RAZONAMIENTO NO MONOTONO: UN BREVE PANORAMA

Luis FARIÑAS DEL CERRO*
Antonio FRIAS DELGADO**

1. Introducción

Desde la obra de Aristóteles, se ha concebido a la lógica como un instrumento para el estudio del razonamiento correcto. Durante muchos siglos, sin embargo, el lenguaje y los recursos empleados para llevar a cabo un proyecto tan ambicioso han sido bastante limitados; y, si se compara el objetivo general y los resultados obtenidos, hay que aceptar que la lógica daba cuenta tan sólo de un pequeño fragmento del razonamiento humano.

La lógica contemporánea, la que arranca de Frege, ha ampliado llamativamente las posibilidades de la lógica; ha ampliado el lenguaje, las técnicas, los campos de aplicación. Durante el siglo XX, la extensión de la lógica tradicional clásica ha sido tan importante, que ésta ya sólo representa un mínimo fragmento de un cuerpo que sigue expandiéndose. En la producción científica actual, medida por los artículos que aparecen en las revistas más importantes del campo, el porcentaje dedicado a la temática más clásica es bastante pequeño, en parte también debido al cierre casi total de la misma.

Ciertamente que la evolución interna de la lógica la ha ido llevando a interesarse cada vez más por nuevos tipos de razonamiento. Pero en este proceso han intervenido también factores externos al núcleo central de la lógica propiamente dicha. Uno de ellos ha sido, sin duda, la Inteligencia Artificial (IA). La razón es obvia: en la medida en que el proyecto de la IA es el de construir artefactos que reproduzcan el comportamiento inteligente, se ha visto obligada a encarar procesos de razonamiento frecuentes en los seres humanos para los que la lógica clásica no tenía suficientes recursos. Aun a riesgo de simplificar, puede decirse que la lógica contemporánea estaba centrada en el razonamiento matemático como modelo; la IA ha llamado poderosamente la atención sobre la importancia del razonamiento ordinario, de 'sentido común'.

Y no es que la IA cuestione o niegue el tipo de razonamiento del que tradicionalmente se ha ocupado la lógica (digamos el razonamiento deductivo apodíctico). Desde luego que un sistema inteligente ha de ser capaz de extraer inferencias de acuerdo a la lógica clásica; si no lo hiciera, no podría ser considerado 'inteligente' en ningún sentido razonable del término. Lo que la IA plantea es que la inteligencia, cualquier cosa que ésta sea aunque no sepamos definirla con exactitud, consiste en más cosas que deducir de este modo; que hay capacidades importantes y

THEORIA - Segunda Época - Vol. X
1995, Nº 23, 7-26

esenciales en lo que llamamos 'inteligencia' que desbordan los limites de la lógica clásica. Y si queremos simular o reproducir comportamientos inteligentes, hemos de contar con mecanismos de tratamiento de la información que amplien la lógica clásica.

Quizás no haya una única cosa que sea 'el razonamiento ordinario'; seguramente es un conjunto heterogéneo de recursos no fácilmente explicitables y muy dependientes de cada contexto concreto en el que se efectúa el razonamiento y en el que la semántica y la pragmática pueden jugar un papel decisivo. Pero lo cierto es que hay regularidades y que nos apoyamos en ellas para inferir información y para comprender lo que los demás hacen o nos transmiten en sus enunciados.

Podríamos señalar algunos grandes rasgos del razonamiento ordinario, por contraposición a la lógica clásica: sus mecanismos son menos explícitos y más difícilmente explicitables; sus conclusiones tienen menos seguridad o son provisionales, dependen más de contextos, tanto semánticos como pragmáticos, que a veces no se pueden formular exhaustivamente; etc. Todo ello hace difícil un tratamiento completo, general y elegante del razonamiento de sentido común. No hay propuestas globales suficientemente satisfactorias; pero sí se han desarrollado en los ultimos decenios varios acercamientos parciales. Aquí nos vamos a ocupar fundamentalmente de uno de ellos: del que trata el tipo de razonamiento que se lleva a cabo en condiciones de información incompleta.

2. Razonamiento no monótono

Comencemos con un ejemplo, manido en la literatura de la no monotonía. Sabiendo que:

i) Los pájaros vuelan,
ii) Tweety es un pájaro,

nos sentimos racionalmente autorizados a extraer la conclusión:

iii) Tweety vuela.

Si lo único que sabemos de Tweety es que es un pájaro, puesto que sabemos que los pájaros vuelan, concluimos que Tweety vuela. Razonamientos de este tipo son muy frecuentes en la vida diaria. Quizás no de esta forma tan explícita, pero sí en formas implícitas; por ejemplo, cuando la información de (iii) explique u organice un conjunto de informaciones. Seguramante, si nos dicen que un sujeto, X, del que sabemos que es alumno de bachillerato, está casado, nos sorprenderá; desde luego que nos sorprende más que si nos dicen que no está casdo. ¿Por qué? Porque, aunque no hubiésemos llevado a cabo un razonamiento explícito sobre el estado civil de X, implícitamente habíamos inferido su soltería, sobre la base de:

iv) los estudiantes de bachillerato están solteros.

Nuestra sorpresa indica que habíamos inferido que X estaba soltero. Repárese en que, en los casos semejantes a los que acabamos de indicar, tanto (i) como (iv) siguen siendo aceptados aunque nos encontremos con excepciones. Lo que ocurrirá es que, en adelante, contaremos con que X está casado o que Tweety no vuela por ser un pingüino; lo incorporaremos como una de las excepciones de (iv) y (i), de modo que en el futuro nos impida usarla como premisa para obtener conclusiones sobre el estado civil de X o la posibilidad de volar de Tweety.

Aunque sea insistir sobre lo evidente, una característica de los "seres" inteligentes (ya se trate de sujetos humanos o de artefactos que quieran reproducir o simular el comportamiento humano) es su posibilidad de efectuar razonamientos y extraer conclusiones aun cuando la información disponible para hacerlo sea incompleta. Las conclusiones de estos razonamientos pueden etiquetarse de "generalmente verdaderas", "razonablemente ciertas" o con expresiones semejantes. La característica distintiva de estas conclusiones es que son provisionales -pueden ser "anulables", se suele decir- en la medida en que la incorporación de nueva información puede hacernos modificar conclusiones extraídas con anterioridad. Esto los diferencia significativamente de aquellos razonamientos en los que la conclusión es absolutamente necesaria, sentadas ciertas premisas.

Por acotar el tema de nuestro estudio, diremos que nos interesan:

a) las conclusiones que pueden extraerse de los enunciados genéricos; es decir, enunciados del tipo "los A son típicamente B", "generalmente los A son B", "si A, generalmente B", "los A son B, con excepciones";

b) cómo reajustar la información global en los casos en los que se produzca una falta de coherencia.

Desde un punto de vista lógico, las características más llamativas de estos enunciados son:

a) Respecto a las inferencias obtenibles a partir de ellos, funcionan como generalizaciones o condicionales normales. Si sabemos que los A son B y que un x concreto es A, nos sentimos autorizados a inferir que ese x concreto es B. Pero

b) Información nueva puede anular la conclusión anterior. Por ejemplo, si descubrimos que x es una de las excepciones, que x no es B, anulamos la conclusión anterior de que lo era. Esta característica llamativa es la que da nombre a este tipo de razonamientos: la pérdida de la monotonía. La monotonía o estabilidad deductiva, consiste en la propiedad siguiente: si A es deducible de un conjunto de premisas, D, A seguirá siendo deducible del conjunto de premisas formado por D y la adición de cualesquiera premisas adicionales. Esto significa que las deducciones monótonas son atemporales y acontextuales; que una vez establecidas no hay nada que nos obligue a revisarlas y a modificar las conclusiones. El razonamiento matemático, por ejemplo, es de esta clase.

c) El hecho de que se cuente simultáneamente con

i) los A son B
ii) x es A

iii) x no es B

no supone ninguna inconsistencia del sistema de conocimiento, como sí la supondría su equivalente clásico monótono.

e) Además de renunciar a la monotonía, hay que renunciar a la validez no restringida, por ejemplo, de la transitividad y la contraposición.

El razonamiento no monótono es frecuente en nuestra vida diaria. Quizás porque supone un cierto equilibrio entre, de una parte, el esfuerzo ingente, y en muchos casos imposible, de conocer todas las excepciones y, de otra parte, el desproporcionado vacío de información necesaria que se produciría de no poder utilizar enunciados generales con excepciones. Parece como si el riesgo de extraer una conclusión que podría luego resultar falsa fuera más aceptable para el sujeto que el vacío provocado por una extrema prudencia.

El razonamiento no monótono está presente, también, en diversos tratamientos de la IA. Por ejemplo, en los modelos de representación de la información que se denominan 'jerarquías de herencia', en ausencia de información específica en contra, suponemos que una clase hereda las propiedades de los niveles superiores de la jerarquía (Minsky 1975). O la asunción de 'mundo cerrado' (Reiter 1978)- en las bases de conocimientos, por la que nos sentimos autorizados a responder "no" a la pregunta por un hecho cuya información no esté explícitamente representada o no sea derivable de la base de conocimientos. O el 'problema de la cualificación' (McCarthy 1977) que tematiza el comportamiento tan frecuente en la vida diaria por el que no consideramos relevantes en una situación particular sino un pequeño número de las muchas precondiciones que facilitan o imposibilitan una acción.

Como puede observarse, lo característico del razonamiento no monótono no reside -o no reside sólo- en un conjunto de nuevas regla de inferencia -que, por decirlo así, extraigan más información-, sino en el modo peculiar en el que las inferencias tienen que supeditarse e interactuar en una contrastación efectiva con el resto de la información relevante disponible para poder ser aplicadas con éxito. Es decir, la aplicación de reglas de inferencia tiene que actuar de consuno con mecanismos que salven al sistema de la contradicción. Para ello hay que tener en cuenta, básicamente, dos tipos de operaciones: una exploración, previa a la aplicación de la regla, para asegurarnos de no producir una incosistencia; y un registro de nuestra conclusión para anularla en el futuro si la ampliación de la información la contradice.

3. Formalismos para el razonamiento no monótono

3.1. Las propiedades intuitivas que hemos considerado, queremos que sean reflejadas en algún formalismo que sea capaz de extraer, de una base de conocimiento, conclusiones basadas en inferencias no monótonas.

Aunque la lógica clásica pueda ser un instrumento general de representación del conocimiento, no puede ofrecer un tratamiento efectivo de la no monotonicidad (la fuerza de dicha imposibilidad reside precisamente en la palabra 'efectivo').

Podríamos creer que, puesto que al fin y al cabo, la no monotonía se debe a la existencia de excepciones a las reglas generales, es posible mantenerse en una lógica clásica suficientemente corregida de modo que se incorporen todas las excepciones. Sin embargo, este propósito no parece realista: algunas excepciones puede que no sean explícitas; o puede que sólo sepamos que hay excepciones, pero sin ser capaces de nombrarlas. Además, habríamos de incorporar todas las desigualdades para estar seguros de que un X concreto al que queremos aplicar una regla es diferente de cualesquiera de las excepciones de la misma.

Hemos empleado la expresión 'formalismo' y no la de 'lógica' por dejar abierta la cuestión de si todos los métodos empleados merecen o no el calificativo de 'lógica'. Seguramente la respuesta, en un sentido o en otro, irá bastante ligada a la filosofía de la lógica que se sostenga.

Así pues, entenderemos por 'formalismos no monótonos' los procedimientos técnicos para efectuar inferencias no monótonas. Los distintos métodos existentes pueden agruparse por similitudes de distinta forma. En Ginsberg (1987), Etherington (1988), Lea Sombé (1990), Lukaszewicz (1990) se presentan diferentes modos de agrupar estos tratamientos.

A la hora de valorar los distintos formalismos, Lea Sombé, por ejemplo, proponen una serie de criterios significativos que permiten un análisis funcional de los diferentes formalismos:

a) Por lo que se refiere a la representación del conocimiento: ¿hay una única o varias maneras posibles de representar en un formalismo una información dada?

b) ¿Es posible contestar a cuestiones generales o sólo las referidas a un individuo concreto?

c) ¿Permite aplicar ciertos principios clásicos -como la transitividad, la contraposición, el silogismo disyuntivo, por ejemplo?

d) Caso de que se puedan obtener conclusiones en conflicto, ¿se aceptan todas -de modo que la elección entre una de ellas se considera una tarea extralógica- o hay criterios de preferencia?

e) ¿Es posible tener en cuenta información nueva sin modificar la representación existente? (modularidad)

f) ¿Es el procedimiento efectivamente calculable?

Si los comparamos respecto a estos criterios, no todos los formalismos son igualmente potentes; unos funcionan mejor en unos casos y otros en otros. En el último decenio hemos asistido a una proliferación de técnicas de las que es imposible dar cuenta con exhaustividad en el marco de un artículo. De todos modos, detrás de muchas de ellas está a veces la misma intuición. Nos limitaremos a presentar algunos de los formalismos que han jugado un papel importante en el tratamiento y la comprensión de la no monotonía -no aludiremos aquí a las lógicas de condicionales, ya que a ellas se dedicará el siguiente articulo. Confiamos en que los aquí expuestos ayuden al lector a introducirse en este campo.

Vaya de entrada, también, el reconocimiento de que los formalismos empleados para dar cuenta del razonamiento no monótono no han carecido de críticas. Reseñemos éstas:

a) Una primera objeción a la totalidad es la de que la lógica tradicionalmente es, por definición, monótona. Una lógica no monótona no pasa de ser una contradicción en los términos.

b) Los formalismos de inferencia no monótonos no son, en general, semidecidibles. Esto acarreará problemas por lo que se refiere a las implementaciones de dichos formalismos.

c) Los críticos a la posición más logicista en IA mantienen que en la no monotonía no hay, en propiedad, un problema de lógica, sino de representación del conocimiento.

d) Los partidarios de la explicación por teoría de probabilidades insisten en la no necesidad de la lógica no monótona. Según ellos, la teoría de probabilidades explica suficientemente el razonamiento no monótono sin necesidad de métodos no numéricos.

Sin embargo, a pesar de las dificultades y de las críticas puntuales como las que acabamos de señalar, parece existir consenso entre los investigadores en torno a algunos puntos:

i) Se necesita alguna extensión no monótona de la lógica de primer orden;

ii) los problemas de la inferencia no monótona precisan de una solución declarativa;

iii) El tratamiento del razonamiento no monótono debiera ser simbólico y no numérico.

Este punto (iii) y la crítica (d) anterior merecen algunas aclaraciones.

Si consideramos el problema de la no monotonía desde el aspecto digamos más psicológico, de los mecanismos que pone en marcha el sujeto cuando razona no monótonamente, es claro que no se trata de un problema de probabilidades. Si sabemos que el 98 por ciento, pongamos por caso, de los estudiantes de bachillerato están solteros, en ausencia de otra información en contra, no concluimos que X tiene el 98 por ciento de probabilidades de ser soltero, sino que X está soltero, sin más. En la vida diaria es muy difícil manejar las probabilidades. Es mucho más fácil organizar un conjunto de información amplio con el dato simple de que X está soltero que con una cuantificación numérica que ha de propagarse de un modo complicado por el conjunto.

Esto no significa, sin embargo, que determinados métodos numéricos no puedan constituir un buen modelo para explicar la no monotonía, con independencia de que sea realmente así o no como los sujetos se las manejan con información incompleta. Aunque generalmente los modelos de la no monotonía son simbólicos, ciertos modelos ordinales de la incertidumbre, como la teoría de la posibilidad o los modelos de probabilidades infinitesimales de Adams, permiten definir una noción de deducción no monótona de modo que es posible extraer informaciones naturales en un contexto de información incompleta.

3.2. Lógicas por defecto

La intuición básica de las lógicas con reglas para defectos es simplemente ésta: podemos extraer conclusiones en ausencia de información explícita en sentido contrario. El modo técnico de hacer operativa esta idea es ya más complicado.

Este tipo de lógica fue introducido por Reiter (1980). Con más propiedad, una teoría con reglas para defectos, T=(D,H), consiste en un conjunto de hechos, H, que son fórmulas cerradas de primer orden y un conjunto de defectos, D, que son reglas específicas de inferencia y que tienen la forma siguiente:

$$\frac{u(x) : v(x)}{w(x)}$$

donde u(x), v(x) y w(x) contienen la variable libre x. La lectura de una regla de este tipo sería: si un individuo determinado, a, tiene la propiedad u y si el que a tenga la propiedad v es consistente con lo que puede deducirse en la teoría, entonces podemos concluir que el individuo a tiene la propiedad w.

Diremos que u(x) es el prerequisito de la regla (lo que ha de contar ya como hecho para poder aplicarla); v(x) es la justificación de la regla (lo que permite o imposibilita su aplicación en un caso concreto) y w(x) es la conclusión de la regla. Su significado intuitivo puede ser: si se conoce u(x) y v(x) es consistente con lo que se conoce, concluir w(x).

Por llevarlo a un plano intuitivo, una regla para defectos se podría entender como una metaregla constituida por una regla normal, cuya premisa es u(x) y cuya conclusión es w(x), más una condición, v(x), que en realidad 'codifica' un complejo conjunto de instrucciones que son las que determinan la aplicabilidad o no de la regla. La complejidad técnica de los formalismos no monótonos reside, en buena medida, en que lo que hay que codificar es algún mecanismo de comprobación de consistencia.

Un defecto se llama normal si v(x)=w(x); es decir, si la justificación y la conclusión coinciden.

Por ejemplo, el enunciado "los estudiantes están solteros" podría representarse así:

$$\frac{E(x) : S(x)}{S(x)}$$

donde E simboliza la propiedad 'ser estudiante' y S simboliza la propiedad 'ser soltero'.

Aplicada a un individuo concreto, X, la regla dice: si 'X es un estudiante' cuenta como un hecho y 'X es soltero' es consistente con el conjunto de los hechos, entonces puede inferirse para que cuente como un nuevo hecho 'X es soltero'.

Se observará inmediatamente una dificultad: tanto el hecho de que X sea estudiante como la consistencia de que X sea soltero no pueden depender única y exclusivamente de los enunciados incluidos explícitamente en el conjunto primitivo H. ¿Por qué? Porque hemos de contar también con los hechos deducibles clásicamente a partir de H o los que puedan obtenerse por aplicación de reglas por defecto. Veámoslo con un ejemplo.

Supongamos que H consta de los dos enunciados siguientes:

i) A o B
ii) No A

Es inmediatamente evidente que, aunque no aparezca como hecho el enunciado

iii) B,

hay que contar con (iii) para usarlo como prerequisito o justificación de posibles reglas.

Para saber si algo es o no deducible no monótonamente de una teoría T necesitamos, pues,

a) expandir nuestra teoría inicial T de modo que el resultado final incluya
a') la base de hechos inicial, H,
a'') las fórmulas deducibles clásicamente, y
a''') las que se obtienen por la aplicación de las reglas por defecto;
b) que el conjunto resultante sea consistente. Esto será especialmente importante con vistas a decidir la aplicación o no de una regla para defectos;
c) comprobar si la fórmula en cuestión pertenece al conjunto así expandido.

Más formalmente: dada una teoría T, una extensión E es un conjunto de fórmulas que tiene las propiedades:

a) Ninguna regla por defecto se puede aplicar consistentemente para obtener una conclusión que no esté ya en la extensión. (Si el prerequisito de una regla, d, está en E y la negación de la justificación no está en E, la conclusión está en E).

b) La extensión es la mínima que cumple la condición anterior.

Es decir:

E es una extensión de T si y solo si $E = \bigcup E_i$, donde

$$E_0 = H$$
$$E_{i+1} = Th(E_i) \cup \{C \mid \frac{A:B}{C} \in D,\ A \in E_i,\ \neg B \notin E\}$$
$$\text{donde } Th(E_i) = \{\phi \mid E_i \vdash \phi\}$$

La extensión E se va definiendo por niveles. En el primer nivel tenemos las fórmulas que simbolizan los hechos con los que contamos en la teoría. Dado un cierto nivel, el inmediatamente siguiente está formado por todas las fórmulas deducibles en los niveles inferiores y las que se puedan obtener por aplicación a las mismas de las reglas para defectos. Se observará que en la definición del nivel i+1 no se exige que la negación de la justificación de la regla a aplicar no pertenezca al nivel i, sino simplemente que no pertenezca a E. Es decir: que si quisiéramos construir efectivamente un nivel, tendríamos que depender de todos ellos, incluso de los aún no construidos. Esta exigencia es crucial a la hora de entender ciertos recursos técnicos.

Una fórmula f es una respuesta a una pregunta a la base de conocimiento representada por T (es decir: ¿es f deducible no monótonamente en T?) si existe al menos una extensión E para T tal que $f \in E$.

La definición de extensión anterior no es en general constructiva, pero puede ser adaptada para que lo sea en el caso de teorías con defectos normales. Las teorías con reglas para defectos pueden tener varias extensiones o ninguna. La existencia de al menos una extensión está asegurada para el caso en que todos los defectos sean normales. Si H es consistente y T tiene una extensión, dicha extensión es consistente.

Veamos algun ejemplo. Supongamos una teoría cuya base de hechos es
H' = Juan es estudiante, los solteros no están casados. (En símbolos: Ea, $\bigwedge x(Sx \rightarrow \neg Cx)$ y cuyas reglas para defectos son: los estudiantes son jovenes, los jovenes están solteros).

$$\frac{E(x) : J(x)}{J(x)} \qquad \frac{J(x) : S(x)}{S(x)}$$

Nuestra teoría tiene una extensión que contiene: Ea -por estar en H'-, J(a) -por aplicación de la primera regla para defectos ya que tenemos el prerequisito, E(a), y J(a) es consistente con las fórmulas que tenemos-, S(a) -por aplicación de la segunda regla para defectos ya que tenemos el prerequisito, J(a), y S(a) es consistente con las fórmulas que tenemos-, $\neg$C(a) -de forma clásica a partir de S(a) y H'. En esta teoría podemos obtener las conclusiones de que Juan es joven y de que Juan está soltero.

Se observará que la transitividad funciona para defectos ejemplificados, aunque no puede construirse un defecto general del tipo

$$\frac{E(x) : S(x)}{S(x)}$$

que exprese que los estudiantes son solteros. La contraposición tampoco es posible, a no ser que se añada como defecto.

Supóngase ahora que añadimos a la base de hechos H' la información de que Juan está casado (en símbolos: Ca). La extensión de la nueva teoría contará con Ea y Ja, por las mismas razones que antes. A partir de la base de hechos puede obtenerse $\neg$S(a) -pues ahora tenemos que Ca. Pero esto nos impide aplicar la segunda regla para defectos, ya que ahora S(a) no es consistente con el conjunto de fórmulas existentes.

Construir una extensión de una teoría no siempre es tan fácil como en los ejemplos anteriores. Schwind (1990) presenta un metodo basado en las tablas semanticas de Smullyan, que es de los más intuitivos y fáciles que existen. La idea básica es ésta:

a) Construir una tabla de las fórmulas de la base de hechos;
b) añadir a cada rama todas las conclusiones de todas las reglas para defectos;

c) revisar cada rama para eliminar:

c') las conclusiones cuyo prerequisito no esté en la rama;

c") las conclusiones de las reglas para las cuales la negación de la justificación esté en la rama.

Las fórmulas existentes en cada rama nos dan las extensiones posibles de la teoría.

Puede ocurrir que haya reglas para defectos que entren en conflicto aplicadas a un caso concreto. Veamos un ejemplo. Consideremos una teoría cuya base de hechos sea:

i) Juan es socialista;

ii) Juan es cristiano.

Las reglas para defectos de la teoría son:

a) Los socialistas son partidarios del aborto;

b) los cristianos no son partidarios del aborto.

¿Es Juan partidario del aborto o no?

En casos como estos de conflicto entre defectos, podríamos optar por varias vías de solución.

a) Hacer explícitas las prioridades. Por ejemplo, aplíquese la regla (i) antes que la regla (j).

b) Introducir un orden entre los defectos -Poole (1985). Si existe una cadena de inferencias del prerequisito (i) al (j) pero no al revés, (i) prevalece sobre (j).

c) Introducir defectos seminormales -Reiter y Criscuolo (1981), Etherington y Reiter (1983)-: Un defecto seminormal tiene la forma

$$\frac{u(x) : v(x) \quad w(x)}{w(x)}$$

La ventaja de un defecto seminormal reside en que podemos utilizar v(x) para forzar una preferencia de ciertos defectos sobre otros, ya que la regla no se aplicará si no son consistentes tanto w(x) como v(x).

El inconveniente principal reside en que no está asegurada la extensión de teorías de defectos seminormales, si bien Etherington ha demostrado que las teorías seminormales finitas ordenadas admiten al menos una extensión.

A modo de conclusión de este apartado, podemos señalar algunas de las características de las teorías con reglas para defectos:

a) La calculabilidad no es simple, incluso aunque nos limitemos a defectos normales. Existen demostradores de teoremas que funcionan en ciertos casos -Besnard, Quiniou y Quinton (1986) para teorías normales sin prerequisitos; Schwind (1990) para teorías normales.

b) No permiten obtener reglas generales -los estudiantes están solteros, por ejemplo-; sólo nos permiten obtener las propiedades de un individuo particular -el estudiante Juan está soltero.

c) No son satisfactorias para la disyunción.

d) La respuesta a las preguntas puede variar según la representación elegida y no existe un procedimiento sistemático que nos permita encontar la representación adecuada.

e) No es posible la contraposición a menos que se añada explícitamente.

3.3. Circunscripción

En las teorías con reglas para defectos podemos establecer si una fórmula dada es una conclusión no monótona de una base de hechos. Pero los procedimientos no son los que esperaría alguien que está familiarizado, básicamente, con los métodos deductivos de la lógica de primer orden. La circunscripción, de alguna manera, se mantiene más próxima a las técnicas clásicas.

La idea intuitiva que subyace al tratamiento de la circunscripción es, precisamente, la de minimizar las excepciones. Por ejemplo, el enunciado 'los jovenes están solteros' se representaría como: un joven, excepto que no sea un caso normal, está soltero.

La representación de esta idea no es difícil. Usando un predicado para la anormalidad, Abn, podríamos escribir:

i) $\bigwedge x(Jx \wedge \neg Abn\ x \rightarrow Sx)$.

La mayor dificultad en este formalismo, introducido por McCarthy (1980, 1986), está en decidir qué es anormal. La idea es que hay que minimizar el numero de objetos anormales, las excepciones, aceptando sólo aquello que se sepa que es anormal. Semánticamente, esto quiere decir que no vamos a considerar todos los modelos de la teoría, sino sólo los modelos mínimos respecto al orden natural de inclusión conjuntista. Sintácticamente, este efecto se consigue añadiendo a la teoría un axioma de circunscripción.

Circunscribir un predicado, en nuestro caso (i) el predicado Abn, quiere decir que no se tendrán en cuenta con vistas a las inferencias más que los modelos en los que los individuos que lo satisfagan sean los mínimos posibles. Nos interesa también que esta circunscripción tenga influencia en los valores que pueda tener S, ya que de lo que se trata es de minimizar la anormalidad de los jovenes que no son solteros.

¿Cómo conseguir por medios sintácticos este efecto? Lo que queremos es poder obtener deductivamente una fórmula a partir de (i) en la que el predicado Abn no aparezca y sea sustituido por un predicado cuya extensión sea la mínima clase de individuos que constituyen las excepciones. Para lograrlo es para lo que se añade un esquema axiomático de circunscripción. Este esquema, bastante farragoso a primera vista, expresa: si sustituimos en la teoría T el predicado de anormalidad Abn por una fórmula cualquiera, f, obteniendo el conjunto de fórmulas T', de modo que resulten derivables a partir de T, tanto T' como el hecho de que la extensión de f está incluida en la de Abn, entonces f y Abn son equivalentes. De otro modo, si se puede demostrar que un conjunto de individuos que satisface T es un subconjunto de las anormalidades, lo consideramos como las únicas anormalidades.

Expresemos estas ideas más formalmente. Supongamos que A es una sentencia de primer orden que contiene el sómbolo de predicado $P(x_1,...,x_n)$ -en notación, $P(\bar{x})$. $A(\phi)$ es el resultado de reemplazar todas las ocurrencias de P en A por la expresión de predicado ϕ.

La circunscripción de P en A(P) es el esquema:

ii) $A(\phi) \wedge \bigwedge \bar{x}(\phi(\bar{x}) \rightarrow P\bar{x} \rightarrow \bigwedge \bar{x}(P\bar{x} \rightarrow \phi(\bar{x}))$

Obsérvese que (ii) puede expresarse en lógica de segundo orden mediante una generalización.

En (ii) $A(\phi)$ expresa la asunción de que ϕ satisface las condiciones satisfechas por P.

$\bigwedge \bar{x}(\phi(\bar{x}) \rightarrow P\bar{x})$ expresa la asunción de que los individuos que satisfacen ϕ son un subconjunto de los que satisfacen P. La conclusión asegura que ϕ y P coinciden.

Expresado de un modo más intuitivo e informal, lo que se está haciendo es lo siguiente. Supongamos que tenemos un conjunto de fórmulas en las que aparece el predicado Abn, sea F(Abn); minimizar la anormalidad equivale a asumir que Abn es el predicado P de menor extensión para el que vale F(P) -puesto que hablar de predicados es hablar de clases de individuos, minimizar la anormalidad es obligar a que los individuos anormales sean los que vienen dados por la clase menor de las que pueden satisfacer a F.

Ahora no necesitamos extender una teoría para comprobar si determinada fórmula pertenece o no a alguna de sus extensiones. Una fórmula, f, es deducible por circunscripción si es deducible del conjunto de fórmulas inicial, T, que son fórmulas cerradas de primer orden, al que añadimos todas las instancias del esquema de circunscripción (ii) para los predicados convenientes.

Veamos un ejemplo. Supongamos que T contiene la fórmula

a) $\bigwedge x(Ex \wedge \neg Abn(x) \rightarrow J(x))$

La circunscripción en T del predicado 'anormal' haciendo variar el predicado 'joven', permite obtener como conclusión: 'todos los estudiantes son jovenes', $\bigwedge x(Ex \rightarrow Jx)$.

En este caso se trataría de elegir los esquemas de circunscripción de tal modo que la anormalidad se reduzca a cero, nadie es anormal, y que todo el mundo sea joven. A modo de ejemplo, veamos cómo podría operarse técnicamente. Sea F una fórmula falsa cualquiera; escogemos F para sustituir a Abn(x) en (a) obteniendo

b) $\bigwedge x(Ex \wedge \neg F \rightarrow Jx)$

Obviamente, (b) es derivable a partir de (a). Aplicando el esquema de circunscripción -es fácil ver cómo la segunda parte del antecedente es también derivable-, obtenemos:

Obviamente, (b) es derivable a partir de (a). Aplicando el esquema de circunscripción -es fácil ver cómo la segunda parte del antecedente es también derivable-, obtenemos:

c) $\bigwedge x(Abn(x) \rightarrow F)$

d) $\bigwedge x(\neg Abn(x))$

A partir de (a) y de (d) se concluye fácilmente $\bigwedge x(Ex \rightarrow Jx)$

Si añadimos a (a)

e) Juan es estudiante: Ea

f) Pedro no es joven: $\neg Jb$

Podemos deducir que todos los estudiantes distintos de Pedro son jovenes:

$\bigwedge x(Ex \wedge x \neq b \rightarrow Jx)$

Para ello elegimos los esquemas de circunscripción de tal modo que todos sean jovenes excepto Pedro y que sería anormal que Pedro fuera estudiante.

En la circunscripción no podríamos deducir, a partir del conjunto de fórmulas considerado, que Juan es joven; sólo la conclusión más débil: si Juan y Pedro son individuos distintos, Juan es joven -$a \neq b \rightarrow Ja$. Ello se debe a que no se pueden deducir igualdades o desigualdades que no sean deducibles del conjunto inicial de fórmulas.

Uno de los atractivos de la circunscripción es que tiene una semántica relativamente clara. Podríamos hacernos una idea semántica de la circunscripción utilizando los modelos preferenciales (Shoham 1988). Para definir la circunscripción se introduce una relación de oreden parcial entre las interpretaciones. Resulta que si una fórmula, f, es deducible de un conjunto T más los esquemas de circunscripción, entonces f es satisfecha por todo modelo minimal de T.

Puede ocurrir que si añadimos un esquema de circunscripción a la teoría, el resultado sea inconsistente. Esto equivale, semánticamente, al hecho de que una teoría puede no tener ningún modelo minimal.

El hecho de que no haya que calcular extensiones de las teorías ha contribuido a que la circunscripción sea bastante estudiada en el campo de los tratamientos de la no monotonía. Como contrapartida, el emplear un esquema axiomático, o una fórmula de segundo orden, la hace poco manejable computacionalmente. Hay pocos métodos para ello, todos los cuales son parciales y de una gran complejidad algorítmica.

3.4. Formalismos basados en la lógica modal

La circunscripción y los defectos son técnicas bastante diferentes y la relación entre ambas es superficial. No ocurre lo mismo con los formalismos basados en lógicas modales, seguramente uno de los métodos más estudiados para dar cuenta de la no monotonía. En las teorías con reglas para defectos, una regla como

$$\frac{E(x) : J(x)}{J(x)}$$

no permite ser formalizada en el propio lenguaje. Bien mirado, el único obstáculo para ello reside en la justificación. Hemos de asumir que J(x) es consistente. En lógica modal, el operador $\Diamond$ intuitivamente expresa la consistencia. Podríamos convertir la anterior regla por defecto en la fórmula

a) $Ex \wedge \Diamond Jx \rightarrow Jx$.

Este es el tratamiento que introducen McDermott y Doyle (1980). La idea sigue siendo la misma: si no es deducible $\neg$Ja -por definición: si Ja es consistente- entonces es deducible Ja. La manera más inmediata de construir una lógica no monótona sería poder contar en este caso con una regla del tipo

(i) $\nvdash_A \neg p \Rightarrow \vdash_A \Diamond p$

que juntamente con (a) nos permitiera efectuar inferencias no monótonas en la práctica. Sin embargo, esta regla es impracticable por ser circular. Nuevamente necesitamos aquí algún equivalente de las extensiones de las teorías con reglas para defectos.

De un modo paralelo a lo que allí ocurría, podemos extender una teoría T por la adición como posibilidades de las fórmulas consistentes, para obtener un conjunto S tal que

$S = Th(T \cup \{\Diamond p \mid p \notin S)$.

Diremos que S constituye un punto fijo de T.

El recurso técnico de McDermott y Doyle consiste en introducir un operador no monótono y caracterizar la inferencia no monótona de una fórmula por su pertenencia a la intersección de todos los puntos fijos de una teoría T. Esta restricción se debe a que T puede que no tenga puntos fijos -por ejemplo: $\Diamond p \rightarrow \neg p$, carece de punto fijo ya que al no ser deducible p, tendríamos $\Diamond p$ y $\neg p$- o que no sean mínimos. Si no los hubiera, por convención, devienen demostrables no monótonamente todas las fórmulas.

En cierto sentido, la lógica epistémica de McDermott y Doyle reproduce los inconvenientes de las teorías con defectos, desde un punto de vista deductivo, ya que la mera formalización de las reglas para defectos no basta para caracterizar de un modo operativo la relación de inferencia no monótona. En general hay que esperar que dondequiera que sea necesaria, de alguna manera, la regla (i) no se puede caracterizar de un modo efectivo una relación de consecuencia como ocurre en la forma normal de definir una lógica.

En vez de recurrir a la posibilidad para expresar la consistencia, podríamos recurrir a la necesidad ($\Box$, L) para expresar lo que sabemos. En este caso estamos razonando sobre lo que sabemos o creemos; de ahí el nombre de 'autoepistémica' a este tipo de lógica introducida por Moore (1985).

En el planteamiento anterior, interpretábamos que si alguien es un estudiante y es consistente que sea joven, es joven:

a) $\bigwedge x(Ex \wedge \Diamond Jx \rightarrow Jx)$

En la lógica autoepistémica se interpreta que si alguien es un estudiante y no es joven, lo sabríamos:

b) $\bigwedge x(Ex \wedge \neg Jx \rightarrow \Box \neg Jx)$

La inferencia no monótona de una fórmula f en la lógica autoepistémica es equiparable a la cuestión: dado un conjunto de fórmulas T que expresan aquello que yo sé o creo, ¿debo creer en f? Ya Stalnaker había señalado que el conjunto T de sentencias que represente las creencias de un agente idealmente racional habría de satisfacer las condiciones siguientes:

i) Si $\phi \in T$ y $\phi \vdash \chi \Rightarrow \chi \in T$
ii) $\phi \in T \Rightarrow \Box\, \phi \in T$
iii) $\phi \notin T \Rightarrow \neg\Box\, \phi \in T$

Del estado de esta teoría se dice que es estable.

La lógica autoepistémica tiene la ventaja de poseer una semántica clara y precisa, existiendo numerosos y potentes resultados, ya desde Moore (1985).

Sin embargo, vemos aquí nuevamente el inconveniente planteado por la condición (iii) para caracterizar una relación de consecuencia. Hemos de recurrir a expansiones estables, S, -es decir: semánticamente completas y correctas respecto a T- de una teoría T; dichas expansiones son aquellas que cumplen la condición:

$$S = Th(T \cup \{\Box p \mid p \in S\} \cup \{\neg\Box p \mid p \notin S\})$$

Para el fragmento proposicional de la lógica autoepistémica existen procedimientos de decisión (Niemelä 1988). El procedimiento consiste en construir las expansiones estables de un conjunto de premisas T usando las tablas analíticas.y comprobar si una fórmula dada está contenida en alguna de ellas. A muy grandes rasgos, los pasos a efectuar son:

i) Construimos la tabla correspondiente a $(T \rightarrow f)$, donde f es la fórmula cuya deducibilidad queremos establecer. Para ello suponemos su falsedad, cerrando las ramas en las que aparezca una contradicción.

ii) Para cada fórmula del tipo Lp ($\Box$p) que aparezca en una rama abierta, construimos la tabla correspondiente a la fórmula $(T \rightarrow p)$.

iii) Repetimos el paso (ii) en tanto que haya fórmulas necesarias en ramas abiertas.

iv) Etiquetamos cada tabla como abierta o cerrada.

v) Añadimos a todas las ramas de todas las tablas Lp si la tabla correspondiente a $(T \rightarrow p)$ está etiquetada de 'cerrada' o $\neg$Lp si lo está como 'abierta'.

vi) Si después de (v) todas las tablas que habían sido etiquetadas de 'abiertas' tienen alguna rama abierta y las que habían sido etiquetadas de 'cerradas' tienen todas las ramas cerradas, diremos que la etiquetación es aceptable.

vii) Repitiendo (v) y (vi) obtendremos todas las etiquetaciones aceptables.

viii) f se deduce no monótonamente de T si hay alguna etiquetación aceptable en la que la tabla correspondiente a $(T \rightarrow f)$ está etiquetada como 'cerrada'.

Las lógicas autoepistémicas y las de defectos están fuertemente relacionadas. Si asociamos a una regla para defectos la fórmula de la lógica autoepistémica

$$\bigwedge x(\Box u(x) \wedge \Diamond v(x) \rightarrow w(x))$$

resulta que toda extensión de una teoría por defecto está contenida en una expansión de la lógica autoepistémica asociada (Konolige 1988).

Existen numerosas propuestas de formalismos basados en la lógica modal que suponen variantes o refinamientos de las ya señaladas.

Halpern y Moses (1985) restringen su estudio a teorías que llaman 'razonables', en el sentido de que cumplen una propiedad de la disyunción restringida:

$$\vdash \Box F_1 \vee \ldots \vee \Box F_n \Rightarrow \vdash \Box F_i$$

donde F_i no tiene operadores modales y la deducibilidad se refiere a S5.

Su finalidad es la de poder obtener extensiones en ciertos casos naturales en los cuales la lógica autoepistémica no tiene extensión.

Levesque (1990) introduce una lógica modal con bastantes refinamientos para capturar todo lo que yo sé sobre algo. Su lógica contiene dos operadores modales, B y N; su axiomatización es el resultado de anadir a K45 para B y N, el axioma:

$$N\phi \rightarrow \neg B\phi$$

para sentencias no modales que sean falsables.

Este axioma, por contraposición, puede entenderse así: si todas las interpretaciones compatibles con mis creencias hacen verdadera a ϕ, alguna de las que no son compatibles la hacen falsa.

La semantica de esta lógica, intuitivamente, es ésta: $B\alpha$ es valida si en todos los mundos accesibles α lo es; $N\alpha$ es valida, si en todos los mundos que no son accesibles α lo es.

Definamos ahora un nuevo operador

$$O\alpha = B\alpha \wedge N\neg\alpha$$

para expresar intuitivamente todo lo que yo sé de α. Obsérvese que, en realidad, lo que estamos haciendo es decir: en todos los mundos compatibles con mis creencias α es verdadera y α es falsa en los otros mundos.

Supongamos que nuestra base de conocimiento, BC, consta del enunciado de que Juan es estudiante, Ea. Representemos que los estudiantes son jovenes así:

g: $\bigwedge x(Ex \wedge \neg B\neg Jx \rightarrow Jx)$

La lógica de Levesque permite deducir que yo sé que Juan es joven -BJa- a partir de la premisa de que todo lo que yo sé son los hechos de la base de conocimiento y g -O(BC $\wedge$ g). Esta lógica es monótona en la deducción, porque la no monotonicidad se ha puesto en el operador O, en todo lo que yo sé.

Meyer y Hoek (1990) presentan un tratamiento monótono del razonamiento no monótono. Si separamos los dos componentes fundamentales del razonamiento no monótono (el hecho de que sobre ciertas bases prefiramos creer algunos enunciados como si fueran verdaderos y el hecho de que cambiamos estas preferencias ante nuevos descubrimientos), el primero es de naturaleza monótona y lo tratan mediante una lógica de preferencias (que añade a los operadores modales clásicos para los que valen los axiomas de S5, un operador de preferencia para el que valen los axiomas de K45, más axiomas propios que relacionan la necesidad y la preferencia). El segundo es tratado como un cambio en los modelos preferidos, como una reasignación de valores de verdad, usando técnicamente una versión de la lógica dinámica.

De una forma general e intuitiva, podríamos decir en terminos autoepistémicos que, así como es fácil dar cuenta de lo que yo sé -pues, en efecto, se trata de usar la regla conocida como introducción de la necesidad:

$\vdash p \Rightarrow \vdash \Box p$

y que no plantea ningun problema desde el punto de vista operativo-, es mucho más difícil dar cuenta de lo que yo ignoro -que se correspondería con la regla

$\nvdash p \Rightarrow \vdash \neg \Box p$

De un modo u otro este es el escollo a salvar.

4. Propiedades de una relación de consecuencia no monótona

Otra forma de atacar el problema de la no monotonía, inaugurado por Gabbay, consiste en definir los principios generales que una noción de deducción debe verificar. Por ejemplo,la relación de consecuencia clásica cumple tres propiedades fundamentales:

i) Reflexividad: $\phi \vdash \phi$

ii) Monotonía: $\Gamma \vdash \phi \Rightarrow \Gamma,\Delta \vdash \phi$

iii) Corte: $\Gamma \vdash \Delta$ y $\Psi,\Delta \vdash B \Rightarrow \Gamma,\Psi \vdash B$

Desde un punto de vista abstracto, podemos considerar las propiedades que debiera cumplir una relación de consecuencia no monótona. Este objetivo es diferente del que hemos analizado en el apartado anterior. En efecto, no se trata aquí de dar un conjunto de reglas que permitan decidir si Juan está o no soltero sabiendo que es joven; sino de exponer aquellas propiedades que esperaríamos que una relación de consecuencia no monótona poseyera. Muy sumariamente, vamos a indicar algunas.

Es evidente, por lo pronto, que las propiedades de monotonía y corte no deben cumplirse. Hemos de introducir formas restringidas de monotonía y transitividad, como

Monotonía racional: $A \not\mid\!\sim B$ y $A \mid\!\sim C \Rightarrow A \wedge B \mid\!\sim C$
Monotonía débil: $\not\mid\!\sim \neg A$ y $\mid\!\sim A \rightarrow B \Rightarrow A \mid\!\sim B$
Acumulación: $A \mid\!\sim B$ y $B \vdash A \Rightarrow (A \mid\!\sim C \Leftrightarrow B \mid\!\sim C)$

La interpretación de los conectores vendrá dada también en nuestro caso por las reglas de introducción de dichos conectores; como por ejemplo:

Y: $A \mid\!\sim B$ y $A \mid\!\sim C \Rightarrow A \mid\!\sim B \wedge C$
O: $A \mid\!\sim C$ y $B \mid\!\sim C \Rightarrow A \vee B \mid\!\sim C$
Condicionalización débil: $A \mid\!\sim B \Rightarrow \mid\!\sim A \rightarrow B$

Hemos de insistir en que estas reglas, y otras que aparecen en la literatura sobre el tema, son principios que una noción de deducción no monótona debe verificar, pero no nos dicen nada sobre lo que debemos hacer en situaciones concretas de deducción, objetivo que era el fundamental en los formalismos presentados en el apartado anterior. En ese sentido y por la forma misma de las reglas, este proyecto es similar al del estudio de los condicionales, que trataremos en el artículo siguiente.

5. Conclusión

a) El esfuerzo desarrollado en la construcción de formalismos no monótonos parece tener su origen en el desafío planteado por Minsky a los logicistas dentro dela IA: la lógica no puede resolver este tipo de problemas. No puede decirse que dichos formalismos sean lógica clásica, pero sí que se construían dentro del espíritu de la lógica.

Al final de nuestro breve recorrido sobre el razonamiento no monótono, quizás sea la ocasión para retomar la cuestión de hasta qué punto estamos aún en la lógica o nos hemos salido simplemente de ella. Este problema puede verse desde distintos ángulos.

El estudio de las propiedades de una relación de consecuencia no monótona (nuestro apartado VI) debe ser considerado sin duda como un tema lógico. Lo que ocurre es que no basta caracterizar abstractamente las propiedades de una relación de consecuencia para contar con un instrumento mediante el que decidir si Juan, que es socialista y cristiano, es partidario del aborto, por ejemplo. Pero lo mismo ocurre en lógica clásica: no basta con la reflexividad, la monotonía y el corte para concluir la mortalidad de Socrates. La objeción más fuerte a la inclusión dentro de la lógica de los formalismos no monótonos puede ser la de que no caracterizan una relación de consecuencia. Nuestra posición puede resumirse informalmente así: si el mejor espíritu de la lógica fue siempre el de dar cuenta del razonamiento, y si la no monotonía es un tipo importante de razonamiento, los instrumentos para dar cuenta del razonamiento no monótono que compartan rasgos comunes con la lógica clásica debieran ser incorporados a la tradición de la lógica.

b) El tratamiento que podríamos decir clásico de la no monotonía es el de considerarla como una lógica clásica con excepciones; una lógica clásica más un mecanismo externo de ajuste de la consistencia. Seguramente, no es éste el único punto de vista posible. Pero las virtudes de un giro copernicano tienen que hacerse valer. Propuestas, aparentemente tan alejadas como la de Gärdenfors (1991), utilizan en el fondo variantes de las mismas ideas.

El trabajo actual en el campo de la no monotonía está centrado en buena medida en el tratamiento de las expectativas y las preferencias, en formalismos para representar órdenes de preferencias. Estos métodos permiten tender puentes hacia los tratamientos numéricos, por cuanto un orden de preferencias puede ser considerado como el lado cualitativo de ciertas teorías numéricas de la incertidumbre, como la teoría de posibilidades o las probabilidades infinitesimales.

c) Seguramente que el razonamiento no monótono es más complejo que la idea que puedan dar los ejemplos que hemos venido usando. Evidentemente, hay grados en la no monotonía. Y hay conflictos no sólo entre conclusiones no monótonas y hechos, sino entre conclusiones no monótonas y otras conclusiones no monótonas, etc. De hecho, esto último es lo esperable en la vida diaria. Las conclusiones revisables, en mayor o menor grado, constituyen buena parte del conjunto de nuestras creencias. Decíamos al comienzo que el miedo a equivocarse parece ser menos fuerte que el miedo al vacío de información. No parece que nos importe demasiado ajustar y reajustar nuestras creencias continuamente, como prueba cualquier diálogo normal en el que los hablantes intentan comunicarse. Como si los mecanismos de reajuste de la información fueran tan básicos que la lógica misma -los mecanismos de inferencia- fuese un subproducto suyo.

*IRIT - U. Paul Sabatier (Toulouse)
**Universidad de Cádiz

BIBLIOGRAFIA

Besnard, P., Quiniou, R. and Quinton, P.: 1983, 'A theorem prover for a decidable subset of default logic', in *AAAI-83*, Washington, DC, 27-30

Etherington, D.W.: 1988, *Reasoning with Incomplete Information: Investigations of Non-Monotonic Reasoning*, Pitman.

Etherington, D.W. and Reiter, R.: 1983, 'On inheritance hierarchies with exceptions', in *AAAI-83*, Washington, DC, 104-108.

Gärdenfors, P.: 1991, 'Nonmonotonic inferences based on expectations', in *Proc. KR'91,* Cambridge, MA, 585-590.

Ginsberg, M.: 1987, *Readings in nonmonotonic reasoning*, Los Altos, Morgan Kaufmann.

Halpern, J.Y. and Moses, Y.: 1985, 'Towards a theory of knowledge and ingnorance', in *Logic and Models of Concurrent Systems*, Springer-Verlag, 459-476.

Konolige, K.: 1988, 'On the relation between default logic and autoepistemic logic', *Artificial Intelligence* 35, 343-382.

Léa Sombé: 1990, 'Reasoning under incomplete information in Artificial Intelligence', *International Journal of Intelligent Systems* 5.

Levesque, H.J.: 1990, 'All I Know: A Study in Autoepistemic Logic', *Artificial Intelligence* 42, 263-309.

Lukaszewicz, W.: 1990, 'Non-monotonic Reasoning - Formalization of Commonsense Reasoning', Ellis Horwood Series in Artificial Intelligence.

McCarthy, J.: 1987, 'Epistemological problems of artificial intelligence', in *Proc. of the Fifth International Joint Conference on Artificial Intelligence*, Cambridge, MA, 1038-1044.

McCarthy, J.: 1980, 'Circumscription - A form of nonmonotonic reasoning', *Artificial Intelligence* 13, 27-39.

McCarthy, J.: 1986, 'Applications of circumscription to formalizing commonsense knowledge', *Artificial Intelligence* 28, 89-116.

McDermott, D. and Doyle, J.: 1980, 'Non-monotonic logic I', *Artificial Intelligence* 13, 41-72.

Meyer, J.J. and van der Hoek, W.: 1991, 'Non-monotonic Reasoning by Monotonic Means', in J. van Eijck (ed.): *Logics in AI. European Workshop JELIA '90*, Berlín, Springer-Verlag, 399-411.

Minsky, M.: 1975, 'A framework for representing knowledge', in Winston, P. (ed.): *The Psychology of Computer Vision*, McGraw-Hill, 211-277.

Moore, R.: 1985, 'Semantical considerations on nonmonotonic logic', *Artificial Intelligence* 25, 75-94.

Niemelä, I.: 1988, 'Decision procedure for autoepistemic logic', in *Proc. 9th Conf. on Automated Deduction*, Springer-Verlag, 676-684.

Poole, D.: 1985, 'On the comparison of theories; preferring the most specific explanation', in *IJCAI-85*, Los Angeles, CA, 144-147.

Reiter, R.: 1978, 'On closed world databases', in Gallaire and Minker (eds.): *Logic and Databases*, New York, Plenum, 55-76.

Reiter, R.: 1980, 'A logic for default reasoning', *Artificial Intelligence* 13, 81-132.

Reiter, R. and Criscuolo, G.: 1981, 'On interacting defaults', in *IJCAI-81*, Vancouver, BC, 270-276.

Shoham, Y.: 1988, *Reasoning about Change*, Cambridge, MIT Press.

Schwind, C.B.: 1990, 'A tableaux-based theorem prover for a decidable subset of default logic', in *CADE-10*.