

## **La mente**

Del mismo autor

*Conceptual revolutions*, Princeton, 1992

*Computational philosophy of science*, Massachusetts, 1993

*Mind readings: Introductory selections on cognitive science*,  
Massachusetts, 1998

*How scientists explain disease*, Princeton, 2000

*Coherence in thought and action (Life and mind: Philosophical issues  
in biology and psychology)*, Massachusetts, 2002

*Hot thought: Mechanisms and applications of emotional cognition*,  
Massachusetts, 2006

*Mente. Una introducción a las ciencias cognitivas*, ha sido publicado en italiano por Guerini, bajo el título: *La mente. Introduzione alla scienza cognitiva*; en alemán por Klett-Cotta, bajo el título: *Kognitionswissenschaft*; en portugués por Artmed como *Mente: introdução à ciência cognitiva*. Hay, igualmente, ediciones en checo y en coreano, y en dos variantes del chino.

Paul Thagard

**La mente**

Introducción a las  
ciencias cognitivas

Traducido por Silvia Jawerbaum y Julieta Barba



**conocimiento**

Primera edición, 2008  
Segunda reimpresión, 2013

© Katz Editores  
Benjamín Matienzo 1831, 10<sup>º</sup> D  
1426-Buenos Aires  
c/Sitio de Zaragoza, 6, 1<sup>ª</sup> planta  
28931 Móstoles-Madrid  
**www.katzeditores.com**

Título de la edición original: *Mind*.  
*Introduction to cognitive science*

© 2005 Massachusetts Institute of Technology,  
The MIT Press, Cambridge, MA

ISBN Argentina: 978-987-1283-63-7  
ISBN España: 978-84-96859-21-0

I. Conocimiento. I. Jawerbaum, Silvia, trad. II. Barba,  
Julieta, trad. III. Título  
CDD 001

El contenido intelectual de esta obra se encuentra  
protegido por diversas leyes y tratados  
internacionales que prohíben la reproducción  
íntegra o extractada, realizada por cualquier  
procedimiento, que no cuente con la autorización  
expresa del editor.

Diseño de colección: tholön kunst

Impreso en España por Romanyà Valls S.A.  
08786 Capellades

Depósito legal: M-34.565-2010

# Índice

9	Agradecimientos
11	Prefacio
	<b>I. ENFOQUES DE LAS CIENCIAS COGNITIVAS</b>
17	1. Representación y computación
47	2. Lógica
77	3. Reglas
101	4. Conceptos
129	5. Analogías
153	6. Imágenes
175	7. Conexiones
205	8. Resumen y evaluación
	<b>II. EXTENSIONES DE LAS CIENCIAS COGNITIVAS</b>
223	9. El cerebro
245	10. Las emociones
265	11. La conciencia
289	12. Cuerpo, mundo y sistemas dinámicos
311	13. Sociedad
327	14. El futuro de las ciencias cognitivas
343	Glosario
349	Bibliografía
381	Índice temático



Para Adam, una megamente



## **Agradecimientos**

Agradezco a los estudiantes de la Universidad de Waterloo que asistieron a los cursos de cuyas discusiones surgió el material para el presente libro. También deseo agradecer a Andrew Brook, Chris Eliasmith, Kim Honeyford, Janet Kolodner, Amy Pierce, Michael Ranney y Eric Steinhart por sus comentarios sobre el texto para la primera edición, y a Chris Eliasmith, Ashok Goel, Michael Ranney y Ethan Toombs por sus aportes para la segunda edición. También quisiera agradecer a Tom Stone, de MIT Press, por su apoyo y sus valiosas sugerencias. Durante la redacción y la revisión de este libro, conté con el subsidio del Consejo de Investigación en Ciencias Naturales e Ingeniería de Canadá. Por último, agradezco a Abninder Litt por haber colaborado en el proceso de corrección del original.



## Prefacio

Las ciencias cognitivas, dedicadas al estudio de la mente y la inteligencia desde un punto de vista interdisciplinario, son fruto de la confluencia entre la filosofía, la psicología, la inteligencia artificial, la neurociencia, la lingüística y la antropología. Sus orígenes intelectuales se remontan a mediados de la década de 1950, cuando investigadores provenientes de distintos campos comenzaron a formular teorías de la mente valiéndose de representaciones y procesos computacionales complejos. En cuanto a los orígenes institucionales, éstos datan de mediados de los años setenta, época en que se fundó la Sociedad de Ciencias Cognitivas en los Estados Unidos y se publicó por primera vez la revista *Cognitive Science*. Desde entonces, más de sesenta universidades estadounidenses y europeas han creado carreras de ciencias cognitivas, y muchas otras universidades dictan materias relacionadas con este campo.

La enseñanza de las ciencias cognitivas no es tarea sencilla, pues los estudiantes llegan con formaciones académicas muy diversas. Yo, personalmente, dicto en la Universidad de Waterloo, desde 1993, una materia muy concurrida titulada Introducción a las Ciencias Cognitivas. Por un lado, se inscriben en mi cátedra estudiantes con conocimientos avanzados de computación, provenientes de carreras de computación científica o ingeniería, que saben poco de psicología o filosofía, pero, por otro lado, vienen

muchos estudiantes que están familiarizados con esas disciplinas pero que de computación entienden bastante poco. Mi intención al escribir este libro ha sido poder contar con un manual que no requiere conocimientos previos en ninguno de los campos de las ciencias cognitivas. También he pretendido mostrar a los estudiantes interesados en los problemas de la mente y la inteligencia que existen muchos enfoques complementarios en la investigación sobre las funciones mentales.

Existen al menos tres formas distintas de presentar las ciencias cognitivas ante un público formado en disciplinas diversas. La primera consiste en centrarse en los distintos campos: la psicología, la inteligencia artificial y las demás disciplinas involucradas. La segunda es organizar el curso según las distintas funciones mentales, como la resolución de problemas, la memoria, el aprendizaje y el lenguaje. En este libro, he decidido tomar otro camino, en el que hago una descripción y una evaluación sistemáticas de las principales teorías de la representación mental formuladas por distintos teóricos, que incluyen la lógica, las reglas, los conceptos, las analogías, las imágenes y las conexiones (redes neuronales artificiales). Con la presentación de esos enfoques teóricos fundamentales he querido mostrar de manera coherente los aportes realizados por los distintos campos de las ciencias cognitivas al conocimiento del funcionamiento de la mente.

Mi propósito al escribir el presente libro ha sido poner a disposición de estudiantes provenientes de distintas disciplinas un material básico que les permita comprender sin problemas un curso introductorio a las ciencias cognitivas. Para cumplir con el objetivo que me había propuesto, debí explicar lógica de forma tal que fuera accesible a los estudiantes de psicología, exponer algoritmos computacionales para que los entendieran los estudiantes de literatura, o presentar debates filosóficos de modo sencillo para que a los estudiantes de ciencias de la computa-

ción les resultaran atractivos.

Aunque este libro está dirigido principalmente a estudiantes de carreras de grado, también ha sido mi intención que fuese de utilidad para los estudiantes de posgrado y los profesores interesados en descubrir cuál es la participación de sus propias disciplinas en los distintos proyectos de las ciencias cognitivas. No se trata, por tanto, de un texto enciclopédico sino de un manual con capítulos breves y concisos, en los que resalto las cuestiones generales en desmedro de las particulares. Dado que abordo las ciencias cognitivas como la intersección y no como la unión entre todos los campos que forman parte de ellas, decidí omitir algunos temas que ya se han tratado en manuales de inteligencia artificial, psicología cognitiva y filosofía de la mente, por nombrar sólo algunas disciplinas sobre las que existen infinidad de textos. Para cerrar cada capítulo, he incluido un resumen y un listado de lecturas recomendadas.

El libro está escrito con gran entusiasmo y respeto por los aportes que han hecho las teorías de la representación mental y los procesos computacionales al conocimiento de la mente, aunque soy consciente de que las ciencias cognitivas tienen un largo camino por recorrer. En la segunda parte, me refiero a las extensiones de los supuestos básicos de las ciencias cognitivas y propongo distintas ideas para trabajar en el futuro.

Estoy muy agradecido a quienes me han hecho llegar ejemplares de la primera edición traducida al italiano, alemán, checo, portugués, japonés, coreano y dos variedades de chino. Para esta segunda edición, he actualizado la primera parte y revisado la segunda, además de añadir capítulos sobre el cerebro, las emociones y la conciencia. Otras novedades son un listado de sitios web al final de cada capítulo y un glosario al final del libro. Para complementar el presente texto, recomiendo la lectura de otro libro del que soy editor: *Mind readings: Introductory selec-*



# I

## Enfoques de las ciencias cognitivas



# 1

## Representación y computación

### EL ESTUDIO DE LA MENTE

¿Quién no se ha preguntado alguna vez cómo funciona la mente? Todos realizamos a diario una gran variedad de tareas mentales, desde resolver problemas en el trabajo o la escuela hasta tomar decisiones vinculadas con nuestra vida personal, encontrar los motivos de las acciones de los demás o adquirir nuevos conceptos, como el de *telefonía celular* o *Internet*. El objetivo principal de la ciencia cognitiva es encontrar las explicaciones de cómo se realizan esas formas de pensamiento. Esta disciplina científica no se limita a describir los distintos tipos de estrategias de aprendizaje y de resolución de problemas, sino que ofrece una explicación acerca de cómo realiza la mente esas operaciones. Además, da cuenta de los casos en los que la mente no funciona con la eficacia deseada; por ejemplo, cuando se toman decisiones equivocadas.

Comprender cómo funciona la mente es importante para muchas actividades prácticas. Si los docentes saben cuáles son los mecanismos de pensamiento de los estudiantes, podrán aplicar los métodos de enseñanza adecuados. Para los ingenieros y otros profesionales dedicados al diseño de productos, es importante saber cuáles son los probables procesos mentales que permitirán que los futuros usuarios hagan un uso eficaz o ineficaz

de esos productos. En el caso de los científicos especializados en informática, el conocimiento de qué factores hacen que las personas sean inteligentes es de utilidad en el diseño de computadoras más inteligentes. Y para los políticos y los responsables de la toma de decisiones, el éxito de su carrera depende de la comprensión de los procesos mentales de las personas con las que interactúan.

Sin embargo, el estudio de la mente no es tarea sencilla, puesto que no es posible tener acceso directo a ella. A lo largo de los siglos, los filósofos y los psicólogos han recurrido a un conjunto de metáforas y comparaciones para referirse a la mente. Se ha afirmado que es como una página en blanco sobre la que se hacen impresiones, o como un instrumento hidráulico en el que se ejercen distintas fuerzas o que es semejante a un conmutador telefónico. En los últimos cincuenta años han surgido nuevas metáforas gracias al desarrollo de nuevas clases de computadoras. Muchos especialistas en ciencia cognitiva, si bien no todos, consideran que el pensamiento es una especie de proceso computacional y utilizan metáforas relacionadas con el campo de la computación para describir y explicar cómo las personas aprenden y resuelven problemas.

### ¿QUÉ SABEMOS?

Cuando empiezan una carrera universitaria, los estudiantes tienen que aprender mucho más que lo que dicen los textos. Según la carrera elegida, variarán los temas, pero en todas es necesario aprender algunos aspectos básicos sobre cómo funciona la universidad. ¿Cómo se hace para anotarse en una materia? ¿A qué hora empiezan las clases? ¿Qué cátedras son buenas y cuántas?

les no? ¿Cuáles son los requisitos para obtener el título? ¿Cuál es el mejor camino para ir de un aula a otra? ¿Cómo son los demás estudiantes? ¿Cuál es el mejor lugar para salir un viernes a la noche?

Las respuestas a esas preguntas se almacenan en algún lugar de la mente, pero ¿dónde? La mayoría de los especialistas coinciden en que el conocimiento consiste en *representaciones mentales*. Todo el mundo sabe qué es una representación no mental, como lo son las palabras escritas en esta página. La frase “esta página” es a su vez una representación de la página que el lector está viendo en este momento. En las instituciones educativas abundan las representaciones gráficas, los mapas, por ejemplo, de modo que los estudiantes están habituados a verlas. Para explicar varios tipos de conocimiento, como qué saben los estudiantes sobre el funcionamiento de la universidad, la ciencia cognitiva propone distintos tipos de representaciones mentales: reglas, conceptos, imágenes y analogías, etc. Los estudiantes conocen, entre otras, la regla que determina *Si quiero recibirme, tengo que cursar diez materias del ciclo de especialización*. También adquieren conceptos nuevos representados por términos como “bird”, “Mickey Mouse” o “gut”, que en las universidades de algunos países angloparlantes se utilizan para referirse a las materias muy fáciles. Para ir de un aula a otra, les resulta útil contar con una representación mental de la distribución espacial de la facultad. Después de completar una materia que les ha parecido buena, es probable que traten de inscribirse en otras similares. Al tener la oportunidad de relacionarse con estudiantes de otras carreras, construyen estereotipos aunque les resulte difícil expresar con precisión qué características tiene cada uno.

El conocimiento sobre la vida en la universidad no se adquiere sólo para acumular información. Los estudiantes se

enfrentan con numerosos problemas: cómo lograr un buen desempeño al cursar las materias, cómo llevar una vida social satisfactoria y cómo obtener un trabajo después de recibirse. Resolver ese tipo de problemas requiere operar con representaciones mentales; por ejemplo, cuántas materias les restan para recibirse o no anotarse nunca más en ninguna materia dictada por el profesor Tedio. Según la ciencia cognitiva, las personas experimentan *procesos* mentales que actúan sobre las representaciones para generar pensamientos y acciones. Los distintos tipos de representaciones mentales, como las reglas y los conceptos, se asocian con distintas clases de procesos mentales. Tomemos por caso las distintas maneras de representar números. La mayoría de las personas conocen la representación arábiga de los números (1, 2, 3, 10, 100...) y los procedimientos para realizar operaciones de adición o multiplicación, entre otras. También pueden utilizar números romanos (I, II, III, X, C...), pero los procedimientos para realizar operaciones aritméticas son distintos. Hagamos la prueba de dividir CIV (104) por XXVI (26).

En la primera parte de este libro presentaremos los distintos enfoques sobre las representaciones y los procesos mentales propuestos en los últimos cuarenta años de investigación en el marco de la ciencia cognitiva. El valor de algunos de esos enfoques ha sido motivo de controversia, y muchos de los principales teóricos de la ciencia cognitiva defendieron con vehemencia la validez del enfoque de su preferencia. En mi caso personal, tiendo a ser más ecléctico, pues creo que las distintas teorías de la representación mental no se excluyen entre sí sino que se complementan. La mente humana es sumamente compleja, y se comprende mejor si se tiene en cuenta el uso de las reglas mencionadas anteriormente y de otras no tan obvias, como las representaciones propuestas por los modelos

“conexionistas” o “de redes neuronales” que comentaremos en el capítulo 7.

## LOS INICIOS

Los primeros intentos por comprender la mente y su funcionamiento se remontan a la antigua Grecia, o quizás antes, con las primeras teorías sobre la naturaleza del conocimiento formuladas por filósofos como Platón y Aristóteles. Platón pensaba que el conocimiento más importante provenía de conceptos como la *virtud*, conceptos innatos que no dependían de la experiencia sensible. Otros filósofos, como Descartes y Leibniz, pensaban que el conocimiento se adquiría gracias al pensamiento y al razonamiento, por lo que su enfoque se denomina *racionalismo*. Por el contrario, Aristóteles abordó el tema del conocimiento recurriendo a reglas que se aprenden gracias a la experiencia, como ocurre con el enunciado *Todos los hombres son mortales*. Esa postura filosófica, que luego abrazaron Locke, Hume y otros pensadores, se denomina *empirismo*. En el siglo XVIII, Kant propuso una combinación del racionalismo y el empirismo, pues sostenía que el conocimiento humano depende de la experiencia sensible y de las capacidades innatas de la mente.

El estudio de la mente permaneció confinado al campo de la filosofía hasta el siglo XIX, época en que nació la psicología experimental. Wilhelm Wundt y sus discípulos diseñaron métodos de laboratorio para el estudio sistemático de las operaciones mentales. Al cabo de unas pocas décadas, la psicología experimental cedió terreno ante el *conductismo*, que negaba la existencia de la mente. Algunos conductistas, John B. Watson (1913), por ejemplo, sostenían que la psicología debía limitarse al estu-

dio de la relación entre estímulos y reacciones observables de la conducta. La conciencia y las representaciones mentales no eran temas de los que la ciencia respetable debiera ocuparse. En los Estados Unidos en particular, el conductismo dominó el campo de la psicología hasta la década de 1950.

Hacia 1956, el panorama intelectual experimentó un cambio drástico. George Miller (1956) presentó una reseña de numerosos estudios que demuestran que la capacidad del pensamiento humano es limitada y que la memoria a corto plazo no es capaz de almacenar más de siete elementos. (Por esa razón, es difícil retener un número de teléfono con sus códigos de área o el del carné de la obra social.) Miller sugirió que las limitaciones de la memoria se superan si la información se recodifica en forma de unidades más pequeñas, es decir, por medio de representaciones mentales que requieren procesos mentales para codificar y decodificar información. Era la época de las primeras computadoras y de pioneros como John McCarthy, Marvin Minsky, Allen Newell y Herbert Simon, fundadores del área del conocimiento conocida como inteligencia artificial. Por su parte, Noam Chomsky (1957, 1959) rechazó los supuestos conductistas sobre el lenguaje como hábito adquirido y afirmó que la capacidad del hombre de comprender el lenguaje se explica en términos de una gramática mental constituida por reglas. No sería injusto considerar a los seis científicos mencionados en este párrafo como los fundadores de la ciencia cognitiva.

En los capítulos siguientes, mostraremos la historia posterior de la ciencia cognitiva vinculándola con distintas teorías de la representación mental. McCarthy fue uno de los principales exponentes del enfoque de la inteligencia artificial basado en la lógica formal (capítulo 2). En la década de 1960, Newell y Simon demostraron el poder de las reglas en su descripción de algunos aspectos de la inteligencia humana, tradición retomada en numero-

tos trabajos posteriores (capítulo 3). En los años setenta, Minsky propuso la noción de marco conceptual como estructura central de las representaciones del conocimiento. Otros investigadores del área de la psicología y la inteligencia artificial plantearon la existencia de estructuras similares a las que denominaron guiones y esquemas (capítulo 4). En esa misma época, los psicólogos también comenzaron a interesarse en las imágenes mentales (capítulo 6). A partir de la década de 1980, se han realizado innumerables trabajos de investigación experimental y computacional centrados en el pensamiento analógico o razonamiento inductivo (capítulo 5). Las propuestas más interesantes de esa década fueron las teorías conexionistas, que establecieron cierto grado de relación entre las representaciones y los procesos mentales y las redes neuronales del cerebro (capítulo 7). Todos esos enfoques han hecho su contribución al conocimiento de la mente, cada uno con sus aciertos y desaciertos (capítulo 8).

La noción fundamental de que la mente debería pensarse en términos de representaciones y procesos mentales ha recibido críticas y sufrido modificaciones, de las que nos ocuparemos en la segunda parte de este libro (capítulos 9 a 14). La década de 1990 fue testigo de la proliferación del uso de técnicas de escaneo cerebral para el estudio de las zonas del cerebro asociadas con el pensamiento y en la actualidad se trabaja con modelos computacionales de la mente cuya estructura refleja la estructura neurológica real. Esos modelos proponen nuevas formas de entender las emociones y la conciencia (capítulos 10 y 11). Las reacciones contrarias al modelo computacional-representacional se apoyan en el argumento de que el organismo y los entornos físico y social desempeñan un papel importante en el pensamiento (capítulos 12 y 13). Finalmente, nos dedicaremos a plantear el futuro de las ciencias cognitivas y a brindar algunas sugerencias para potenciales trabajos de carácter interdisciplinario (capítulo 14).

## MÉTODOS DE LAS CIENCIAS COGNITIVAS

Las ciencias cognitivas no son sólo el quehacer de un conjunto de personas provenientes de distintas disciplinas que se reúnen en un bar a conversar sobre la mente. Para apreciar cuáles son las ideas unificadoras sobre las que se construyen las ciencias cognitivas es necesario tener un panorama de los métodos que los investigadores de las distintas disciplinas utilizan en el estudio de la mente y la inteligencia.

Si bien en la actualidad abundan la teorización y los modelos computacionales en la psicología cognitiva, el método principal de investigación es la experimentación con sujetos humanos. Las personas que participan en los estudios –por lo general, estudiantes universitarios– acuden al laboratorio, donde se las estudia en un entorno con condiciones controladas. Algunos de los trabajos realizados, como veremos en distintos capítulos de este libro, consisten en analizar los errores que se cometen al hacer razonamientos deductivos, la forma en que se crean y aplican conceptos, la velocidad con que se piensa utilizando imágenes mentales y el desempeño durante la resolución de problemas mediante analogías. Las conclusiones sobre cómo funciona la mente deben trascender el “sentido común” y la introspección, porque como muchas de las operaciones mentales no son accesibles a nivel consciente, pueden darnos una idea general equivocada. Los experimentos psicológicos que intentan desentrañar las operaciones mentales desde distintos ángulos son fundamentales para que las ciencias cognitivas tengan carácter científico.

Es cierto que la teoría sin experimentación está vacía, pero también es cierto que la experimentación sin teoría es ciega. El tratamiento de las cuestiones esenciales de la naturaleza de la mente requiere que los experimentos psicológicos puedan inter-

pretarse dentro de un marco teórico que postule la existencia de representaciones y procesos mentales. Una de las mejores maneras de construir marcos teóricos consiste en crear y poner a prueba modelos computacionales análogos a las operaciones mentales. Para complementar los experimentos psicológicos centrados en el razonamiento deductivo, la formación de conceptos, las imágenes mentales y la resolución de problemas por analogía, los investigadores han creado modelos computacionales que simulan distintos aspectos del desempeño humano. El diseño y la construcción de modelos computacionales y la experimentación subsiguiente conforman las actividades primordiales de la inteligencia artificial, la rama de las ciencias de la computación que se ocupa de los sistemas inteligentes. En un contexto ideal, en el campo de las ciencias cognitivas, los modelos computacionales y la experimentación psicológica van de la mano, pero muchos estudios fundamentales en el campo de la inteligencia artificial sobre la relevancia de distintos abordajes de la representación del conocimiento no han tenido relación con la psicología experimental.

En lingüística, a veces se realizan experimentos psicológicos o se desarrollan modelos computacionales; sin embargo, los métodos utilizados en esta disciplina suelen ser otros. Los lingüistas que siguen la tradición chomskiana se dedican a buscar los principios gramaticales que permiten explicar la estructura básica del lenguaje. La detección de esos principios se realiza mediante la comparación de las diferencias sutiles existentes entre oraciones gramaticales y agramaticales. En español, por ejemplo, “Juan pateó la pelota” o “¿Qué pateó Juan?” son gramaticales; en cambio, “Juan la pateó pelota” y “¿Qué patearon Juan?” son agramaticales. Una gramática del español debe poder explicar por qué las primeras oraciones son aceptables y las segundas no lo son. En los capítulos siguientes daremos otros

ejemplos de la investigación teórica y empírica que llevan a cabo los lingüistas chomskianos y los pertenecientes a otras tradiciones lingüísticas.

Al igual que en el campo de la psicología cognitiva, en neurociencia se realizan experimentos controlados, pero las observaciones son muy distintas, pues el objeto de estudio de la neurociencia es el cerebro en sí. Cuando usan sujetos no humanos en los experimentos, los investigadores pueden insertar electrodos y registrar así el estímulo de neuronas individuales. Esa técnica es imposible de usar en humanos porque es demasiado invasiva; no obstante, en los últimos años se han empezado a utilizar la resonancia magnética y la tomografía por emisión de positrones, que permiten observar la actividad de distintas zonas del cerebro en personas a las que se solicita que realicen diversas tareas mentales. Por ejemplo, gracias a la tomografía computada, se determinó qué zonas del cerebro participan en la formación de imágenes mentales y la interpretación de palabras. Otras evidencias del funcionamiento del cerebro surgen de la observación de pacientes con determinados daños cerebrales. Un accidente cerebrovascular que afecta a zonas asociadas con el lenguaje puede ocasionar deficiencias tales como la incapacidad de emitir oraciones. Como la psicología cognitiva, la neurociencia es una disciplina teórica y experimental a la vez, en la que gran parte de la teoría se construye a partir de modelos computacionales del comportamiento de grupos neuronales.

La antropología cognitiva se ocupa de analizar el pensamiento humano con relación a distintos entornos culturales. El estudio de la mente, claro está, no se restringe a saber cómo piensan los hablantes de una lengua determinada sino a considerar las posibles diferencias en los modos de pensar en las distintas culturas. En los capítulos 12 y 13, analizaremos el creciente interés que despierta entre los especialistas en ciencias cognitivas

el funcionamiento de la mente en entornos físicos y sociales particulares. Para la antropología cultural, es primordial el método etnográfico, que requiere que los investigadores convivan e interactúen con los miembros de una determinada cultura durante el tiempo necesario para llegar a conocer su sistema social y cognitivo. Los antropólogos cognitivos han investigado, por ejemplo, las similitudes y las diferencias entre las palabras que denotan colores en distintas culturas.

Salvo contadas excepciones, los filósofos no se dedican a la observación empírica ni diseñan modelos computacionales; aun así, la filosofía sigue siendo esencial para las ciencias cognitivas, porque se ocupa de temas fundamentales que constituyen los pilares de los enfoques experimentales y computacionales del estudio de la mente. Temas abstractos como la naturaleza de la representación y la computación no necesariamente se tienen en cuenta en la práctica cotidiana de la psicología o la inteligencia artificial, pero es inevitable tratarlos cuando los investigadores reflexionan sobre su tarea. La filosofía también se ocupa de cuestiones generales como la relación entre la mente y el cuerpo, y de aspectos metodológicos como la naturaleza de las explicaciones proporcionadas por las ciencias cognitivas. Además de la descripción de cómo piensan las personas, la filosofía se dedica a cuestiones normativas sobre cómo *deberían* pensar. Las ciencias cognitivas añaden al objetivo teórico de la comprensión del pensamiento humano el objetivo práctico de mejorarlo, proceso que requiere una reflexión normativa sobre qué esperamos del pensamiento. La filosofía de la mente no utiliza un método en particular y sería deseable que compartiera el interés por los resultados empíricos con las teorías elaboradas en otras disciplinas.

En su forma más laxa, las ciencias cognitivas son la suma de las disciplinas que hemos mencionado: la psicología, la inteligencia artificial, la lingüística, la neurociencia, la antropología

y la filosofía. El trabajo interdisciplinario es mucho más interesante cuando hay una convergencia teórica y experimental respecto de las conclusiones sobre la naturaleza de la mente. En los capítulos siguientes ilustraremos esa convergencia, es decir, mostraremos el trabajo de las ciencias cognitivas en áreas de intersección de varias disciplinas. Por ejemplo, la psicología y la inteligencia artificial confluyen en los modelos computacionales que describen la conducta de los sujetos experimentales. La mejor manera de comprender la complejidad del pensamiento humano consiste en utilizar múltiples métodos, en especial, en combinar experimentos psicológicos y neurológicos con modelos computacionales. En términos teóricos, la idea más fructífera ha sido el enfoque computacional-representacional de la mente.

#### EL MODELO COMPUTACIONAL-REPRESENTACIONAL DE LA MENTE

La hipótesis central de las ciencias cognitivas es que la mejor forma de estudiar el pensamiento es entenderlo en términos de estructuras de representaciones mentales sobre las que operan procesos computacionales. Si bien no hay consenso en cuanto a la naturaleza de las representaciones y los procesos computacionales que constituyen el pensamiento, la hipótesis central es tan general que abarca todos los enfoques actuales de las ciencias cognitivas, incluso las teorías conexionistas. Denominaremos MCRM al modelo basado en esa hipótesis central, como sigla de modelo computacional-representacional de la mente.

El MCRM podría ser incorrecto. En la segunda parte de este libro, presentaremos objeciones al modelo según las cuales las

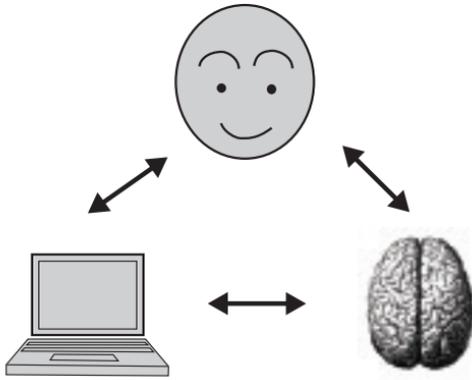
ideas sobre la representación y la computación no serían adecuadas para explicar algunas cuestiones fundamentales referidas a la mente. Con todo, al evaluar los aciertos de las distintas teorías de la representación del conocimiento, advertiremos el gran aporte que ha hecho el MCRM. Sin duda, el modelo es el mejor a nivel teórico y experimental entre todos los enfoques del estudio de la mente que se han desarrollado hasta el momento. A pesar de que no todos los que trabajan en ciencias cognitivas están de acuerdo con los postulados del MCRM, la cantidad de artículos publicados en las principales revistas especializadas en psicología y otras disciplinas demuestra que, hoy en día, es el enfoque dominante.

Gran parte del éxito del MCRM se debe al hecho de que emplea una rica analogía del campo de la computación. Como veremos en el capítulo 5, muchas veces las analogías enriquecen las ideas científicas. En el tema que nos ocupa, la comparación de la mente con las computadoras permite un abordaje más fructífero del estudio de la mente que el de las analogías anteriores, como la del conmutador telefónico, por ejemplo. El lector que sepa de ciencias de la computación estará familiarizado con la caracterización de los programas como estructuras de datos y algoritmos. Los lenguajes de programación modernos incluyen una variedad de estructuras de datos formados por cadenas de letras, “abc”, por ejemplo, y números, además de por estructuras más complejas como las listas (A B C) y los árboles. Un algoritmo es un procedimiento mecánico que opera en varias clases de estructuras. Por ejemplo, en la escuela primaria los alumnos aprenden un algoritmo para dividir números de varias cifras. Otro algoritmo sencillo consiste en invertir el orden de una lista, es decir, cambiar el orden (A B C) por (C B A), procedimiento que se lleva a cabo mediante otros procedimientos más simples que consisten en tomar un

elemento de una lista y añadirlo al inicio de otra, de modo que la computadora arma una lista invertida y forma (A) en primer lugar, luego (B A) y por último (C B A). De modo similar, el MCRM adopta el supuesto de que en la mente existen representaciones mentales análogas a estructuras de datos y procesos computacionales semejantes a algoritmos tal como se muestra en el siguiente esquema:

<i>Programa</i>	<i>Mente</i>
estructuras de datos	representaciones mentales
+ algoritmos	+ procesos computacionales
= ejecución de programas	= pensamiento

Ésta ha sido la analogía dominante en las ciencias cognitivas, si bien dio un giro cuando se introdujo otro elemento: el cerebro. Desde el conexionismo se propusieron ideas novedosas sobre la representación y la computación que se valen de las neuronas y sus conexiones como modelo de estructuras de datos, y del estímulo de las células nerviosas y la diseminación de la actividad neuronal como modelo de algoritmos. Así, el MCRM recurre a una compleja analogía tripartita que vincula la mente, el cerebro y las computadoras, como se observa en la figura 1.1. La mente, el cerebro y la computación por separado sirven para sugerir nuevas ideas referidas a los otros dos elementos. Existe más de un modelo computacional de la mente, ya que las distintas clases de computadoras y de lenguajes de programación proporcionan distintos modelos de cómo podría funcionar la mente. Las computadoras a las que estamos acostumbrados cuentan con procesadores en serie, es decir, realizan de a una tarea por vez, pero el cerebro y algunas computadoras modernas tienen procesadores en paralelo, capaces de realizar varias operaciones en simultáneo.



**Figura 1.1.** Triple analogía entre la mente, la computadora y el cerebro.

Para quienes saben mucho de computación, concebir la mente en términos computacionales resulta algo natural, incluso si no están de acuerdo con la noción de que la mente es comparable a una computadora. Los que nunca han creado un programa de computación pero están acostumbrados a consultar libros de cocina pueden recurrir a otra analogía. Por lo general, las recetas constan de dos partes: una lista de ingredientes y una serie de instrucciones que indican cómo combinarlos. El plato final se obtiene cuando se siguen las instrucciones usando los ingredientes de la lista; del mismo modo, un programa ejecutable se obtiene cuando se aplican algoritmos a las estructuras de datos —números y listas— y el pensamiento es, de acuerdo con el MCRM, el resultado de la aplicación de procesos computacionales a las representaciones mentales. La analogía entre la receta y el pensamiento es algo débil, porque los ingredientes no son representaciones y las instrucciones para cocinar requieren a alguien que las interprete. En los capítulos 2 a 7 ofreceremos ejemplos de procesos computacionales que se ajustan mejor a las operaciones mentales.

## TEORÍAS, MODELOS Y PROGRAMAS

Los modelos de computación suelen ser útiles para la investigación teórica de los procesos mentales. Para comprender los modelos formulados por las ciencias cognitivas es necesario apreciar las distinciones y las relaciones entre cuatro elementos fundamentales: teoría, modelo, programa y plataforma. Una *teoría* cognitiva postula una serie de estructuras representacionales y un conjunto de procesos que operan sobre ellas. Un *modelo* computacional vuelve más precisos esos procesos y estructuras a través de interpretaciones que consisten en analogías con programas de computación formados por estructuras de datos y algoritmos. Las ideas vagas sobre las representaciones pueden complementarse con ideas computacionales precisas sobre estructuras de datos a la vez que los procesos mentales se definen por medio de algoritmos. La puesta a prueba del modelo se realiza implementándolo en un *programa* de software creado con un lenguaje de programación, como el LISP o el Java. El programa se ejecuta en distintas *plataformas* de hardware; por ejemplo, Macintosh, Sun Workstation o computadoras personales de IBM, aunque también pueden ejecutarse en plataformas especialmente diseñadas para un tipo particular de hardware con varios procesadores que funcionan en paralelo. Muchas clases de estructuras y procesos se investigan de esta manera, como es el caso de las reglas y las estrategias de búsqueda de determinados tipos de inteligencia artificial y el de las representaciones distribuidas y la propagación de los procesos de activación de las ideas conexionistas más recientes.

Imaginemos que deseamos entender cómo aprenden los niños a sumar. Utilizaremos el ejemplo de la operación de suma  $13 + 28$ . Una teoría cognitiva explicaría cómo representan los niños esos números y cómo procesan las representaciones para hacer

la suma. Además, determinaría si la representación del número 13 consiste en una estructura simple, una estructura combinada como *10 más 3* o un complejo de estructuras como las de las neuronas. La teoría también postularía procesos que operan sobre las estructuras para obtener el resultado, 41, y la operación que transforma la estructura  $30 + 11$  en 41. Un modelo computacional especificaría la naturaleza de las representaciones y los procesos mediante la caracterización de las estructuras y los algoritmos programables que serían análogos a las representaciones y los procesos mentales utilizados en la adición. Para evaluar la teoría y el modelo, se crea un programa de computación en un lenguaje como el LISP, y luego se ejecuta el programa para comparar su funcionamiento con el de las personas cuando realizan una suma. Es preciso que el programa obtenga el resultado correcto, que cometa los mismos tipos de error que las personas y que su ejecución pueda realizarse en plataformas comunes o en alguna diseñada especialmente para un tipo de computadora en particular; por ejemplo, una que imite la estructura neuronal del cerebro.

La analogía entre la mente y las computadoras resulta útil en las tres etapas del desarrollo de una teoría cognitiva: descubrimiento, modificación y evaluación. Por lo general, las ideas computacionales acerca de distintos tipos de programas sugieren nuevas clases de estructuras y procesos mentales. La elaboración de teorías y la creación de modelos y programas van de la mano, porque la escritura de un programa normalmente se acompaña de la creación de estructuras de datos y algoritmos nuevos, que a su vez se incorporan en el modelo y tienen su análogo en la teoría. Por ejemplo, cuando un programador escribe un programa de computación que simula la operación de suma, puede pensar en una clase de estructuras de datos que dé origen a nuevas ideas sobre la manera en que los niños represen-

tan los números. Del mismo modo, la evaluación de la teoría, el modelo y el programa a veces se realiza en conjunto, dado que la confianza en la teoría depende de la validez del modelo, observable en el desempeño del programa. Si el programa de computación diseñado para simular la adición no es capaz de sumar o si comete menos errores que los esperados, entonces la teoría cognitiva correspondiente es inadecuada.

El programa ejecutable puede contribuir a la evaluación del modelo y la teoría de tres modos distintos. En primer lugar, permite demostrar que las representaciones y los procesos postulados son realizables en el plano computacional, lo cual es importante, puesto que muchos algoritmos que a primera vista parecen razonables no se adecuan a los problemas de mayor envergadura que deben resolver las computadoras reales. En segundo lugar, para demostrar la posibilidad de llevar a la práctica una teoría y su plausibilidad psicológica, el programa debe probarse cualitativamente con distintos ejemplos de pensamiento. El programa de adición, por ejemplo, debería dar los mismos tipos de respuestas correctas e incorrectas que los niños. Y en tercer lugar, para demostrar que la teoría se adecua al pensamiento humano, el programa debe evaluarse cuantitativamente, es decir, tiene que generar predicciones detalladas sobre el pensamiento humano comparables con los resultados de los experimentos psicológicos. Si existen experimentos que demuestran que los niños dan un determinado porcentaje de respuestas correctas a cierta clase de problemas de adición, el programa de computación tendría que acertar las respuestas en porcentajes similares. Las teorías cognitivas por sí solas no tienen la precisión que permite hacer predicciones cuantitativas; para eso se precisan los modelos y los programas, que vinculan la teoría con los datos observados.

**Recuadro 1.1.** Criterios para la evaluación de teorías de la representación mental

---

- (1) Capacidad representacional
  - (2) Capacidad computacional
    - (a) Resolución de problemas
      - (i) Planificación
      - (ii) Decisión
      - (iii) Explicación
    - (b) Aprendizaje
    - (c) Lenguaje
  - (3) Plausibilidad psicológica
  - (4) Plausibilidad neurológica
  - (5) Aplicabilidad práctica
    - (a) Educación
    - (b) Diseño
    - (c) Sistemas inteligentes
    - (d) Enfermedades mentales
- 

EVALUACIÓN DE LOS ENFOQUES  
DE LA REPRESENTACIÓN MENTAL

Una vez presentados los lineamientos generales, seremos más específicos en cuanto a qué se espera de una teoría de la representación mental. En el recuadro 1.1 se muestran cinco criterios complejos para la evaluación de una determinada explicación de las representaciones y las operaciones computacionales con la cual se pretende explicar el pensamiento. En los capítulos 2 a 7 se ilustra el uso de esos criterios en la evaluación de seis enfoques distintos de la representación mental: la lógica, las reglas, los conceptos, las imágenes, los casos y las conexiones (redes neuronales artificiales).

Cada uno de esos enfoques propone un tipo distinto de representación y un conjunto de procesos computacionales propio.

El primer criterio –la capacidad representacional– se vincula con la cantidad de información que es capaz de expresar un tipo particular de representación. Por ejemplo, en una universidad se advierte: “Atención ingresantes: deben inscribirse en las materias antes del inicio de las clases”. Los estudiantes que tomen en serio la advertencia deberán contar con una representación interna que les permita hacer inferencias posteriores y llegar a la conclusión de que deben ir a la secretaría de admisiones e inscribirse en las materias del cuatrimestre siguiente. Veremos que los distintos tipos de representación mental propuestos varían en cuanto a su capacidad representacional.

Las representaciones mentales son importantes no sólo por lo que expresan sino también por lo que puede hacerse con ellas. La capacidad computacional de una representación mental se evalúa considerando cómo explica las tres clases de pensamiento superior. La primera es la resolución de problemas: una teoría de la representación mental debería explicar cómo razonan las personas para lograr sus objetivos. Hay al menos tres tipos de resolución de problemas que necesitan explicación: planificación, toma de decisiones y justificación. En la planificación se recurre al razonamiento para resolver cómo se pasa de un estado inicial a un estado final atravesando varios estados intermedios. Para ello, deben tenerse en cuenta cuestiones prácticas: una persona que va a tomar un vuelo considerará cómo llegar al aeropuerto a horario y un estudiante que debe resolver ejercicios de un libro o de un examen pensará cómo llegará a la solución leyendo la información proporcionada en el planteo del problema. El estado inicial comprende los conocimientos previos del estudiante y la información que figura en el planteo, y el estado final implica tener una respuesta. Para ello hay que encontrar la solución mediante la construcción de una secuencia de cálculos adecuada.

En la toma de decisiones, las personas tienen a su disposición distintos medios, de los cuales deben elegir el mejor para lograr sus objetivos. Pensemos, por ejemplo, en un estudiante que está por recibirse y tiene que elegir entre buscar un trabajo, comenzar el doctorado o seguir otros cursos de posgrado. No es una decisión fácil, ya que requiere definir objetivos y pensar en el curso de acción más adecuado. En la planificación, la tarea consiste en encontrar una secuencia de acciones; en cambio, en la toma de decisiones se trata de elegir entre una cantidad de acciones posibles para preparar el mejor plan.

La justificación implica descubrir *por qué* ha ocurrido cierto suceso. Las preguntas pueden ser desde una tan cotidiana como por qué un amigo al que estamos esperando para cenar aún no ha llegado, hasta una enmarcada dentro de las ciencias; por ejemplo, por qué evoluciona el lenguaje humano. Cualquier persona dotada de una inteligencia básica es capaz de planificar, tomar decisiones y encontrar justificaciones. Una teoría cognitiva debe tener la capacidad computacional suficiente para brindar explicaciones posibles de cómo se resuelve ese tipo de problemas.

La capacidad computacional de un sistema de representaciones y procesos no se limita a cuánto puede computar un sistema; también se debe considerar cuán eficiente es. Pensemos en un proceso cuya primera aplicación tarda un segundo, la segunda tarda el doble que la primera, la tercera tarda el doble que la anterior, y así sucesivamente. La vigésima vez que se aplique tardaría  $2^{20}$  segundos, más de los que han transcurrido desde que surgió el universo, hace unos 15.000 millones de años. Los sistemas inteligentes naturales y artificiales deben tener una velocidad tal que les permita operar con eficacia en su contexto particular.

Cuando alguien resuelve un problema, por lo general aprende de la experiencia; así, si se enfrenta otra vez a una situación semejante, le resultará más fácil encontrar la solución. Por ejemplo,

a los estudiantes que se anotan por primera vez en una materia todo les parece confuso porque no saben cuáles son los trámites que deben hacer ni cuáles son las mejores cátedras. Con el tiempo, todo se simplifica. En parte, ser inteligente consiste en aprender de la experiencia; entonces, una teoría de la representación mental debe tener una capacidad computacional que explique cómo aprenden los humanos. Los distintos enfoques de la representación mental dan cuenta de varias clases de aprendizaje, desde la adquisición de nuevos conceptos, como *anotarse* en una materia, y reglas como *No anotarse en cursos que empiecen a las 8.30*, hasta otras clases de adaptaciones sutiles.

Además de la resolución de problemas y el aprendizaje, una teoría cognitiva general debe explicar el uso del lenguaje humano. Nuestra especie es la única capaz de hacer un uso complejo del lenguaje. Es probable que eso se deba a los principios generales aplicados en la resolución de problemas y en el aprendizaje, pero también podría ser que el lenguaje fuera una capacidad cognitiva única y que hubiese que considerarla por separado. Es necesario explicar al menos tres aspectos del uso del lenguaje: la capacidad de comprensión, de generar oraciones y la capacidad universal de adquisición del lenguaje durante la infancia. Los distintos enfoques de la representación del conocimiento dan respuestas muy diferentes a cómo funcionan esos procesos.

Como rama de la ingeniería, la inteligencia artificial desarrolla modelos computacionales para la resolución de problemas, el aprendizaje y el uso del lenguaje en los que no se tiene en cuenta cómo se llevan a cabo esos procesos en la mente humana. Lo que importa es lograr que las computadoras los lleven a cabo. En cambio, las ciencias cognitivas intentan comprender la cognición *humana*, de modo que es fundamental que una teoría de la representación mental refleje cómo piensan los humanos, más allá de tener una gran capacidad representacional y

computacional. En consecuencia, el tercer criterio para evaluar una teoría de la representación mental es la plausibilidad psicológica, es decir, es necesario dar cuenta no sólo de las capacidades cualitativas de los humanos sino también de los resultados cuantitativos de los experimentos psicológicos por medio de los cuales se estudian esas capacidades. Los experimentos en cuestión evalúan las mismas tareas superiores que comentamos para la capacidad computacional: la resolución de problemas, el aprendizaje y el lenguaje. La diferencia radica en que una teoría cognitiva de la representación mental debe demostrar cómo es posible realizar una tarea en el plano computacional, pero además tiene que poder explicar cómo la realizan los humanos.

Del mismo modo, como el pensamiento se produce en el cerebro, una teoría de la representación mental debe corresponderse con los resultados de los experimentos de la neurociencia. Hasta no hace mucho tiempo, las técnicas neurológicas, como el registro electroencefalográfico de las ondas cerebrales, poco decían de la cognición superior, pero en los últimos veinte años han surgido nuevas técnicas de escaneo cerebral que permiten detectar el momento y las zonas del cerebro en que se llevan a cabo ciertas tareas cognitivas. Así, la neurociencia cognitiva se ha convertido en una disciplina científica de gran importancia para la reflexión sobre el funcionamiento de la mente. Por lo tanto, si bien aún no hay información suficiente sobre cómo opera el cerebro a nivel cognitivo, es necesario evaluar los enfoques de la representación del conocimiento en términos de plausibilidad neurológica (capítulo 9).

El quinto y último criterio a la hora de evaluar teorías de la representación mental es la aplicabilidad práctica. Aunque el objetivo principal de las ciencias cognitivas es llegar a comprender la mente, se espera que sus resultados sean aplicables. En este libro presentaremos las consideraciones de los distintos enfoques de