

Agentes inteligentes (racionales)

La observación de que algunos agentes se comportan mejor que otros nos lleva naturalmente a la idea de agente racional, aquel que se comporta lo mejor que puede. La forma de actuar del agente depende de la naturaleza del medio, algunos hábitats son más complejos que otros.

Un agente es cualquier cosa capaz de percibir su medio ambiente con la ayuda de **sensores** y actuar en ese medio utilizando **actuadores**.

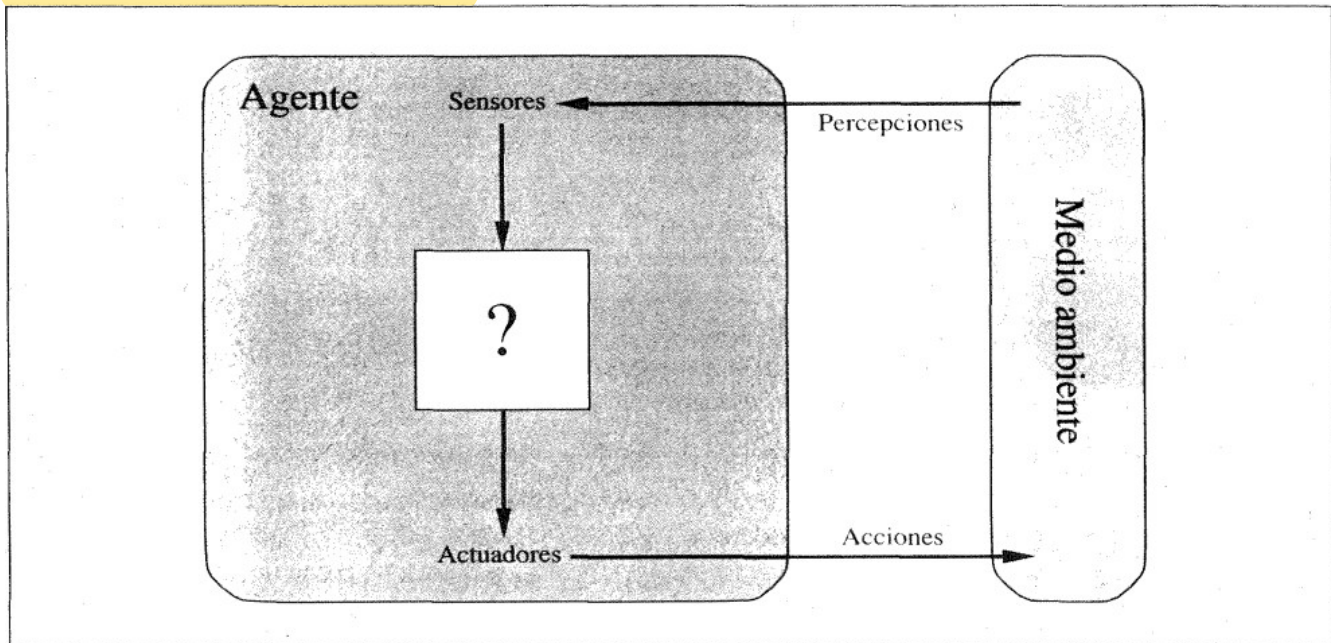


Figura 2.1 Los agentes interactúan con el medioambiente mediante sensores y efectores.

Un agente humano tiene ojos, oídos y otros órganos sensoriales además de manos, piernas, boca y otras partes del cuerpo para actuar. Un agente robot recibe pulsaciones del teclado, archivos de información y paquetes vía red a modo de entradas sensoriales y actúa sobre el medio con mensajes en el monitor, escribiendo ficheros y enviando paquetes por la red.

El término **percepción** se utiliza en este contexto para indicar que el agente puede recibir entradas en cualquier instante. La secuencia de percepciones de un agente refleja el historial completo de lo que el agente ha recibido. En general, un agente tomará una decisión en un momento dado dependiendo de la secuencia completa de percepciones hasta ese instante.

La función que describe el comportamiento de un agente se puede presentar en forma de tabla, en la mayoría de los casos esta tabla sería muy grande (hasta infinita).

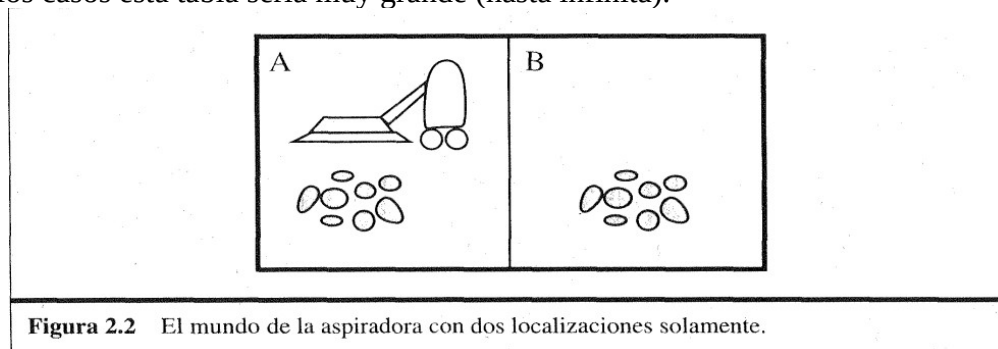


Figura 2.2 El mundo de la aspiradora con dos localizaciones solamente.

Secuencia de percepciones	Acción
[A, Limpio]	Derecha
[A, Sucio]	Aspirar
[B, Limpio]	Izquierda
[B, Sucio]	Aspirar
[A, Limpio], [A, Limpio]	Derecha
[A, Limpio], [A, Sucio]	Aspirar
—	—
—	—
—	—
[A, Limpio], [A, Limpio], [A, Limpio]	Derecha
[A, Limpio], [A, Limpio], [A, Sucio]	Aspirar
—	—
—	—
—	—

Figura 2.3 Tabla parcial de una función de agente sencilla para el mundo de la aspiradora que se muestra en la Figura 2.2.

Revisando la tabla, se aprecia que se pueden definir varios agentes para el mundo de la aspiradora simplemente rellorando la columna de la derecha de formas distintas. La pregunta obvia, entonces es: ¿Cual es la mejor forma de rellorar la tabla? En otras palabras, ¿qué hace que un agente sea bueno o malo, inteligente o no?

Un agente racional es aquel que hace lo correcto, en términos conceptuales, cada elemento de la tabla que define la función del agente se tendría que rellorar correctamente. Pero ¿qué significa hacer lo correcto? Como primera aproximación, se puede decir que lo correcto es aquello que permite al agente obtener un resultado mejor. Por tanto, se necesita determinar una forma de medir el éxito.

Medidas de rendimiento

Las medidas de rendimiento incluyen los criterios que determinan el éxito en el comportamiento del agente. Cuando se sitúa un agente en un medio, este genera una secuencia de acciones de acuerdo con las percepciones que recibe. Esta secuencia de acciones hace que su hábitat pase por una secuencia de estados. Si las acciones son las deseadas, entonces el agente habrá actuado correctamente. Obviamente, no hay una única medida adecuada para todos los agentes. Se puede preguntar al agente por su opinión subjetiva acerca de su propia actuación, pero muchos agentes serían incapaces de contestar, y otros podrían engañarse a sí mismos. Por tanto hay que insistir en la importancia de utilizar medidas de rendimiento objetivas, que normalmente determinará el diseñador encargado de la construcción del agente.

Si retomamos el ejemplo de la aspiradora, se puede proponer utilizar como medida de rendimiento la cantidad de suciedad limpiada en un período de tiempo. Sin embargo, un agente racional puede maximizar su medida de rendimiento limpiando la suciedad, tirándola al suelo, limpiándola de nuevo, y así sucesivamente. Una medida de rendimiento más adecuada recompensaría al agente por tener el suelo limpio.

La selección de la medida de rendimiento no siempre es fácil de elegir. Por ejemplo la noción del suelo limpio en el ejemplo anterior está basada en un nivel de limpieza promedio a lo largo del tiempo. Además, dicho nivel promedio de limpieza se puede alcanzar de diferentes formas, por ejemplo, llevando a cabo una limpieza mediocre pero continua, o bien limpiar en profundidad y tomar largos descansos. La manera más adecuada de hacerlo podría ser tomada de la opinión de un encargado de limpieza profesional. Pero en realidad es una cuestión filosófica con fuertes implicaciones. ¿Qué es mejor, una vida temeraria con altos y bajos, o una existencia segura pero aburrida?

Racionalidad

La racionalidad en un momento determinado depende de cuatro factores:

1. La medida de rendimiento que define el criterio de éxito
2. El conocimiento previo del agente sobre el medio en el que habita
3. Las acciones que el agente puede llevar a cabo
4. La secuencia de percepciones del agente hasta ese momento

Esto nos lleva a la definición de un agente racional:

En cada posible secuencia de percepciones, un agente racional deberá emprender aquella acción que se espera maximice su medida de rendimiento, basándose en las evidencias aportadas por las percepciones y el conocimiento que el agente tiene almacenado.

Regresando al caso de la aspiradora, considere lo siguiente:

- La medida de rendimiento premia con un punto al agente por cada recuadro limpio en un periodo de tiempo concreto.
- La geografía del medio se conoce a priori, pero la distribución de la suciedad y la posición inicial del agente se desconoce. Las acciones izquierda y derecha, mueven al agente hacia la izquierda y derecha respectivamente, excepto en el caso en que esto lleve al agente fuera del medio.
- Las únicas acciones permitidas son: izquierda, derecha, aspirar y NoOp (no hacer nada).
- El agente percibe correctamente su localización, y si ésta contiene o no suciedad.

Bajo estas circunstancias se puede afirmar que el agente actúa de manera racional. Sin embargo el agente se puede volver irracional en algunas circunstancias. Por ejemplo, cuando toda la suciedad se haya eliminado el agente oscilara innecesariamente de izquierda a derecha. Un agente más eficiente no haría nada si está seguro de que todas las cuadrículas están limpias. El agente haría revisiones periódicas para identificar si alguna cuadrícula se ha ensuciado y limpiarla.

Omnisciencia, aprendizaje y autonomía

La definición propuesta de racionalidad no requiere omnisciencia, ya que la elección racional depende sólo de la secuencia de percepción hasta la fecha. Es necesario asegurarse de no haber permitido, por descuido, que el agente se dedique decididamente a llevar a cabo acciones poco inteligentes. Por ejemplo, si el agente no mirase a ambos lados de la calle antes de cruzar una calle muy concurrida, entonces su secuencia de percepción no le indicaría que se está acercando un gran camión a gran velocidad. ¿La definición de racionalidad nos está indicando que está bien cruzar la calle? ¡Todo lo contrario! Primero, no sería racional cruzar la calle sólo teniendo una secuencia de percepciones incompleta: el riesgo de accidente al cruzar sin mirar es demasiado grande. Segundo, un agente racional debe elegir la acción de mirar antes de intentar cruzar la calle, ya que el mirar maximiza el rendimiento esperado. Llevar a cabo acciones con la intención de modificar percepciones futuras, proceso comúnmente denominado recopilación de información, es una parte importante de la racionalidad.

La definición propuesta implica que el agente racional no sólo recopile información, sino que aprenda lo máximo posible de lo que está percibiendo. La configuración inicial del agente puede reflejar un conocimiento preliminar del entorno, pero a medida que el agente adquiere experiencia éste puede modificarse y aumentar.

Los agentes con éxito dividen las tareas de calcular la función del agente en tres periodos diferentes: cuando se está diseñando el agente, cuando está pensando en la siguiente operación, y cuando está aprendiendo de la experiencia.

Se dice que un agente carece de autonomía cuando se apoya más en el conocimiento inicial que le proporciona su diseñador que en sus propias percepciones. Un agente racional debe ser autónomo, debe

saber aprender a determinar cómo tiene que compensar el conocimiento incompleto o parcial inicial. Por ejemplo, el agente aspiradora que aprenda a prever dónde y cuando aparecerá suciedad adicional lo hará mejor que otro que no aprenda. De la misma forma que la evolución proporciona a los animales sólo los reflejos necesarios para que puedan sobrevivir lo suficiente para aprender por ellos mismos, sería razonable proporcionar a los agentes que disponen de inteligencia artificial un conocimiento inicial, así como de la capacidad de aprendizaje.

Especificación del entorno de trabajo

En la discusión de la racionalidad de un agente aspiradora simple, hubo que especificar las medidas de rendimiento, el entorno, y los actuadores y sensores del agente. Todo ello forma lo que se llama el entorno de trabajo. En el diseño de un agente, el primer paso debe ser siempre especificar el entorno de trabajo de la forma más completa posible.

Considérese ahora un problema más complejo: un taxista automático. Primero, ¿cuál es la **medida de desempeño** que queremos que tenga el taxista automático? Dentro de las cualidades deseables que debería tener se incluyen el que llegue al destino correcto, que minimice el consumo de combustible, que minimice el tiempo de viaje y/o coste, que minimice el número de infracciones de tránsito y de molestias a otros conductores, que maximice la seguridad, la comodidad del pasajero y el beneficio. Obviamente algunos de estos objetivos entran en conflicto, por lo que habrá que llegar a acuerdos.

Siguiente, ¿cuál es el **entorno** en el que se encontrará el taxi? Cualquier taxista debe estar preparado para circular por distintas carreteras, desde caminos rurales y calles urbanas hasta autopistas de 12 carriles. En los caminos se pueden encontrar con tráfico, peatones, animales, obras, coches, charcos y baches. El taxi puede operar en California del Sur, donde la nieve es un problema ocasional, o en Alaska, donde la nieve es un problema común. Puede conducir siempre por la derecha, o puede ser lo suficientemente flexible como para que circule por la izquierda cuando se encuentre en el Reino Unido o en Japón. Obviamente, cuanto más restringido sea el entorno, más fácil será el problema del diseño.

Los **actuadores** disponibles en un taxi automático serán más o menos los mismos que los que tiene a su alcance un conductor humano: el control del motor a través del acelerador, el control sobre la dirección y los frenos. Además, necesitará tener una pantalla de visualización o un sintetizador de voz para responder a los pasajeros, y quizá algún mecanismo para comunicarse con otros vehículos.

Sus **sensores** básicos deben incluir una o más cámaras de video dirigidas. Para controlar el vehículo adecuadamente, debe conocer el estado mecánico del vehículo, de forma que necesitará sensores que controlen el motor y el sistema eléctrico. Debe tener un GPS y sensores infrarrojos o sonares para detectar las distancias con respecto a otros coches y obstáculos. Finalmente, necesitará un teclado o micrófono para que el pasajero le indique su destino.

Tipo de agente	Medidas de rendimiento	Entorno	Actuadores	Sensores
Taxista	Seguro, rápido, legal, viaje confortable, maximización del beneficio	Carreteras, otro tráfico, peatones, clientes	Dirección, acelerador, freno, señal, bocina, visualizador	Cámaras, sónar, velocímetro, GPS, tacómetro, visualizador de la aceleración, sensores del motor, teclado

Figura 2.4 Descripción REAS del entorno de trabajo de un taxista automático.

Otros ejemplos:

Tipo de agente	Medidas de rendimiento	Entorno	Actuadores	Sensores
Sistema de diagnóstico médico	Pacientes sanos, reducir costes, demandas	Pacientes, hospital, personal	Visualizar preguntas, pruebas, diagnósticos, tratamientos, casos	Teclado para la entrada de síntomas, conclusiones, respuestas de pacientes
Sistema de análisis de imágenes de satélites	Categorización de imagen correcta	Conexión con el satélite en órbita	Visualizar la categorización de una escena	Matriz de pixels de colores
Robot para la selección de componentes	Porcentaje de componentes clasificados en los cubos correctos	Cinta transportadora con componentes, cubos	Brazo y mano articulados	Cámara, sensor angular
Controlador de una refinería	Maximizar la pureza, producción y seguridad	Refinería, operadores	Válvulas, bombas, calentadores, monitores	Temperatura, presión, sensores químicos
Tutor de inglés interactivo	Maximizar la puntuación de los estudiantes en los exámenes	Conjunto de estudiantes, agencia examinadora	Visualizar los ejercicios, sugerencias, correcciones	Teclado de entrada

Figura 2.5 Ejemplos de tipos de agentes y sus descripciones REAS.

Propiedades de los entornos de trabajo

Totalmente observable vs. parcialmente observable.

Si los sensores del agente le proporcionan acceso al estado completo del medio en cada momento, entonces se dice que el entorno de trabajo es totalmente observable. Un entorno de trabajo es, totalmente observable si los sensores detectan todos los aspectos que son relevantes en la toma de decisiones; la relevancia, en cada momento, depende de las medidas de rendimiento. Entornos totalmente observables son convenientes ya que el agente no necesita mantener ningún estado interno para saber qué sucede en el mundo. Un entorno puede ser parcialmente observable debido al ruido y a la existencia de sensores poco exactos o porque los sensores no reciben información de parte del sistema, por ejemplo, un agente aspiradora con sólo un sensor de suciedad local no puede saber si hay suciedad en la otra cuadrícula, y un taxi automatizado no puede saber que están pensando otros conductores.

Determinista vs. Estocástico.

Si el siguiente estado del medio está totalmente determinado por el estado actual y la acción ejecutada por el agente, entonces se dice que el entorno es determinista, de otra forma es estocástico. En principio, un agente no se tiene que preocupar de la incertidumbre en un medio totalmente observable y determinista. Sin embargo, si el medio es parcialmente observable entonces puede ser estocástico. Esto es particularmente cierto si se trata de un medio complejo, haciendo difícil el mantener constancia de todos los aspectos observados. Así, a menudo es mejor tratar a los escenarios complejos como estocásticos. El agente taxi es claramente estocástico en este sentido, ya que no se puede predecir el comportamiento del tráfico exactamente, más aún una rueda se puede reventar o el motor puede fallar sin previo aviso. Si el medio es determinista, excepto por las acciones de otros agentes, decimos que el medio es estratégico.

Episódico vs. secuencial.

En un entorno de trabajo episódico, la experiencia del agente se divide en episodios atómicos. Cada episodio consiste en la percepción del agente y la realización de una única acción posterior. Es muy importante tener en cuenta que el siguiente episodio no depende de las acciones que se realizaron en episodios previos. En los medios episódicos la elección de la acción en cada episodio depende sólo del episodio en sí mismo. Muchas tareas de clasificación son episódicas. Por ejemplo, un agente que tenga que seleccionar partes defectuosas en una cadena de montaje basa sus decisiones en la parte que está evaluando en cada momento, sin tener en cuenta decisiones previas. En entornos secuenciales, por otro lado, la decisión actual puede afectar a decisiones futuras. El ajedrez y el taxista son secuenciales: en ambos casos, las acciones que se realizan a corto plazo pueden tener consecuencias a largo plazo. Los medios episódicos son más simples que los secuenciales porque el agente no necesita pensar a futuro.

Estático vs. dinámico.

Si el entorno puede cambiar cuando el agente está deliberando, entonces se dice que el entorno es dinámico para el agente, de otra forma se dice que es estático. Los medios estáticos son fáciles de tratar ya que el agente no necesita estar pendiente del mundo mientras está tomando la decisión sobre una acción, ni necesita preocuparse sobre el paso del tiempo. Los medios dinámicos, por el contrario, mantienen continuamente al agente tomando decisiones sobre qué hacer. Si el entorno no cambia con el paso del tiempo, pero el rendimiento del agente sí, entonces se dice que el medio es semidinámico. El taxista es claramente dinámico: tanto los otros coches como el taxi se están moviendo mientras el algoritmo que guía la conducción indica qué es lo próximo a hacer. El ajedrez, cuando se juega con un reloj, es semidinámico. Los crucigramas son estáticos.

Discreto vs. continuo.

La distinción entre discreto y continuo se puede aplicar al estado del medio, a la forma en la que se maneja el tiempo y a las percepciones y acciones del agente. Por ejemplo, el ajedrez tiene un conjunto discreto de percepciones y acciones. El taxista conduciendo define un estado continuo y un problema de tiempo continuo: la velocidad y la ubicación del taxi y de los otros vehículos pasan por rango de valores continuos a lo largo del tiempo.

Agente individual vs. multiagente.

La distinción entre el entorno de un agente individual y el de un sistema multiagente puede parecer suficientemente simple. Por ejemplo, un agente resolviendo un crucigrama por sí mismo está claramente en un entorno de agente individual, mientras que un agente que juega al ajedrez está en un entorno con dos agentes. Sin embargo hay algunas diferencias sutiles. Primero, se ha descrito que una entidad puede percibirse como un agente, pero no se ha explicado que entidades deben considerarse agentes. ¿El agente A (por ejemplo el agente taxista) tiene que tratar un objeto B (otro vehículo) como un agente, o puede tratarse meramente como un objeto con un comportamiento estocástico, como las olas de la playa o las hojas que mueve el viento? La distinción clave está en identificar si el

comportamiento de B está mejor descrito por la maximización de una medida de rendimiento cuyo valor depende del comportamiento de A . Por ejemplo, en el ajedrez, la entidad oponente B intenta maximizar su medida de rendimiento, la cual, según las reglas, minimiza la medida de rendimiento del agente A. Por tanto, el ajedrez es un entorno multiagente competitivo. Por otro lado, para el agente taxista, el evitar colisiones maximiza la medida de rendimiento de todos los agentes, así pues es un entorno multiagente parcialmente cooperativo.

Como es de esperar, el caso más complejo es el parcialmente observable, estocástico, secuencial, dinámico, continuo y multiagente. Un agente taxista es un problema complejo en todas las propiedades.

Entornos de trabajo	Observable	Determinista	Episódico	Estático	Discreto	Agentes
Crucigrama Ajedrez con reloj	Totalmente Totalmente	Determinista Estratégico	Secuencial Secuencial	Estático Semi	Discreto Discreto	Individual Multi
Póker Backgammon	Parcialmente Totalmente	Estratégico Estocástico	Secuencial Secuencial	Estático Estático	Discreto Discreto	Multi Multi
Taxi circulando Diagnóstico médico	Parcialmente Parcialmente	Estocástico Estocástico	Secuencial Secuencial	Dinámico Dinámico	Continuo Continuo	Multi Individual
Análisis de imagen Robot clasificador	Totalmente Parcialmente	Determinista Estocástico	Episódico Episódico	Semi Dinámico	Continuo Continuo	Individual Individual
Controlador de refinería Tutor interactivo de inglés	Parcialmente Parcialmente	Estocástico Estocástico	Secuencial Secuencial	Dinámico Dinámico	Continuo Discreto	Individual Multi

Figura 2.6 Ejemplos de entornos de trabajo y sus características.

Estructura de los agentes

A pesar de que implementar una tabla con las funciones deseadas para el agente funciona, el desafío clave de la IA es encontrar la forma de escribir programas, que en la medida de lo posible, reproduzcan un comportamiento racional a partir de una pequeña cantidad de código en vez de a partir de una tabla con un gran número de entradas. Existen bastantes ejemplos que muestran que se puede hacer con éxito en otras áreas: por ejemplo, las grandes tablas de las raíces cuadradas utilizadas por ingenieros y estudiantes antes de 1970 se han reemplazado por un programa de cinco líneas que implementa el método de Newton en las calculadoras electrónicas. La pregunta es, en el caso del comportamiento inteligente general, ¿puede la IA hacer lo mismo que Newton con las raíces cuadradas? Se cree que la respuesta es afirmativa.

Existen cuatro tipos básicos de programas para agentes que encarnan los principios que subyacen en casi todos los sistemas inteligentes.

- Agentes reactivos simples.
- Agentes reactivos basados en modelos.
- Agentes basados en objetivos.
- Agentes basados en utilidad.

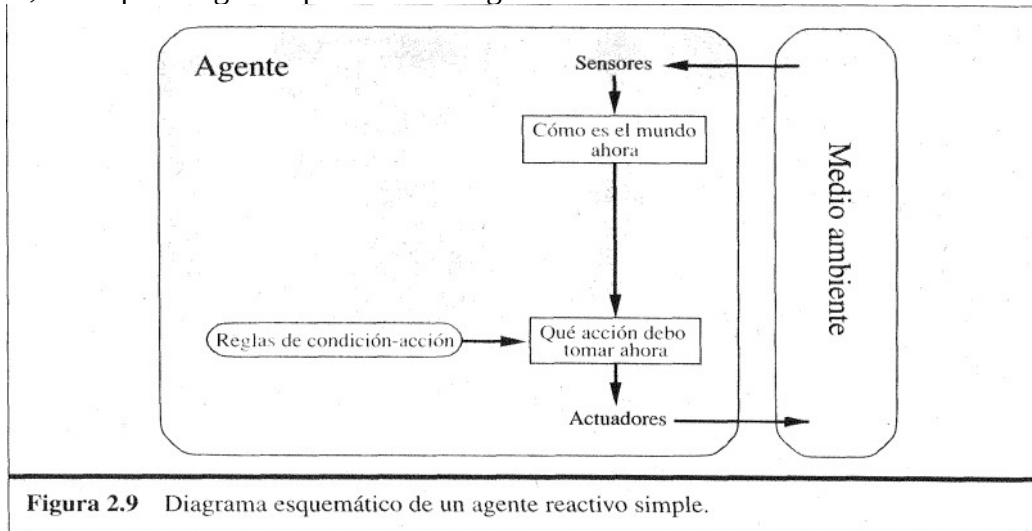
Agentes reactivos simples

Estos agentes seleccionan las acciones sobre la base de las percepciones actuales, ignorando el resto de las percepciones históricas. Por ejemplo, el agente aspiradora es un agente reactivo simple porque toma sus decisiones sólo con base en la localización actual y si ésta está sucia. Por ejemplo:

```
función AGENTE-ASPIRADORA-REACTIVO([localización, estado]) devuelve una acción
    si estado = Sucio entonces devolver Aspirar
    de otra forma, si localización = A entonces devolver Derecha
    de otra forma, si localización = B entonces devolver Izquierda
```

Figura 2.8 Programa para el agente aspiradora de reactivo simple en el entorno definido por las dos cuadrículas. Este programa implementa la función de agente presentada en la Figura 2.3.

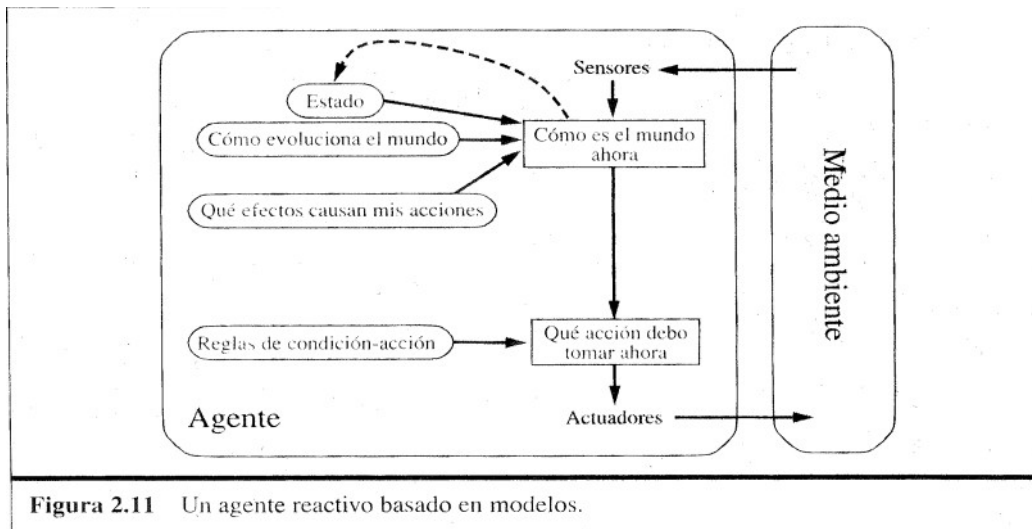
En general, éste tipo de agentes presentan la siguiente estructura:



Agentes reactivos basados en modelos

La forma más efectiva que tienen los agentes de manejar la visibilidad parcial es almacenar información de las partes del mundo que no pueden ver. O lo que es lo mismo, el agente debe mantener algún tipo de estado interno que dependa de la historia percibida y que de ese modo refleje por lo menos alguno de los aspectos no observables del estado actual.

La actualización de la información del estado interno según pasa el tiempo requiere codificar dos tipos de conocimiento en el programa del agente. Primero, se necesita alguna información acerca de cómo evoluciona el mundo independientemente del agente. Segundo, se necesita información sobre cómo afectan al mundo las acciones del agente. Este conocimiento acerca de "cómo funciona el mundo" se denomina modelo del mundo. Un agente que utilice este modelo es un agente basado en modelos.



Agentes basados en objetivos

El conocimiento sobre el estado actual del mundo no es siempre suficiente para decidir que hacer. Por ejemplo, en un cruce de carreteras, el taxista puede girar a la izquierda, derecha o seguir hacia adelante. La decisión correcta depende de dónde quiere ir el taxi. En otras palabras, además de la descripción del estado actual, el agente necesita algún tipo de información sobre su meta, por ejemplo,

llegar al destino propuesto por el pasajero. El programa del agente se puede combinar con información sobre los resultados de las acciones posibles para elegir las acciones que permitan alcanzar el objetivo. La siguiente figura muestra la estructura del agente basado en objetivo.

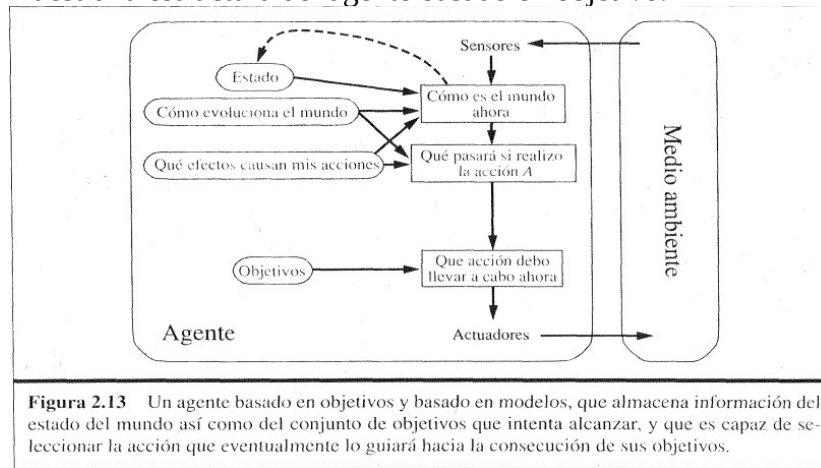


Figura 2.13 Un agente basado en objetivos y basado en modelos, que almacena información del estado del mundo así como del conjunto de objetivos que intenta alcanzar, y que es capaz de seleccionar la acción que eventualmente lo guiará hacia la consecución de sus objetivos.

En algunas ocasiones, la selección de acciones basadas en objetivos es directa, cuando alcanzar los objetivos es el resultado inmediato de una acción individual. En otras ocasiones, puede ser más complicado, cuando el agente tiene que considerar secuencias complejas para encontrar el camino que le permita alcanzar el objetivo. Búsqueda y planificación son los subcampos de la IA centrados en encontrar secuencias de acciones que permitan a los agentes alcanzar sus metas.

Agentes basados en utilidad

Las metas por sí solas no son realmente suficientes para generar comportamiento de gran calidad en la mayoría de los entornos. Por ejemplo, hay muchas secuencias de acciones que llevarán al taxi a su destino (y por tanto a alcanzar su objetivo), pero algunas son más rápidas, más seguras, o más baratas que otras. Una medida de eficiencia más general debería permitir una comparación entre estados del mundo diferentes de acuerdo al nivel exacto de utilidad que el agente alcance cuando se llegue a un estado u otro.

Una función de utilidad permite tomar decisiones racionales en dos tipos de casos en los que las metas son inadecuadas. Primero, cuando haya objetivos conflictivos, y solo se puedan alcanzar algunos de ellos (por ejemplo, velocidad y seguridad), la función de utilidad determina el equilibrio adecuado. Segundo, cuando haya varios objetivos por los que se pueda guiar el agente, y ninguno de ellos se pueda alcanzar con certeza, la utilidad proporciona un mecanismo para ponderar la probabilidad de éxito en función de la importancia de los objetivos.

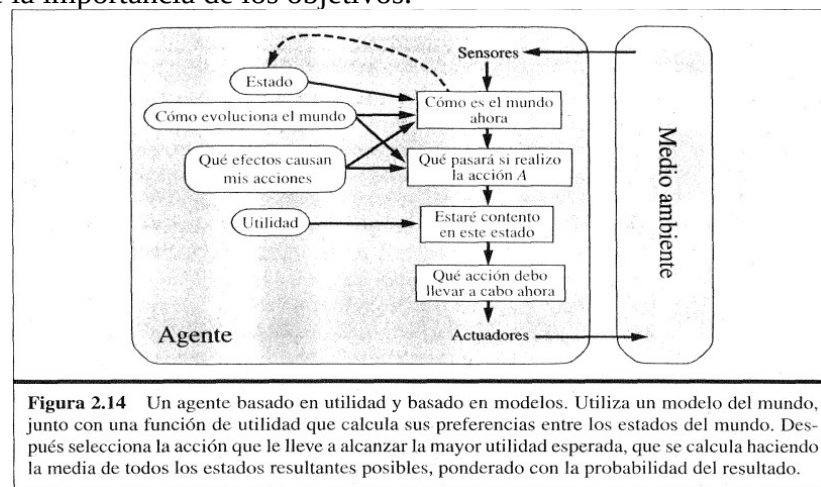


Figura 2.14 Un agente basado en utilidad y basado en modelos. Utiliza un modelo del mundo, junto con una función de utilidad que calcula sus preferencias entre los estados del mundo. Después selecciona la acción que le lleve a alcanzar la mayor utilidad esperada, que se calcula haciendo la media de todos los estados resultantes posibles, ponderado con la probabilidad del resultado.

Agentes que aprenden

Se han descrito programas para agentes que poseen varios métodos para seleccionar acciones. Turing (1950), en su temprano y famoso artículo, consideró la idea de programar sus máquinas inteligentes a mano. Estimó cuánto tiempo podía llevar y concluyó, que sería deseable utilizar algún método más rápido. El método que propone es construir máquinas que aprendan y después enseñarles.. En muchas áreas de IA, éste es el método más adecuado para crear sistemas novedosos. El aprendizaje tiene otras ventajas, como se ha explicado anteriormente: permite que el agente opere en medios inicialmente desconocidos y que sea más competente que si sólo utilizase su conocimiento inicial.

Un agente que aprende se puede dividir en cuatro componentes conceptuales. La distinción más importante entre el elemento de aprendizaje y el elemento de actuación es que el primero tiene la responsabilidad de hacer mejoras y el segundo de la selección de acciones. El elemento de actuación es lo que anteriormente mencionábamos como agente completo. El elemento de aprendizaje se alimenta constantemente de las críticas sobre la actuación del agente y determina cómo se debe modificar el elemento de actuación para proporcionar mejores resultados.

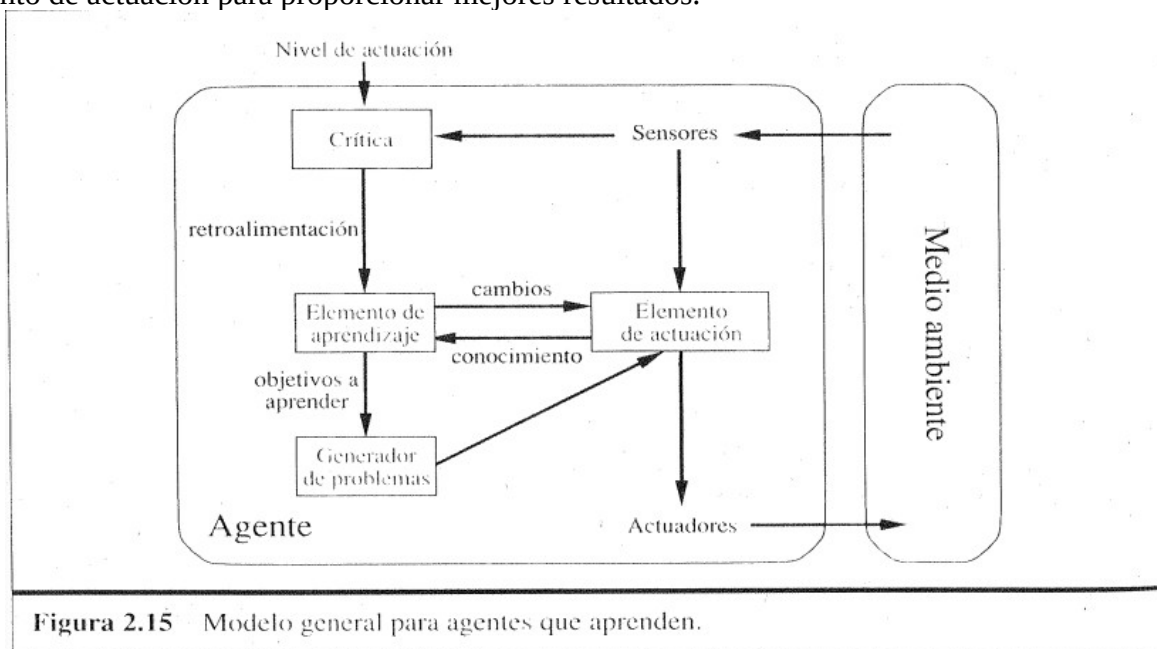


Figura 2.15 Modelo general para agentes que aprenden.

El último componente del agente con capacidad de aprendizaje es el generador de problemas. Es responsable de sugerir acciones que guiarán al agente hacia experiencias nuevas e informativas. Lo interesante es que si el elemento de actuación sigue su camino puede continuar llevando a cabo las acciones que sean mejores, dado su conocimiento. Pero si el agente está dispuesto a explorar un poco, y llevar a cabo algunas acciones que no sean óptimas a corto plazo, puede descubrir acciones mejores a largo plazo.