Modelos de razonamiento abductivo[1]

ÁNGEL NEPOMUCENO FERNÁNDEZ
Universidad de Sevilla

I. INTRODUCCIÓN

El concepto de *abducción* se debe al filósofo norteamericano C. S. Peirce, quien considera que, además de las dos formas de inferencia tradicionalmente aceptadas, la *deducción* y la *inducción*, es bastante común, en la vida ordinaria y en la investigación científica, una tercera forma de inferencia a la que identifica con aquel término, también denominada *retroducción* o *generación de hipótesis*. Como se indica en Aliseda (1998), para este autor la deducción es de carácter analítico; en efecto, la información de las premisas contiene la información de la conclusión. Sin embargo, la inducción es ampliativa, la información de las premisas se amplía para pasar a la conclusión, aunque no se pueda considerar que es meramente sintética. Finalmente, la abducción es una forma de inferencia distinta de la deducción y de la inducción.

De los tres tipos de inferencia, la deducción es la de la completa certeza, pues la conclusión es consecuencia lógica (en sentido clásico) de las premisas, se trata, por así decir, de una proposición necesaria. En la inducción, en la que dada la verdad de las premisas "probablemente" sea verdadera la conclusión, no contamos con esta propiedad, mientras que en la formulación de hipótesis, primera forma de referirse Peirce a la abducción, es la de menor certeza, ya que se viene a sugerir que "podría ser el caso". Considerados estos tres modos de inferencia como otras tantas etapas en la indagación lógica, para Peirce la hipótesis, denominada ya abducción, viene a ser la primera de ellas; ahora se trata de construir hipótesis explicativas: observado un hecho expresado en la proposición C, si A fuera una proposición verdadera, entonces C sería algo corriente; por ello, tenemos razones

1 Este trabajo se ha realizado en el marco de las actividades del Grupo HUM609 del Plan Andaluz de Investigación y se ha beneficiado del Proyecto HUM2004-01255 del Ministerio de Educación y Ciencia.

para suponer que A es verdadera. De esta manera aparece un doble carácter de la abducción: el intuitivo y el racional, lo cual hace más difícil su interpretación por parte de los exegetas del pensamiento peirceano, hasta el punto de haber sido tomada en muchas ocasiones como una forma de inducción.

En este trabajo estudiamos la abducción como la forma de inferencia científica por excelencia, centrándonos en algunas propuestas significativas desde un punto de vista lógico. Presentamos, en primer lugar, un breve estudio de las características de la abducción siguiendo diversos planteamientos (Aliseda, 1997), (Thagard, 1988) y (Hintikka, 1999) y, tras ocuparnos de la noción de sistema lógico, abordamos las propuestas más relevantes de modelos lógicos de la inferencia abductiva, en concreto los basados en el propio Peirce, en tablas semánticas (Aliseda, 1997), algunas variantes de tablas semánticas (Aliseda, Nepomuceno, 2002) y otras (Soler, 2004), (Nepomuceno, 2004) y, por último, la basada en la concepción informacional de los sistemas lógicos (Corcoran, 1998).

II. Características de la abducción

Los planteamientos de la abducción como forma de inferencia llevan a la formulación de cuestiones parejas a las surgidas al tratar los conceptos de razonamiento, argumentación y otros relacionados. El doble aspecto puesto de manifiesto en la división peirceana no es el único. En efecto, se presentan varios aspectos a tener en cuenta para determinar qué sea la abducción, de entre los cuales destacamos tres que tratan de dar respuesta a las siguientes cuestiones (Aliseda, 1997; p. 6 y ss.)

1. ¿Producto o proceso? Hay razones para considerar ambas caras, apoyadas en consideraciones de metodología de la ciencia. Tras la distinción entre contextos de justificación y de descubrimiento, se ha centrado el interés "lógico" en el primero, en detrimento del segundo, aunque éste es de suma importancia para establecer, si ello es posible, una lógica del descubrimiento. Esta perspectiva encuentra su justificación en la historia de algunas disciplinas, como la Inteligencia Artificial, asimismo en la autoridad de algunos lógicos influyentes, entre los que destacan Beth y, más recientemente, J. van Benthem, D. Gabbay, etc. Prestando atención al lado "producto", el mayor interés estará en las condiciones que hacen que cierta información se considere explicativa, mientras que del lado "proceso" se centra la atención en los procedimientos que producen tales explicaciones.
2. ¿Construcción o selección? Si se trata de obtener hipótesis, tal vez aparezca un abanico de éstas, de las cuales solamente una será la mejor o más adecuada. Así la abducción consistirá no sólo en hallar hipótesis explicativas, sino también en proporcionar criterios de preferencia acerca de las hipótesis encontradas. No obstante, cabe distinguir las dos fases, una primera de búsqueda de hipótesis y una segunda de aplicación de determinados criterios de selección.

3. Abducción *versus* inducción. La tradicional distinción entre inferencia deductiva e inductiva había tomado como básica la primera, definiendo negativamente la segunda, en cuya clase se incluiría toda forma de inferencia no deductiva. Ahora bien, como subraya nuestra autora, proviniendo de J. S. Mill tal clasificación, éste reconoce varios métodos de descubrimiento y demostración de relaciones causales. En cualquier caso, la abducción difícilmente es asimilable en todas sus formas a ciertos tipos de inducción: desde la afirmación de un hecho cabe obtener abductivamente una explicación, mientras que en las generalizaciones inductivas se parte de ejemplos o casos concretos para pasar a una afirmación general, que suele adoptar el formato de proposición categórica universal.

Si hemos de tener en cuenta diversos aspectos en la abducción, ello es debido a que una delimitación absolutamente nítida de la noción es una empresa que no resulta nada fácil. Thagard (1988; p. 52 y ss.), señala cuatro clases de abducción a tener en cuenta en la elaboración de su programa "PI" para el estudio "computacional" de los procesos de investigación científica: abducción *simple*, *existencial*, de *formación de reglas* y *analógica*. El modo de obtener hipótesis explicativas a partir de la observación de un hecho, tal como se ha mencionado en la Introducción, sería una abducción simple en expresión de carácter proposicional. En términos de un lenguaje de predicados, dada una regla condicional de la forma "todo S es P" y observado el hecho "a es P", tenemos razones para afirmar hipotéticamente "a es S". En la abducción existencial la hipótesis obtenida es una afirmación de carácter existencial; así, por ejemplo, si se observa el hecho de que un cierto objeto metálico cambia su posición en un plano, expresable como "el objeto metálico *a* cambia su posición", y contamos con la regla "todo objeto metálico ante la presencia cercana de un objeto imantado en un mismo plano cambia su posición", se afirmará la hipótesis "existe un objeto cercano a *a* que es imantado". La formación de reglas se plantea a veces como una forma de generalización inductiva: desde "a_1 es S", "a_2 es S", ..., "a_n es S" se alcanza por generalización inductiva que "todo *a* es S", sin embargo, para la obtención de una regla mediante abducción se propone ésta como hipótesis explicativa del hecho de que se trate: si hemos de explicar por qué "*a* es P" y sabemos que "*a* es S", una hipótesis explicativa sería que "todo S es P", independientemente de la mayor o menor fuerza que tal afirmación pueda tener en su contexto. Por último, la abducción analógica aplica la analogía para la formulación de hipótesis que explican otros fenómenos conocidos: dado el hecho C_1, que viene perfectamente explicado por la hipótesis H, dado que el hecho C_2 fuera análogo al C_1, se abduce H como una explicación de C_2.

En todos los casos indicados se detecta que a partir de las observaciones y hechos conocidos pueden ser abducidas más de una hipótesis, es decir, tal vez nos encontremos en muchas situaciones con la existencia de múltiples explicaciones para un mismo hecho, si bien una de ellas será la explicación preferida atendiendo a determinadas razones. De aquí que la abducción haya sido considerada no ya como la inferencia de una explicación potencial, sino como la mejor explicación posible (Niiniluoto: 1999; p. 442). La denominada "inferencia de la mejor explicación" (abreviadamente "IBE", por las iniciales correspondientes en lengua inglesa), dicho sucintamente, recomienda la inferencia desde la evidencia de un hecho hasta cierta hipótesis cuando ésta es mejor explicación del hecho que cualquier otra hipótesis que pueda ser considerada (un estudio de la noción de IBE, propuesta inicialmente por Harman, se halla en Lipton, 1991).

Si nos centramos en las clases de abducción indicadas, tanto en la abducción simple como en la existencial, realizadas en ciertos contextos cognoscitivos, lo más natural será contar con más de una regla, lo cual permitiría la formulación de afirmaciones que constituyen otras tantas hipótesis alternativas. Asimismo, en el caso de la abducción de reglas explicativas, planteado por qué "*a* es P", además de conocer que "*a* es S", tal vez sean conocidos los hechos "*a* es Q_1", "*a* es Q_2", ..., "*a* es Q_n", siendo abducibles también las reglas "todo Q_1 es P", "todo Q_2 es P" y así sucesivamente. Por último, dado un hecho y varias hipótesis explicativas del mismo, si otro hecho es análogo al primero, en función del tipo de analogía, será preferible una hipótesis u otra como explicación del segundo.

Para presentar, de manera resumida, una caracterización general de la abducción que nos sirva como marco de referencia de los distintos procesos abductivos nos basamos en Hintikka (1999, p. 91), quien, siguiendo los planteamientos de Kapitan relativos a la noción estudiada por Peirce, propone las cuatro tesis siguientes:

1) Tesis inferencial. La abducción es, o por lo menos incluye, un proceso inferencial, o varios de ellos. En un proceso inferencial, haciendo uso de determinado lenguaje especializado, aparecen razonamientos que concluyen en una sentencia determinada; ésta, de acuerdo con el tipo de razonamiento, es la conclusión deducida cuando se trata de razonamiento deductivo, o bien la sentencia inducida si el razonamiento es inductivo y, si es el abductivo, entonces habría sido abducida la sentencia en cuestión. En definitiva, como en el caso de la deducción y de la inducción, en la abducción se requiere un proceso inferencial.
2) Tesis de objetivos. Los objetivos que corresponden a la abducción científica se pueden resumir en dos: a) generar nuevas hipótesis en el curso de la investigación científica, y b) seleccionar hipótesis para su posterior examen o verificación.

3) Tesis de comprehensión. Todas las operaciones por las cuales se engendran las teorías deberían ser incluidas en la abducción científica.
4) Tesis de autonomía. Tratándose de un proceso inferencial, la abducción incorpora razonamientos que son distintos e irreductibles a las otras formas de inferencia, es decir, es netamente distinta de la deducción y de la inducción.

En el estudio de las lógicas no clásicas, especialmente en su formulación secuencial, se presta atención a las reglas estructurales, las cuales no son relativas a las constantes lógicas sino a la relación de consecuencia correspondiente. Aunque ello no es incompatible, independientemente de este análisis estructural, cada modelo de abducción será analizable como un sistema en el que se deben dar, de alguna manera, las cuatro tesis indicadas.

III. Sistemas lógicos

El primer componente de un *sistema lógico* es un determinado lenguaje, generalmente un lenguaje formal, que consta de un conjunto no vacío de símbolos y una sintaxis que permite definir la clase de las fórmulas o expresiones bien formadas del mismo. Las reglas (sintácticas) se establecen recursivamente, de manera que, dada una cadena de símbolos, se cuenta con un procedimiento efectivo para determinar si dicha cadena constituye una fórmula o no. El vocabulario tiene, al menos, tres clases de símbolos:

1) Símbolos no lógicos. Según el tipo de lenguaje de que se trate, serán variables proposicionales, signos predicativos, etc.
2) Constantes lógicas, habitualmente *negador*, *disyuntor*, *conjuntor*, *implicador*, *cuantificador existencial* y *cuantificador universal*. Si se trata de un lenguaje proposicional, no posee los dos últimos; otros lenguajes pueden contener otros símbolos lógicos (como un *descriptor*, un operador de *abstracción*, etc.).
3) Los signos auxiliares como, por ejemplo, paréntesis, izquierdo y derecho, corchetes, etc.

La semántica de un lenguaje formal se define en términos de la *teoría de modelos*. Para ello se consideran estructuras abstractas del mismo tipo que el lenguaje; así, una estructura del mismo tipo que un lenguaje formal L, es decir, una L-estructura, viene dada por un dominio no vacío y una asignación de valores semánticos a los términos de L como la denotación de los mismos en tal L-estructura. Los signos lógicos tienen una interpretación constante en todas las L-estructuras (y en toda L^*-estructura, para todo lenguaje L^* que tenga los mismos signos lógicos que L) y decimos que se evalúan según tal interpretación. Asimismo se asignan valores de verdad a las fórmulas de L, o bien a una clase especial de ellas, a saber, las *sentencias*, fórmulas en las cuales no ocurren variables libres. Una L-estructura M *satisface* una fórmula A cuando, de acuerdo con las denotaciones de los términos que ocurren en A y la evaluación de sus signos lógicos, A es verdadera en M, también decimos que M es un modelo de A. La relación (semántica) de *consecuencia lógica* (en sentido clásico) se define entre conjuntos de fórmulas y fórmulas: una fórmula A es consecuencia lógica de un conjunto de fórmulas P si todo modelo de P (simultáneamente, de cada fórmula de P) es modelo de A, simbólicamente, $P \models A$. Aquellas fórmulas que son consecuencia lógica de cualquier conjunto de fórmulas (incluso del vacío), se dice que son *universalmente válidas*, verdaderas en virtud de su forma; $\models A$ expresa que A es universalmente válida, lo que equivale a afirmar que toda L-estructura la satisface.

Dado el lenguaje L, un sistema lógico o, abreviadamente, una *lógica*, es el par $(L, \models)$. Este es el punto de vista semántico, por así decir. También es definible una operación de consecuencia de carácter sintáctico, como es el caso de los sistemas de cálculo lógico axiomático o los deductivo-naturales, en los que se establece un sistema de inferencias o procedimiento de transformaciones de unas fórmulas en otras. Esta operación es expresable mediante $\vdash$: para el conjunto de fórmulas P y la fórmula A, $P \vdash A$ indica que la fórmula A se ha obtenido a partir de P en el sistema (clásico) correspondiente. Por las características de las operaciones de consecuencia, la de carácter sintáctico y la semántica, en algunos casos, éstas son intercambiables; más específicamente, en los sistemas de lógica clásica, por ejemplo, para una fórmula A y un conjunto de fórmulas P se verifica:

$$P \vdash A \text{ si y sólo si } P \models A,$$

por lo cual se dice que el sistema formal es *correcto* y *completo*. Dada la corrección y la completitud de los sistemas clásicos, una lógica vendrá dada también por el par $(L, \vdash)$.

Nos interesan las teorías científicas desde el punto de vista de las lógicas mediante las cuales pueden ser analizadas. En general, una teoría científica T se expresa en un lenguaje especializado, tal vez correlacionable con un lenguaje formal, en el sentido de que las denotaciones de los términos del lenguaje de T pueden hacerse corresponder con las denotaciones de un lenguaje formal en cierta estructura de su mismo tipo, así como sentencias del lenguaje de T con fórmulas del lenguaje formal, de manera que una clase de fórmulas sería el correlato formal o formalización de T, en cuya determinación puede jugar un papel importante cierto sistema lógico. Así, por ejemplo, si T es una teoría de matemática clásica, debidamente axiomatizada, su formalización (un conjunto de fórmulas del correspondiente lenguaje formal) constará de un conjunto de fórmulas que son correlatos de los axiomas específicos de *T* y de las fórmulas que se obtienen a partir de éstas aplicando el sistema inferencial, es decir, la lógica correspondiente.

IV. Modelos de lógica abductiva

IV.1. *El modelo inicial de Peirce*

El primer modelo a considerar es el propio modelo de Peirce (diversos pasajes, Peirce, 1958), acerca del cual se han dado algunas indicaciones en la introducción. El trabajo epistemológico de este autor tiene diversos aspectos relevantes, destacando su apuesta por el razonamiento de carácter ampliativo frente a la deducción en sentido clásico, de marcado carácter analítico. Su *lógica de la indagación* se compone de tres modos de razonamiento, a saber, deducción, inducción e hipótesis, siendo cada uno de ellos un proceso de prueba que se corresponde con una forma silogística, como se ilustra en el tradicional ejemplo de las alubias:

a) Deducción.

Regla:	Todas las alubias de esta bolsa son blancas,
Caso:	Estas alubias son de esta bolsa,
Resultado:	Estas alubias son blancas.

b) Inducción.

Caso:	Estas alubias son de esta bolsa,
Resultado:	Estas alubias son blancas,
Regla:	Todas las alubias de esta bolsa son blancas.

c) Hipótesis.

Regla:	Todas las alubias de esta bolsa son blancas,
Resultado:	Estas alubias son blancas,
Caso:	Estas alubias son de esta bolsa.

Los tres modos de inferencia, deductivo, inductivo y abductivo, están presentes en el proceso de descubrimiento. Dicho sucintamente, el patrón de inferencia peirceano tiene como punto de partida los hechos observados, un conjunto de enunciados a partir de los cuales se formulan hipótesis por abducción; desde las hipótesis se deducen nuevos enunciados que se contrastan con los datos iniciales, evaluando de este modo las hipótesis formuladas y, en su caso, se procede por inducción para la obtención de nuevas hipótesis.

Desde un punto de vista lógico, el planteamiento de Peirce viene a ser el de considerar que tenemos una regla general desde la que se debería deducir cierto hecho, si bien la regla no es suficiente, falta añadir alguna premisa para poder completar la deducción. Extendiendo el planteamiento, desde un conjunto de "premisas" no podemos deducir determinado enunciado que, sin embargo, sería deducible con una premisa adicional. Partiendo de una base de conocimiento **B**, si se produce el hecho sorprendente *B*, abductivamente se obtiene *A*, de tal manera que **B** junto con *A* constituyen las premisas desde las que se deduce *B*, es decir **B**, $A \models B$. Simbólicamente, como una operación lógica, es definible

$$(\mathbf{B}, B) \models_{\mathrm{AB}} A \text{ si y sólo si } \mathbf{B}, A \models B.$$

Las cuatro tesis señaladas anteriormente se verifican en el modelo peirceano, máxime cuando estas tesis han sido formuladas con un primer propósito de caracterizar las propuestas de Peirce. En efecto, la base de datos y el hecho constituyen el punto de partida, pero se requiere que este último sea "sorprendente" en el sentido de que cabía esperar que fuera deducible de la base de datos aunque no lo es realmente; puesto que cabe pensar en conexiones entre los enunciados correspondientes, se pretende "inferir" una premisa adicional necesaria, lo que viene exigido por la tesis inferencial. La tesis de objetivos está incluida en el patrón de inferencia esquematizado, ya que en el caso de Peirce se trata de la búsqueda de hipótesis. La inclusión de las operaciones de la teoría, en este caso la base de datos, no aparece especificada en el esquema propuesto, no obstante la exigencia misma de que se trate de explicar un hecho sorprendente lleva aparejada la necesidad de cumplir la tesis de comprehensión. La tesis de autonomía, por último, viene satisfecha por la especificidad del silogismo abductivo y la regla de razonamiento abductiva por excelencia, a saber de *Si A, entonces C* y de *C*, se abduce *A*.

IV.2. *Modelo de Aliseda*

Entre los modelos de sistemas abductivos destaca el modelo proposicional de tablas semánticas presentado en Aliseda (1997, p. 48 y ss.). Esta autora distingue varias versiones de razonamiento abductivo como forma de obtención de hipótesis explicativas; por su interés, más allá del modelo específico, ya que sienta las bases de los estudios lógicos de la abducción, las presentamos a continuación. Para mayor facilidad, consideramos que dada una teoría, operamos con su formalización *T*. Sean *T* una teoría, en este caso un conjunto de fórmulas de un lenguaje proposicional, y una fórmula *q*, que constituyen un problema abductivo, lo que representamos mediante (T, q); entonces la fórmula

p es una explicación abductiva o solución abductiva (del tipo indicado por el nombre previo) si se verifica que:

Plana

(i) $T, p \models q$

Consistente

(i) $T, p \models q$
(ii) T, p son consistentes

Explicativa

(i) $T, p \models q$
(ii) No es el caso que $T \models q$
(iii) No es el caso que $p \models q$

Mínima

(i) $T, p \models q$
(ii) p es la explicación más débil

Preferencial

(i) $T, p \models q$
(ii) p es la mejor explicación de acuerdo con un orden preferencial.

Dado que la abducción es una operación por la cual se obtiene una solución a un problema abductivo, se pueden considerar a los componentes de éste como las "premisas" de una operación lógica de consecuencia y a la solución correspondiente su "conclusión", tratándose de una operación bien distinta de $\vdash$ o de $\models$ en sentido clásico. Dado el problema abductivo (T, q), la operación de abducción se expresa $(T, q) \models_{AB} p$, como hemos indicado anteriormente. Nótese que, de acuerdo con las definiciones, en las cuales se hace uso de la operación clásica $\models$, se trata de una operación de carácter semántico.

En el caso de la abducción plana, la más elemental por así decir, planteado el problema abductivo (T, q), fácilmente se obtiene una solución: basta tomar la propia q, ya que $T, q \models q$; asimismo si r no es consistente con T, se tiene que $T, r \models q$, por lo cual $(T, q) \models_{AB} r$. Se trata, en definitiva de soluciones triviales. La mínima y la preferencial establecen una segunda cláusula que podría ser especificada en un sentido lógico; así, por ejemplo, un requisito especial se exige en la siguiente cláusula, en la que la operación de abducción (ahora selectiva en cuanto a los posibles resultados de su aplicación) se representa mediante $\models_{AB*}$: $(T, q) \models_{AB*} p$ si $(T, q) \models_{AB} p$ y para toda r sintácticamente relacionada y distinta de p tal que $(T, q) \models_{AB} r$ se verifica que $T, r \models p$. Es decir, de entre varias soluciones, obtenidas mediante $\models_{AB}$ que están sintácticamente relacionadas, se opta como solución (nueva operación de abducción, $\models_{AB*}$) por aquélla que es consecuencia de la conjunción de T con cada una de ellas.

Las más interesantes son tal vez la abducción consistente y la explicativa, desde el punto de vista de la aplicación de las tablas semánticas, aunque la autora hace un pormenorizado estudio de todas y cada una de las propuestas. Repasemos brevemente tal procedimiento. El método de las *tablas* o *árboles semánticos* fue iniciado por Beth (Beth, 1969) tratando la búsqueda sistemática de un contraejemplo a una supuesta relación de consecuencia lógica, tomada ésta en sentido clásico, aunque se han desarrollado tablas para otras lógicas (Nepomuceno, 2003).

A partir de un lenguaje proposicional, decimos que una fórmula A es *constituyente* de otra fórmula B si A es una subfórmula de B o de otra fórmula C equivalente (en sentido clásico) a B. Cada fórmula atómica o negación de una atómica es llamada *literal* (una variable proposicional y la negación de ésta, se dice que son literales complementarios). Se consideran tres clases de fórmulas no literales:

1) Negación de una subfórmula que es a su vez una negación.
2) Clase α. Son una conjunción de fórmulas, o la negación de una disyunción de fórmulas o negación de una implicación entre fórmulas,
3) Clase β. Son una disyunción de fórmulas, o la negación de una conjunción de fórmulas o bien una implicación entre fórmulas.

Para un conjunto (finito, no vacío) de fórmulas R, alguna de las cuales no es un literal, se define la tabla semántica $S(T)$ como una clase de sucesiones de fórmulas del lenguaje en cuestión; a cada sucesión se denomina *rama* y consta de las fórmulas de R, el conjunto inicial o *raíz*, y de aquéllas que se obtienen por aplicación de las reglas de formación a cada fórmula no literal de la rama en la que no se haya actuado; el proceso continúa hasta que aparece un par de contradicción, es decir dos literales complementarios (contradictorios, por tanto), en cuyo caso la rama es *cerrada*, o bien hasta que no quede ninguna fórmula no literal sin haberse aplicado a ella la regla correspondiente. A continuación presentamos las reglas, indicando bajo la línea la fórmula con que prosigue la rama en cuestión, o la división de la rama en su caso, al aplicar la regla sobre la fórmula anotada sobre la línea:

1) Regla de la doble negación:

$$\frac{\neg\neg A}{A}$$

2) Regla para fórmulas de la clase α :

$$\frac{A \wedge B}{\begin{matrix} A \\ B \end{matrix}}; \frac{\neg(A \vee B)}{\begin{matrix} \neg A \\ \neg B \end{matrix}}; \frac{\neg(A \rightarrow B)}{\begin{matrix} A \\ \neg B \end{matrix}};$$

3) Regla para fórmulas de la clase β :

$$\frac{A \vee B}{A \mid B}; \frac{\neg(A \wedge B)}{\neg A \mid \neg B}; \frac{A \rightarrow B}{\neg A \mid B};$$

De acuerdo con la propiedad fundamental de las tablas, el conjunto raíz es satisfactible si y sólo si la tabla es abierta, de donde, por contraposición, la raíz es no satisfactible si y sólo si la tabla es cerrada. De aquí que, en general, si P es un conjunto de sentencias y C es una sentencia tales que $P \models C$, entonces

$$S(P \cup \{\neg C\})$$

es una tabla cerrada. Estos resultados son de aplicación en el estudio de la inferencia abductiva. En efecto, sea el problema abductivo (T, q); naturalmente, si se obtiene una fórmula p, de acuerdo con la noción de abducción, se verificará $T, p \models q$, de manera que la siguiente tabla deberá ser cerrada

$$S(T \cup \{p\} \cup \{\neg q\})$$

Ahora bien,

$$T, p \models q \text{ si y sólo si } T, \neg q \models \neg p$$

por lo que podemos redefinir la operación $\models_{AB}$ en términos de las tablas semánticas. Para ello, consideramos que dado un conjunto de fórmulas P, *un conjunto de cierre* de la tabla $S(P)$ es un conjunto de fórmulas $\{A_1,...,A_n\}$ tal que al añadir a cada rama abierta R alguna de estas fórmulas (en su caso, obtenida desde ellas por composición proposicional), se produce un par de contradicción y, por tanto, un cierre de la misma; $R + A_i$ indica que se añade A_i a la rama R, para alguna i menor o igual que n. Si el cierre de todas las ramas se produce con una de estas fórmulas, se indicará $S(P) + A_i$ (en su caso, $S(P) + A_i + A_j$ +...). De este modo,

$$(T, q) \models_{AB} p \text{ si y sólo si } S(T \cup \{\neg q\}) + p \text{ es cerrada}$$

Veamos un sencillo ejemplo. Sea la siguiente teoría

$$T = \{p \rightarrow q, q \rightarrow r\}$$

y se trata de explicar el hecho r. Tomamos como raíz la propia teoría junto con la negación del hecho, dando lugar a una tabla con las siguientes ramas:

R1=$\{p \rightarrow q,\ q \rightarrow r,\ \neg r,\ \neg p,\ \neg q\}$,
R2=$\{p \rightarrow q,\ q \rightarrow r,\ \neg r,\ \neg p,\ \neg r\}$,
R3=$\{p \rightarrow q,\ q \rightarrow r,\ \neg r,\ q,\ \neg q\}$,
R4=$\{p \rightarrow q,\ q \rightarrow r,\ \neg r,\ q,\ r\}$.

Las ramas R3 y R4 son cerradas ya que aparece en cada una de ellas un par de contradicción, es decir, dos literales complementarios, q y $\neg q$ en R3 y r y $\neg r$ en R4. Las ramas R1 y R2 son abiertas, por lo que a partir de las mismas es definible (extensionalmente) el conjunto de cierre: $\{r, p, q\}$, que contiene todos los abducibles. La solución definitiva al problema abductivo (T, r) se tomará en función del tipo de abducción que estemos considerando. En el caso de la plana, es suficiente cualquiera de los tres literales abducibles, lo mismo que si se trata de la abducción consistente. Sin embargo, para la explicativa tenemos que únicamente p y q son soluciones, ya que r, que sería una solución trivial, no cumple el requisito (iii): $r \models r$. En este ejemplo no cabe buscar la explicación más débil para hablar de la abducción mínima, de manera que las inferencias abductivas posibles son

$$(T, r) \models_{AB} p \text{ y } (T, r) \models_{AB} q.$$

Desde el punto de vista de la abducción preferencial, aunque en el ejemplo que nos ocupa no hemos indicado específicamente un orden preferencial, atendiendo a la tesis de comprehensión propuesta entre las tesis características de la abducción de Hintikka referidas más arriba, sería preferible p sobre q, pues el primer literal pone en juego, por así decir, toda la teoría, es decir, todas las fórmulas de T, en el sentido de que r se deduce a partir de p y de todas las fórmulas de T (en el caso de q, no se necesita más que una fórmula de las dos que integran T).

En este modelo se estudia la operación $\models_{AB}$ (la cual se simboliza con otra notación) desde el punto de vista de su caracterización estructural. A partir del análisis de la relación de consecuencia lógica clásica, $\models$, cuyas reglas estructurales son *reflexividad, contracción, permutación, monotonía* y *corte* (transitividad), en el caso de la abducción consistente, se comprueba que, en general, no es reflexiva, monótona ni transitiva. No obstante, se dan algunas de éstas con ciertas condiciones, en concreto, *reflexividad condicional, corte simultáneo* y *consistencia de la conclusión* (Aliseda, 1997: 56). Se plantea entonces si la abducción explicativa consistente, tal vez la más interesante para la metodología de la ciencia, es realmente una lógica y la respuesta estará en función del concepto de lógica asumido. Si nos atenemos a lo definido anteriormente, que un sistema lógico viene dado por un lenguaje formal y una operación de consecuencia, la abducción mediante tablas semánticas descrito tiene como base un sistema inferencial claramente aceptado desde distintas perspectivas (filosóficamente, como búsqueda sistemática de contraejemplos, computacionalmente, con cierto atractivo para su implementación, etc.), en el que se contempla la adición de literales a ramas abiertas. Por otra parte, algunas de las formas de abducción contempladas verifican las tesis de Kapitan-Hintikka como, por ejemplo, la abducción explicativa consistente con algún criterio preferencial. En efecto, se trata de un proceso de inferencias (tesis inferencial), genera hipótesis explicativas (tesis de objetivos), entran en juego todas las "operaciones de la teoría" (tesis de comprehensión) y, por último, es distinta de la mera deducción y de la inducción por cuanto al procedimiento de tablas se añade la obtención de conjuntos de cierre y adición de literales (tesis de autonomía).

IV.3. *Abducción y tablas de primer orden*

Aunque el análisis de las características de las leyes científicas requiere una atención que no cabe prestar en este trabajo, desde un punto de vista sintáctico se suelen presentar como expresiones condicionales generalizadas universalmente (Nagel, 1978: p. 56 y ss.). Asimismo, la explicación mediante la abducción de reglas, como hemos visto, hace uso de este formato: "para todo x, si x es S, entonces x es P"; sin embargo, en el modelo anterior se ha hecho uso de un lenguaje proposicional, claramente insuficiente para expresar proposiciones de esta clase. En primer término, bastaría con adoptar un lenguaje de predicados (de primer orden), manteniendo lo esencial del modelo: el uso de las tablas semánticas y el estudio de las ramas abiertas y las condiciones requeridas para obtener su cierre, tomando como raíz la teoría y el hecho a explicar, debidamente formalizados.

En Mayer, Pirri (1993) se presenta una primera tentativa de estudio de la abducción mediante cálculo de secuentes y tablas semánticas en primer orden. Para ampliar el método de las tablas a primer orden, a las reglas antes indicadas para el nivel proposicional, hay que añadir ahora aquéllas que tratan de sentencias de las clases γ y δ , fórmulas con cuantificadores universales y existenciales (o negaciones de éstos, en orden inverso) como signo principal, respectivamente. Con el mismo criterio para la presentación de las reglas a nivel proposicional (encima de la línea la fórmula sobre la que se actúa y bajo ella las fórmulas que se añaden a la rama correspondiente), las nuevas reglas son:

4) Regla para fórmulas de la clase γ :

$$\frac{\forall xB}{\begin{array}{c} B(a_1/x) \\ B(a_2/x) \\ \vdots \\ B(a_n/x) \end{array}}$$

donde las fórmulas $B(a_i/x)$ representan las obtenidas por sustitución, con todas las constantes que ocurren en la rama de que se trate. También son de esta clase las fórmulas $\neg\exists xB$, que dan lugar a fórmulas $\neg B(a_i/x)$.

5) Regla para fórmulas de la clase δ :

$$\frac{\exists xB}{B(a_{n+1}/x)}$$

donde $B(a_{n+1}/x)$ representa la fórmula resultante de sustituir en la fórmula B cada ocurrencia libre de la variable x por la constante a_{n+1}, que ha de ser nueva en la rama. También son de esta clase las fórmulas $\neg\forall xB$, que dan lugar a $\neg B(a_{n+1}/x)$

Veamos un ejemplo sencillo. $T = \{\forall x(Px \rightarrow Qx); \forall x(Qx \rightarrow Rx)\}$; hecho Ra. Tomando como raíz las fórmulas de T junto con $\neg Ra$, se genera una tabla con las siguientes ramas:

R1 = $\{\forall x(Px \rightarrow Qx), \forall x(Qx \rightarrow Rx), \neg Ra, Pa \rightarrow Qa, Qa \rightarrow Ra, \neg Pa, \neg Qa\}$,
R2= $\{\forall x(Px \rightarrow Qx), \forall x(Qx \rightarrow Rx), \neg Ra, Pa \rightarrow Qa, Qa \rightarrow Ra, \neg Pa, Ra\}$,
R3= $\{\forall x(Px \rightarrow Qx), \forall x(Qx \rightarrow Rx), \neg Ra, Pa \rightarrow Qa, Qa \rightarrow Ra, Qa, \neg Qa\}$,

$$\text{R4} = \{\forall x(Px \rightarrow Qx), \forall x(Qx \rightarrow Rx), \neg Ra, Pa \rightarrow Qa, Qa \rightarrow Ra, Qa, Ra\}.$$

Las ramas R2, R3 y R4 son cerradas, mientras que R1 es abierta; respecto de esta última, $\{Ra, Pa, Qa\}$ constituye un conjunto de cierre y, de manera análoga al caso proposicional, la mejor explicación sería Pa.

En una simple extensión como la ejemplificada, no obstante, se nos plantea un problema inexistente a nivel proposicional, como se indica a continuación. Para mayor facilidad consideramos la abducción explicativa y consistente. Sea un problema abductivo (T, A), donde T representa la teoría, como en el modelo proposicional, aunque ahora se trata de fórmulas de primer orden, A es también una fórmula de primer orden; entonces B es una solución si se verifica que $T, B \models A$. En términos de tablas semánticas, al darse dicha relación de consecuencia lógica, la tabla semántica $S(T \cup \{\neg A\} \cup \{B\})$ será cerrada; al construir la tabla $S(T \cup \{\neg A\})$ hallaremos alguna rama abierta desde la que se establece un conjunto de cierre y, por tanto, el conjunto de abducibles, como hemos visto en el ejemplo. Ahora bien, los sistemas de primer orden son semidecidibles en el sentido de que si, en general, es definible un procedimiento para determinar que $A_1,A_2,\ldots,A_n \models B$, tal procedimiento no sería efectivo para establecer que no se da esta consecuencia lógica. En el caso de las tablas semánticas en primer orden, ello se traduce en la posibilidad de ramas infinitas: Si $\{A_1,A_2,\ldots, A_n, \neg B\}$ no es satisfactible, la tabla semántica es cerrada, pero si $\{A_1,A_2,\ldots, A_n, \neg B\}$ es satisfactible, podría ocurrir que la tabla en cuestión tuviera una rama infinita (Nepomuceno, 2003: p. 113 y ss.).

Veamos un ejemplo en el que se plantea este problema. Sea $T = \{\forall x\exists yRxy\}$ y el hecho $\neg Ra_1a_1$. Tomando como raíz la fórmula de la teoría y $\neg\neg R\, a_1a_1$, se genera una única rama que se expande con las fórmulas $\forall x \exists yRxy$, $\neg\neg R\, a_1a_1$, $R\, a_1a_1$, $\exists yRa_1y$, $R\, a_1a_2$, $\exists yRa_2y$, $R\, a_2a_3$, ... (la primera constante aparecía en la raíz, pero la aparición de cada constante nueva al aplicar la regla δ obliga a volver a la regla γ repetidamente, dando lugar a una nueva fórmula de la clase δ, por la cual aparecerá otra constante nueva y así sucesivamente). Sin embargo, la fórmula inicial junto con $\neg\neg R\, a_1a_1$ es satisfactible en un universo de discurso unitario.

Una variante de la regla para fórmulas de la clase δ consiste en dividir la rama en tantas ramas como constantes ocurrían previamente más una, es decir formando nuevas ramas continuando la precedente con $B(a_1/x)$, $B(a_2/x)$, ..., $B(a_n/x)$, $B(a_{n+1}/x)$. En este caso tendremos lo que denominaremos una *tabla DB* (iniciales de Díaz-Boolos, en Aliseda, Nepomuceno, 2002); estas tablas modificadas permiten definir modelos mínimos de las fórmulas de la raíz. De esta manera, para algunas sentencias que son satisfactibles en dominios finitos y cuyas tablas tenían ramas abiertas infinitas, se hace posible la construcción de tablas DB con ramas finitas acabadas y abiertas. Esta modificación del procedimiento de las tablas no soluciona el problema de la semidecidibilidad de los sistemas de predicados de primer orden, pero hace más amplia la clase de las sentencias acerca de las cuales podemos establecer ciertos resultados de una manera efectiva. Así, en el ejemplo comentado, obtendríamos una tabla DB con una rama acabada y abierta integrada por las fórmulas $\forall x \exists y Rxy$, $\neg\neg R\ a_1 a_1$, $R\ a_1 a_1$, $\exists y R a_1 y$, $R\ a_1 a_1$; como el conjunto de cierre para esta rama tiene $\neg R\ a_1 a_1$ como único literal de cierre, no cabe una explicación (distinta de la afirmación del propio hecho, lo que sería trivial): dado que es definible un modelo con tal universo de discurso, interpretando R como la relación "identidad", éste satisface la teoría, pero no satisface el hecho en cuestión. En cuanto a universos de discurso de mayor cardinal, se iría comprobando que en los conjuntos de cierre siempre aparecerían $\neg R\ a_i a_i$, para cada i finito, lo que llevaría a proponer, finalmente, $\forall x \neg Rxx$ con la restricción de que la teoría únicamente tiene sentido en dominios no unitarios (y, obviamente, distintos de vacío).

Así pues, el modelo abductivo de Aliseda haciendo uso de las tablas semánticas es ampliable mediante un lenguaje de predicados de primer orden. Cuando se trabaja con una teoría que sea finitamente satisfactible, el procedimiento de tablas será el de las tablas DB, las cuales, de acuerdo con la definición (modificada) de la regla para fórmulas de la clase δ , contienen las tablas estándar en el sentido de que cada rama de una tabla estándar es también una rama de la tabla DB de la misma raíz que la anterior. Esta ampliación también da lugar a diversas formas de abducción, manteniéndose las características del modelo precedente por lo que respecta a las tesis de Kapitan-Hintikka.

IV.4. *Procedimientos de resolución*

A partir de los planteamientos de Aliseda, haciendo uso del método de resolución en lugar de del de las tablas semánticas, también se procede sistemáticamente en la búsqueda de soluciones a los problemas abductivos. En efecto recordemos brevemente este método a nivel proposicional. Una *cláusula* no es más que un conjunto de literales que viene a ser equivalente a una fórmula en forma disyuntiva o *disyunción elemental* (disyunción de literales), mientras que un conjunto de cláusulas representa (en expresión clausal) una forma normal conjuntiva, es decir, una conjunción de disyunciones elementales. Una valoración de los literales satisface una cláusula si al menos un literal de la cláusula tiene asignado el valor de verdad "verdadero". Un conjunto de cláusulas será verdadero para una asignación de valores a los literales si cada una de las cláusulas es satisfecha por la asignación.

El cálculo de resolución tiene como única regla el denominado *principio de resolución*, de acuerdo con el cual, si **A** y **B** representan dos cláusulas de manera que en la primera ocurre un literal l y en la segunda un literal l' complementario del primero, entonces se obtiene una nueva cláusula resultante de unir las anteriores excepto estos dos literales, es decir, desde las cláusulas **A** y **B** se obtiene la cláusula $(\mathbf{A}\text{-}\{l\}) \cup (\mathbf{B}\text{-}\{l'\})$. Como propiedad fundamental, se tiene que un conjunto de cláusulas es inconsistente si mediante resolución (aplicando sucesivamente esta regla) se obtiene la cláusula vacía, la cual, por definición, carece de literales.

Así pues, planteado un problema abductivo (T, A), dado que nos interesa obtener B tal que $T, \neg A \models \neg B$, ello equivale a que $T \cup \{\neg A\} \cup \{\neg\neg B\}$, es decir, $T \cup \{\neg A\} \cup \{B\}$, sea inconsistente; mediante resolución, partiendo de T y $\neg A$, expresadas en forma clausal, se determinará qué fórmulas son necesarias para obtener la cláusula vacía y de aquí el conjunto de abducibles.

Una interesante propuesta es la basada en un procedimiento de *resolución dual*, definible como dual del conocido cálculo basado en el principio de resolución (Soler, 2004). En la resolución dual, cada cláusula equivale a una conjunción elemental, es decir, $\{p, \neg q, r\}$, por ejemplo, no es más que la expresión clausal (dual) de la fórmula $p \wedge \neg q \wedge r$. Un conjunto de cláusulas es ahora equivalente a una fórmula en forma normal disyuntiva, es decir, una disyunción de conjunciones elementales. El formato del principio de resolución se mantiene. Sin embargo, a diferencia de lo que sucedía en la resolución normal, una cláusula es satisfecha por una valoración de los literales si todos éstos tienen asignados el valor de verdad "verdadero", mientras que un conjunto de cláusulas es verdadero si al menos una de las cláusulas es verdadera.

El *principio de resolución dual*, en cuanto a su presentación, se expresa como el ya estudiado. Sin embargo, la obtención de la cláusula vacía indica que el conjunto inicial de cláusulas es siempre verdadero. Soler (2004, p. 393) propone una aplicación de resolución dual al estudio de problemas abductivos; para ello, hemos de tener en cuenta que dado un problema abductivo (T, A), se trataría de hallar B tal que $T, B \models A$, lo cual equivale a afirmar la siguiente validez universal: $\models T \wedge B \rightarrow A$; en este contexto y en otros análogos, nos tomaremos la licencia de que T exprese la conjunción de las fórmulas de la propia teoría T ¿Qué es necesario para que $T \rightarrow A$, una vez se haya completado, llegue a ser una expresión universalmente válida? Se trabajará con la forma $\neg T \vee A$, equivalente a la anterior, aplicándose el mencionado principio de resolución dual. Veamos el ejemplo proposicional: $T = \{p \rightarrow q, q \rightarrow r\}$; sea r el hecho a explicar; $\neg T$ es la fórmula $\neg(p \rightarrow q) \vee \neg(q \rightarrow r)$, equivalente a $(p \wedge \neg q) \vee (q \wedge \neg r)$, y al considerar la disyunción con r, resulta el siguiente conjunto en forma clausal (dual): $\{p, \neg q\}, \{q, \neg r\}, \{r\}$; por sucesiva aplicación del principio de resolución (dual) se obtienen $\{q\}$ y $\{p\}$. Las cláusulas unitarias definen el conjunto de abducibles, para lo que se toman directamente los literales (no sus complementarios): C=$\{r, q, p\}$; naturalmente, $\models \neg(T \wedge X) \vee A$ o, lo que es equivalente, $\models T \wedge X \rightarrow A$, para cualquier X de C. La opción definitiva por uno de estos literales también estará en función del tipo de abducción que se esté considerando; tanto $\{q\}$ como $\{p\}$ cumplen con los requisitos de la explicativa consistente y en cuanto a la adopción de un criterio preferencial, el autor propone que éste se tome de lo que denomina "historia de las cláusulas", lo cual, aplicado a este ejemplo, nos da como mejor explicación p ya que su historia incluye la de q (la cláusula $\{q\}$ apareció antes que $\{p\}$, siendo necesaria, por así decir, para la obtención de esta última).

Tanto la resolución como la resolución dual son, en principio, ampliables a primer orden, independientemente de que se hayan de tener en cuenta ciertas dificultades, ya indicadas en el apartado anterior. Dado que este modelo también se fundamenta en las características de la relación de consecuencia lógica en sentido clásico, un somero examen del modo en que se obtienen las hipótesis explicativas pone de manifiesto que también en este caso se verifican las tesis de Kapitan-Hintikka.

IV.5. *Los cálculos deductivo-naturales*

Los cálculos deductivo-naturales, a diferencia de las tablas semánticas, son de carácter sintáctico y definen también una operación de consecuencia (sintáctica) en sentido clásico, representable como $\vdash$. La obtención de una solución a un problema abductivo mediante aplicación de estos sistemas tiene como punto de partida las propiedades de esta relación de consecuencia. En efecto, para un conjunto P de fórmulas de un lenguaje formal y las fórmulas A y B, se verifica que

$$P, A \vdash B \text{ si y sólo si } P, \neg B \vdash \neg A.$$

De este modo, dada una teoría T y el problema abductivo (T, B), la búsqueda de una solución consistirá en hallar una fórmula A tal que $T, A \vdash B$; es decir, la operación de abducción $\vdash_{AB}$, ahora de carácter sintáctico, es definible en estos nuevos términos:

$$(T, B) \vdash_{AB} A \text{ si y sólo si } T, A \vdash B,$$

y, teniendo en cuenta la equivalencia anterior,

$$(T, B) \vdash_{AB} A \text{ si y sólo si } T, \neg B \vdash \neg A.$$

Así pues, planteado el problema abductivo, se trata de obtener en primer término una deducción a partir de la teoría y la negación del hecho a explicar.

Ahora bien, se ha afirmado anteriormente, lo que constituye la tesis de autonomía de Kapitan-Hintikka, que la (operación de) abducción debe ser netamente distinta tanto de la inducción como de la deducción, aunque, de acuerdo con la tesis inferencial, es (o incluye) algún proceso inferencial. En definitiva, estamos prácticamente ante los mismos presupuestos a partir de los cuales se define el modelo de Aliseda estudiado más arriba.

Dado un lenguaje formal y planteado un problema abductivo, en este modelo la abducción de una "hipótesis explicativa" se define como una sucesión de fórmulas que comienza con las fórmulas de la "teoría" como premisas, seguidas de la negación de la fórmula que expresa el "hecho", el cierre de la misma por una fórmula determinada, la fórmula resultante de aplicar el teorema de la deducción (con tal negación como antecedente y la fórmula que cierra la hipótesis como consecuente) y concluye con una fórmula como resultado de aplicación de la *afirmación de hipótesis abductiva* (abreviadamente, AHA), la cual constituye la solución al problema abductivo (Nepomuceno, 2004). Esta AHA se expresa con una serie de condiciones antecedentes, en función del tipo de abducción que estemos considerando, según la clasificación propuesta por Aliseda (1997; p. 48 y ss.). Para la abducción explicativa consistente, tendremos como AHA

$$\frac{(T, B);\ \vdash \neg B \rightarrow \neg A;\ \text{no-}(T \vdash A);\ \text{no-}(T, A \vdash \perp);\ A \neq B}{A}$$

no-(T, A $\vdash$ $\perp$) expresa que A es consistente con T (ya que no es el caso que de T, A se deduzca $\perp$, la constante de falsedad).

Veamos los ejemplos ya estudiados, aplicando ahora este tipo de cálculo. En el caso proposicional, para $T = \{p \rightarrow q, q \rightarrow r\}$ y r como hecho a explicar:

1. $p \rightarrow q$	Premisa,
2. $q \rightarrow r$	Premisa,
3. $\neg r$	Hipótesis auxiliar
4. p	Hipótesis auxiliar
5. q	*Modus ponens* 1, 4,
6. r	*Modus ponens* 2, 5,
7. $r \wedge \neg r$	Producto lógico 6, 3,
8. $\neg p$	Absurdo 4,
9. $\neg r \rightarrow \neg p$	Teorema de la deducción (deducción subsidiaria 3-8),
10. p	AHA, 1, 2, 9.

El paso dado para alcanzar 10 es justamente la regla mencionada, puesto que se dan todas las condiciones antecedentes.

La regla AHA ha sido enunciada de manera que los esquemas involucrados pueden representar fórmulas de un lenguaje proposicional o de uno de predicados de primer orden. Por ello, en principio, el procedimiento es también definible para este último tipo de lenguajes. No obstante, hemos de tener en cuenta que en las condiciones antecedentes se incluyen indicaciones acerca de la operación de consecuencia sintáctica, $\vdash$, por lo que no debemos olvidar el problema de la indecidibilidad de los sistemas de cálculo de primer orden. En cualquier caso, nos ocupamos de otro ejemplo ya estudiado. Sea el problema abductivo (T, A), de modo que la teoría $T = \{\forall x(Px \rightarrow Qx); \forall x(Qx \rightarrow Rx)\}$ y Ra sea el hecho a explicar; se obtiene:

1. $\forall x(Px \rightarrow Qx)$	Premisa,
2. $\forall x(Qx \rightarrow Rx)$	Premisa,
3. $\neg Ra$	Hipótesis auxiliar,
4. $Pa \rightarrow Qa$	Eliminación cuantificador 1,
5. $Qa \rightarrow Ra$	Eliminación cuantificador 2,
6. $Pa \rightarrow Ra$	Silogismo hipotético 4, 5,
7. $\neg Pa$	*Modus Tollens* 3, 6,
8. $\neg Ra \rightarrow \neg Pa$	Teorema de la deducción (deducción subsidiaria 3-7),
9. Pa	AHA, 1, 2, 8.

Tal como ha sido definida la regla AHA, la obtención de la hipótesis explicativa se ajusta a los requisitos de las tesis de Kapitan-Hintikka. Así, la tesis inferencial se cumple en ambos ejemplos, dado que contienen procesos inferenciales, tal es el caso de la deducción subsidiaria. Se ha generado una hipótesis que, en el caso de la investigación científica en concreto, sería posteriormente sometida a los procesos de contrastación (tesis de objetivos). La utilización de "toda" la teoría, aunque no aparece en el antecedente de la regla, podría plantearse como la exigencia de involucrar todos los postulados de la teoría en la obtención de la hipótesis, lo que viene a ser una forma de verificar la comprehensión. Por último, la naturaleza de AHA es claramente distinta de la mera deducción, sin que se trate de un proceso estrictamente inductivo, es decir, verifica lo establecido en la tesis de autonomía.

IV.6. *El punto de vista informacional*

La distinción entre una concepción *transformacional* y otra *informacional* de la lógica aparece en Corcoran (1998). El punto de vista informacional estudia la argumentación a partir de la noción primitiva de información. A este respecto, hemos de recordar que una argumentación se presenta como una sucesión de proposiciones, ordenada de manera no arbitraria, en la que se distinguen un conjunto inicial, las *premisas*, una cadena de razones y, finalmente, una última proposición denominada *conclusión*. Expresadas las proposiciones en un lenguaje dado, la argumentación adopta la forma de una terna (P, K, C), donde P es el conjunto de las sentencias que expresan las premisas, K es la cadena de razones y C la conclusión; eventualmente, alguna de las primeras podría ser vacía, mientras que la conclusión es un elemento indispensable. A los extremos de una argumentación, es decir, al par (P, C) se le suele denominar *argumento* de la misma.

El estudio de la corrección de una argumentación se reduce al estudio de la *validez* del argumento en cuestión. En una aproximación informacional a la lógica, se investiga la validez (o invalidez) argumental en términos de contenido de información de proposiciones y de conjuntos de proposiciones. La información entonces se concibe con respecto a un dominio de investigación determinado, dándose una información "total" del dominio de investigación. Cabe también hablar de la información "nula". Cada proposición tiene un contenido informativo que no se debe confundir con el valor semántico de la sentencia que la expresa en un lenguaje, aunque se pueda alcanzar el contenido informativo de una proposición mediante el análisis de la correspondiente sentencia.

Corcoran caracteriza su aproximación informativo-teórica a la lógica mediante seis observaciones que pueden ser entendidas como otras tantas reglas, las cuales permiten establecer la relación entre el concepto lógico de consecuencia y la noción de contenido informativo. En general, $P \models_I B$, representa que la proposición que la sentencia B significa es una *consecuencia lógica* (en el sentido de la concepción informacional) del conjunto de las proposiciones expresadas por las sentencias del conjunto P, o, lo que es lo mismo, que las proposiciones expresadas por las sentencias del conjunto P implican lógicamente (desde el punto de vista informacional) la proposición significada por B. Si se trabaja con un lenguaje formal, las sentencias serán fórmulas de dicho lenguaje, usadas como medio de expresión de proposiciones; teniendo esto en cuenta, para abreviar, podemos referirnos al contenido informativo de una sentencia, bien entendido que se trata del contenido informativo de la proposición significada por tal sentencia. Las reglas en cuestión son las siguientes:

1) $P \models_I B$ si el contenido informativo de B está incluido en el contenido informativo de la unión de los contenidos informativos de las sentencias que integran P.
2) B no es consecuencia lógica (informativa) de P, es decir, P no implica lógicamente B, si las sentencias de P no comparten información con B.

3) Si B es una tautología, entonces su contenido informativo es nulo.
4) Si B es una contradicción, entonces su contenido informativo es la información total.
5) Una proposición y su negación no comparten información alguna y la unión de ambos contenidos informativos da lugar a la información total.
6) La disyunción de dos proposiciones tiene como contenido informativo la información que comparten ambas proposiciones.

Como regla adicional se considera que dos proposiciones son equivalentes cuando comparten exactamente el mismo contenido informativo. Asimismo, se desprende de las reglas que la conjunción de dos proposiciones contiene la "unión" de la información de cada una de ellas. En Corcoran (1998: p. 116 y ss.) aparecen varios ejemplos de pasajes históricos para identificar esta aproximación, en concreto textos de Boole, De Morgan, Jevons y Venn; por otro lado, las perspectivas de la teoría de modelos, teoría de conjuntos y sustitucional podrían ser establecidas no como aproximaciones informativo teóricas, sino de carácter transformacional, de acuerdo con la cual un argumento es válido si ninguna transformación da lugar a un argumento en el que todas las transformadas de las premisas son verdaderas y la transformada de la conclusión es falsa.

La noción de inferencia deductiva, desde el punto de vista informacional, queda claramente establecida por las reglas 1) y 2). Veamos la abducción desde este mismo punto de vista. Dado un problema abductivo (T, A), lo que se plantea es completar un proceso que de estar concluido daría como resultado una relación de consecuencia informativa; el problema abductivo se plantea simbólicamente como T, ¿? $\models_I A$, de manera que se deberá hallar una B que pueda ocupar el hueco indicado por ¿?. Corcoran (2004) señala que la condición necesaria y suficiente para tal B es que (la proposición significada por) ésta implique lógicamente (la proposición significada por la expresión condicional) $T \rightarrow A$, ya que la información de B está en la información del consecuente pero no en el antecedente, es decir, de acuerdo con la regla 6), dicha información está contenida en la información que comparten $\neg T$ y A, dado que tal expresión condicional equivale a $\neg T \vee A$. La definición de abducción plana en estos términos, dado el problema abductivo (T, A), sería la siguiente:

$$(T, A) \models_{AB} B \text{ si y sólo si } B \models_I T \rightarrow A,$$

o, lo que resulta equivalente,

$$(T, A) \models_{AB} B \text{ si y sólo si } B \models_I \neg T \vee A.$$

La conexión entre la teoría T y el hecho a explicar A es una relación entre los contenidos informativos de las proposiciones correspondientes. A este respecto, si la información que comparten la teoría y el hecho es la información nula, es decir, si la teoría y el hecho no comparten información en absoluto, entonces el intento de explicación tendrá un resultado más bien pobre, por así decir; en este caso la información que correspondería a B sería justamente la información correspondiente a A: al no compartir A información ninguna con T, toda la información de A está contenida en la información de $\neg T$, por lo que la información de $\neg T \vee A$ coincide con la información de A (como se ha indicado, la información de una disyunción es la información compartida por cada proposición de la disyunción).

Por otra parte, las reglas 3) y 4) permiten establecer criterios de relevancia y resultan útiles para el estudio de algunas clases de abducción. En efecto, desde el punto de vista de la argumentación científica, carecen de interés aquellas inferencias deductivas que parten de premisas inconsistentes, es decir, que contienen una contradicción, dado que, en tal caso, se parte de la información total del dominio de investigación de que se trate y la información de cualquier proposición pertinente estará contenida en aquélla. Asimismo, ninguna tautología tiene interés como "teorema", dado que su información es nula y, en consecuencia, está contenida en la información proporcionada por cualquier proposición que se tome como premisa. Si lo tenemos en cuenta para el estudio de la abducción, en el caso de la abducción consistente, se exigirá que B junto con T no proporcionen toda la información del dominio de investigación de la teoría, pues si su conjunción es contradictoria, entonces implican cualquier proposición. En cuanto a la abducción explicativa, cabe excluir el caso en que el hecho y la teoría no compartan información, como se ha indicado antes; además, la segunda condición de relevancia misma exige que B no coincida con la propia A, pues si $B = A$, es evidente que $T, A \models_{\mathrm{I}} A$, es decir, $A \models_{\mathrm{I}} T \rightarrow A$.

V. Consideraciones finales

Los supuestos básicos de la abducción, como hemos señalado anteriormente, aparecen en el pensamiento de Peirce. Desde un punto de vista lógico, dado un conjunto de sentencias T de determinado lenguaje (la "teoría") y una sentencia A del mismo (el "hecho"), se plantea un problema abductivo como el de hallar alguna sentencia B (la "hipótesis explicativa") tal que A sea una consecuencia del conjunto de sentencias que contiene tanto a las de T como a B. Aunque los modelos lógicos referidos hacían uso de un lenguaje determinado, la caracterización de (las formas de) la abducción es independiente del lenguaje en el cual se presenta el correspondiente problema. Las distintas formas o versiones de la abducción dadas por Aliseda constituyen un ineludible punto de partida para la elaboración de un modelo lógico, asumiendo, en cualquier caso, que al encontrar una solución, es decir, al alcanzar una hipótesis explicativa, hemos realizado una "inferencia", una operación lógica, aunque de naturaleza distinta a la relación de consecuencia en sentido clásico, ya sea ésta de carácter semántico o sintáctico.

Las aplicaciones de estos modelos son diversas, entre las que destacan las relacionadas con los estudios lingüísticos y con las ciencias de la computación. La comprensión del lenguaje natural, sobre todo el estudio de las presuposiciones y la interpretación de oraciones ambiguas, es analizable en términos de explicaciones a partir de una teoría. Por lo que respecta al segundo campo, en Kakas, Kowalski, Toni (1993) se mencionan como aplicaciones en inteligencia artificial la asimilación de conocimiento, ya se trate de actualización de bases de datos, problemas de planificación, diagnóstico médico u otros, además del razonamiento por defecto y se destaca la programación lógica, de manera que la programación lógica abductiva es un área de cierta relevancia en este campo.

En general, un programa lógico viene dado por un conjunto de reglas o cláusulas, que se presentan de la forma $l \leftarrow l_1, l_2, ..., l_n$, cuya interpretación es que se da l, si $l_1, l_2, ..., l_n$, donde tanto l, la *cabeza*, como cada l_i, que constituyen el *cuerpo*, son literales, se expresan en un lenguaje proposicional o en uno de predicados de primer orden y se hace uso del método de resolución o alguna variante (como SLD, abreviatura de "Selective Linear resolution for Definite clauses"). Dado un programa, se plantean preguntas que no siempre pueden ser respondidas por el mismo; por ejemplo, dado el programa

$$[p \leftarrow q; p \leftarrow q, r; p \leftarrow r, s],$$

p no se puede deducir a menos que consideremos un cuerpo de las cláusulas (o más de uno en este caso). El proceso de abducción consiste en la búsqueda de los datos que permiten deducir p, y se trata de computar el conjunto de estas "explicaciones"; en este ejemplo se codificarán tres posibles abducibles para la observación p, a saber, $\{q\}$, $\{q, r\}$ y $\{r, s\}$. Kakas, Kowalski, Toni (1993: 6-7) señalan que las tareas abductivas en este contexto pueden ser verdaderamente intratables, aunque se requieran explicaciones mínimas, si bien son tratables para clases restringidas de programas lógicos. Así pues, se requiere un conjunto de abducibles y, para evitar una excesiva complejidad, estos abducibles han de verificar ciertas condiciones a las que se les suele denominar *restricciones de integridad* ("integrity constraints"), de manera que un *marco* abductivo se define como la terna (T, S, I), donde T representa la teoría, es decir, el programa, S es el conjunto de los abducibles, mientras que I expresa el conjunto de las restricciones de integridad.

Además de estas aplicaciones y las indicadas en Aliseda (1997), se presentan otras posibilidades. La lógica clásica elemental, empero, no es el único soporte para la construcción de modelos lógicos de abducción. El planteamiento de problemas abductivos en un contexto modal, por ejemplo, sugiere la modelación de la inferencia abductiva en términos de esta extensión de lógica clásica y lo mismo cabe decir de contextos epistémicos, doxásticos, temporales, etc. Cabe por ello esperar nuevas modelaciones tomando como punto de partida las correspondientes relaciones de consecuencia.

Son varias las líneas de investigación para proseguir el estudio de modelos lógicos de la abducción. Entre éstas se hallan, por ejemplo, el método de tablas semánticas con fórmulas con variables libres, la resolución dual con lenguajes de primer orden, la obtención de hipótesis más complejas (no solamente literales, sino también conjunciones, disyunciones, expresiones condicionales y cuantificaciones), los problemas de corrección y completitud, etc. Hay, pues, razones suficientes para continuar la investigación en este campo.

Referencias

ALISEDA, A. 1997, *Seeking Explanations: Abduction in Logic, Philosophy of Science and Artificial Intelligence*. Amsterdam: ILLC Dissertation Series.

ALISEDA, A. 1998, "La abducción como cambio epistémico: C. S. Peirce y las teorías epistémicas en inteligencia artificial". *Analogía* 12:125-144.

ALISEDA, A.; NEPOMUCENO, A. 2002, "Abduction in First Order Semantic Tableaux". Contributed Talks International Congress on Causation and Explanation in Natural and Social Sciences, Ghent (Abstracts).

BETH, E. W. 1969, "Semantic Entailment and Formal Derivability", en Hintikka, J. (ed.) *The Philosophy of Mathematics*. London: Oxford University Press: 9-41 (Versión en español "Entrañamiento semántico y derivabilidad formal" *Cuadernos Teorema* n. 18).

CORCORAN, J. 1998, "Information-theoretic logic", en Martínez, C., Rivas, C., Villegas, L. (eds.) *Truth in perspective*. Aldashgate, England and Brookfield, Vermont: Ashgate Publishing Co.: 113-135.

CORCORAN, J. 2004, "Abduction: An Information-Theoretic Approach (abstract)", Comunicación personal (pendiente de publicación).

HINTIKKA, J. 1999, "What is abduction? The Fundamental Problem of Contemporary Epistemology ". En J. Hintikka, *Inquiry as Inquiry: A logic of Scientific Discover*. Dordrecht: Kluwer Academic Publishers: 91-113.

KAKAS, A.. C.; KOWALSKI, R. A.; TONI, F. 1993, *The Role of Abduction in Logic Programming* (revision del trabajo "Abductive logic programming" en *Journal of Logic and Computation* 2(6): 719-770): 1-80.

LIPTON, P. 1991, *Inference to the Best Explanation*. London, Routledge.
MAYER, M. C.; PIRRI, F. 1993, "First order abduction via tableau and sequente calculi". En *Bulletin of the IGPL*, Max Planck Institute: 99-117.
NAGEL, E. 1978, *La estructura de la ciencia*. Trad. N. Míguez. Buenos Aires: Paidós.
NEPOMUCENO, A. 2003, *El método de las tablas semánticas*. Sevilla: Editorial Kronos.
NEPOMUCENO, A. 2004, "Un cálculo abductivo natural", en A. Vicente y otros (eds.) *Actas del IV Congreso de la Sociedad de Lógica, Metodología y Filosofía de la Ciencia en España*. Valladolid, pp. 382-384.
NIINILUOTO, I. 1999: "Defending abduction". *Philosophy of Science*. Proceeding of the 1998 Biennial Meeting of the Philosophy of Science Association. Supplement to volume 66, n. 3: S436-S451.
PEIRCE, C. S. 1958, *Collected Papers of Charles Sanders Peirce*. Vols. 1-6. Cambridge: Harvard University Press.
SOLER, F. 2004, "Cálculo de δ -resolución proposicional", en A. Vicente y otros (eds.) *Actas del IV Congreso de la Sociedad de Lógica, Metodología y Filosofía de la Ciencia en España*. Valladolid, pp. 392-395.
THAGARD, P. 1988, *Computational Philosophy of Science*. The MIT Press, Cambridge, MA.