

Inferencia Ampliativa y Razonamiento no Monotónico

Claudio Delrieux

Departamento de Ingeniería Eléctrica

Universidad Nacional del Sur — Alem 1253 — (8000) Bahía Blanca

claudio@acm.org

1 Razonamiento no monotónico

La necesidad de contar con modelos de razonamiento no monotónico en los sistemas inteligentes fue rápidamente reconocida en la comunidad del KR&R [1], especialmente para poder manejar los problemas originados al aplicar deducción en teorías incompletas. Entendemos aquí por no monotónico a todo aquel sistema de razonamiento que carezca de la propiedad de aditividad o monotonía. Cualquier sistema de razonamiento que utilice reglas ampliativas de inferencia tiene necesariamente la propiedad de ser no monotónico. Es decir, esta propiedad no surge solamente por el uso de reglas o condicionales derrotables, sino también por el uso de otras reglas o patrones de inferencia (por ejemplo, inducción, abducción, analogía, probabilidades, etc.).

Las reglas derrotables en particular, y la inferencia ampliativa en general, pueden considerarse como un mecanismo heurístico de *construcción* de nuevo conocimiento a partir del ya disponible. Este conocimiento generado, las conclusiones de nuestro sistema, es de naturaleza tentativa, es decir, es aceptable en tanto que no se modifique el contexto dentro del cual fue producido. Esto también muestra por qué la no monotonicidad no va necesariamente de la mano con la inferencia ampliativa. Podemos tener sistemas de razonamiento no ampliativos (que no construyen nuevo conocimiento) y que al mismo tiempo sean no monotónicos. Esto es lo que sucede, por ejemplo, con el razonamiento plausible y con la revisión de creencias.

Los primeros sistemas formales propuestos para razonar en teorías incompletas no buscaron un apartamiento radical hacia reglas de inferencia no deductiva utilizadas en otro tipo de problemas de razonamiento. En general se buscó la manera de respetar la *forma* de la deducción por medio de esquemas de inferencia similares a los utilizados en la deducción, especialmente *modus ponens*. De esa manera, representamos con una premisa *Normalmente, cuando a sucede, entonces b también sucede* la predisposición a realizar la inferencia no monotónica de *b* cuando en el contexto de la teoría es posible demostrar *a* pero no es posible demostrar $\neg b$. Dicha inferencia será no monotónica porque, en cuanto $\neg b$ pueda genuinamente demostrarse en el contexto, entonces la conclusión de la inferencia queda *bloqueada*.

En base a estas propuestas iniciadoras, el área del razonamiento no monotónico experimentó en la década del '80 un explosivo surgimiento de sistemas. Los mismos tuvieron el objetivo de solucionar el problema del *encadenamiento* (sea de implicaciones, reglas o condiciones de no anormalidad) de una línea de razonamiento y de su aceptación tentativa frente a la posibilidad de otras líneas conflictivas. Sin ser exhaustivos, podemos citar entre los continuadores al sistema de Poole [7], basado en un comparador de argumentos o teorías por especificidad, la lógica autoepistémica de Moore [5], que completa la lógica no monotónica de McDermott y Doyle, los sistemas de Nute [6] y Etherington [2], basados en operadores condicionales, el sistema de Geffner [3], que utiliza nociones de irrelevancia, y los sistemas argumentativos de Loui [4], y Simari-Loui [8], donde se representan las premisas tentativas como relaciones metalingüísticas entre conjuntos de literales, y se establece un mecanismo de razonamiento no monotónico basado en una generalización de la noción de demostración como secuencia, en la cual se permite la inclusión de las premisas tentativas. Estas demostraciones generalizadas pasaron a denominarse *argumentos o teorías*.

2 Otros esquemas de inferencia ampliativa

Sin embargo, parte del objetivo inicial del razonamiento no monotónico pareció perderse en el momento en el que la investigación se centró en la generación, comparación y aceptación de argumentos contruídos a partir de conocimiento válido y reglas derrotables. Otros problemas de gran importancia, como la extensión del contexto de la teoría por medio de la aceptación de evidencia tentativa o el uso de otros esquemas ampliativos de inferencia (abducción, inducción o analogía, por ejemplo), quedaron fuera de consideración de una manera definitiva. Cada uno de estos patrones de inferencia cumple con objetivos diferentes y en cierto punto complementarios al uso de reglas derrotables. Si el razonamiento no monotónico busca extender el conjunto de conclusiones de una teoría por medio de un conjunto de reglas tentativas conocidas y aceptadas *a priori*, entonces podemos pensar que el razonamiento inductivo *genera* este tipo de reglas a partir de la observación de un conjunto de casos particulares en los que se observa una regularidad. La abducción, por su parte, busca explicar un resultado no predicho por una teoría en función de alguna evidencia que lo justifique y que ha sido pasado por alto. En el razonamiento por analogía se busca aplicar o proyectar un esquema de inferencia adecuado en un dominio en otro dominio con el que es posible encontrar similitudes estructurales. El razonamiento plausible, por último, busca extender el dominio de una teoría por medio de la aceptación de información u observaciones que provienen de diversas fuentes, incluida la formación de hipótesis y la conjetura.

¿Es posible generar y comparar argumentos que produzcan conclusiones tentativas a partir de dichos esquemas? La respuesta a esta pregunta tiene su trascendencia, dado que propone un tratamiento igualitario entre las reglas derrotables y la inferencia ampliativa. En el primer caso, las conclusiones que se obtengan son tentativas porque se utiliza como premisa una regla derrotable, la cual puede tener excepciones. En el segundo caso, la conclusión también es tentativa, pero porque el esquema de inferencia mismo puede tener excepciones. El objetivo de este trabajo consiste en investigar los modelos de integración del razonamiento no monotónico con algunos o todos los patrones ampliativos de inferencia aquí planteados. Para ello es necesario generalizar los mecanismos de generación, comparación y aceptación de teorías o argumentos, y encontrar en cada caso los mecanismos de resolución de conflictos y las estrategias computacionales asociadas.

La generación de argumentos puede establecerse de una manera similar a la demostración como secuencia, donde además se permite utilizar cada uno de los esquemas de inferencia vistos. La comparación de argumentos, sin embargo, plantea toda una serie de dificultades metodológicas, dado que no siempre es posible comparar el soporte asertivo de una conclusión obtenida con reglas derrotables (por ejemplo) con otra obtenida por abducción. Una estrategia posible consiste en utilizar un mecanismo similar al del razonamiento plausible para representar el soporte asertivo de la información provista por diversas fuentes. Otra estrategia complementaria consiste en investigar los distintos mecanismos de comparación que puedan existir como derrotadores de estos esquemas de inferencia.

3 Argumentos con otros esquemas de inferencia ampliativa

Utilizaremos en este trabajo la definición de sistemas argumentativos abstractos de Vreeswijk [9], por ser lo suficientemente amplia como para acomodar nuestros objetivos. En Vreeswijk, un sistema argumentativo es un lenguaje, un conjunto de reglas de inferencia, y una relación de orden parcial entre argumentos. El lenguaje \mathcal{L} no tiene ninguna estructura especial, es más, no requiere el uso de conectivos lógicos. La única restricción es la inclusión del símbolo \perp como indicación de contradicción. Las reglas de inferencia son o bien estrictas (similares a la implicación material) o bien derrotables, donde el formato general es

$$s_1, s_2, \dots, s_n \rightarrow t,$$

siendo cada elemento del antecedente y el consecuente un miembro del lenguaje.

El encadenamiento de reglas origina los argumentos. En Vreeswijk un argumento (básico) es un elemento de \mathcal{L} , o bien una fórmula

$$a_1, a_2, \dots, a_n \rightarrow b,$$

donde cada miembro del antecedente es un (sub)argumento, y existe una regla

$$s_1, s_2, \dots, s_n \rightarrow b,$$

tal que el consecuente de a_1 es s_1 , etc. con la restricción de que b no aparece en ningún subargumento.

En nuestra propuesta de agregar otros esquemas de inferencia, el trabajo consiste en representar cualquiera de ellos como regla de inferencia en el marco propuesto por Vreeswijk. Por ejemplo, para representar la inferencia abductiva podemos utilizar una regla

$$\frac{\begin{array}{l} b(t) \\ \not\sim b(t) \\ a(X) \vdash b(X) \end{array}}{a(t)},$$

es decir, *si se produce la observación de un hecho $b(t)$ el cual no es predicho, pero encontramos que $a(X)$, agregado como antecedente, permite generar un argumento para $b(X)$, entonces inferimos $a(t)$* . El siguiente ejemplo nos permite ilustrar de qué manera podrían combinarse la generación de argumentos y el uso de abducción.

EJEMPLO 3.1 *En un sistema de diagnóstico ocupacional encontramos que quien tiene un trabajo (t) normalmente percibe un ingreso (i), normalmente no estudia (e), y paga aporte jubilatorio (j). Quienes estudian normalmente trabajan, y quienes estudian y ganaron una beca (b) perciben un ingreso pero no pagan aporte jubilatorio:*

$$\begin{array}{l} t(X) \rightarrow i(X) \\ t(X) \rightarrow j(X) \\ t(X) \rightarrow \neg e(X) \\ e(X) \rightarrow t(X) \\ b(X) \rightarrow e(X) \\ b(X) \rightarrow i(X) \end{array}$$

¿Qué se puede afirmar de juan, de quién sabemos que paga su aporte jubilatorio?

Por inferencia abductiva, de $j(\text{juan})$ se sigue $t(\text{juan})$ sin contradicción, sigue también $\neg e(\text{juan})$. Por último, aplicando abducción nuevamente, se sigue $\neg b(\text{juan})$.

¿Qué podemos afirmar de ana, de quién sabemos que percibe un ingreso?

*En estas condiciones, es posible encontrar abductivamente una justificación en $t(\text{ana})$, y otra en $b(\text{ana})$. Con respecto la primera justificación, se sigue además que $j(\text{ana})$ y que $\neg e(\text{ana})$ y por consiguiente $\neg b(\text{ana})$, es decir, *ana percibe un ingreso por trabajar y paga su jubilación, pero no estudia y por consiguiente no está becada. De la segunda justificación se sigue $e(\text{ana})$, y además $t(\text{ana})$ y $j(\text{ana})$, es decir, *ana está becada y por consiguiente estudia, pero además ana trabaja y por consiguiente paga su jubilación. Adoptando una visión escéptica, la justificación abductiva será aquella que pertenezca a todas las interpretaciones posibles (o permanecer indeterminado si no hay justificación común). Es decir, se acepta que ana trabaja como explicación***

de su ingreso, y además se predice que ana debe normalmente pagar una jubilación, hecho que habrá que corroborar.

¿Qué podemos afirmar de pedro, de quién sabemos que percibe un ingreso pero que no paga aporte jubilatorio?

En este caso, las justificaciones del ingreso de pedro son idénticas a las justificaciones del ingreso de ana, pero el hecho adicional de que pedro no paga jubilación bloquea la primera, y por lo tanto se concluye que pedro es un estudiante becado.

References

- [1] Matthew L. Ginsberg (ed.). *Readings in Nonmonotonic Reasoning*. Morgan Kaufmann Publishers, Los Altos, California, 1987.
- [2] David W Etherington. *Reasoning with Incomplete Information*. Morgan Kaufmann Publishers, Los Altos, CA, 1988.
- [3] Héctor Geffner and Judea Pearl. Conditional Entailment: Bridging Two Approaches to Default Reasoning. *Artificial Intelligence*, 53(2-3):209–244, 1992.
- [4] Ronald P. Loui. Defeat Among Arguments: A System of Defeasible Inference. *Computational Intelligence*, 3(3), 1987.
- [5] Robert C. Moore. Semantical Considerations of Nonmonotonic Logic. *Artificial Intelligence*, 25(1):75–94, 1985.
- [6] Donald Nute. Defeasible Reasoning. In James H. Fetzer, editor, *Aspects of Artificial Intelligence*, pages 251–288. Kluwer Academic Publishers, Norwell, MA, 1988.
- [7] David L. Poole. On the Comparison of Theories: Preferring the Most Specific Explanation. In *Proceedings of the Ninth International Joint Conference on Artificial Intelligence*, pages 144–147, Los Altos, CA, 1985. International Joint Conference on Artificial Intelligence, Morgan Kaufmann Publishers.
- [8] Guillermo R. Simari and Ronald P. Loui. A Mathematical Treatment of Defeasible Reasoning and its Implementation. *Artificial Intelligence*, 53(2-3):125–158, 1992.
- [9] G. A. W. Vreeswijk. Abstract Argumentation Systems. *Artificial Intelligence*, 90(2):225–279, 1997.