesariamente conocidas, por lo que se espera que el sistema retracte o revise las conclusiones obtenidas con dicho conocimiento cada vez que aparece una excepción. Esto significa que se espera un comportamiento no monotónico en el sistema.

El modelo consta de dos partes: el razonamiento con premisas *de re*, que cumplen la función de generalizaciones accidentales, y el razonamiento con premisas *de dicto* que cumplen la función de los postulados hipotéticos. El primero de los aspectos es adecuadamente formalizado con la lógica *default* de Reiter [8]. La misma representa el conocimiento tentativo por medio de esquemas de *reglas default*. Dichos esquemas de regla pueden tener excepciones sin que las mismas estén explícitamente representadas. Esto significa que sus conclusiones pueden revisarse en caso de que exista nueva evidencia. De ese modo tendremos que la regla $\frac{a(X) : b(X)}{c(X)}$ es un esquema que representa el proceso de derivar una conclusión (o consecuente) $c(X)$ dados el conocimiento cierto (o prerequisite) de que $a(X)$ y la consistencia del conocimiento (o justificación) de que $b(X)$.

DEFINICIÓN 4.1 [8] Una Teoría Default T es un par $T = \langle \mathcal{K}, \Delta \rangle$ donde \mathcal{K} es un conjunto de sentencias cerradas en un lenguaje de primer orden \mathcal{L} y Δ es un conjunto de reglas *default*.

Intuitivamente, \mathcal{K} es el conocimiento deductivo de la teoría, brindando un contexto a la misma. Δ intenta producir un conjunto de extensiones consistentes a partir de \mathcal{K} . La noción de extensión intenta capturar la idea del conjunto (consistente) de creencias que se puede extraer a partir de una teoría dada. Las extensiones pueden ser múltiples o pueden no existir. Cada una de dichas extensiones es una posible visión del mundo compatible con el conjunto original de creencias que AR originalmente poseía.

DEFINICIÓN 4.2 [8] Dada una teoría $T = \langle \mathcal{K}, \Delta \rangle$, para cualquier conjunto \mathcal{A} de sentencias sea $\Gamma(\mathcal{A})$ el menor conjunto que satisfice

1. $\mathcal{K} \in \Gamma(\mathcal{A})$,
2. $Th(\Gamma(\mathcal{A})) = \Gamma(\mathcal{A})$,
3. Si existe una regla $\frac{a:b}{c}$ en Δ tal que $a \in \Gamma(\mathcal{A})$, $\neg b \notin \Gamma(\mathcal{A})$, entonces $c \in \Gamma(\mathcal{A})$.

Entonces un conjunto E de sentencias $E \subseteq \mathcal{L}$ es una extensión para Δ si y sólo si $\Gamma(E) = E$, es decir, E es un punto fijo del operador Γ .

Un resultado importante es que una teoría T tiene una extensión E inconsistente si y sólo si \mathcal{K} es inconsistente, y E es la única extensión de la teoría. Sin embargo, cuando existen dos o más reglas cuyos consecuentes son mutuamente inconsistentes dado \mathcal{K} , entonces la lógica no es lo suficientemente fuerte como para elegir una extensión entre las posibles.

EJEMPLO 4.1 Sea la teoría $T = \langle \{ave(opus), herido(opus)\}, \left\{ \frac{ave(X):vuela(X)}{vuela(X)}, \frac{herido(X):\neg vuela(X)}{\neg vuela(X)} \right\} \rangle$. Tenemos dos extensiones: $E^1 = \{ave(opus), herido(opus), vuela(opus)\}$ y $E^2 = \{ave(opus), herido(opus), \neg vuela(opus)\}$.

Consideraremos ahora la segunda parte de la formalización presentada, la cual incorpora el razonamiento con premisas *de dicto* que cumplen la función de los postulados hipotéticos. En principio, los mismos pueden incorporarse como premisas en tanto no se genere una contradicción. En cambio, si ocurre una contradicción es necesario abandonar alguna de las premisas *de dicto* incorporadas, es decir, no permitir que el mecanismo de inferencia la considere. Una estrategia para restaurar la consistencia consiste en abandonar la mínima cantidad de premisas, y de las mismas, aquellas que por alguna razón signifiquen la menor pérdida de información valiosa. Cada premisa, entonces, estará acompañada de un índice que determine su importancia desde un punto de vista epistemológico. *Plausibilidad* es un término adecuado para referirse a la importancia de una premisa fáctica. La plausibilidad de una premisa es una medida de su grado de soporte asertivo, *i.e.*, de su valoración epistemológica [1].

Describiremos el sistema \mathcal{P} de razonamiento plausible de un agente razonador ideal AR. El mismo está inspirado en un tratamiento propuesto por Rescher [10]. El conocimiento \mathcal{K} que AR tiene, puede pensarse como un conjunto consistente expresado en un lenguaje de primer orden \mathcal{L} . Pero además AR tiene conocimiento adquirido a través de un conjunto de informantes $\mathcal{I} = \{I_1, I_2, \dots, I_n\}$, donde cada I_k es una fuente de información. Cada I_k , entonces, está capacitado para informar a AR de un conjunto de literales de base. La unión de dichos conjuntos no es necesariamente consistente, por lo que AR utilizará la plausibilidad para restaurar la consistencia cuando los conjuntos de información den origen a contradicción.

El agente AR tiene establecido en \mathcal{I} un orden parcial al que nos referiremos como *relación de plausibilidad* \mathcal{P} . Esto significa que, dados dos informantes I_p y I_q , AR puede dar preferencia a la información de I_p , o bien a la de I_q , o bien no dar preferencia a ninguno. Utilizamos la notación $I_i \prec_{\mathcal{P}} I_j$ cuando $I_i \preceq_{\mathcal{P}} I_j$ y $I_j \not\preceq_{\mathcal{P}} I_i$. Se completa el conjunto \mathcal{I} con AR mismo, considerado como un "informante" I_{\top} cuya plausibilidad es máxima, y con todo el grupo de informantes "desconocidos" considerándolos como un único informante I_{\perp} al cual se le asigna una plausibilidad menor que la de cualquier otro informante conocido. Podemos decir entonces que \mathcal{I} es un conjunto reticulado bajo la relación $\prec_{\mathcal{P}}$.

DEFINICIÓN 4.3 Una estructura de informantes es un par $(\mathcal{I}, \prec_{\mathcal{P}})$, donde \mathcal{I} es un conjunto de informantes $\{I_1, I_2, \dots, I_n\}$ y $\prec_{\mathcal{P}}$ es un orden parcial sobre \mathcal{I} , llamado relación de plausibilidad. \mathcal{I} contiene un elemento I_{\top} tal que $\forall I_i \in \mathcal{I}. I_i \prec_{\mathcal{P}} I_{\top}$, y un elemento I_{\perp} tal que $\forall I_i \in \mathcal{I}. I_{\perp} \prec_{\mathcal{P}} I_i$. Cada informante provee un conjunto de ítem de información $\langle l, I_i \rangle$. Un conjunto \mathfrak{S}_k de ítem tiene como cota de plausibilidad a su menor cota superior bajo \mathcal{P} , es decir, al conjunto $\text{lub}\{\langle l_i, I_i \rangle : \langle l_i, I_i \rangle \in \mathfrak{S}_k\}$. La cota de un conjunto \mathfrak{S}_k de ítem es mayor a la cota de otro conjunto \mathfrak{S}_j si y solo si cada elemento de la cota de \mathfrak{S}_j es a lo sumo tan plausible como cada elemento de la cota de \mathfrak{S}_k , pero existe por lo menos un elemento de la cota de \mathfrak{S}_k estrictamente más plausible que todos los elementos de la cota de \mathfrak{S}_j . La unión de todos los ítem informados es el conjunto de información \mathfrak{S} .

Dado un conjunto de información \mathfrak{S} , ¿cuál es el subconjunto $\mathfrak{S}_{\mathcal{P}}$ del mismo, consistente con \mathcal{K} , que AR aceptará? La solución propuesta por Rescher es considerar los subconjuntos máximamente consistentes (dado \mathcal{K}) de \mathfrak{S} . Una variante escéptica propuesta por Roos [12] consiste en considerar la intersección de todos los conjuntos generados bajo distintas extensiones lineales de \mathcal{P} . Esta solución es semánticamente adecuada pero computacionalmente inaceptable, por lo que proponemos una solución semánticamente equivalente a la de Roos, pero computacionalmente tratable. La misma consiste en verificar si una determinada sentencia pertenece al conjunto $\text{Th}(\{l : \langle l, I_i \rangle \in \mathfrak{S}_{\mathcal{P}}\} \cup \mathcal{K})$ de conclusiones plausibles.

DEFINICIÓN 4.4 Sea un conjunto de conocimiento lógico \mathcal{K} , un conjunto de información \mathfrak{S} , y una sentencia p indecidible con respecto a \mathcal{K} (es decir, $\mathcal{K} \not\vdash p, \mathcal{K} \not\vdash \neg p$).

- (Fundamento) p tiene fundamento si y solo si existe un conjunto $\mathfrak{S}_i \subseteq \mathfrak{S}$ tal que $\{l : \langle l, I \rangle \in \mathfrak{S}_i\} \cup \mathcal{K} \vdash p$.
- (Duda) p está en duda si y solo si su negación tiene fundamento, es decir, existe un conjunto $\mathfrak{S}_j \subseteq \mathfrak{S}$ tal que $\{l : \langle l, I \rangle \in \mathfrak{S}_j\} \cup \mathcal{K} \vdash \neg p$.
- (Aceptación) p es aceptada si y solo si tiene fundamento y no está en duda, o si tiene fundamento en \mathfrak{S}_i , está en duda a causa de \mathfrak{S}_j , pero la cota de plausibilidad de \mathfrak{S}_i es mayor que la cota de \mathfrak{S}_j .

Esta definición es computacionalmente tratable, ya que incurre en una doble recursión, y las invocaciones a la demostrabilidad, al realizarse por encadenamiento hacia atrás, se computan con técnicas estándar de programación en lógica.

EJEMPLO 4.2 Sea el caso en que $\mathfrak{S} = \{\langle p, I_i \rangle, \langle \neg p, I_j \rangle, \langle \neg q, I_j \rangle\}$, con $I_j \prec_p I_i$, y en \mathcal{K} encontramos $\{p \Rightarrow q, q \Rightarrow r\}$. En este caso, p es aceptado porque su fundamento es más plausible que el fundamento de $\neg p$. Al mismo tiempo, la aceptación de p permite derivar q , cuyo fundamento es más plausible que el de $\neg q$ y por lo tanto es aceptado. Por último, la aceptación de q permite derivar r , el cual es aceptado por tener fundamento y no estar en duda.

Las razones para utilizar un razonamiento escéptico se pueden comprender a partir del siguiente ejemplo.

EJEMPLO 4.3 (Paradoja de la lotería cualitativa [5, 6]). Sea el caso en que $\mathfrak{S} = \{\langle \neg p, I \rangle, \langle \neg q, I \rangle, \langle \neg r, I \rangle\}$, y en \mathcal{K} encontramos $\{(p \vee q \vee r)\}$. En este caso, $\neg p$ no es aceptado porque su fundamento es tan plausible como el fundamento de p , el cual es obtenido a partir del razonamiento $\{\neg q, \neg r, (p \vee q \vee r)\} \vdash p$. Lo mismo sucede con $\neg q$ y con $\neg r$.

5 Una teoría computacional de la ciencia

Normalmente una teoría es utilizada dentro de dos contextos claramente diferenciados: el *contexto de explicación* y el *contexto de descubrimiento* [7]. En esta sección presentaremos algunos ejemplos ilustrativos del sistema dentro del contexto de explicación, dejando para la próxima sección los ejemplos correspondientes al uso del sistema dentro del contexto del descubrimiento. Comenzamos con un caso donde interviene exclusivamente el conocimiento deductivo. Como vimos, existen los principios generales (expresados en la sentencia P a continuación), las leyes L que se siguen lógicamente de los principios, pero que pragmáticamente tienen un uso más adecuado, y la evidencia, es decir, los fenómenos directamente constatables.

EJEMPLO 5.1 Nuestro conocimiento acerca de gravitación se reduce a:

- P : Existe una fuerza que atrae a los objetos (masivos) entre sí.
 $\forall X, Y. o(X) \wedge o(Y) \Rightarrow a(X, Y)$
- L : Los objetos son atraídos hacia la tierra.
 $\forall X. o(X) \Rightarrow a(\text{tierra}, X)$
- e_1 : Esta piedra es atraída hacia la tierra.
 $a(\text{tierra}, p)$
- e_2 : Este globo de hidrógeno no es atraído hacia la tierra.
 $\neg a(\text{tierra}, g)$

La primer sentencia es un principio, la segunda es una ley, y las dos últimas son evidencia. La ley se deduce como caso particular del principio, y es por lo tanto utilizada por su mayor valor pragmático. Sin embargo, si bien la ley permite explicar (o predecir) la tercer sentencia, fracasa con la cuarta. Es decir, tenemos $P \vdash L$, $L \vdash e_1$, pero $L \not\vdash e_2$. Esto, sin embargo, no lleva a abandonar la ley, sino a buscar razones por las cuales fracasa en la explicación de este caso. Es decir, la teoría se modifica de modo tal que la justificación tome la forma $L, C \vdash E$, donde E cubre tanto a e_1 como a e_2 , y C expresa las condiciones particulares a cada objeto (en este caso, su propiedad de ser más pesado que el aire).

La propuesta para sistematizar un conjunto de evidencia es modificar el *explanans* Hempeliano para incluir un conjunto de conocimiento plausible \mathcal{P} y un conjunto Δ de reglas *default*. De ese modo tenemos que $\mathcal{K}, \Delta, \mathfrak{S}_{\mathcal{P}}$ justifica e . En este caso, $\mathcal{K} \cup \Delta$ reemplaza al núcleo L Hempeliano, y $\mathfrak{S}_{\mathcal{P}}$ al conjunto C de condiciones particulares.

EJEMPLO 5.2 (Ejemplo 5.1 revisitado).

P : Existe una fuerza que atrae a los objetos masivos entre sí.

$$\forall X, Y. o(X) \wedge o(Y) \Rightarrow a(X, Y)$$

L : Los objetos tienden a caer hacia la tierra.

$$o(X) : a(\text{tierra}, X)$$

$$\frac{}{a(\text{tierra}, X)}$$

e_1 : Esta piedra cae hacia la tierra.

$$a(\text{tierra}, p)$$

e_2 : Este globo de hidrógeno no cae hacia la tierra.

$$\neg a(\text{tierra}, g)$$

En este ejemplo, L , por estar sujeta a excepciones de origen desconocido, se utiliza como regla tentativa. De ese modo e_1 está justificada dentro de la teoría, mientras que e_2 constituye una excepción. Más adelante, una amiga que ha leído a Arquímedes y a Torricelli nos sugiere la presencia de una atmósfera como causa de algunas de las excepciones en nuestra teoría:

P : Existe una fuerza que atrae a los objetos masivos entre sí.

$$\forall X, Y. o(X) \wedge o(Y) \Rightarrow a(X, Y)$$

L_1 : Los objetos más pesados que el aire tienden a caer hacia la tierra.

$$o(X) \wedge p(X) : a(\text{tierra}, X)$$

$$\frac{}{a(\text{tierra}, X)}$$

L_2 : Los objetos menos pesados que el aire no tienden a caer hacia la tierra.

$$o(X) \wedge \neg p(X) : \neg a(\text{tierra}, X)$$

$$\frac{}{\neg a(\text{tierra}, X)}$$

c_1 : Esta piedra es más pesada que el aire.

$$p(p)$$

c_2 : Este globo de hidrógeno es menos pesado que el aire.

$$\neg p(g)$$

e_1 : Esta piedra cae hacia la tierra.

$$a(\text{tierra}, p)$$

e_2 : Este globo de hidrógeno no cae hacia la tierra.

$$\neg a(\text{tierra}, g)$$

En este caso, la teoría sistematiza mejor la evidencia, aunque aún está sujeta a excepciones (aviones en vuelo, por ejemplo).

6 Contexto de descubrimiento

Hasta el momento, hemos encontrado una correspondencia entre una idealización del razonamiento científico y un sistema de razonamiento automatizado. Dentro del contexto de explicación la utilidad de esta correspondencia es la de una representación formal. El paso importante consiste en automatizar la búsqueda y la revisión de justificaciones alternativas dentro del contexto de descubrimiento. Dicho contexto está caracterizado por la generación de hipótesis condicionales, las cuales –si están confirmadas por la experiencia– prestan apoyo a la teoría. Es decir, se generan conjeturas con bases firmes, seguido de pruebas rigurosas. El proceso de conjetura normalmente es asistemático, interviniendo la intuición e imaginación dentro de los límites de la disciplina. Habitualmente se espera que las hipótesis de trabajo en el contexto de descubrimiento respondan a criterios como contrastabilidad, repetibilidad, simplicidad y coherencia. Una hipótesis que genere experimentos que se confirman en algunos casos y fracasan en otros no es en principio desechada.

Estas características hacen del razonamiento científico un campo de muy difícil sistematización y formalización, dado que el patrón de razonamiento va de lo particular a lo general (lo opuesto a la deducción). La forma de expresar las teorías científicas expresada en este trabajo, junto con la importancia epistémica asignada a cada parte del conocimiento de las mismas, permite avanzar hacia el diseño de experimentos hipotéticos y su subsecuente análisis. Analizaremos primero la forma de tratar el razonamiento hipotético, para luego establecer una manera de generar hipótesis por inducción y contrastarlas por medio de dicho razonamiento.

La manera de implementar el razonamiento hipotético en este marco consiste en tratar a la premisa hipotética (problemática o contrafáctica) como un conocimiento particular de máxima plausibilidad. Por lo tanto el conflicto que se produce al introducirla se establece a partir de otras piezas de conocimiento particular. Podemos citar tres casos de conflicto. En el primero, la premisa hipotética es contradictoria con la consecuencia de un razonamiento *default* basado en conocimiento particular (el cual puede ser plausible).

EJEMPLO 6.1 Conocemos $\text{pingüino}(\text{opus})$ y $\frac{\text{pingüino}(X): \neg \text{vuela}(X)}{\neg \text{vuela}(X)}$, y nos planteamos la hipótesis “¿Qué sucedería si *opus* volara?”. En este caso la conclusión que se prefiere es que “Si *opus* volara entonces es una excepción a la regla de que normalmente los pingüinos no vuelan”. Esto es así porque si hay una razón plausible para aceptar $\text{pingüino}(\text{opus})$, entonces a dicha razón se le otorga mayor importancia epistémica que a la regla *default*.

Un segundo caso ocurre cuando la premisa hipotética es contradictoria con un conjunto de conocimiento particular. En este caso ocurre una ambigüedad contextual, es decir, existen distintas posibilidades de restaurar la consistencia abandonando distintas partes del conocimiento particular.

EJEMPLO 6.2 Conocemos $\text{pingüino}(\text{opus})$ y $\text{canario}(\text{tweety})$, y nos planteamos la hipótesis “¿Qué sucedería si *opus* y *tweety* fuesen de la misma especie?”. En este caso existe un conflicto entre sentencias particulares. La conclusión que se obtiene depende de la existencia de una preferencia epistémica (i.e., plausibilidad) entre $\text{pingüino}(\text{opus})$ y $\text{canario}(\text{tweety})$. Es decir, si tenemos razones para confiar más en el reporte $\text{pingüino}(\text{opus})$ que en $\text{canario}(\text{tweety})$, preferiremos “Si *opus* y *tweety* fuesen de la misma especie, entonces *tweety* sería un pingüino” y viceversa.

Un tercer caso ocurre cuando la premisa hipotética es contradictoria con la consecuencia de un razonamiento deductivo basado en conocimiento particular. Este es otro tipo de ambigüedad contextual, proveniente del conflicto entre hechos y reglas.

EJEMPLO 6.3 Conocemos $\text{pingüino}(\text{opus})$ y $\forall X. \text{pingüino}(X) \Rightarrow \text{ave}(X)$, y nos planteamos la hipótesis “¿Qué sucedería si opus no fuese un ave?”. En este pesa el hecho de que se asigna mayor importancia epistémica al conocimiento general que al conocimiento particular. Esto es similar al criterio nómico de la teoría de la ciencia, según el cual las leyes tienen prevalencia sobre los postulados particulares. La conclusión que se obtiene, entonces, es que “Si opus no fuese un ave entonces no sería un pingüino”.

Estos ejemplos abarcan las posibilidades de razonamiento hipotético para todos los tipos de conocimiento presentados. Veremos entonces de qué manera la generación de una hipótesis por inducción puede ser manejada dentro del marco de razonamiento hipotético presentado. Un sencillo esquema de inducción puede ser $\frac{a:b, b}{a}$, es decir, la inferencia como búsqueda de una explicación.

EJEMPLO 6.4 Conocemos $\frac{\text{ave}(X): \text{vuela}(X)}{\text{vuela}(X)}$ y $\text{vuela}(\text{opus})$. Entonces podemos generar la hipótesis inductiva $\text{ave}(\text{opus})$, la cual confirma la evidencia $\text{vuela}(\text{opus})$. Pero si luego se agrega a nuestro conocimiento que $\text{pingüino}(\text{opus})$, $\text{pingüino}(X) \Rightarrow \text{ave}(X)$, $\frac{\text{pingüino}(X): \neg \text{vuela}(X)}{\neg \text{vuela}(X)}$, $\frac{\text{en_avion}(X): \text{vuela}(X)}{\text{vuela}(X)}$, entonces se genera otra hipótesis inductiva para explicar $\text{vuela}(\text{opus})$, a saber, $\text{en_avion}(\text{opus})$. En este punto tenemos dos razonamientos explicativos: “ opus vuela porque es un ave” y “ opus vuela porque está en un avión”. De ambas hipótesis, la primera está falsada porque $\text{pingüino}(\text{opus})$ determina que opus sea ave sin que necesariamente tenga la capacidad de volar.

7 Conclusiones y trabajo futuro

Se ha presentado una correspondencia entre los aspectos formales de la teoría de la ciencia y un sistema de razonamiento revisable. El mismo incorpora conocimiento lógico, reglas *default* y conocimiento particular tentativo. El sistema permite representar los distintos tipos de conocimiento existentes en una teoría científica, y permite implementar facilidades como la generación y el razonamiento a partir de hipótesis y su verificación o falsificación.

Una característica importante que escapa al sistema propuesto en este trabajo es la posibilidad de incorporar un *ranking* entre las reglas *default*. En nuestra analogía, esto corresponde a la elección entre distintas explicaciones en función de la estructura interna de cada una. Por ejemplo, “El postulado P es preferible a la ley L fuera del contexto C ”. Por lo tanto frente a evidencia que no incluye a C , un razonamiento que utiliza a L contradictorio con un razonamiento que utiliza a P será derrotado por este último. Este razonamiento es no monotónico, porque si nueva evidencia produce como condición que el experimento está dentro del contexto C , entonces la conclusión anterior debe ser abandonada. Otra propuesta para investigar consiste en utilizar el sistema MTD de Simari y Loui [14] por contar con un comparador de teorías basado en especificidad. Este comparador permitiría implementar principios internos de comparación entre teorías científicas.

Estas posibilidades plantean una serie de cuestiones epistemológicas que es necesario precisar. La más importante de ellas es qué tipo de orden se establece entre dicho *ranking* entre las reglas *default*, la importancia epistémica y la especificidad. Una respuesta conservadora (dentro de nuestra presentación) es permitir que las reglas derrotadas siempre sean epistémicamente menos importantes que el conocimiento plausible. Una respuesta generalizada establecería un único orden parcial para todo el conocimiento, cuyo supremo es el conocimiento válido, de modo que toda pieza de conocimiento -regla derrotada o literal plausible- esté referida a un nodo dentro de dicho ordenamiento.

Agradecimiento: Algunas de las ideas aquí presentadas fueron discutidas con Fernando Tohmé, Guillermo Simari, Juan Manuel Torres y Jorge Roetti. Deseo además agradecer a un revisor anónimo de las 24 JAIIO.

Referencias

- [1] Roderick M. Chisholm. *Theory of Knowledge*. Prentice Hall, Englewood Cliffs, New Jersey, 1977.
- [2] Claudio Delrieux. Hacia una Teoría Computacional de la Ciencia. En *24 JAIIO - Jornadas Argentinas de Informática e Investigación Operativa*, Buenos Aires, 1995.
- [3] Carl G. Hempel. *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. The Free Press, New York, 1965.
- [4] Carl G. Hempel y Paul Oppenheim. The Logic of Explanation. *Philosophy of Science*, 15:135–175, 1948.
- [5] Donald Perlis. On The Consistency of Commonsense Reasoning. *Computational Intelligence*, 2(1):180–190, 1986.
- [6] David Poole. The Effect of Knowledge on Belief: Conditioning, Specificity and the Lottery Paradox in Default Reasoning. *Artificial Intelligence*, 49(1,2,3):281–307, 1991.
- [7] Karl Popper. *The Logic of Scientific Discovery*. Hutchinson, London, 1959.
- [8] Raymond Reiter. A Logic for Default Reasoning. *Artificial Intelligence*, 13(1,2):81–132, 1980.
- [9] Nicholas Rescher. *Hypothetical Reasoning*. North Holland, Amsterdam, 1974.
- [10] Nicholas Rescher. *Plausible Reasoning*. Van Gorcum, Dordrecht, 1976.
- [11] Nicholas Rescher. *Scientific Explanation*. McGraw-Hill, New York, 1969.
- [12] Nico Roos. A Logic for Reasoning with Inconsistent Knowledge. *Artificial Intelligence*, 57(1):69–104, 1992.
- [13] Guillermo R. Simari y Claudio A. Delrieux. Plausibilidad y Razonamiento no Monotónico. En *XIX CLEI, Conferencia Latinoamericana de Informática, Tomo II*, páginas 81–96, 22 JAIIO, Jornadas Argentinas de Informática e Investigación Operativa, SADIO, Buenos Aires, 1993.
- [14] Guillermo R. Simari y Ronald P. Loui. A Mathematical Treatment of Defeasible Reasoning and its Implementation. *Artificial Intelligence*, 53(2-3):125–158, 1992.