



Agentes lógicos

Donde diseñaremos agentes que pueden construir representaciones del mundo, utilizar un proceso de inferencia para derivar nuevas representaciones del mundo, y emplear éstas para deducir qué hacer.

En este capítulo se introducen los agentes basados en conocimiento. Los conceptos que discutiremos (la *representación* del conocimiento y los procesos de *razonamiento* que permiten que éste evolucione) son centrales en todo el ámbito de la inteligencia artificial.

De algún modo, las personas conocen las cosas y realizan razonamientos. Tanto el conocimiento como el razonamiento son también importantes para los agentes artificiales, porque les permiten comportamientos con éxito que serían muy difíciles de alcanzar mediante otros mecanismos. Ya hemos visto cómo el conocimiento acerca de los efectos de las acciones permiten a los agentes que resuelven problemas actuar correctamente en entornos complejos. Un agente reflexivo sólo podría hallar un camino de Arad a Bucarest mediante la suerte del principiante. Sin embargo, el conocimiento de los agentes que resuelven problemas es muy específico e inflexible. Un programa de ajedrez puede calcular los movimientos permitidos de su rey, pero no puede saber de ninguna manera que una pieza **no** puede estar en dos casillas diferentes al mismo tiempo. Los agentes basados en conocimiento se pueden aprovechar del conocimiento expresado en formas muy genéricas, combinando y recombinando la información para adaptarse a diversos propósitos. A veces, este proceso puede apartarse bastante de las necesidades del momento (como cuando un matemático demuestra un teorema o un astrónomo calcula la esperanza de vida de la Tierra).

El conocimiento y el razonamiento juegan un papel importante cuando se trata con entornos parcialmente observables. Un agente basado en conocimiento puede combinar

el conocimiento general con las percepciones reales para inferir aspectos ocultos del estado del mundo, antes de seleccionar cualquier acción. Por ejemplo, un médico diagnóstica a un paciente (es decir, infiere una enfermedad que no es directamente observable) antes de seleccionar un tratamiento. Parte del conocimiento que utiliza el médico está en forma de reglas que ha aprendido de los libros de texto y sus profesores, y parte en forma de patrones de asociación que el médico no es capaz de describir explícitamente. Si este conocimiento está en la cabeza del médico, es su conocimiento.

El entendimiento del lenguaje natural también necesita inferir estados ocultos, en concreto, la intención del que habla. Cuando escuchamos, «John vió el diamante a través de la ventana y lo codició», sabemos que «lo» se refiere al diamante y no a la ventana (quizá de forma inconsciente, razonamos con nuestro conocimiento acerca del papel relativo de las cosas). De forma similar, cuando escuchamos, «John lanzó el ladrillo a la ventana y se rompió», sabemos que «se» se refiere a la ventana. El razonamiento nos permite hacer frente a una variedad virtualmente infinita de manifestaciones utilizando un conjunto finito de conocimiento de sentido común. Los agentes que resuelven problemas presentan dificultades con este tipo de ambigüedad debido a que su representación de los problemas con contingencias es inherentemente exponencial.

Nuestra principal razón para estudiar los agentes basados en conocimiento es su flexibilidad. Ellos son capaces de aceptar tareas nuevas en forma de objetivos descritos explícitamente, pueden obtener rápidamente competencias informándose acerca del conocimiento del entorno o aprendiéndolo, y pueden adaptarse a los cambios del entorno actualizando el conocimiento relevante.

En la Sección 7.1 comenzamos con el diseño general del agente. En la Sección 7.2 se introduce un nuevo entorno muy sencillo, el mundo de *wumpus*, y se muestra la forma de actuar de un agente basado en conocimiento sin entrar en los detalles técnicos. Entonces, en la Sección 7.3, explicamos los principios generales de la **lógica**. La lógica será el instrumento principal para la representación del conocimiento en toda la Parte III de este libro. El conocimiento de los agentes lógicos siempre es *categorico* (cada proposición acerca del mundo es verdadera o falsa, si bien, el agente puede ser agnóstico acerca de algunas proposiciones).

La lógica presenta la ventaja pedagógica de ser un ejemplo sencillo de representación para los agentes basados en conocimiento, pero tiene serias limitaciones. En concreto, gran parte del razonamiento llevado a cabo por las personas y otros agentes en entornos parcialmente observables se basa en manejar conocimiento que es *incierto*. La lógica no puede representar bien esta incertidumbre, así que trataremos las probabilidades en la Parte V, que sí puede. Y en la Parte VI y la Parte VII trataremos otras representaciones, incluidas algunas basadas en matemáticas continuas como combinaciones de funciones Gaussianas, redes neuronales y otras representaciones.

En la Sección 7.4 de este capítulo se presenta una lógica muy sencilla denominada **lógica proposicional**. Aunque es mucho menos expresiva que la **lógica de primer orden** (Capítulo 8), la lógica proposicional nos permitirá ilustrar los conceptos fundamentales de la lógica. En las secciones 7.5 y 7.6 describiremos la tecnología, que está bastante desarrollada, para el razonamiento basado en lógica proposicional. Finalmente, en la sección 7.7 se combina el concepto de agente lógico con la tecnología de la lógica proposicional para la construcción de unos agentes muy sencillos en nuestro ejemplo