EPISTEMOLOGÍA E HISTORIA DE LA CIENCIA

SELECCIÓN DE TRABAJOS DE LAS VII JORNADAS
1997

Patricia Morey
José Ahumada
Editores

ÁREA LOGICO-EPISTEMOLÓGICA DE LA ESCUELA DE FILOSOFÍA
CENTRO DE INVESTIGACIONES DE LA FACULTAD DE FILOSOFÍA Y HUMANIDADES
UNIVERSIDAD NACIONAL DE CÓRDOBA

Javier Legris
Universidad Nacional de Buenos Aires/CONICET

EL SUEÑO DE LEIBNIZ. SOBRE DEDUCCIÓN AUTOMÁTICA Y RAZONADORES ARTIFICIALES

Gottfried Wilhelm Leibniz fue uno de los primeros en tener la idea de automatizar los procedimientos inferenciales. Se necesitaron casi doscientos años para que esta idea fuera tomada realmente en serio y casi un siglo más para que se obtuvieran algunos resultados. Estos resultados constituyen el núcleo de lo que se llama actualmente "deducción automática", "razonamiento automático", "demostración automática de teoremas" y, más en general, "lógica computacional". Estos campos o áreas de investigación forman una parte importante, a su vez, de la disciplina tecnológica llamada inteligencia artificial, la cual ha tenido en los últimos años un inmenso desarrollo y un enorme impacto (directo e indirecto) en la sociedad. En lo que sigue se ofrecerá un breve panorama de la manera en que se está realizando "el sueño de Leibniz", indicando algunas de sus propiedades y limitaciones. Pero más específcamente, se discutirá, en primer lugar, el papel que la deducción automática tiene en el desarrollo actual de la lógica matemática y, en segundo lugar, la importancia que tiene en la justificación filosófica de la inferencia deductiva. La investigación sobre la deducción automática aporta fundamentaciones "constructivas" de la lógica y un análisis de la inferencia lógica como un mecanismo para obtener conocimiento.

1. De Leibniz a los demostradores artificiales

Entre las numerosas ideas que la humanidad debe al genio de Leibniz está la de la automatización del razonamiento. Leibniz fue más allá del Ars Magna de Raimundo Lullio y la Computatio sive Logica de Thomas Hobbes al observar que la automatización del razonamiento requería un lenguaje con un alfabeto básico y completo y un sistema combinatorio. Esto es lo que él llamó respectivamente lingua characteristica y calculus ratiocinator, un lenguaje formal universal y un cálculo o mecanismo de computación en ese lenguaje. Este cálculo podría ser automatizable, tal como sucedía con la aritmética gracias a la máquina calculadora que él había desarrollado en 1671. Así, la justificación racional de las propias afirmaciones adoptaría la forma de un cálculo y, más aun, toda controversia entre los seres humanos se solucionaría tomando el lápiz y diciendo "¡Calculemos!" (véase GP VII, p. 200).

Este programa de Leibniz sería retomado, en toda su extensión, dos siglos más tarde por Gottlob Frege con su Begriffschrift, o escritura conceptual, que fue también el punto de partida de la lógica matemática, la forma moderna de la lógica, y ofreció, a la vez, un modelo básico a seguir para la deducción automática. En el Begriffschrift se encuentra tanto un lenguaje formal, que es lo que llamamos hoy lenguaje de primer orden, como un procedimiento formal de derivación para ese lenguaje, que da lugar a la lógica de primer orden o lógica elemental.

Posteriormente, la sistematización de la lógica encarada en los Principia Mathematica de Russell y Whitehead y el programa metamatemático de David Hilbert parecían conducir a una pronta realización del sueño leibniciano. Sin embargo, los resultados de incompletud e indecidibilidad obtenidos en la década de 1930, entre otros por Kurt Gödel y Alonzo Church en la década de 1930, lo convirtieron en algo imposible, al menos en su versión absoluta y universal, pues mostraba la imposibilidad de hallar procedimientos computables para determinar si un enunciado es una ley de la lógica de primer orden o no. Y lo mismo sucederá en el caso de cualquier teoría construida sobre la lógica de primer orden.

La aparición de la deducción automática debe verse contra el trasfondo de estas limitaciones. Si bien resultaba imposible construir algoritmos de decisión para la lógica de primer orden en su totalidad, quedaba abierta la posibilidad de construir algoritmos para buscar una demostración de cualquier ley lógica. Como se verá más adelante, esta idea sí era realizable en virtud de ciertos resultados de la teoría de la demostración. Por lo tanto, se propuso construir programas efectivos para demostrar teoremas. Estos programas son los demostradores de teoremas.

En pocas palabras, la deducción automática surge de la conjunción de dos hechos, uno tecnológico y otro científico: en primer lugar, la existencia de máquinas computadoras suficientemente veloces y, en segundo lugar, el corpus teórico proporcionado por la lógica matemática a través de la disciplina llamada teoría de la demostración. Ahora bien, a partir de los avances en la construcción de máquinas computadoras, en la década de 1950, comenzaron los primeros intentos de desarrollar demostradores de teoremas lógicos (debidos, entre otros, a Newell, Shaw y Simon y a Hao Wang; para la los detalles históricos véase, por ejemplo, Loveland 1984). El trabajo siguió dos líneas diferentes. De un lado, debía avanzarse en la eficiencia y rapidez de los programas. De otro lado, los programas debían imitar o simular la forma en que los "razonadores naturales" (de hecho, los seres humanos) efectuaban deducciones, convirtiéndose, por lo tanto, en "razonadores artificiales". Este último aspecto iba cobrando cada vez más importancia a medida que se consolidaba la idea de desarrollar la inteligencia artificial, y el equilibrio entre ambos aspectos es un desideratum de la investigación actual. Aquí también aparece el interés por los diferentes métodos o sistemas de deducción desarrollados por los lógicos matemáticos y se encuentra un campo interdisciplinario de trabajo para expertos en computación, lógicos, matemáticos y lingüistas.

2. Deducción automática y teoría de la demostración

La teoría de la demostración nació en la década de 1920 en el contexto del programa de Hilbert para la fundamentación de la matemática. El objetivo central de este programa consistía en demostrar la consistencia de teorías matemáticas mediante procedimientos aceptables desde "el punto de vista finito" o "finitismo", tales como el empleo de inducción completa de la manera más sencilla posible sobre el proceso de derivación de los teoremas. El cumplimiento de este objetivo requería un análisis de las demostraciones en estas teorías, las cuales debían ser formuladas en términos de sistemas formales en los que no se tomaba en cuenta el contenido o significado de sus expresiones. La relación de deducción se analizaba, entonces, como una relación puramente sintática entre fórmulas del sistema en cuestión, la cual recibía el nombre de derivabilidad formal.

Debido al carácter esencialmente finitista del programa, toda derivación en un sistema formal debía consistir en un número finito de pasos, cada uno de los cuales resultaba de la aplicación de reglas de derivación admitidas en el sistema, que debían tener un claro sentido finitista. En el desarrollo posterior de la teoría de la demostración, esta exigencia de un carácter finito de las reglas fue dejada de lado, siendo a menudo reemplazada por la exigencia de un carácter constructivo.

Ahora bien, hay dos resultados fundamentales de la teoría de la demostración, relativos a la lógica de primer orden, que hacen posible la deducción automática. Estos dos resultados son el teorema de Herbrand y el Hauptsatz de Gentzen.

El teorema de Herbrand fue formulado en la tesis doctoral de su autor, Jacques Herbrand escrita en 1929 y publicada un año después (véase Herbrand 1930). El teorema, en su versión original, afirma que dada una fórmula del lenguaje de primer orden, puede generarse una secuencia infinita de fórmulas sin cuantificadores, tal que la fórmula original es un teorema de la lógica de primer orden si, y sólo si, hay una fórmula de la secuencia generada que es teorema de la lógica de enunciados.

Este resultado es la base para el método de Herbrand para demostrar la validez de una fórmula válida de la lógica de primer orden, que es un método puramente mecánico y que puede verse como un sustituto constructivo de la teoría de modelos (véase van Heijenoort 1992). El método puede interpretarse del modo siguiente recurriendo al concepto de interpretación: dada una fórmula se busca una interpretación que la falsifique; si la fórmula es válida, entonces se determinará en un número finito de pasos que no existe tal interpretación falsificadora. Esta propiedad de la lógica de primer orden ha sido bautizada en el ámbito de la computación como semidecidibilidad y es una condición necesaria para la deducción automática: si un teorema es válido, su validez puede demostrarse en un número finito de pasos. La semidecidibilidad es lo que permite la automatización de la deducción (y, por lo tanto, la realización -si bien parcial- del "sueño de Leibniz").

El Hauptsatz fue formulado por Gentzen también en su tesis doctoral, presentada en 1933 (véase Gentzen 1935). El teorema afirma, entre otras cosas, que existen demostraciones normales para la lógica de primer orden; es decir, para todo teorema de la lógica de primer orden existe una demostración canónica y en cierto sentido "mínima", que "no da rodeos". Gentzen formulaba este teorema respecto de un sistema formal construido sobre la base de secuentes, afirmando que una regla del sistema, que puede interpretarse como la afirmación de la transitividad de la relación de derivabilidad, es suprimible en el sistema. Como consecuencia, se obtenía la llamada "propiedad de la subfórmula": Toda fórmula que aparece en la derivación de un teorema es subfórmula del mismo, de modo que ninguna otra fórmula desaparece en la derivación. (Sobre la importancia del Hauptsatz y de la propiedad de la subfórmula, véase Legris 1995.) La existencia de sistemas que permiten demostraciones normales es esencial para la deducción automática, pues son demostraciones de este tipo las que puede encontrar el programa de búsqueda.

En su misma tesis de doctorado, Gentzen formuló además lo que sería otra de las piezas fundamentales de la teoría de la demostración: el sistema de deducción natural. Debido a sus reglas de introducción y de eliminación de símbolos lógicos (que proporcionan a éstas un significado constructivo) a la inclusión de derivaciones a partir de supuestos y a la existencia de subderivaciones, el sistema de deducción natural parecía reflejar los mecanismos y estrategias deductivos empleados por los "razonadores naturales" que somos los seres humanos. (Y esto fue expresamente señalado por su autor, véase Legris 1995.) Más tarde pudo demostrarse que este sistema satisfacía también una versión del Hauptsatz: toda derivación de una ley lógica en deducción natural podía convertirse en una derivación normal que no contiene aplicaciones superfluas de reglas de introducción o eliminación ("no da rodeos", véase Prawitz 1965). De este modo, se disponía de una formalización intuitivamente adecuada de los mecanismos deductivos empleados por los seres humanos, apta para su adopción en deducción automática.

Importante es que la deducción natural originó también el estudio de la estructura de las demostraciones consideradas no ya como meras derivaciones en sistemas formales, sino directamente como objetos de análisis. Como consecuencia, la teoría de la demostración pasó a ser una teoría acerca de demostraciones. De este modo, se ponía de relieve un aspecto característico de la teoría de la demostración, ya presente en sus inicios: su interés en cómo se efectúa una demostración, en las estrategias que ésta adopta y en los pasos o mecanismos que sigue.

Los diferentes sistemas deductivos formalizan, entre otras cosas, diferentes maneras en que los seres humanos deducen (por ejemplo, mediante métodos directos o refutatorios). Estos sistema deductivos son analizados en la deducción automática con el fin de especificar medios de algoritmizar el proceso de demostración y de búsqueda de demostraciones. Así, por ejemplo, se supone actualmente que el análisis de sistemas de secuentes y de la deducción natural puede llevar a la construcción de demostradores que sean a la vez razonablemente veloces, fáciles de depurar o mejorar y que simulen aceptablemente las demostraciones efectuadas por los demostradores humanos (véase, por ejemplo, Tennant 1992). De este modo, la deducción automática no es más que la versión computacional de la teoría de la demostración.

3. La construcción de demostradores

Las máquinas inteligentes pueden cumplir alguna de estas funciones: procesamiento de conocimiento; representación de conocimiento o asimilación de nuevo conocimiento. Aunque las tres funciones tienen cierto grado de dependencia mutua, los demostradores cumplen el primer tipo de función: sirven para procesar el conocimiento dado, obteniendo un conocimiento "nuevo" en cierto sentido.

La construcción de demostradores tiene tres etapas. En primer lugar, se debe elegir el sistema deductivo adecuado. En segundo lugar, debe construirse un algoritmo para la búsqueda de demostraciones respecto del sistema elegido. En tercer lugar, el algoritmo obtenido debe ser implementado en un lenguaje de computación de la manera más eficiente posible. Obviamente, las tres etapas están interrelacionadas.

En la primera etapa (la elección del sistema lógico) debe tomarse en cuenta la complejidad computacional del sistema, su justificación filosófica en relación con el significado de las constantes lógicas y su aplicabilidad a diferentes áreas de las ciencias. El problema no reside únicamente en elegir un sistema determinado para la lógica de primer orden, sino también en la elección de la lógica, por ejemplo la lógica clásica, o alguna lógica no clásica (lógica intuicionista, lógica minimal o lógica relevante).

En la segunda etapa, si el sistema es decidible (como sucede, por ejemplo, en subsistemas de la lógica de primer orden), pueden construirse algoritmos de decisión. En todo caso, el problema central es el del control del espacio de búsqueda de demostraciones y es aquí donde deben aparecer métodos heurísticos (véase Wos et al. 1985, p. 27). En general, los seres humanos exploran sólo una porción del espacio de búsqueda y adoptan los caminos más relevantes para construir las demostraciones. La tarea consiste en trasladar de algún modo estas habilidades a los demostradores artificiales.

En la construcción de demostradores, propiedades estudiadas por la teoría de la demostración son de especial importancia. Así sucede con la propiedad de separabilidad, que permite tratar por separado constantes lógicas o conjuntos de ellas y, dado un sistema deductivo, determinar subsistemas referidos a un conjunto de constantes lógicas. (Un sistema de reglas es separable, si para demostrar un teorema, en el cual aparece la constante lógica *, se necesitan únicamente reglas para *.) Pero también aparecen problemas nuevos como la tratabilidad de los algoritmos, determinada por su complejidad computacional, y la eficiencia de la implementación, es decir que consiga obtener la demostración en el menor número de pasos. Ambos conceptos son relativos (conviene hablar, por ejemplo, de algoritmos de búsqueda más tratables que otros o de implementaciones más eficientes que otras). En general, todo problema tratable es decidible, pero no vale la conversa. Los dos problemas estaban ausentes en la investigación que se hacía dentro de la teoría de la demostración (o al menos eran irrelevantes) pero, en el contexto de la deducción automática, son decisivos para adoptar una lógica o un tipo de presentación formal de la lógica.

Importa aquí el método que sigue el sistema deductivo elegido para construir una demostración, sea de manera directa o de manera indirecta. En el primer caso el demostrador procede directamente dando las razones suficientes para afirmar el enunciado a demostrar. En el segundo caso, el demostrador procede por refutación, buscando contraejemplos para el enunciado a demostrar. En el caso en que a partir del sistema pueden construirse tanto algoritmos de demostración directa como de refutación, se está frente a demostradores que dan razones completas para decir que el enunciado es demostrable. Si el demostrador indica además cuál es la demostración más breve (más eficiente), entonces constituye lo que Tennant ha llamado un racionauta, demostradores óptimos (véase Tennant 1992, p. 33).

4. La deducción automática y la fundamentación de la inferencia deductiva

El desarrollo de la computación y la inteligencia artificial ha producido cambios en la lógica matemática al punto de hablarse de un "giro computacional" en lógica. Dejando de lado aspectos contextuales, la investigación en deducción automática ha conferido importancia, como se acaba de señalar, al estudio de la complejidad algorítmica y la eficiencia de los mecanismos de demostración. Estas propiedades pasaban inadvertidas en la investigación puramente teórica, pero en la deducción automática se convierten en criterios para aceptar un sistema de deducción. Por ejemplo, aparece en primer plano el problema de hacer explícitas las estrategias de demostración y establecer el "control" de la búsqueda de demostraciones y tanto las estrategias como los procedmientos de control pasan a formar parte de la caracterización de las demostraciones correctas. Estos aspectos han conducido a la formulación de una semántica procedimental que tiene como concepto básico el de regla de computación.

Por lo demás, los métodos constructivos característicos de muchas investigaciones dentro de la teoría de la demostración han extendido su ámbito de aplicación y, en relación con el problema de la eficiencia, se ha formulado un concepto cuantitativo de constructibilidad, en el sentido de grados de constructibilidad aplicados a demostraciones. En general, la situación es comparable con la que existe en el ámbito de la matemática, donde se habla de una "matemática computacional", en la que, por ejemplo, es importante determinar aquellas funciones computables que sean "factibles" (véase, por ejemplo, Beeson 1992). El concepto mismo de constructibilidad ha debido ser reconsiderado

Otros cambios más fundamentales tienen que ver con la justificación filosófica de la inferencia deductiva. En primer lugar, el "giro computacional" en la lógica, al que ya se ha hecho referencia, privilegia una filosofía constructivista de la lógica. En efecto, una concepción constructiva del significado de las constantes lógicas toma en cuenta la capacidad de efectuar cómputos o de realizar procedimientos efectivos. Históricamente, los antecedentes de esta posición se encuentran en la "lógica de problemas" de Andrei Kolmogorov, en la cual los enunciados que contienen constantes lógicas significan un tipo particular de problemas expresables mediante funciones (véase Kolmogorov 1932). Pero también desarrollos semejantes se encuentran en la "lógica operativa" de Paul Lorenzen, donde se estudian las propiedades de reglas efectivas, concebidas como "operaciones esquemáticas", por medio de las cuales se caracterizan las constantes lógicas (véase Lorenzen 1969). Estos conceptos ofrecen una base para una semántica procedimental como la recién mencionada.

Así, resulta natural que quien trabaja en deducción automática preste especial atención a fundamentaciones constructivistas de la lógica, que vinculan el significado de las constantes lógicas a procedimientos constructivos o "efectivos", y deje en principio de lado los métodos no constructivos que proporciona la teoría de modelos y las ideas filosóficas subyacentes a los mismos.

En segundo lugar, el giro computacional lleva a ver a la inferencia deductiva en el marco del estudio de los procesos de conocimiento y de las maneras de representarlo y formalizarlo. Así, este giro computacional es al mismo tiempo un "giro cognitivo". La inferencia deductiva es, en varios sentidos y por varias razones, la base de todo tipo de razonamiento y de toda argumentación. (Esto es así en parte porque ella es el núcleo central de todo razonamiento y de toda forma de argumentación.) Los razonadores artificiales son máquinas concebidas para obtener conocimiento a partir de un conocimiento dado. En el caso concreto de los demostradores de teoremas, las operaciones que dan como resultado un teorema representan una demostración en el sentido usual del término y, por lo tanto, una justificación efectiva del conocimiento que representa el teorema.

De esta manera, la investigación en deducción automática es compatible con la idea de que demostrar un enunciado es una forma de conocer su verdad, otorgándole a las demostraciones un valor cognoscitivo. En este aspecto, la deducción automática retoma implícitamente la concepción aristotélica que considera a la lógica como teoría del conocimiento demostrativo y vincula naturalmente la teoría de la demostración con el problema de la justificación epistémica.

En síntesis, en la deducción automática se hacen más evidentes ciertas propiedades ya presentes en la teoría de la demostración y se agregan nuevos factores para evaluar sistemas deductivos y una nueva metodología para tratarlos. Sin embargo, desde el punto de vista conceptual, la deducción automática no es más que la aplicación tecnológica de la teoría de la demostración. Así pues, es la misma concepción de la inferencia deductiva aquella que propuso Leibniz, reapareció con Frege, subyace a la investigación sobre la teoría de la demostración y se continúa en la deducción automática.

Referencias bibliográficas

Beeson, Michael. 1993. "Constructivity in The Nineties". En Philosophy of Mathematics. Proceedings of the 15th. International Wittgenstein'Symposium, comp. por Johannes Czermak. Viena, Hölder-Pichler-Tempsky, pp. 127-144.

Gentzen, Gerhard. 1935. "Untersuchungen über das logische Schliessen". En Mathematische Zeitschrift 39, pp. 176-210 y 405-431.

Herbrand, Jacques. 1930. Recherches sur la théorie de la démonstration. Reimpreso en Ecrits Logiques compilados por Jean van Heijenoort. París, Presses Universitaires de France, 1968.

Kolmogorov, Andrei N. 1932. "Zur Deutung der intuitionistischen Logik". En Mathematische Zeitschrift 35, pp. 58-65.

Legris, Javier. 1995. "Demostraciones de consistencia y derivabilidad formal: A 60 años de las 'Investigaciones sobre la deducción lógica' de G. Gentzen". En Epistemología e Historia de la Ciencia, compilado por Alberto Moreno. Córdoba (Arg.), Universidad Nacional de Córdoba.

Leibniz, Gottfried Wilhelm. GP. Die philosophischen Schriften von G.W. Leibniz. 7 vols. Compilado por C.I. Gerhardt, Berlín, 1875-1890.

Lorenzen, Paul. 1969. Einführung in die operative Logik und Mathematik. 2da. ed., Berlín-Heidelberg-N.York, Springer.

Loveland, Donald W. 1984. "Automated Theorem-Proving: A Quarter-Century Review". En Automated Theorem Proving: After 25 Years comp. por W.W. Bledsoe y D.W. Loveland. Providence (RI), American Mathematical Society, 1984.

Prawitz, Dag. 1965. Natural Deduction. A Proof-Theoretical Study. Estocolmo, Almqvist & Wiksell.

Tennant, Neil. 1992. Autologic. Edimburgo, Edinburgh University Press.

van Heijenoort, Jean. 1992. "Historical Developement of Modern Logic". En Modern Logic 2, 242-255.

Wos, Larry et al. 1985. "An Overview of Automated Reasoning and Related Fields". Journal of Automated Reasoning 1, 5-41.