

Inteligencia Artificial

Tema 2. Agentes Inteligentes

José Manuel Fuertes García

`jmf@ujaen.es`

Departamento de Informática
Universidad de Jaén

16 de febrero de 2017

Objetivos

- ▶ Entender los conceptos de agente, función de agente y programa de agente
- ▶ Conocer las características de los agentes racionales
- ▶ Entender y saber especificar los elementos que integran un agente inteligente (REAS)
- ▶ Saber especificar un entorno de trabajo a través de sus propiedades
- ▶ Conocer los diferentes tipos de agentes, estructura y limitaciones

Índice

Introducción a los agentes inteligentes

El entorno

Estructura de un agente

Tipos de agentes

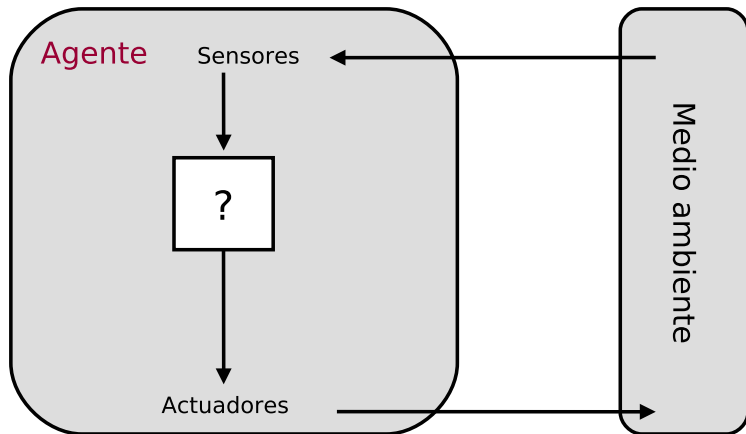
Subáreas dentro de la Inteligencia Artificial

Agente

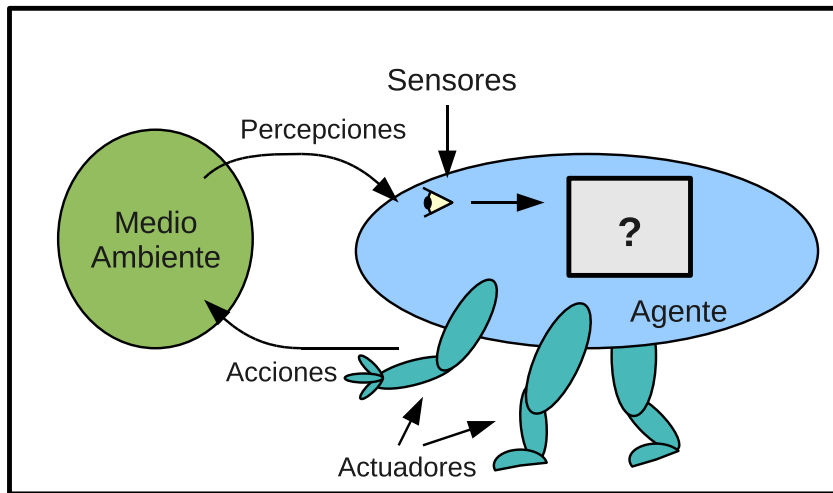
Ente activo embebido en un entorno que es capaz de percibir su medio ambiente con la ayuda de sensores y actuar en ese medio utilizando actuadores

- ▶ *Cuerpo:*
 - ▶ Percibe el entorno por medio de sensores
 - ▶ Actúa sobre el entorno por medio de efectores o actuadores
- ▶ *Mente:*
 - ▶ Determina las acciones a partir de las percepciones
 - ▶ Existe una medida de rendimiento que guía dicho proceso

Estructura de un agente



Estructura de un agente



Percepción y acción

Percepción:

- ▶ 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
- ▶ Un agente tomará una decisión en un momento dado dependiendo de la secuencia completa de percepciones hasta ese instante
- ▶ El comportamiento del agente viene dado por la llamada **función del agente**, que proyecta una percepción dada en una acción
- ▶ Esta función se puede representar como una tabla (infinita en la mayor parte de los casos)

Función y programa del agente

Función del agente:

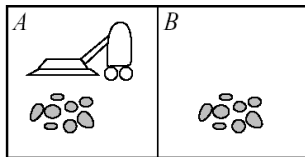
- ▶ Descripción matemática abstracta que se implementará mediante el programa del agente

Programa del agente:

- ▶ Implementación concreta de la función del agente, que se ejecuta sobre la arquitectura del mismo

Ejemplo de agente

El mundo de la aspiradora:



- **Percepciones:** localización y contenidos.
Ej: [A, Sucio]
- **Acciones:** Izquierda, Derecha, Aspirar, NoOp

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
...	...

Función `Agente_aspiradora_reactivo` (localización, estado)

```

si estado = Sucio
    entonces devolver Aspirar
si no, si localización = A
    entonces devolver Derecha
si no, si localización = B
    entonces devolver Izquierda
  
```

Tipos de agentes

Agentes naturales:

- ▶ Cuerpo biológico y entorno natural
- ▶ Sensores: ojos, oídos, lengua, etc.
- ▶ Actuadores: piernas, brazos, manos, etc.
- ▶ Medida de rendimiento: sobrevivir, reproducirse, ...

Agentes artificiales:

- ▶ Agentes *hardware* (robots):
 - ▶ Interactúan directamente con un entorno físico. Poseen *cuerpo*
 - ▶ Sensores: cámaras, telémetros, infrarrojos, etc.
 - ▶ Actuadores: ruedas/piernas, manipuladores, etc.
- ▶ Agentes *software* (softbots):
 - ▶ Actúan en entornos virtuales (Ej: Internet)
 - ▶ Todo *software*: no necesitan manipular físicamente el entorno
 - ▶ Sensores y actuadores: dependientes del entorno

Agentes racionales

Un agente racional...

- ▶ ... hace lo correcto
- ▶ ... actúa de forma racional en su entorno
- ▶ Las **medidas de rendimiento** incluyen los criterios que determinan el éxito en el comportamiento del agente
- ▶ No hay una única medida adecuada para todos los agentes
- ▶ Regla general: diseñar medidas de utilidad de acuerdo con lo que se quiere para el entorno más que de acuerdo a cómo se cree que debe comportarse el agente

Racionalidad

Determinantes del comportamiento racional:

- ▶ Medida de *rendimiento*: define el grado de éxito del agente
- ▶ Secuencia de *percepciones*: la experiencia del agente
- ▶ *Conocimientos a priori* sobre su entorno
- ▶ *Acciones* que el agente puede realizar (capacidades)

En cada posible secuencia de percepciones, un **agente racional** deberá emprender aquella acción que supuestamente maximice su medida de rendimiento, basándose en las evidencias aportadas por la secuencia de percepciones y en el conocimiento que el agente mantiene almacenado

Racionalidad *versus* Omnisciencia

Racionalidad \neq Omnisciencia

Racionalidad \neq Clarividencia

Racionalidad \neq Éxito

- ▶ Un agente omnisciente conoce el resultado de su acción y actúa de acuerdo con él
- ▶ En la realidad la omnisciencia no es posible
- ▶ La selección racional de acciones sólo se basa en la información disponible
- ▶ La racionalidad maximiza el resultado esperado mientras que la perfección maximiza el resultado real

Características de los agentes racionales

Recopilación de información:

Realizar acciones con la intención de modificar percepciones futuras

Aprendizaje:

El agente racional no sólo recopila información sino que aprende lo máximo posible de lo que percibe

Autonomía:

Un agente tiene autonomía cuando se apoya en sus propias percepciones tanto o más que en el conocimiento inicial que le proporciona el diseñador

Características de los agentes racionales

- ▶ Los conocimientos *a priori* de un agente compilan la inteligencia del diseñador
- ▶ Un agente que no prestase atención a sus percepciones:
 - ▶ No sería inteligente
 - ▶ Sólo podría actuar en entornos extremadamente simples
 - ▶ Fracasaría en situaciones no anticipadas
- ▶ Un agente es más autónomo...
 - ▶ ... cuanto más rige su comportamiento por su propia experiencia
 - ▶ ... cuanto menos depende de sus conocimientos *a priori*

Agente inteligente = comportamiento racional + aprendizaje + autonomía

Elementos que integran un agente inteligente

Cuestiones que hay que concretar:

¿Cuál es la medida de Rendimiento? $\implies R$

¿Cuál es el Entorno? $\implies E$

¿Cuáles son los Actuadores? $\implies A$

¿Cuáles son los Sensores? $\implies S$

Ejemplo de agente inteligente

Taxi automático

- ▶ **Medida de Rendimiento:** Seguridad, destino, beneficios, legal, viaje confortable, ...
- ▶ **Entorno:** Carreteras, tráfico, peatones, meteorología, ...
- ▶ **Actuadores:** Dirección, acelerador, freno, bocina, ...
- ▶ **Sensores:** Cámara, sensor de aceleración, indicador de la gasolina, sensores del motor, teclado, GPS, ...

Ejemplo de agente inteligente

Compra en Internet

- ▶ **Medida de Rendimiento:** Precio, calidad, adecuación, eficiencia, ...
- ▶ **Entorno:** Sitios web actuales y futuros, vendedores, transportistas, ...
- ▶ **Actuadores:** Mostrar al usuario, seguir una URL, rellenar formulario, ...
- ▶ **Sensores:** Páginas HTML (texto, gráficos, scripts), ...

Índice

Introducción a los agentes inteligentes

El entorno

Estructura de un agente

Tipos de agentes

Subáreas dentro de la Inteligencia Artificial

El entorno

El entorno de trabajo es el **problema** para el que el agente racional es la **solución**

- ▶ ¿Cómo especificar el entorno de trabajo?
- ▶ ¿Qué propiedades puede tener un entorno de trabajo?

Especificación del entorno

- ▶ Es difícil categorizar el rango de los entornos de trabajo
- ▶ Se pueden señalar algunas propiedades que determinan, hasta cierto punto, el diseño más adecuado para el agente y la utilización de algunas técnicas para la implementación del mismo
- ▶ Clasificaremos los entornos considerando cada propiedad en cuestión desde el punto de vista del agente

Características de los entornos

- ▶ **Totalmente observable**: Si los sensores detectan todos los aspectos que son relevantes en la toma de decisiones
- ▶ **Parcialmente observable**: El agente no dispone de toda la información para tomar la decisión
 - ▶ Ruido, sensores inexactos, o que no reciben información sobre la totalidad del entorno
- ▶ **Determinista**: Cuando el siguiente estado del medio está totalmente determinado por el estado actual y la acción ejecutada por el agente
 - ▶ Ajedrez, agente *software* en entorno simulado, ...
- ▶ **Estocástico**: Si el estado siguiente resulta impredecible a partir de la información disponible
 - ▶ Gestión de tráfico, robots en entornos reales, ...

Características de los entornos

- ▶ **Episódico**: Si la acción que el agente ha de realizar no depende de acciones pasadas ni afecta a acciones futuras
- ▶ **Secuencial**: La decisión presente viene determinada por acciones pasadas y/o puede afectar a decisiones futuras
- ▶ **Dinámico**: El entorno puede cambiar cuando el agente está deliberando (no hacer nada se convierte en una opción)
 - ▶ Gestión de tráfico (cambios imprevisibles)
- ▶ **Estático**: El entorno no cambia mientras el agente decide qué hacer
 - ▶ Agente *software* en un laberinto simulado (el entorno no cambia)
- ▶ **Semidinámico**: El entorno no cambia con el tiempo, pero el rendimiento sí
 - ▶ Ajedrez (cuando se juega con reloj)

Características de los entornos

- ▶ **Discreto** o **continuo**:
 - ▶ Aplicable al estado del medio, a la forma en que se maneja el tiempo y a las percepciones y acciones del agente
- ▶ **Individual**: Sólo hay un agente
- ▶ **Multiagente**: Hay varios agentes. El rendimiento de cada uno depende de las acciones de los otros.
 - ▶ **Competitivo**: Ajedrez
 - ▶ **Cooperativo**: Robots de enjambre
 - ▶ **Mixto**: Taxi

Ejemplos de entornos de trabajo

	Observable	Determinista	Episódico	Estático	Discreto	Agentes
Crucigrama	Totalmente	Determinista	Secuencial	Estático	Discreto	Individual
Ajedrez con reloj	Totalmente	Estratégico	Secuencial	Semi	Discreto	Multi
Póker	Parcial	Estratégico	Secuencial	Estático	Discreto	Multi
Backgammon	Totalmente	Estocástico	Secuencial	Estático	Discreto	Multi
Taxi	Parcial	Estocástico	Secuencial	Dinámico	Continuo	Multi
Diagnóstico Médico	Parcial	Estocástico	Secuencial	Dinámico	Continuo	Individual
Análisis Imagen	Totalmente	Determinista	Episódico	Semi	Continuo	Individual
Robot Clasificador	Parcial	Estocástico	Episódico	Dinámico	Continuo	Individual
Control Refinería	Parcial	Estocástico	Secuencial	Dinámico	Continuo	Individual
Tutor Interactivo	Parcial	Estocástico	Secuencial	Dinámico	Discreto	Multi

- **Mundo real:** Parcialmente observable, estocástico, secuencial, dinámico, continuo y multiagente

Algunos ejemplos:

- ▶ Vacuum cleaner
- ▶ Micromouse robot competition
- ▶ Micromouse-competition
- ▶ Micromouse classic
- ▶ Mouse Run

Índice

Introducción a los agentes inteligentes

El entorno

Estructura de un agente

Tipos de agentes

Subáreas dentro de la Inteligencia Artificial

Estructura de los agentes

En un agente se distinguen dos componentes:

- ▶ Arquitectura
- ▶ Programa

Programa del agente

- ▶ *Software* que determina el comportamiento del agente
- ▶ Implementa la función de percepción-acción
- ▶ Debe ser apropiado para la arquitectura
- ▶ Es la parte que se encarga de la I.A.

Programa del agente

- ▶ **Estructura general de un programa de agente:** recibe las percepciones actuales como entradas y devuelve una acción para los actuadores
- ▶ **Función del agente \neq programa del agente**
 - ▶ El programa del agente toma la percepción actual como entrada
 - ▶ La función del agente recibe la percepción histórica completa
- ▶ Si la acción del agente depende de la secuencia completa de percepciones, el agente debe recordarlas

Función Agente_Dirigido_Mediante_Tabla (percepción) **devuelve** una acción

estático: *percepciones*, una secuencia, vacía inicialmente

tabla, una tabla de acciones, indexada por las secuencias de percepciones,
totalmente definida inicialmente

añadir la *percepción* al final de las *percepciones*

acción \leftarrow Consulta(*percepciones*, *tabla*)

devolver (*acción*)

¿Esta propuesta es práctica?

Arquitectura de un agente

- ▶ Hace que las percepciones de los sensores estén disponibles para el programa
- ▶ Ejecuta los programas
- ▶ Controla los actuadores
- ▶ Elementos mínimos que la constituyen:
 - ▶ Componente de percepción
 - ▶ Componente de selección de acciones
 - ▶ Componente de acción

Índice

Introducción a los agentes inteligentes

El entorno

Estructura de un agente

Tipos de agentes

Subáreas dentro de la Inteligencia Artificial

Tipos de agentes

- ▶ Existen cuatro tipos, de menor a mayor generalidad:
 - ▶ Agentes reactivos simples
 - ▶ Agentes reactivos basados en modelos
 - ▶ Agentes basados en objetivos
 - ▶ Agentes basados en utilidad
- ▶ Todos ellos pueden convertirse en **agentes que aprenden**

Agentes reactivos simples

Seleccionan las acciones en base a las percepciones actuales, ignorando el historial de percepciones anteriores

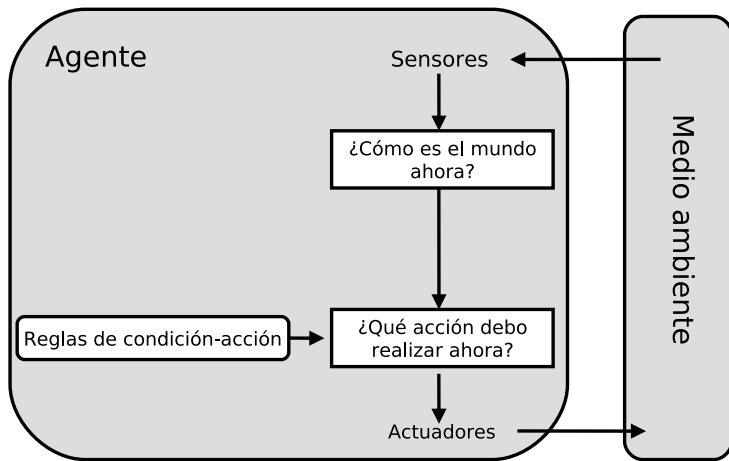
Ejemplo:

- Programa para el agente *aspiradora*:

Función Agente_Aspiradora_Reactivo (localización, estado) devuelve una acción

```
Si estado == sucio
    entonces devolver Aspirar
si no Si localización == A
    entonces devolver Derecha
    si no devolver Izquierda
```

Agentes reactivos simples



Agentes reactivos simples

Función Agente_Reactivo_Simple (percepción) **devuelve** una acción

estático: *reglas*, un conjunto de reglas condición-acción

estado \leftarrow Interpretar_Entrada(*percepción*)

regla \leftarrow Regla_Coincidencia(*estado*, *reglas*)

acción \leftarrow Regla_Acción[*regla*]

devolver *acción*

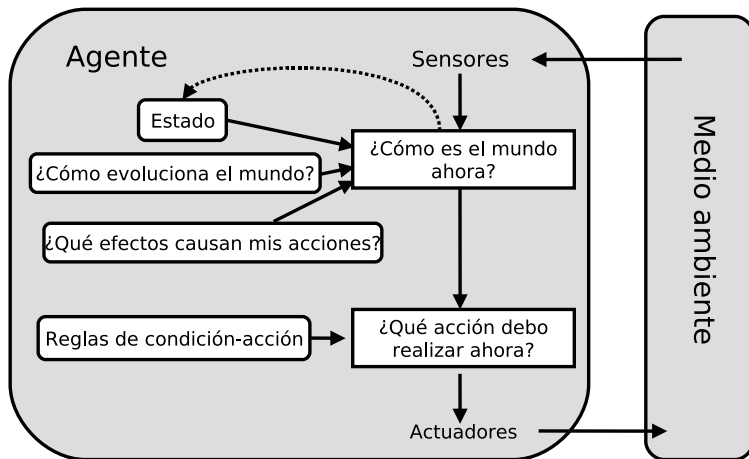
Limitaciones de los agentes reactivos simples

- ▶ Los agentes de este tipo sólo pueden tomar la decisión correcta bajo la base de la percepción actual, lo cual sólo es posible si el entorno es totalmente observable
- ▶ En ocasiones generan ciclos infinitos
- ▶ Salir de estos ciclos infinitos es posible si los agentes pueden seleccionar sus acciones aleatoriamente

Agentes reactivos basados en modelos

- ▶ Para manejar la visibilidad parcial, almacenan información sobre las partes del mundo que no se pueden ver
- ▶ Mantienen un estado interno que depende de la historia percibida y que refleja alguno de los aspectos no observables
- ▶ Este tipo de agentes mantiene un *modelo simbólico del mundo*
 - ▶ Información sobre cómo evoluciona el mundo independientemente del agente
 - ▶ Información sobre cómo afectan al mundo las acciones del agente

Agentes reactivos basados en modelos



Función de un agente reactivo basado en modelos

función Agente_reactivo_con_estado (percepción) **devuelve** una acción
estático:

estado (descripción actual del estado del mundo)

reglas (conjunto de reglas condición-acción)

acción (acción más reciente. Inicialmente ninguna)

estado \leftarrow Actualizar_estado (*estado*, *acción*, *percepción*)

regla \leftarrow Regla_coincidencia (*estado*, *reglas*)

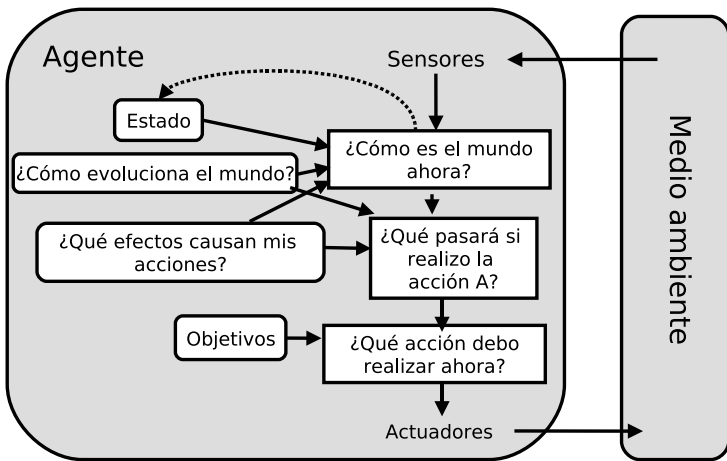
acción \leftarrow Regla_acción[*regla*]

devolver *acción*

Agentes basados en objetivos

- ▶ El conocimiento sobre el estado actual del mundo no siempre es suficiente para decidir qué hacer
- ▶ Además de la descripción del estado actual es necesaria información sobre su meta
- ▶ La selección basada en objetivos puede ser directa si al realizar una acción individual se tiene como resultado directo la consecución del objetivo
- ▶ A veces el agente tiene que considerar secuencias complejas para encontrar el camino que le permita alcanzar el objetivo
 - ▶ Búsqueda
 - ▶ Planificación

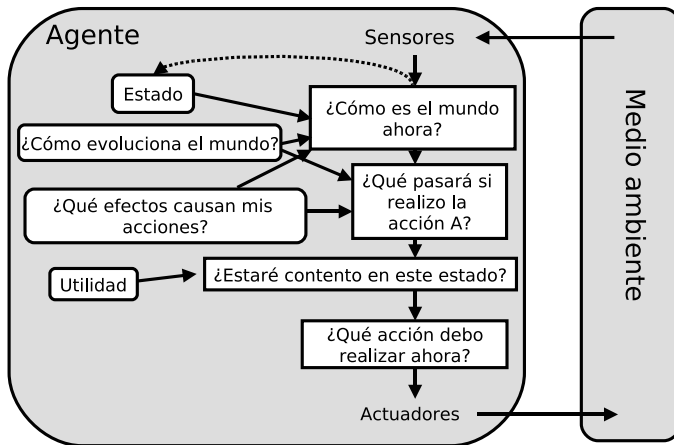
Agentes basados en objetivos



Agentes basados en utilidad

- ▶ En ocasiones las metas no son suficientes para generar un comportamiento gran calidad
- ▶ Es necesaria una función de utilidad
- ▶ Permite tomar decisiones racionales cuando las metas son inadecuadas:
 - ▶ Cuando haya objetivos conflictivos y solo se pueda alcanzar algunos de ellos
 - ▶ Cuando haya varios objetivos por los que se pueda guiar el agente y ninguno de ellos se pueda alcanzar con certeza
- ▶ Se basa en la toma de decisiones

Agentes basados en utilidad



Agentes que aprenden. Componentes

Elemento de actuación:

- ▶ Se encarga de seleccionar acciones externas
- ▶ Es lo que hasta ahora hemos considerado el agente completo

Elemento de aprendizaje:

- ▶ Se encarga de realizar mejoras
- ▶ Se realimenta con las críticas sobre la actuación del agente
- ▶ Determina las modificaciones del elemento de actuación para mejorar los resultados

Agentes que aprenden. Componentes

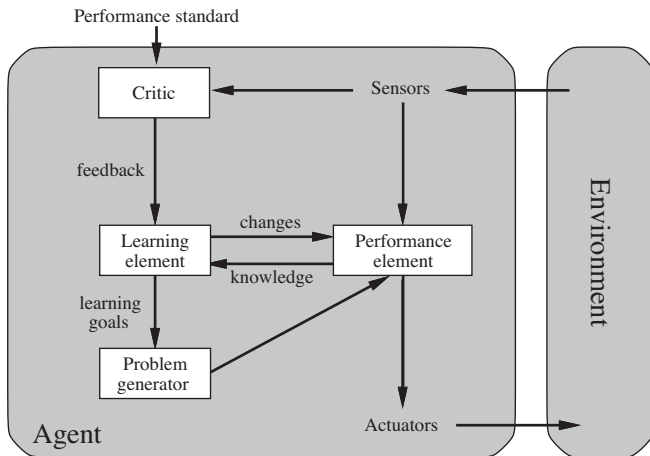
Críticas:

- ▶ Indica al elemento de aprendizaje la calidad del agente respecto a un nivel de actuación fijo
- ▶ Son necesarias porque las percepciones por sí solas no indican necesariamente el éxito del agente

Generador de problemas:

- ▶ Es responsable de sugerir acciones que guíen al agente hacia experiencias nuevas e informativas
- ▶ Permite la exploración y la realización de acciones no totalmente óptimas a corto plazo, pero mejores a largo plazo

Agentes que aprenden



Índice

Introducción a los agentes inteligentes

El entorno

Estructura de un agente

Tipos de agentes

Subáreas dentro de la Inteligencia Artificial

Subáreas de la Inteligencia Artificial

- ▶ **Resolución de problemas mediante búsqueda:**
 - ▶ Actuar de forma racional en entornos bien definidos: espacios de estado
 - ▶ Entornos totalmente observables, deterministas, estáticos y discretos
- ▶ **Representación del conocimiento y razonamiento:**
 - ▶ Combatir la complejidad: estructurar la representación del entorno
 - ▶ Entornos parcialmente observables/no deterministas
 - ▶ Razonamiento no monótono
- ▶ **Planificación:**
 - ▶ Combatir la complejidad: representación estructurada + inferencia especializada
 - ▶ Entornos no deterministas: planificación condicional
 - ▶ Entornos dinámicos: replanificación

Subáreas de la Inteligencia Artificial

▶ Aprendizaje:

- ▶ Combatir la complejidad: aprender a actuar más rápido
- ▶ Mejorar el rendimiento: aprender a actuar mejor
- ▶ Mejorar la autonomía: reducir la dependencia de conocimientos *a priori*

▶ Incertidumbre:

- ▶ Entornos parcialmente observables/no deterministas/continuos: creencias (lógica difusa, redes bayesianas, etc.)
- ▶ Medidas de rendimiento basadas en la utilidad específica (Teoría de la Utilidad, Inferencia basada en la Teoría de la Decisión)

Subáreas de la Inteligencia Artificial

► Comunicación:

- Entornos en los que el agente interactúa de forma flexible con humanos (procesamiento del lenguaje natural, actos de habla, etc.)
- Entornos multiagente (razonamientos sobre otros agentes, coordinación, lenguajes de comunicación entre agentes, etc.)

► Percepción y actuación:

- Robótica: entornos físicos y agentes *hardware*
- Visión y audición computacional (reconocimiento de patrones, procesamiento de la señal, etc.)

Resumen

- ▶ Un **agente** es algo que percibe y actúa en un medio
- ▶ La **función del agente** describe lo que hace el agente bajo cualquier circunstancia
- ▶ La **medida de rendimiento** evalúa el comportamiento del agente en un medio
- ▶ Un **agente racional** actúa con la intención de maximizar la medida de rendimiento
- ▶ El **programa del agente** implementa la función del agente. Existen diferentes definiciones del programa de un agente. Difieren en eficiencia, solidez y flexibilidad
- ▶ Para describir el entorno hay que especificar la descripción **REAS**
- ▶ Los entornos se caracterizan con diferentes dimensiones: *observable, determinista, episódico, discreto, agente simple*
- ▶ Existen diferentes **tipos de arquitecturas de agente**: *reactivo, basado en modelos, en objetivos, en utilidad o agentes que aprenden*