

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



Esta obra está bajo una Licencia Creative Commons atribución NoComercial-SinDerivadas 2.5 Argentina



HACIA UNA TAXONOMÍA DE LA INVESTIGACIÓN EN INTELIGENCIA ARTIFICIAL

Introducción: ¿Qué tipo de disciplina es la Inteligencia Artificial?

La primera cuestión epistemológica que se plantea respecto de la Inteligencia Artificial (IA) es determinar de qué tipo de disciplina se trata. El problema, de pertinencia para la filosofía, no es ajeno a la disciplina misma: se manifiesta en la diversidad de objetivos y metodologías de las investigaciones utilizadas bajo su nombre. Al recorrer los índices de las publicaciones especializadas en IA pueden encontrarse trabajos sobre: representación de conocimiento, aprendizaje de conceptos, formación de hipótesis, técnicas de programación lógica, etc. Cabe preguntarse: ¿es el objeto de estudio de estos trabajos el mismo en todos los casos? ¿es la metodología utilizada en cada investigación la misma, o, siendo distintas, guardan alguna relación?

Para responder a estas preguntas no basta con ver los trabajos incluidos en alguna de las áreas mencionadas y abstraer los aspectos en común, puesto que pueden encontrarse contribuciones tan variadas, respecto tanto de los objetivos buscados como de las metodologías empleadas, que para un neófito resultaría difícil de creer que tratan un problema común de fondo. La creación de programas es, por supuesto, el fin último de la IA. Pero la creencia de que éste es el *único* fin es, quizás, una de las causas de la confusión reinante acerca de la disciplina. Un síntoma claro de esta confusión es que, en muchos casos, *referees* de revistas y congresos especializados rechazan reportes de investigación por no brindar lineamientos para la implementación programativa de las ideas desarrolladas, o por ser excesivamente matemáticos o lógicos y hacer escasa referencia a aspectos computacionales concretos. Por el contrario, otras veces se rechazan trabajos por no incluir proposiciones acerca de la generalidad de los problemas que supuestamente un sistema puede manejar, esto es, el fundamento de porqué un sistema funciona tan bien como su autor pregona.

La necesidad de un análisis que descomponga la IA en líneas de investigación diferenciadas se ha vuelto imperiosa, puesto que frecuentemente se manifiesta una confusión metodológica que hace difícil apreciar cuándo se produce un avance, y en qué aspecto se lo logra. Para una aproximación al problema, en lo que sigue brindaremos, primero, un ejemplo de estudio en representación de conocimiento, a saber, el desarrollo de

la lógica *default* de Raymond Reiter (Reiter[80]), con el fin de mostrar cómo pueden diferenciarse distintas líneas de trabajo en un programa de investigación. En segundo lugar expondremos una taxonomía de la investigación en IA propuesta por Alan Bundy (Bundy[90]), para apreciar si ésta permite dar cuenta de la distinción. El resultado negativo del análisis nos llevará a buscar una taxonomía más apropiada para el objetivo.

Mostraremos que el grado de especialización del conocimiento requerido en cada aspecto de un problema de IA permite diferenciar en la disciplina distintas líneas de investigación, así como clarificar sus respectivos objetivos y metodologías. Tal criterio surgirá como resultado de un refinamiento del concepto de *tecnología operativa* de M. Bunge (Bunge[83]).

Una muestra de investigación en IA: la lógica *default*

Tomemos para analizar un caso paradigmático de investigación en representación de conocimiento (RC), como lo es la lógica *default* de Reiter¹. Ésta constituye un caso típico de investigación en IA básica, más precisamente, en fundamentos lógicos de la IA. El sistema describe un formalismo para representar conocimiento de sentido común e inferencias no-monotónicas. Reiter introduce reglas de inferencia llamadas *default* que permiten establecer conclusiones tentativas (o inferencias *prima facie*), a partir del conocimiento representado. Pueden describirse formalmente todas las extensiones posibles de una teoría de acuerdo a las reglas *default* que contenga. Sin embargo, no hay nada en el sistema mismo que diga cómo ha de ser implementado en un programa que permita dar con estas extensiones. Más aún, el interés es no hacer referencia a ningún tipo de procedimiento, para que el sistema sirva a propósitos generales. Todo lo concerniente a la viabilidad de tales implementaciones es una cuestión relacionada con las técnicas computacionales que se utilicen para tal fin. Por ejemplo, el desarrollo de las técnicas computacionales eficientes para computar todas las extensiones posibles de una teoría (si esto es posible) es una tarea en la que deben tenerse en cuenta la complejidad y tratabilidad computacionales del problema. Luego, esas técnicas quedarán a disposición de los diseñadores de programas específicos (que requieran inferencias no-monotónicas, en este caso), quienes las aplicarán o no, de acuerdo a sus propósitos.

Podemos, entonces, distinguir distintas líneas de trabajo en la investigación:

a. *Línea formal*: es la tomada por Reiter en la elaboración de la lógica *default*. Comprende

(i) la creación de un lenguaje para representar el conocimiento. Esto puede hacerse desde un punto de vista puramente formal, por ejemplo, extendiendo los lenguajes lógicos conocidos de un modo apropiado para expresar lo deseado;

(ii) la caracterización de las inferencias correctas del conocimiento representado. Tal caracterización es independiente del o los algoritmos que se creen para dar con las

¹ Una breve exposición de este sistema se brinda en el apéndice.

inferencias deseadas (el sistema se esquematiza como caja negra). Respecto de esto, puede decirse que el formalismo cumple un rol normativo.

b. Línea implementativa: de desarrollo posterior al de la línea formal, se ocupa de

(i) desarrollar lenguajes de programación (como Prolog, Lisp, o extensiones de éstos) que permitan expresar lo mejor posible lo que permite expresar el lenguaje lógico o cuasi-lógico descripto. Aquí deben tenerse en cuenta los aspectos computacionales, pues la traducción de un lenguaje formal en un lenguaje de programación suele conducir a una pérdida de poder expresivo a costas de la eficiencia computacional, o viceversa;

(ii) desarrollar técnicas computacionales que permitan manipular el conocimiento representado para obtener las inferencias deseadas de un modo eficiente. El aporte de los conocimientos derivados de las ciencias de la computación son claves para desarrollar estructuras de datos adecuadas, o para reducir la complejidad de los algoritmos y alcanzar tratabilidad computacional de los razonamientos buscados.

c. Línea de aplicación: Encargada de utilizar las técnicas obtenidas en la línea implementacional para construir programas que resuelvan problemas específicos, como los sistemas expertos.

En la siguiente sección veremos cómo se acomodan estas líneas en la taxonomía de la investigación en IA dada por Bundy.

IA básica y aplicada

Alan Bundy propone una división de la IA en tres partes: (1) las ciencias cognitivas, encargadas de elaborar modelos de la mente con ayuda de técnicas computacionales; (2) la IA básica, encargada de crear técnicas para modelar procesos inteligentes, y las relaciones que se dan entre éstas, sin referencia a aplicaciones concretas; y (3) la IA aplicada, encargada de producir programas para la resolución de problemas concretos, mediante la utilización de las técnicas desarrolladas por la IA básica. Centraremos nuestra atención en los dos últimos ítems, primordialmente, en la IA básica (no tocaremos en este trabajo el tema de las ciencias cognitivas, que no consideramos dentro de la IA, ni su relación con ésta).

Bundy compara a las ciencias de la computación con la matemática pura, poniendo la IA básica y la IA aplicada en los roles de la matemática aplicada y la ingeniería, respectivamente. La diferencia metodológica entre la IA aplicada y la IA básica no requiere un gran análisis: obviamente, no es el mismo el método seguido para crear un programa, que el seguido para crear un lenguaje formal o de programación, o para analizar la complejidad computacional de un algoritmo. Pero donde se vuelve necesario un análisis algo más fino es dentro de la IA básica, puesto que, según la entiende Bundy, las distintas líneas de investigación suelen confundirse en ella. El método general que señala para la IA básica consiste en: programación exploratoria-análisis de los programas logrados-generalización de las técnicas utilizadas-reconstrucción racional de los algoritmos. Pero esta metodología, que es aplicable a las dos últimas líneas de trabajo, excluye las tareas que son previas a la programación, y que están comprendidas en lo que llamamos la línea

formal. En nuestro ejemplo, esa metodología da cuenta de la búsqueda de las técnicas computacionales para implementar la lógica *default*, pero no de la búsqueda de la lógica en sí (Bundy no habla explícitamente de la lógica *default*, pero sí de la circunscripción de McCarthy[80], que es también un formalismo²). Entonces vemos que, si bien la IA básica abarca problemas que pueden ir desde las propiedades lógicas de un sistema formal hasta cuestiones netamente computacionales, no hay una metodología general que sirva para cubrir a todos ellos.

En el caso de la RC, dentro de la investigación básica tenemos, como vimos, (1) los fundamentos de la RC, en particular, los *fundamentos lógicos* del conocimiento (lo cual abarca la línea formal), y (2) el desarrollo de *técnicas computacionales* apropiadas para crear programas que incorporen y manipulen el conocimiento (línea implementacional). Cada una de estas líneas de investigación requiere de una metodología particular, y una taxonomía apropiada debe dar cuenta de sus diferencias. El próximo paso que daremos será buscar una división en la IA básica que permita diferenciar esos distintos aspectos, lo cual sugeriría la adopción de un criterio para evaluar los aportes producidos.

Conocimiento especializado en la investigación en IA

Puesto que el fin de la IA es eminentemente práctico, sus teorías son de carácter tecnológico. Según Bunge[83], existen dos tipos de teorías tecnológicas: las *sustantivas*, y las *operativas* o *no-sustantivas*. Las primeras se apoyan fuertemente en el conocimiento de las ciencias fácticas, mientras que las segundas no, ó en muy poca medida. Consideramos que la IA es una tecnología operativa, teniendo en cuenta lo siguiente. Por una parte, es claro que las investigaciones en fundamentos lógicos de la IA, resultan en teorías no-sustantivas³, puesto que el único conocimiento en el que se fundan proviene de las ciencias formales. Pero por otra parte, las técnicas computacionales para la implementación de formalismos se basan, además, en la teoría de la computación, la cual se sustenta a su vez en conocimiento matemático y, en parte, lingüístico. En este punto haremos una digresión importante.

La teoría de la computación se funda en la idea de una máquina abstracta (máquina de Turing) que, como tal, no se diferencia de las entidades matemáticas. Luego, la teoría de la computación es una teoría formal. Como el desarrollo de técnicas computacionales en IA tiene como fundamento una teoría formal, se trata, entonces, de tecnología no-sustantiva. Nuestra opinión es que dentro de las teorías tecnológicas no-sustantivas hay distintos grados o niveles, y que es en esta nivelación donde en realidad las líneas de investigación en IA pueden ser distinguidas.

Consideremos que el grado de operatividad (por llamarlo así) de una tecnología está dado por la especialización del conocimiento (científico o tecnológico) que supone.

² cf. A. Bundy, op. cit., pp. 216 y 221.

³ Como aclara Bunge, las teorías de este tipo no difieren de las científicas sino en el fin, que es más práctico que cognoscitivo (cf. Bunge, ibid., p. 685).

Así, toda la tecnología de implementación computacional de sistemas de IA (que cubre el desarrollo de lenguajes de programación, estructuras de datos, algoritmos, estrategias de búsqueda, etc.) se funda en un conocimiento especializado (teoría de la computación) que ignoran los sistemas de IA puramente formales (como el de Reiter). A su vez, la línea de aplicación es la más operativa de las tres, ya que supone el conocimiento más especializado (además de conocer las técnicas de programación, un programador conoce las técnicas de aplicación de esas técnicas).

Esta distinción en grados de operatividad permite explicar que una persona versada en lógica o matemática pueda hacer importantes aportes a la IA, aunque no conozca en profundidad las ciencias de la computación. Por supuesto, tales aportes están restringidos a los aspectos lógico-formales de las investigaciones que antes mostramos como la línea formal. Así se entiende que, por ejemplo, los sistemas que formalizan cambios de teorías, como el de Alchourrón, Gärdenfors y Makinson -AGM[85]-, u otros de lógicas no-standard, como las lógicas autoepistémicas -e.g., Moore[85]-, se consideren a menudo incluidos en el campo de la IA a la vez que en el de la lógica. En cambio, las investigaciones que, dentro de la IA, se ocupan del desarrollo de técnicas computacionales adecuadas para la implementación de formalismos, o la búsqueda de algoritmos de baja complejidad computacional -cuestiones propias de la línea implementacional-, constituyen tecnología de un grado más alto de operatividad, puesto que se valen del conocimiento especializado producido en teoría de la computación (además de la lógica matemática). Este conocimiento es el que guía al investigador para lograr lenguajes de programación, estructuras de datos, o algoritmos lo más eficientes posible.

Esta clasificación ofrece consecuencias metodológicas importantes. Permite mostrar que las cuestiones de implementación son ajenas a los fundamentos lógicos de la IA y, por esto, exigencias en ese aspecto a los investigadores del tema son irrelevantes. Por otra parte, sí es de exigir a los autores de formalismos resultados acerca del comportamiento de sus sistemas con ejemplos claves, demostraciones de que obedecen determinadas especificaciones, o la relación que guardan con otros formalismos conocidos, como por ejemplo, mostrar que un sistema puede reducirse a otro, etc.

A menudo la IA se ve desprestigiada porque gran parte de los investigadores trabajan con una metodología deliberadamente confusa. La distinción dada permite diferenciar distintas líneas de trabajo para que las investigaciones enfoquen los problemas de la forma más nítida posible. Asimismo, sugiere criterios para la evaluación de los trabajos que no confundan las líneas de investigación, eviten exigencias irrelevantes, y exijan los resultados que las subdisciplinas requieran.

Resumen comparativo de los criterios vistos

En lo expuesto podemos ver que en la clasificación de Bundy las líneas formal e implementacional constituyen investigación básica, mientras que la línea de aplicación corresponde, obviamente, a la investigación aplicada. Sin embargo, no ofrece un quiebre entre el tipo de investigación de la línea formal, que es la comprendida por la lógica *default*

de Reiter, y el de la línea implementacional. Este quiebre es captado por la distinción de grados en la concepción de Bunge de tecnología operativa. En la línea formal el trabajo se emparenta más con la lógica matemática que con las ciencias de la computación. De hecho, el conocimiento de computación necesario para hacer un aporte en estos terrenos es escaso o nulo. En cambio, en la etapa implementacional la investigación descansa ampliamente tanto en el conocimiento teórico desarrollado en las etapas anteriores como en el brindado por las ciencias de la computación.

Conclusión

La IA comprende distintas líneas de investigación, cada una de las cuales enfoca un problema parcial en la búsqueda de un objetivo común: producir programas que simulen inteligencia. La metodología de trabajo utilizada en estas líneas depende de las características del conocimiento específico en que se apoyan, de sus técnicas, y del alcance de éstas. Tener en cuenta esas características puede ayudar a evaluar los trabajos y detectar en qué aspectos particulares se producen los progresos de la disciplina.

Apéndice: Lógica *Default* de Reiter

Brevemente, la lógica *default* de Reiter consiste en un sistema construido con el lenguaje de la lógica de primer orden, más un conjunto de reglas metalingüísticas, llamadas

“reglas *default*”, de la forma $\frac{\alpha, \beta}{\gamma}$, (donde ‘ α ’, ‘ β ’ y ‘ γ ’ son sentencias del lenguaje),

interpretadas intuitivamente como “si ‘ α ’ (prerrequisito) es verdadera y la afirmación de ‘ β ’ (justificación) no produce una inconsistencia, entonces se afirma ‘ γ ’ (conclusión)”. Una teoría *default* T es un par $\langle S, D \rangle$, donde S es un conjunto de sentencias en el lenguaje de primer orden, y D es un conjunto de reglas *default*. Una extensión E de T es un conjunto consistente, lógicamente cerrado, que contiene a S , y para cualquier regla *default* en D , si su prerrequisito está en la extensión y no está la negación de la justificación, entonces su conclusión está en la extensión. Una teoría puede tener ninguna, una, o múltiples

extensiones. Por ejemplo, la teoría en la que $S = \emptyset$ y $D = \left\{ \frac{\neg\alpha}{\beta}, \frac{\neg\beta}{\alpha} \right\}$, tiene dos

extensiones, una conteniendo a ‘ α ’ pero no a ‘ β ’, y la otra conteniendo a ‘ β ’ pero no a ‘ α ’.

Referencias

AGM[85] Alchourrón, C., P. Gärdenfors and D. Makinson; "On the logic of theory change: partial meet functions for contraction and revision." *Journal of Symbolic Logic* 50, 1985, 510-530.

Bundy[93] Bundy, A.; "What Kind of Field is Artificial Intelligence?"; en D. Partridge y Y. Wilks (eds.), *The Foundations of Artificial Intelligence: A Sourcebook*; Cambridge University Press, 1993 (reprint), 215-222.

Bunge[83] Bunge, M.; *La Investigación Científica*; 2º ed., 1983, Ariel, Barcelona.

McCarthy[80] McCarthy, J.; "Circumscription -A form of Non-monotonic Reasoning"; *Artificial Intelligence* 13 (1-2), 1980, 27-39, 171-172.

Moore[85] Moore, R.; "Semantical Considerations of Nonmonotonic Logic"; *Artificial Intelligence* 24 (1), 1985, 75-94.

Reiter[80] Reiter, R.; "A Logic For *Default Reasoning*", *Artificial Intelligence* 13 (1-2), 1980, 81-132.