

Computational Intelligence: a Logical Approach

David Poole

Alan Mackworth

Randy Goebel

Oxford University Press (1998)

Por:

Matías Alvarado*

Centro de Investigación en Computación

Instituto Politécnico Nacional

Inteligencia Computacional

La obra que aquí se reseña es un estudio –en forma de libro de texto– de la ciencia de la Inteligencia Artificial a la luz de la metodología, teoría y desarrollo de *Sistemas de Agentes*, en boga en los años recientes. El nombre que los autores utilizan para esta disciplina es Inteligencia Computacional (*Computational Intelligence*), fundamentalmente, por las siguientes razones (Capítulo 1).

- Hacer énfasis, desde el nombre, del uso intensivo que se hace de la computación.
- Evitar las confusiones a las cuales ha dado y da lugar el nombre Inteligencia Artificial.

La primera razón es esencial para la Inteligencia Artificial (AI), la cual comparte, como objeto de estudio, con la psicología, la filosofía y las ciencias cognitivas en general, la comprensión de lo que es la inteligencia y el conocimiento. La diferencia estriba en que la IA, que en lo sucesivo llamaremos Inteligencia Computacional (IC), es una ciencia que se aboca, en gran medida, al diseño y puesta en marcha de artefactos y dispositivos, *agentes*, que hacen un uso intensivo de la computación, y cuyo comportamiento pueda catalogarse como *inteligente*.

La segunda razón de utilizar Inteligencia Computacional en vez de Inteligencia Artificial, tiene que ver con que la investigación y desarrollo en la IC, no conlleva, necesariamente, la comprensión de lo que es la inteligencia humana, los cuales son objetivos declarados de la psicología, la epistemología y las neurociencias en general. Más aún, tampoco se asume el que en los artefactos y dispositivos creados en la IC se tenga que hacer una emulación de la inteligencia humana, lo cual no

significa que no se asuman y utilicen algunos resultados que hayan emergido de estas ciencias. El punto es que la IC no asume los límites propios de los objetivos de esas disciplinas en lo que se refiere a la comprensión de la inteligencia, sino que en la IC se diseñan y ponen en marcha agentes computacionales cuya interacción con el mundo real pueda considerarse inteligente. Ello sin reparar en la similitud que puedan tener con los procesos de la inteligencia humana.

Resumiendo, en la Inteligencia Computacional, se pretende el diseño e implantación de agentes computacionales capaces de incluir las siguientes características:

- Capacidad de percibir el entorno.
- Capacidad de razonar sobre la información que proviene del entorno y sobre su propia información.
- Capacidad de aprender de experiencias previas.
- Capacidad de actuar oportunamente en el entorno.

Sin duda, un dispositivo que reúne las características puntualizadas puede considerarse inteligente.

Agentes

Para la Inteligencia Computacional, un agente es un ente, ya sea natural, artificial o un híbrido natural-artificial, que se relaciona en un entorno. De acuerdo con esto, un agente debe ser capaz de recibir y/o percibir la información del entorno, procesarla y razonar a partir de ella, para después actuar de manera correcta. A la información que posee un agente para actuar en su entorno y realizar sus objetivos, en la IC se le llama conocimiento.

Más allá de los agentes desarrollados la IC, los hay desde muy simples tales como un termostato o un gusano, hasta muy complejos como un avión o el ser humano. Típicamente,

* Agradezco a Francisco Cantú Ortiz el haberme recomendado este libro del Centro de Inteligencia Artificial del ITESM, México.

entre más complejo es un agente, mayor grado de inteligencia muestra. Compárese, por ejemplo, un gusano respecto a cualquier vertebrado superior, y cualesquiera de estos con respecto al hombre. También hay agentes sociales tales como una colonia de hormigas o una sociedad humana, los cuales son más complejos e inteligentes que cualesquiera de los agentes individuos que las componen: una colonia de hormigas o una sociedad humana tienen mayor capacidad de elaboración, desarrollo y adaptación respecto al entorno que cualquiera de los individuos.

Para facilitar la comprensión de los conceptos, metodologías y técnicas en la IC, de manera gradual se desarrolla en el libro la implementación de los siguientes agentes, los cuales actúan en distintos entornos:

- Un robot que realiza entregas.
- Un sistema que apoya la realización de diagnósticos médicos.
- Un robot que colecta información en el World Wide Web (infobot).

El robot que realiza entregas (Capítulo 4), ya sea de cartas, documentos y café, actúa en el entorno físico que corresponde a un conjunto de oficinas donde trabajan personas. Por tanto, debe ser capaz de preguntar, recibir respuestas, trazar trayectorias, evitar obstáculos, etc. Este robot es un agente que realiza tareas sencillas para un ser humano (el reparto de café a solicitud expresa), pero requiere de una actualización continua sobre las condiciones del entorno físico para realizar su labor adecuadamente y sin causar daños a otros o a él mismo. Necesita por tanto de sensores, mecanismos de razonamiento y dispositivos para coger y depositar objetos.

El sistema que apoya la realización de diagnósticos, es un Sistema Experto que posee información sobre enfermedades, causas que las originan y terapias que puedan aplicarse para la recuperación de pacientes que las padezcan. Este sistema debe ser capaz de interactuar con el médico al que auxilia en la determinación del diagnóstico, proporcionando una explicación razonada y coherente de los motivos que lo llevan a dar un diagnóstico y las terapias que sugiere para la curación. Este agente trabaja sobre el dominio especializado de las enfermedades y esta subordinado a la supervisión que el experto médico humano realiza de él de manera continua; la interacción con el humano es a un nivel simbólico netamente.

Finalmente, el infobot es un agente que busca información en el WWW sobre algún tema especificado por un usuario y actúa en un ambiente netamente computacional —el web, sin contacto con el mundo físico. A diferencia del sistema experto, sus actuaciones no están supervisadas por un experto humano: el usuario de este agente, sólo recoge la información que el infobot ha recolectado en el web.

Cada uno de estos agentes está especializado para actuar en un dominio particular y resulta del todo inútil en un dominio

diferente para el cual fue diseñado *ex profeso*. La cuestión que surge inmediatamente es si conviene diseñar e implementar agentes universales, capaces de actuar inteligentemente en cualquier ambiente. Dado el estado de desarrollo de la IC, donde el gasto de recursos computacionales, para cualesquiera de los robots mencionados es muy alto, la pretensión de construir un agente universal es impracticable, actualmente. Hoy en día la tendencia es construir agentes especializados en determinada labor; si bien, hay principios y métodos generales para el diseño y puesta en marcha de cualquier agente, sobre los cuales se comenta enseguida.

Los temas que se comentan son la Representación y Razonamiento del Conocimiento, Búsqueda y Planificación, Lógica Modal, Conocimiento con Incertidumbre y Razonamiento basado en Suposiciones; finalmente, hay una sección dedicada a Aprendizaje Automático. Temas relacionados con estos, se comentan a lo largo del libro. Todo este material, clásico en Inteligencia Computacional, es estudiado desde la perspectiva de los Sistemas Multi-Agentes.

Representación y Razonamiento

Para la implantación de un agente con las características antes puntualizadas, y tal que una parte o el total del mismo esté basado en sistemas computacionales, se requiere lo siguiente:

- Una representación del conocimiento.
- Mecanismos de razonamiento.

El estudio de las herramientas desarrollados para la representación y razonamiento, al nivel conceptual y teórico, corresponde al aspecto científico de la IC y se comentan en el Capítulo 5, en tanto que las propias del desarrollo y puesta en marcha de los agentes, correspondientes por tanto a la ingeniería y tecnología de la IC, se desarrollan en el Capítulo 6.

Inicialmente, se asume un agente tal que su conocimiento del mundo es completo y el mundo no cambia (Capítulos 2 y 3). Bajo tales condiciones, el conocimiento del agente es definido, sin ambigüedades. Un caso de estudio es el de un agente encargado del mantenimiento de la instalación eléctrica de una casa, en la cual la información sobre los causantes de fallos y soluciones a ellos es información bien definida. Las operaciones sobre una base de datos, tales como unión, proyección, etc., es un ejemplo más de actuación de un agente bajo estos supuestos. Una parte del formalismo lógico del Cálculo de Predicados, a saber, el uso de predicados con variables acotadas y de formulas conjuntivas afirmadas, resulta adecuada para la representación, tanto sintáctica como semántica, de este tipo de situaciones.

Para la representación de conocimiento definido, se supone que los objetos del dominio tienen un nombre único (unique name assumption); asimismo, que no hay más objetos en el

dominio que los del mundo actual (closed world assumption). Los formalismos para la representación y razonamiento bajo estos supuestos de nombre único y mundo cerrado fueron desarrollados a finales de la década de los setenta e inicios de la de los ochenta (Capítulo 7).

Sin embargo, es frecuente que los objetos no sean nombrados de manera unívoca y, mucho más importante, que los objetos del dominio no sólo sean los del mundo actual. De hecho un agente percibe, razona y actúa a través del tiempo, y las condiciones del entorno (el mundo) y sus propios conocimientos y creencias cambian conforme transcurre el tiempo (Capítulo 7). Esto hace que un agente no necesariamente tenga conocimiento definido sobre el entorno ni sobre las consecuencias de sus acciones. Por tanto, sus mecanismos de representación y razonamiento han de ser aptos para manipular información no totalmente definida, incierta, ambigua. Las formulas disyuntivas del cálculo de predicados y el uso de las variables libres, fueron las primera herramientas formales utilizada para la representación de conocimiento ambiguo, dado que permiten la expresión de alternativas, y la posibilidad de representar de manera genérica los elementos del dominio que satisfacen una fórmula, respectivamente.

Para la representación de agentes dinámicos que actúan y cambian conforme transcurre el tiempo, los formalismos lógicos más conocidos son el Cálculo de Situaciones (*Situational Calculus*), el Cálculo de Eventos y el lenguaje STRIPS (Capítulo 8). En el Cálculo de Situaciones y STRIPS, la representación de las acciones y cambios realizados por el agente, se basa en la traslación entre estados: una acción lleva de un estado a otro. En el Cálculo de Eventos, el tiempo, continuo o discreto, se considera de manera explícita y permite modelar acciones y eventos externos pero concurrentes al accionar del agente. El agente reactivo que se desarrolla en el Capítulo 12 del libro está basado en el Cálculo de Eventos.

Búsqueda y Planificación

El conjunto de estados y eventos que soportan la modelación de las acciones y cambios de un agente, y como consecuencia los cambios del mundo en el que actúa, conforman espacios de *Búsqueda y Planificación*. La búsqueda es sobre el espacio de posibles soluciones a los problemas planteados al agente en el logro de sus objetivos. Hay un conjunto de métodos y técnicas, clásicos la mayoría, para optimizar la búsqueda en el espacio de soluciones, los cuales se comentan en el Capítulo 4 del libro. Los métodos clásicos de búsqueda, usualmente mantienen un único estado actual, en tanto que métodos recientes son capaces de mantener varios estados actuales, aunque sin interacción entre ellos. Métodos basados en principios estocásticos son capaces de modelar búsquedas con varios estados actuales que interactúan entre sí. La

interacción es similar a la de *Algoritmos Genéticos* en cuanto a la combinación aleatoria de estados actuales (genes padres) que dan lugar a nuevos estados (genes hijos), cuyas variables hereden combinaciones de valores de las de los padres.

Para la obtención de los objetivos (sobre todo de mediano y largo alcance), se requiere la elaboración de planes, es decir, determinar la sucesión de acciones a realizar para lograrlos por parte del agente (Capítulo 8). Usualmente, un plan se lleva a cabo a partir de un estado inicial del cual se siguen trayectorias de estados, dependiendo de las acciones realizadas de un estado a otro. Para la obtención de una meta (objetivo) puede haber planes dados de antemano, o más de un plan, en cuyo caso se hace un ordenamiento parcial de las distintas sucesiones de acciones mediante los cuales la meta puede lograrse.

Lógica Modal

Si bien el lenguaje utilizado en los comienzos de la IC fue el de la lógica de proposiciones y de predicados, dada la complejidad de los dominios de estudio de la IC surgieron otros lenguajes con mejor *capacidad expresiva*, es decir, capaces de expresar con mayor claridad y precisión las relaciones entre los objetos de un dominio. El lenguaje de la lógica modal, de uso frecuente en IC para fines de representación, incorpora al lenguaje del cálculo de predicados *operadores modales*, los cuales hacen posible el expresar *modalidades de la verdad*, a saber, necesidad y posibilidad, conocimiento y creencia, o nociones temporales tales como siempre o algunas veces. La semántica de mundos posibles, que es la propia de la lógica modal, hizo posible el hablar no sólo del mundo (escenario, estado) actual (real), sino de cualquiera posible a partir de él. Esta capacidad expresiva de la semántica de mundos posibles ha hecho que la lógica modal sea muy utilizada para fines de representación y razonamiento en IC.

Conocimiento con Incertidumbre

La representación y razonamiento sobre conocimiento con incertidumbre donde interviene el azar (Capítulo 10), se basa en la Teoría de Probabilidades. Más aún, se propone una *semántica probabilística de mundos posibles*. Este híbrido formal aprovecha las ventajas expresivas de la Semántica de Mundos Posibles de Kripke y Hintikka, junto con la sólida fundamentación matemática de la Teoría de Probabilidades para la representación de información con incertidumbre o aleatoria. Las *redes de creencias*, también llamadas *causales* o *bayesianas*, utilizan la probabilidad condicional para determinar la independencia (dependencia) local y bajo ciertas condiciones de una variable aleatoria con respecto a otras. Es común la aplicación de estas técnicas en el desarrollo de sistema experto, como el del asistente para la elaboración de diagnóstico.

Razonamiento Basado en Suposiciones

Es común en la vida diaria hacer inferencias a partir de información *segura* junto con suposiciones basadas en evidencias no definitivas: no está una persona si no se le ha visto, o si todas las aves que conocemos vuelan concluimos que todas las aves vuelan, etc. Una característica del razonamiento usado en suposiciones por *default reasoning* -presente en prácticamente cualquier razonamiento de sentido común- es su no monotonía, es decir, no necesariamente al añadir información las conclusiones se mantienen. En los ejemplos anteriores respectivamente nos enteramos que la persona a quien no hemos visto estaba en otra oficina; se conoce a una garza o un pingüino, aves que no vuelan. En la IC, un agente inteligente debe incluir mecanismos que le permitan realizar conclusiones plausibles a partir de información incompleta, y por otro lado, la posibilidad de retractarse de ellas en presencia de información adicional que las refute (capítulo 9).

El llamado *razonamiento evidencial*, a partir de observaciones, busca establecer una teoría sobre ellas; tal tipo de razonamiento es el que se aplica en la elaboración de diagnósticos. Una formalización del mismo es por *abducción*, la cual busca establecer las explicaciones y causas a partir de los efectos. Los mecanismos de razonamiento híbridos, utilizan ambos tipos de razonamiento, el evidencial seguido del causal; este último es el que obtiene conclusiones, por deducción, a partir de una teoría.

Aprendizaje

La capacidad de aprender es una característica esencial de un comportamiento inteligente. En Inteligencia Computacional, el aprendizaje consiste en el aprovechamiento de experiencia previa para mejorar el rendimiento de un agente (Capítulo 11). Varios paradigmas de aprendizaje se han desarrollado en Inteligencia Computacional. Un enfoque clásico es considerar el aprendizaje un problema de clasificación.

El agente computacional aprende, si a partir de una clasificación dada es capaz de clasificar objetos que no estén actualmente considerados en las clases. La clasificación se implementa de la siguiente manera: los objetos de un dominio, definidos mediante parejas de atributo-valor, conforman una estructura arborecente, donde un atributo es un nodo del árbol y los valores para el atributo son las ramas que parten de ese nodo. Las hojas son las clases en que se distribuyen los objetos. A partir de un árbol inicial, llamado *árbol de decisión*, los nuevos objetos se distribuyen en las clases de acuerdo a los valores de sus atributos. El punto esencial del aprendizaje radica en si el sistema de reglas (una para cada clase) asociado al árbol de decisión, es capaz de generar nuevas reglas (nuevas clases) para poder clasificar objetos no considerados en la clasificación actual. Este paradigma, de gran eficacia, está en el núcleo de diversos sistemas que involucran actividades de clasificación y aprendizaje. Una

generalización de los árboles de decisión son las *redes neuronales*. Otro paradigma importante es el aprendizaje con razonamiento basado en casos (*case-based reasoning*), en el cual, las soluciones dadas a un problema, se archivan y, cuando se presenta un nuevo problema, se revisan las soluciones dadas a problemas similares al actual y a partir de ellas se elabora la que pueda resolver el problema en curso. Tal tipo de aprendizaje se aplica en agentes que apoyan actividades en áreas tan diversas como medicina, ingeniería o composición musical.

El aprendizaje en un dominio de información con incertidumbre utiliza métodos de la teoría de probabilidades y la implementación de agentes está basada en redes bayesianas (causales). Otros paradigmas de aprendizaje son el *basado en explicaciones* y el de *espacio de versiones*.

Con esto se finalizan los comentarios sobre los principales temas que se abordan en el libro.

Conclusiones

Una gran virtud del libro es que el material teórico que sirve de fundamento a los conceptos centrales, se introduce conforme la necesidad lo hace evidente: las formulas afirmadas y conjuntivas del cálculo de predicados para formalizar la representación y razonamiento sobre conocimiento definido; las fórmulas disyuntivas del mismo cálculo cuando los agentes razonan con conocimiento ambiguo; la Teoría de Probabilidades cuando se aborda el uso de conocimiento con incertidumbre; los conceptos de lógica modal cuando los agentes están inmersos en mundos que involucran nociones epistémicas, temporales, dinámicas, etcétera.

Es notable el equilibrio que se mantiene al tratar los aspectos científicos respecto a los tecnológicos de la Inteligencia Computacional: se desarrollan los conceptos teóricos en lo que se refiere a la comprensión de lo que es la inteligencia y el conocimiento; de manera paralela se comentan las técnicas en la IC para el desarrollo de agentes que apoyan la solución de problemas concretos. De hecho, los tres agentes que principalmente ejemplifican los conceptos, son una muestra de ello.

La obra se recomienda ampliamente, para una introducción didáctica y moderna a la tradicionalmente llamada Inteligencia Artificial. Algunas obras que en lo general complementan los temas tratados son las siguientes:

Bibliografía

- Russell, S. & Norvig, P. *Artificial Intelligence: a Modern Approach*, Series in Artificial Intelligence, Prentice Hall 1995.
Huns, M. & Readings in Agents, Morgan Kaufmann 1997.

