

Tema 2: Agentes



Objetivos

- Comprender que el desarrollo de sistemas inteligentes pasa por el diseño de agentes capaces de representar conocimiento y resolver problemas y que puede orientarse a la construcción de sistemas bien completamente autónomos o bien que interactúen y ayuden a los humanos.
- Conocer el concepto de agente inteligente y el ciclo de vida “percepción, decisión y actuación”.
- Adquirir las habilidades básicas para construir sistemas capaces de resolver problemas mediante técnicas de IA.

Estudia este tema en...

- Nils J. Nilsson, *“Inteligencia Artificial: Una nueva síntesis”*, Ed. Mc Graw Hill, 2000. pp. 17-32, 63-74, 103-122, 147-162
- Sobre el concepto de agente inteligente consultar además:
 - Michael Wooldridge, Nicholas R. Jennings, Intelligent Agents: Theory and Practice, Knowledge Engineering Review 10 115-152, 1995
- Sobre el concepto de agente racional consultar además:
 - S. Russell, P. Norvig, Artificial Intelligence: A modern Approach, Tercera Edición, Ed. Pearson, 2016.

Contenido

- **Agentes inteligentes**
- **Arquitecturas de agentes**
- **Agentes reactivos**

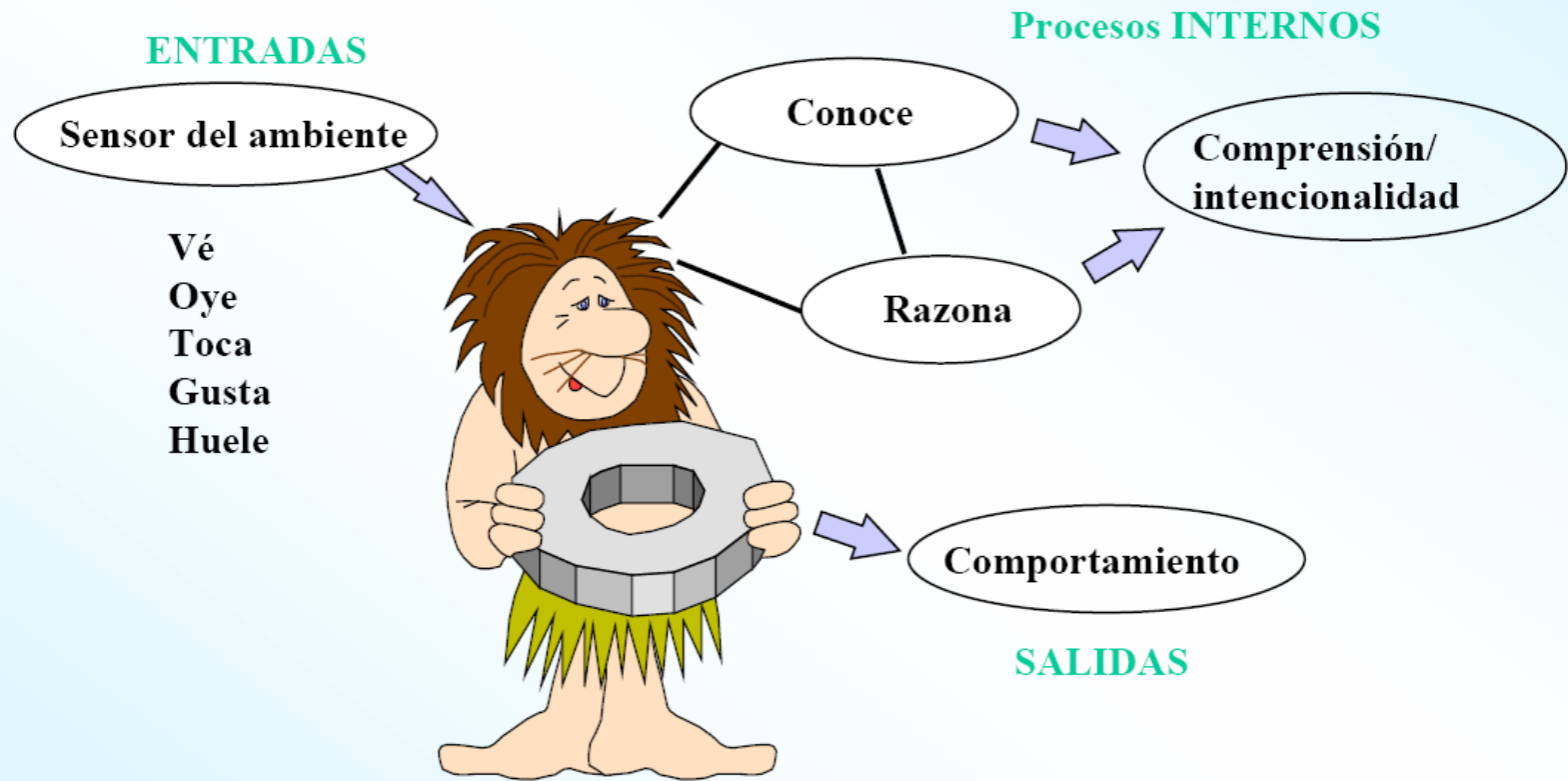
Contenido

- **Agentes inteligentes**
- **Arquitecturas de agentes**
- **Agentes reactivos**

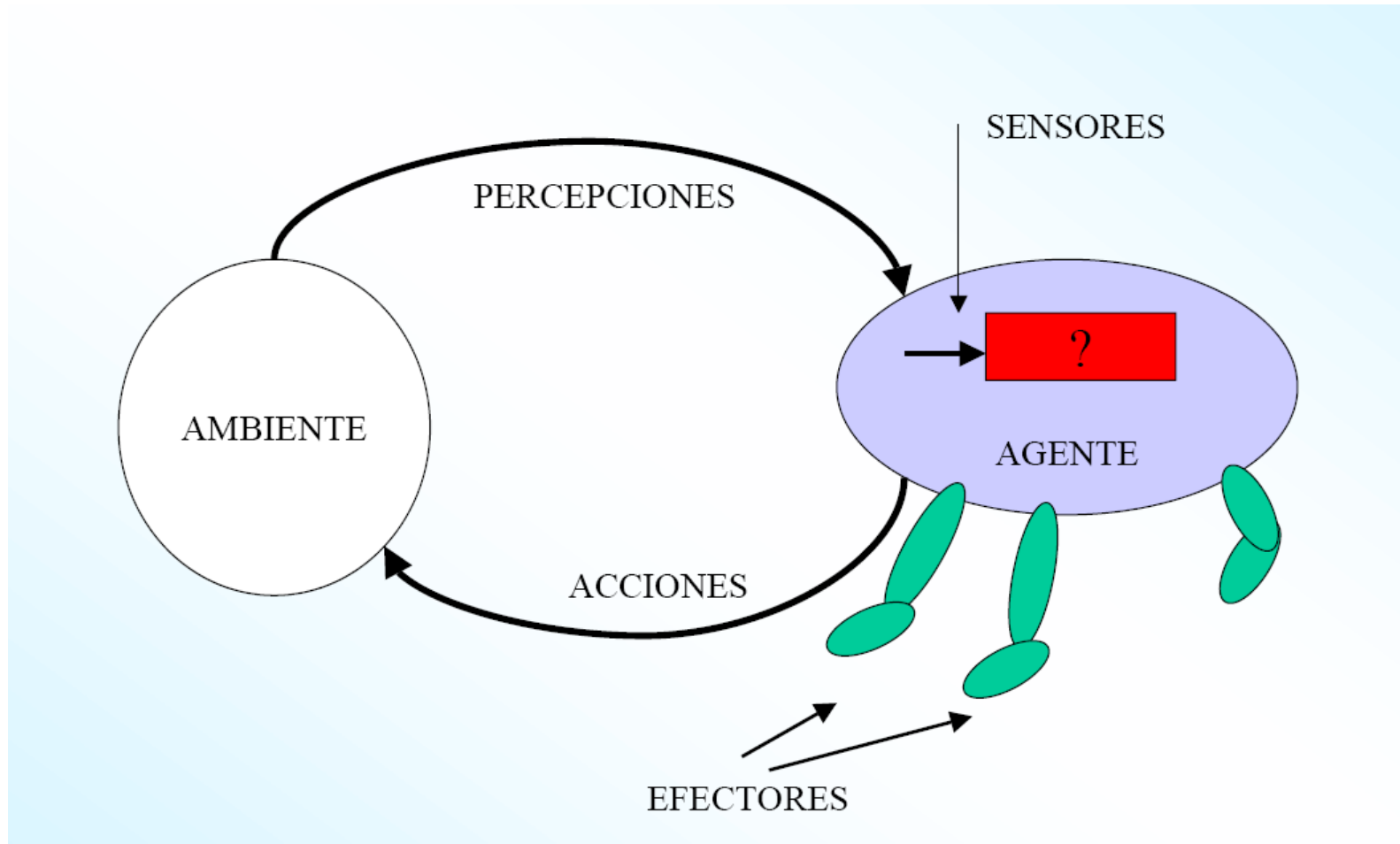
Concepto de Agente inteligente

- Un **Agente inteligente** es un sistema de ordenador, **situado** en algún entorno, que es capaz de realizar acciones de forma **autónoma** y que es **flexible** para lograr los objetivos planteados.
 - **Situación:** el agente recibe entradas sensoriales de un entorno en donde está situado y realiza acciones que cambian dicho entorno
 - **Autonomía:** el sistema es capaz de actuar sin la intervención directa de los humanos y tiene control sobre sus propias acciones y estado interno

Agentes inteligentes



Agentes



Agentes e Inteligencia Artificial

Los agentes ofrecen una perspectiva innovadora para **representar la Inteligencia Artificial**, permitiendo abordar sus diversas áreas desde un **enfoque basado en agentes** que ejecutan acciones en función de percepciones procesadas a través de mecanismos avanzados de análisis y toma de decisiones.

Flexibilidad

“capaz de realizar acciones de forma **autónoma** y que es **flexible** para lograr los objetivos planteados.”

- **Reactivo:** el agente debe percibir el entorno y responder de una forma temporal a los cambios que ocurren en dicho entorno
- **Pro-activo:** los agentes no deben simplemente actuar en respuesta a su entorno, deben de ser capaces de exhibir comportamientos dirigidos a lograr objetivos que sean oportunos, y tomar la iniciativa cuando sea apropiado
- **Social:** los agentes deben de ser capaces de interactuar, cuando sea apropiado, con otros agentes artificiales o humanos para completar su propio proceso de resolución del problema y ayudar a otros con sus actividades

Sistemas basados en agentes

- Un Sistema Basado en Agentes será un sistema en el que la abstracción clave utilizada es precisamente la de agente
- **Sistemas multi-agente:** un sistema diseñado e implementado con varios agentes interactuando
- Los sistemas multi-agente son interesantes para representar problemas que tienen
 - múltiples formas de ser resueltos,
 - múltiples perspectivas y/o
 - múltiples entidades para resolver el problema

Interacción entre agentes

- **Cooperación:** trabajar juntos para resolver algo
- **Coordinación:** organizar una actividad para evitar las interacciones perjudiciales y explotar las beneficiosas
- **Negociación:** llegar a un acuerdo que sea aceptable por todas las partes implicadas

Sistemas Multi-Agente

- **Inteligencia Artificial Distribuida**
 - Resolución de Problemas Distribuida
 - Sistemas Multi-Agente
 - SMA: una red más o menos unida de resolutores de problemas que trabajan conjuntamente para resolver problemas que están más allá de las capacidades individuales o del conocimiento de cada resolutor del problema
 - Resolutor=agente (autónomo y de naturaleza heterogénea)

Características de un SMA

- Cada agente tiene información incompleta, o no todas las capacidades para resolver el problema, así cada agente tiene un punto de vista limitado.
- No hay un sistema de control global.
- Los datos no están centralizados.
- La computación es asíncrona.

Cooperación y Negociación

- Cooperación: herramienta fundamental en la formación de equipos (p.e. ROBOCUP)



- Negociación: coordinación y resolución de conflictos

Contenido

- **Agentes inteligentes**
- **Arquitecturas de agentes**
- **Agentes reactivos**

Arquitecturas de Agentes

- Arquitecturas deliberativas
- Arquitecturas reactivas
- Arquitecturas híbridas

Arquitecturas deliberativas

- **Sistema de símbolos físicos:** un conjunto de entidades físicas (símbolos) que pueden combinarse para formar estructuras, y que es capaz de ejecutar procesos que operan con dichos símbolos de acuerdo a conjuntos de instrucciones codificadas simbólicamente
- La **hipótesis de sistema de símbolos físicos** dice que tales sistemas son capaces de generar acciones inteligentes
- **Agente deliberativo:** aquel que contiene un modelo simbólico del mundo explícitamente representado, y cuyas decisiones se realizan a través de un razonamiento lógico basado en emparejamientos de patrones y manipulaciones simbólicas. Evalúa diferentes opciones de acción **considerando sus objetivos** y selecciona la acción que considera más adecuada según su razonamiento lógico.

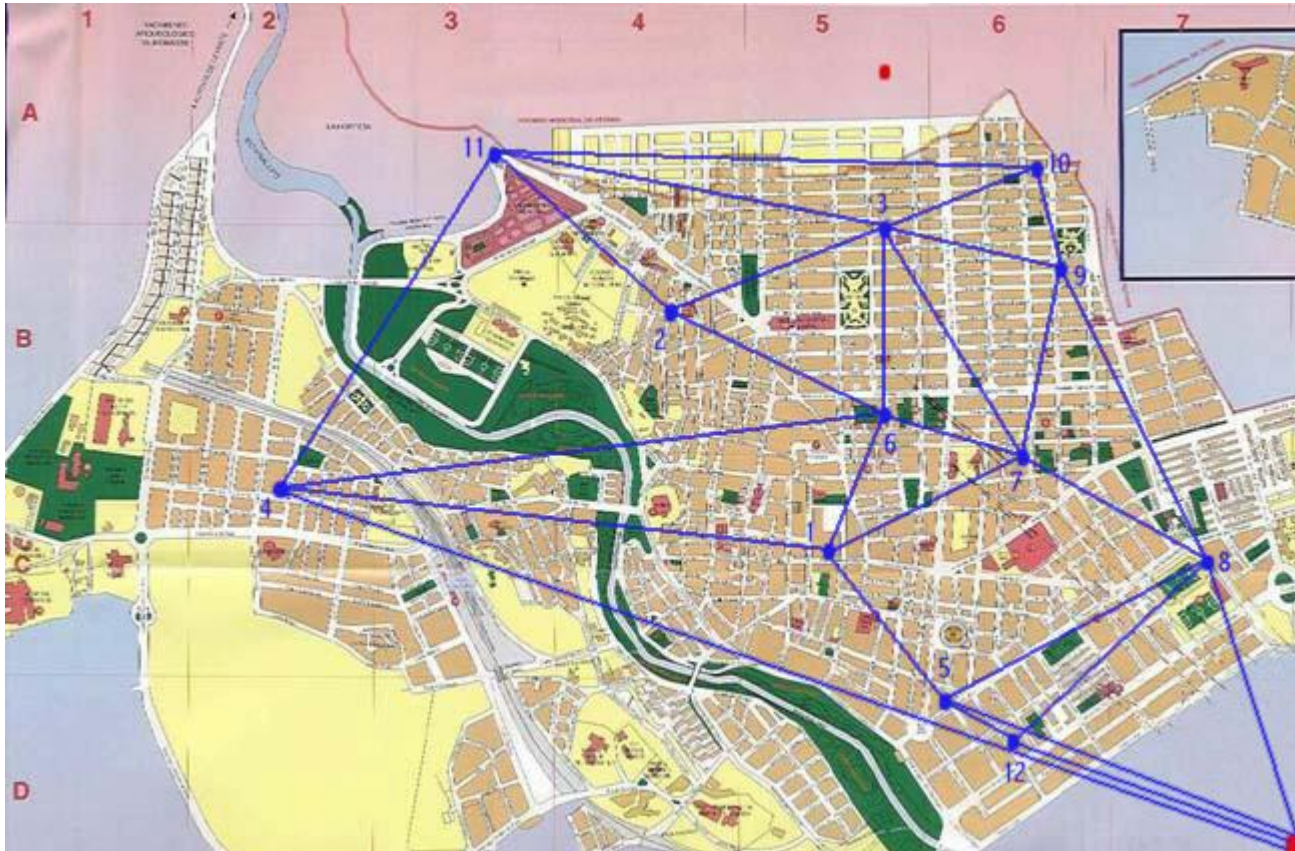
Características: Representación del mundo, Toma decisiones a través de razonamiento lógico. **Planes a largo plazo y cumple sus objetivos.**

Este proceso incluye planificación, predicción de consecuencias y evaluación de alternativas

Arquitecturas deliberativas

- El problema de trasladar en un tiempo razonable para que sea útil el mundo real en una descripción simbólica precisa y adecuada
- El problema de representar simbólicamente la información acerca de entidades y procesos complejos del mundo real, y como conseguir que los agentes razonen con esta información para que los resultados sean útiles

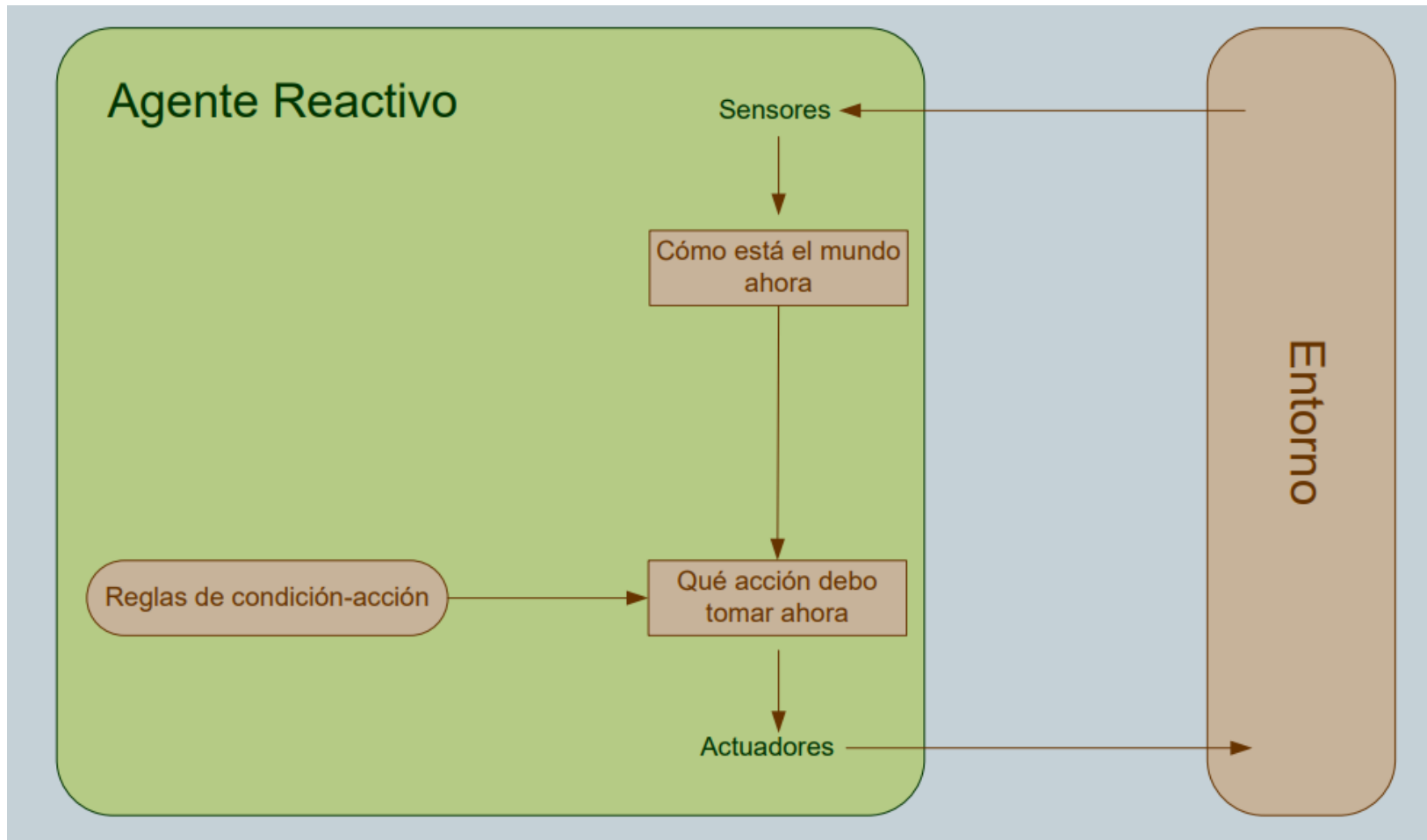
Ejemplo de agente deliberativo: Problema del viajante de comercio



Arquitecturas Reactivas

- Una arquitectura reactiva es aquella que no incluye ninguna clase de modelo centralizado de representación simbólica del mundo, y no hace uso de razonamiento complejo
 - El comportamiento inteligente puede ser generado sin una representación explícita ni un razonamiento abstracto explícito de la clase que la IA simbólica propone.
 - La inteligencia es una propiedad emergente de ciertos sistemas complejos
 - El comportamiento “inteligente” surge como el resultado de la interacción del agente con su entorno.

Arquitecturas Reactivas

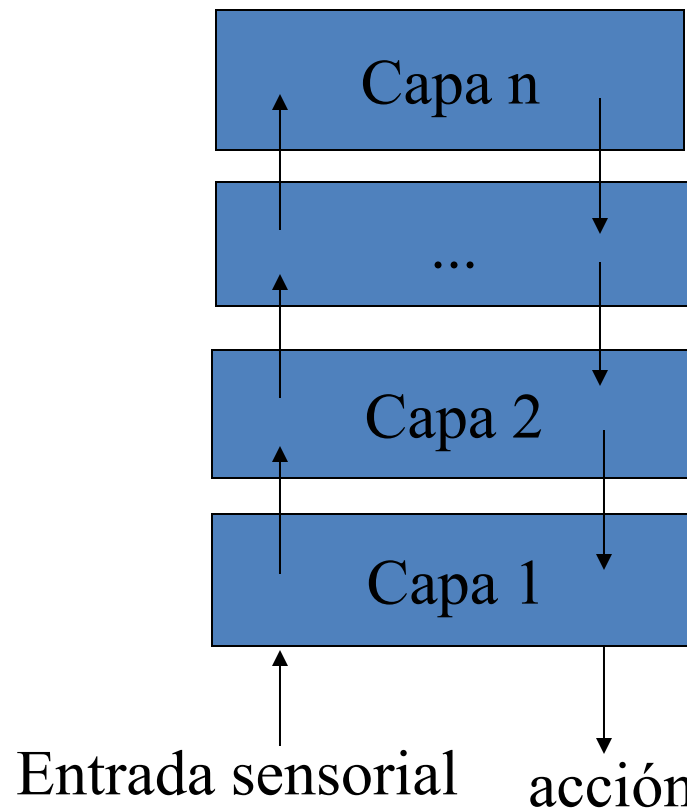


Ejemplo de agentes reactivo: un robot que recorre un pasillo



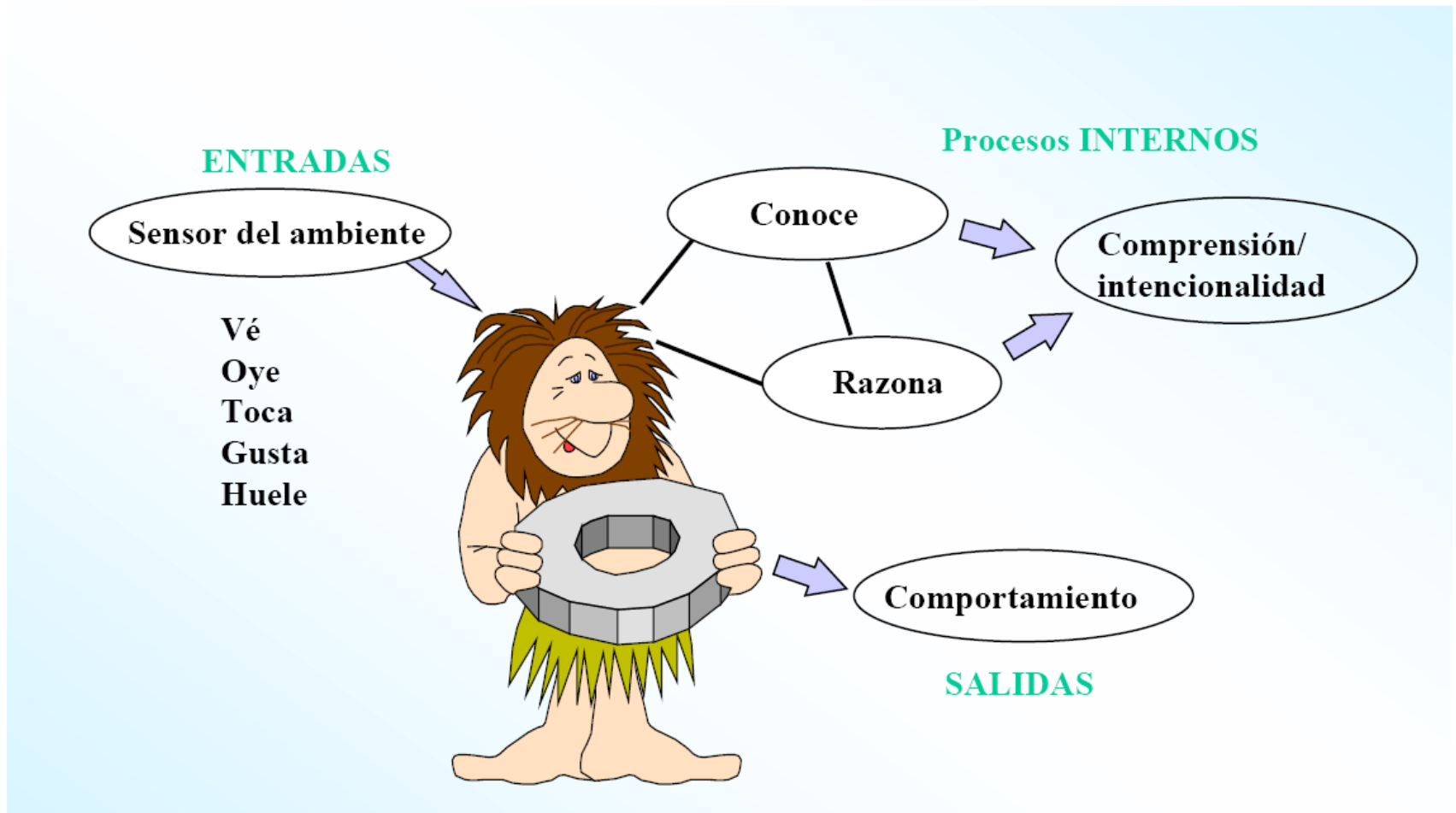
Arquitecturas Híbridas

- Estructura vertical

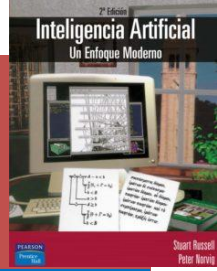


Tiene características de agentes deliberativos y reactivos. Complementan unas cualidades con otras.

Agentes inteligentes (Russell, Norving)



Agentes inteligentes/racionales



La racionalidad en un momento determinado depende de cuatro factores:

- La medida de rendimiento que define el criterio de éxito.
- El conocimiento del medio en el que habita acumulado por el agente.
- Las acciones que el agente puede llevar a cabo.
- La secuencia de percepciones del agente hasta este momento.

Esto nos lleva a la **definición de agente racional**:

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.

Agentes inteligentes/racionales

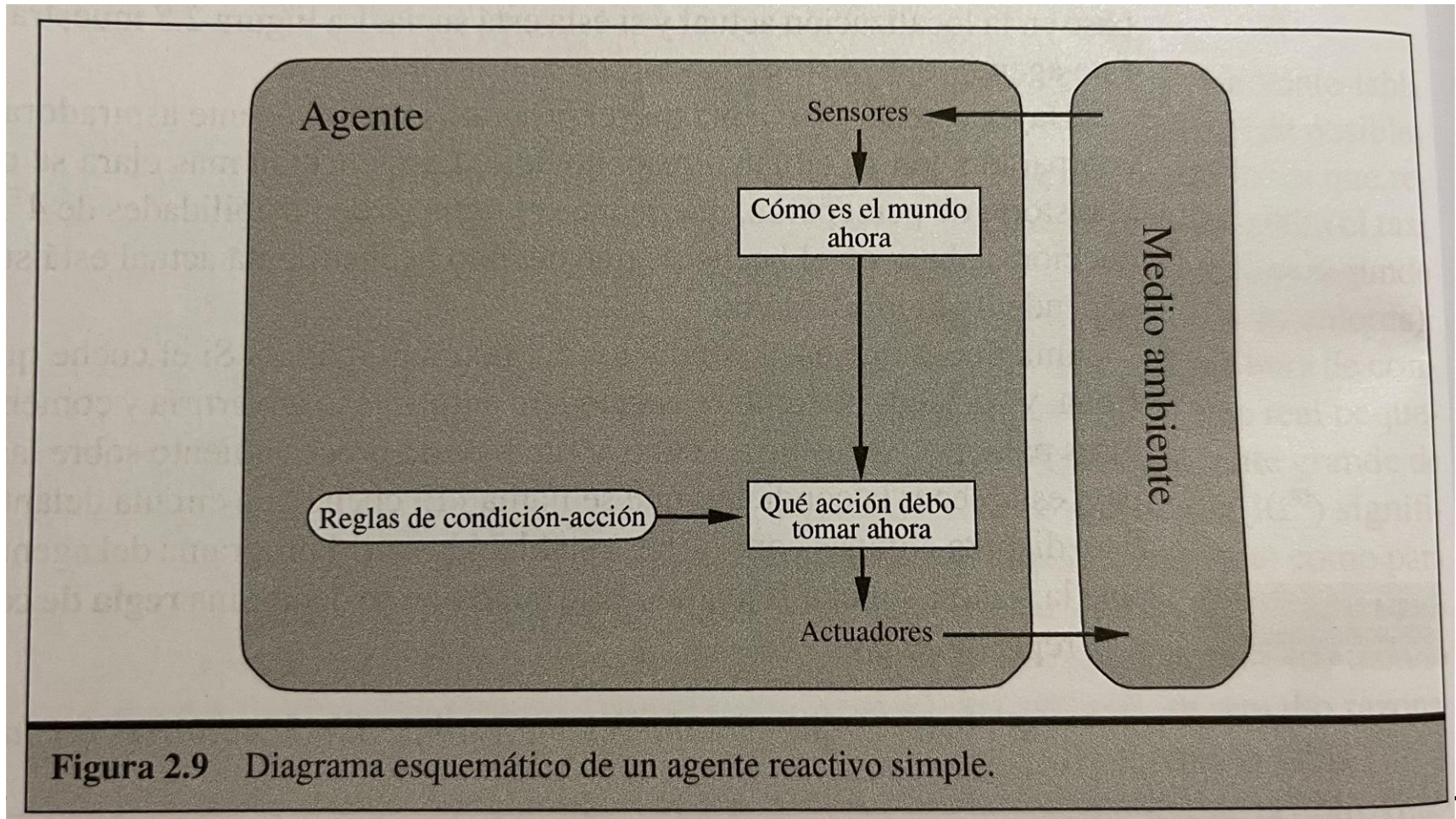
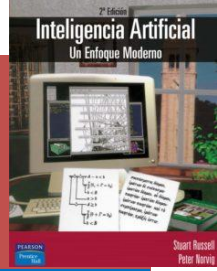


Figura 2.9 Diagrama esquemático de un agente reactivo simple.

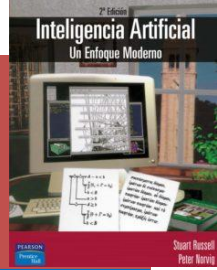
Agentes inteligentes/racionales



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.

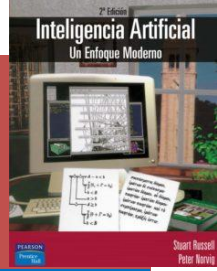
Agentes inteligentes/racionales



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.

Agentes inteligentes/racionales



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*

Figura 2.10 Un agente reactivo simple, que actúa de acuerdo a la regla cuya condición coincida con el estado actual, definido por la percepción.

Agentes inteligentes/racionales

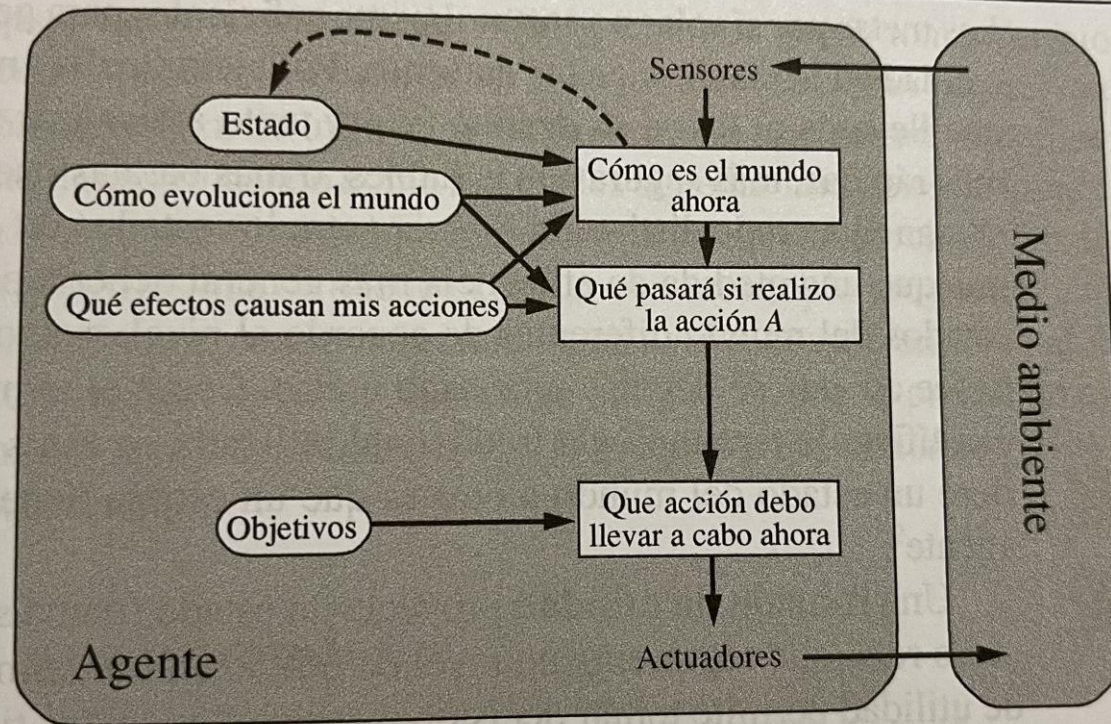
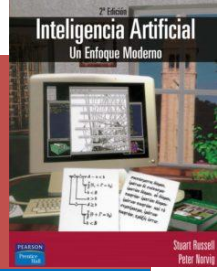


Figura 2.13 Un agente basado en objetivos y basado en modelos, que almacena información del estado del mundo así como del conjunto de objetivos que intenta alcanzar, y que es capaz de seleccionar la acción que eventualmente lo guiará hacia la consecución de sus objetivos.

Agentes inteligentes/racionales

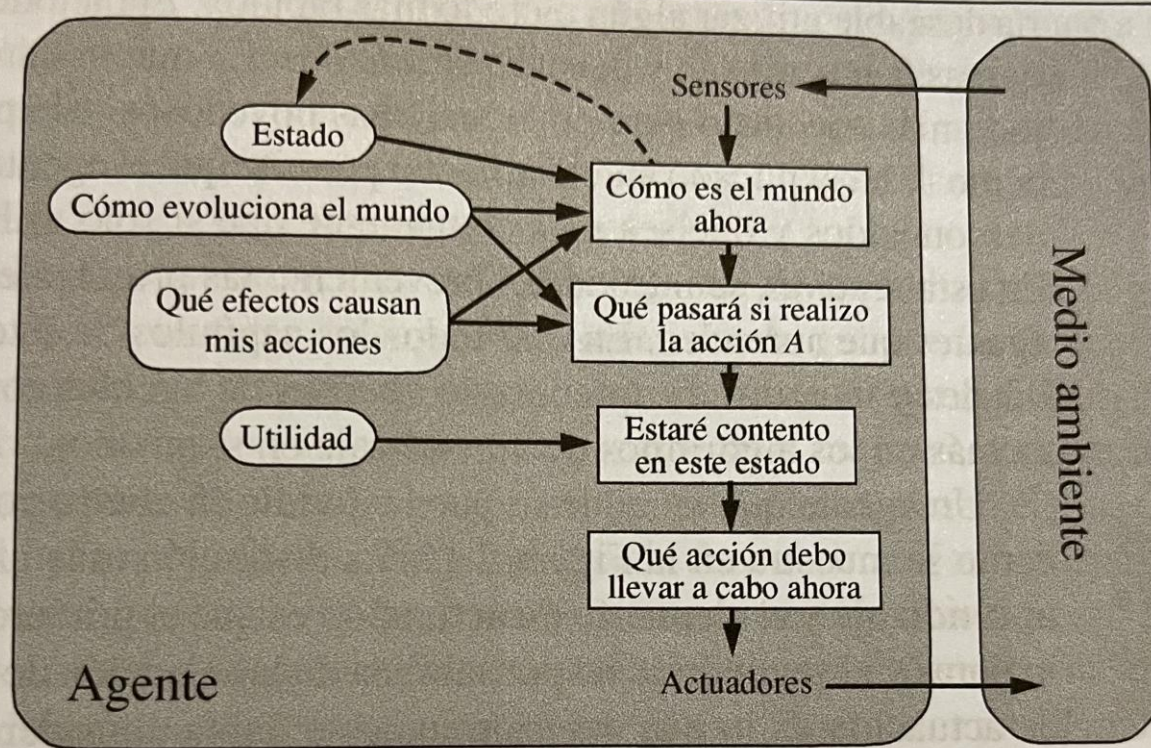
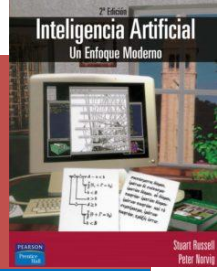


Figura 2.14 Un agente basado en utilidad y basado en modelos. Utiliza un modelo del mundo, junto con una función de utilidad que calcula sus preferencias entre los estados del mundo. Después selecciona la acción que le lleve a alcanzar la mayor utilidad esperada, que se calcula haciendo la media de todos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

Agentes inteligentes/racionales

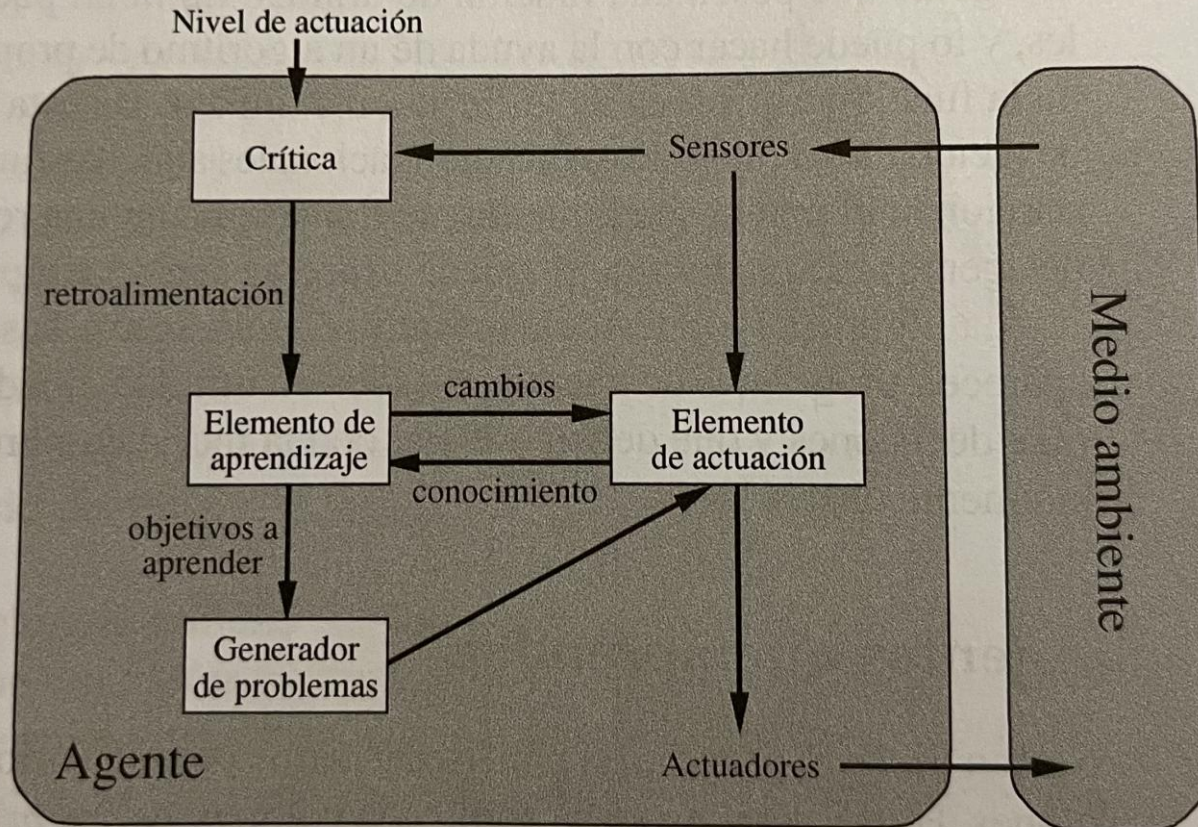


Figura 2.15 Modelo general para agentes que aprenden.

Contenido

- Agentes inteligentes
- Arquitecturas de agentes
- **Agentes reactivos**

Agentes reactivos

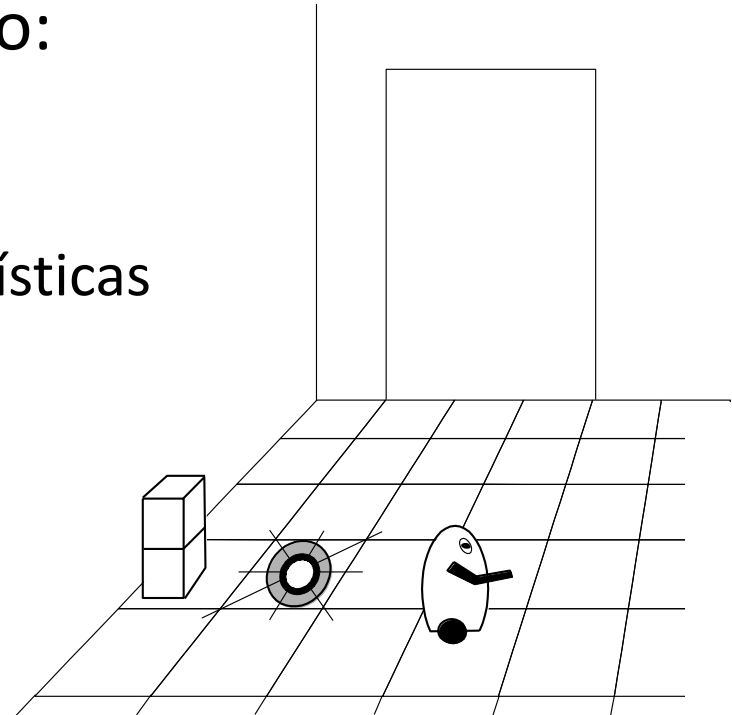
La característica principal de un agente reactivo es su capacidad de responder directamente a estímulos del entorno sin recurrir a un modelo interno o un proceso de planificación. Este tipo de agente actúa de manera inmediata basándose en reglas predefinidas que relacionan percepciones con acciones, lo que lo hace eficiente en entornos dinámicos, aunque limitado en situaciones que requieran razonamiento complejo.

Agentes reactivos

- Representaciones del mundo
- Diseño de un agente reactivo: arquitecturas de agentes
- Agentes reactivos con memoria

Representaciones del mundo

- Representaciones del mundo:
 - modelos icónicos
 - modelos basados en características



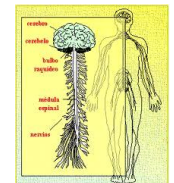
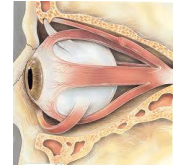
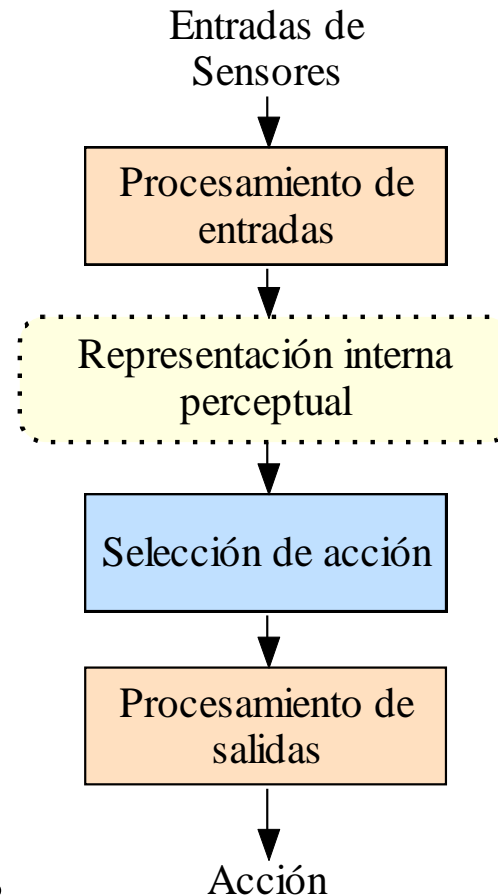
© 1998 Morgan Kaufmann Publishers

El mundo espacial cuadriculado

Diseño de un agente reactivo

• Percepción y Acción:

- El agente reactivo percibe su entorno a través de sensores.
- Procesa la información percibida y hace una representación interna de la misma.
- Escoge una acción, entre las posibles, considerando la información percibida.
- Transforma la acción en señales para los actuadores y la realiza.

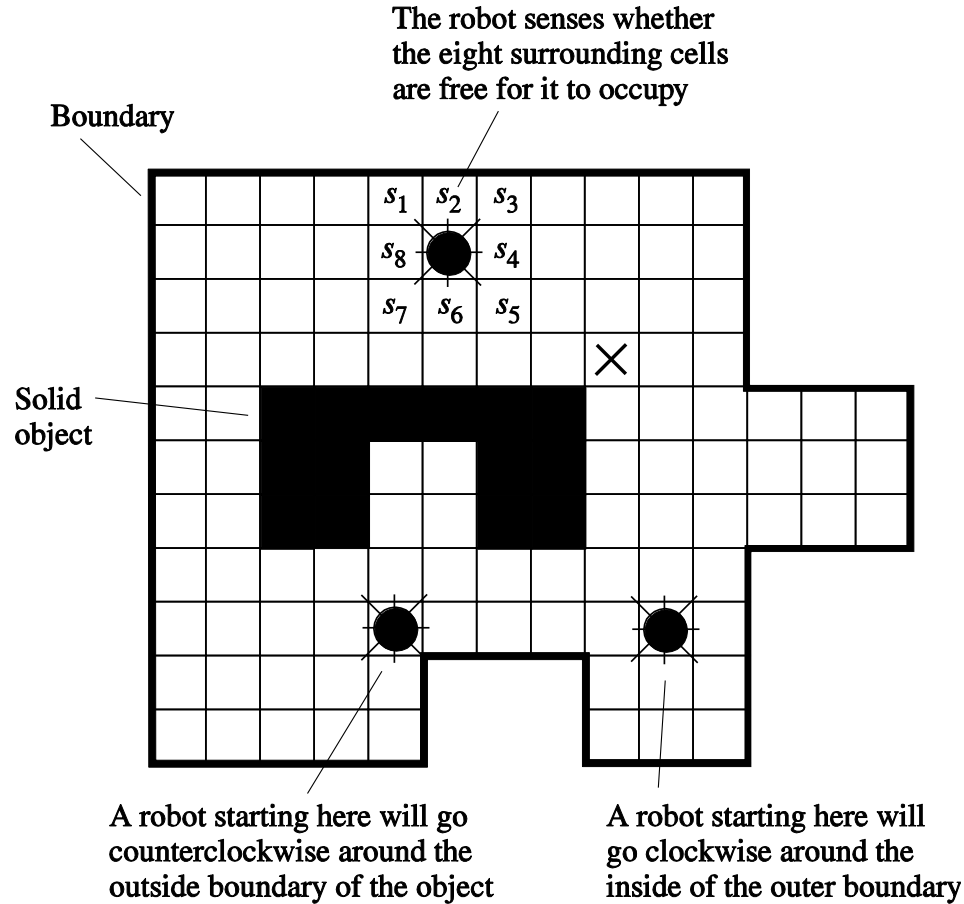


Diseño de un agente reactivo

• Ejemplo:

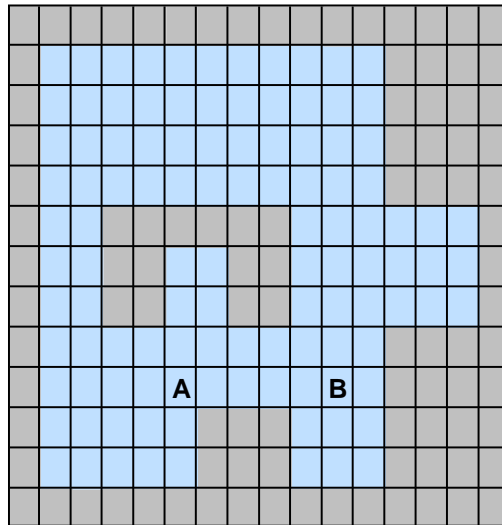
- Supongamos un robot en un mundo dividido en cuadrículas.
- El robot puede percibir si las 8 casillas vecinas están libres o no, con un sensor s_i por cada casilla i .
- El objetivo del robot es ir a una pared y seguir su perímetro indefinidamente.
- Tiene 4 posibles movimientos (de 1 casilla cada uno): Ir a Norte, Sur, Este u Oeste.
- No se permite que el entorno contenga pasillos estrechos (aquellas casillas rodeadas por dos o más obstáculos a ambos lados).

Diseño de un agente reactivo




© 1998 Morgan Kaufmann Publishers

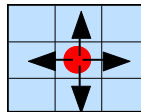
Representación



Sensores:

S_1	S_2	S_3
S_8		S_4
S_7	S_6	S_5

Movimientos:



Usaremos un vector de 8 componentes.

Cada componente i vale 0 si el sensor s_i no detecta obstáculo y vale 1 si lo detecta.

Ejemplo posición **A**:

$A =$

0	0	0	0	1	0	0	0
---	---	---	---	---	---	---	---

Movimientos posibles

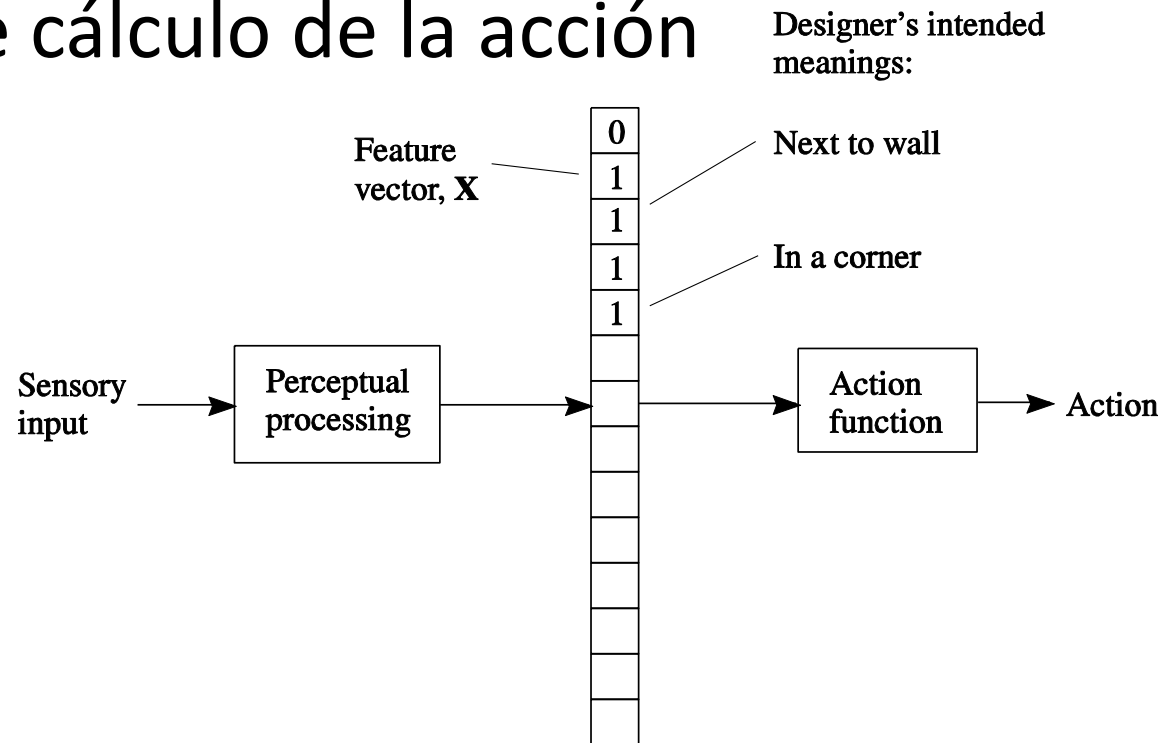
- NORTE: mueve el robot una celda hacia arriba
- ESTE: mueve el robot una celda a la derecha
- SUR: mueve el robot una celda hacia abajo
- OESTE: mueve el robot una celda a la izquierda

TRABAJO DEL DISEÑADOR:

desarrollar una función definida sobre las entradas sensoriales que seleccione la acción apropiada en cada momento para llevar a cabo con éxito la tarea del robot.

Proceso en dos fases

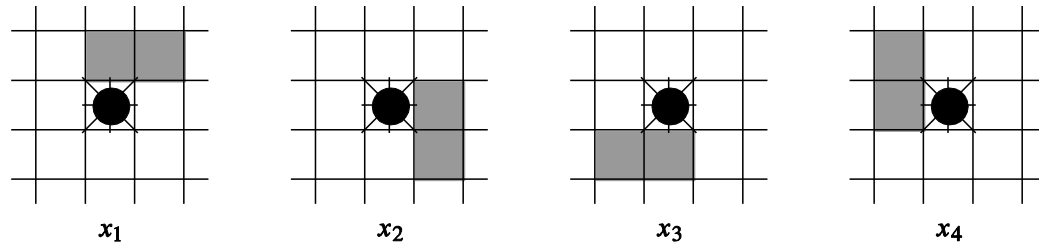
- Procesamiento perceptual
- Fase de cálculo de la acción



© 1998 Morgan Kaufmann Publishers

Percepción y acción

- Percepción:



In each diagram, the indicated feature has value 1 if and only if at least one of the shaded cells is *not* free.

© 1998 Morgan Kaufman Publishers

- Acción:

- si todas las características son cero, moverse al norte
- si $x_1=1$ y $x_2=0$, moverse al este
- si $x_2=1$ y $x_3=0$, moverse al sur
- si $x_3=1$ y $x_4=0$, moverse al oeste
- si $x_4=1$ y $x_1=0$, moverse al norte

Arquitecturas de agentes reactivos

- Sistemas de producción
- Redes
- Arquitecturas de subsunción

Sistemas de Producción

$$c_1 \rightarrow a_1$$

$$c_2 \rightarrow a_2$$

...

$$c_i \rightarrow a_i$$

...

$$c_m \rightarrow a_m$$

en donde C_i es una función booleana definida sobre el vector de características, habitualmente una conjunción de literales booleanos.

Tarea de seguimiento de bordes

Ejemplo de proceso sin fin

$x_4\overline{x_1} \rightarrow$ norte

$x_3\overline{x_4} \rightarrow$ oeste

$x_2\overline{x_3} \rightarrow$ sur

$x_1\overline{x_2} \rightarrow$ este

$1 \rightarrow$ norte

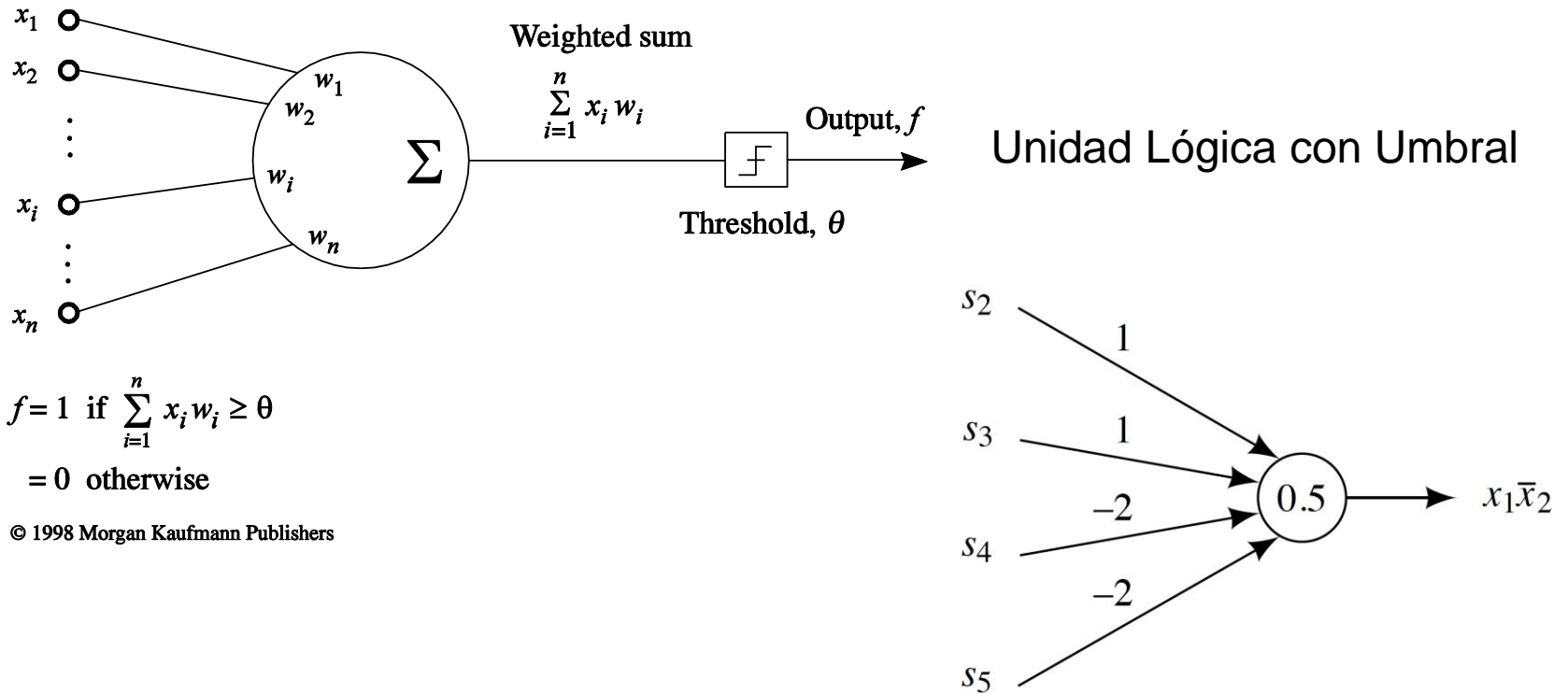
Tarea llevar al robot a una esquina cóncava

$c \rightarrow \text{nil}$

$l \rightarrow s\text{-}b$

Ejemplo de proceso con objetivo

Redes

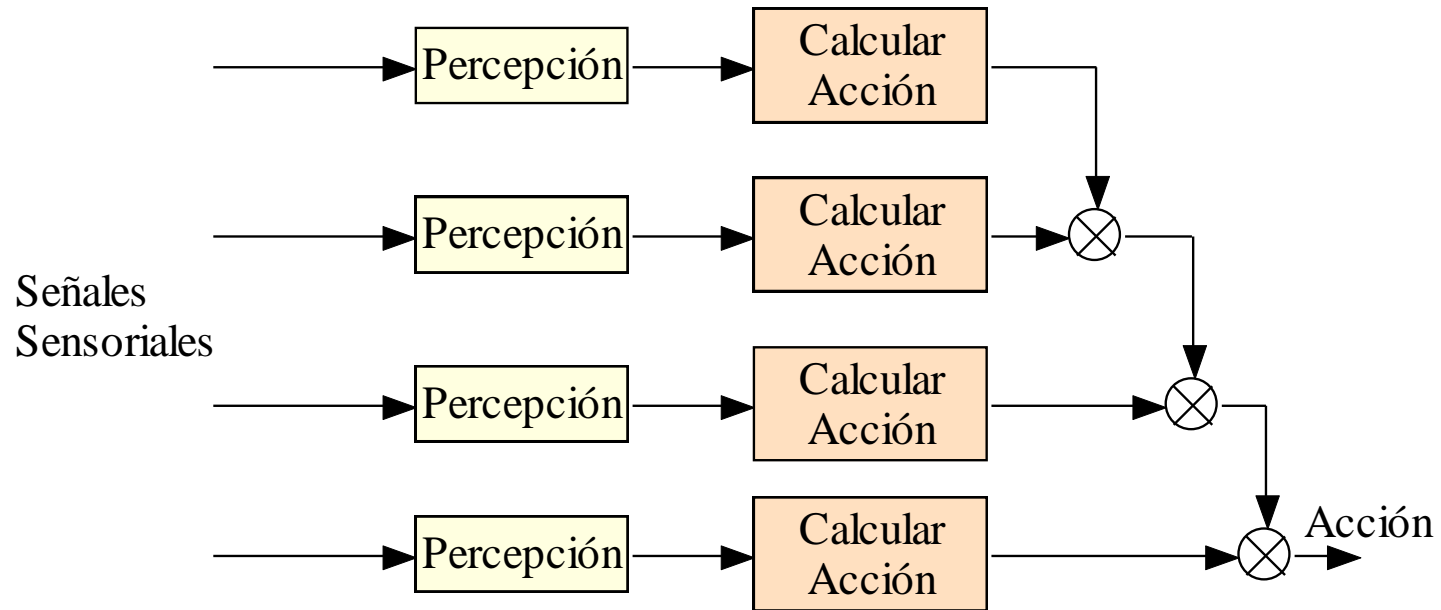


Red neuronal: red de unidades lógicas con umbral

Arquitectura de subsunción

- **La arquitectura de subsunción** consiste en agrupar **módulos de comportamiento**.
- Cada módulo de comportamiento tiene una acción asociada, recibe la percepción directamente y comprueba una condición. Si esta se cumple, el módulo devuelve la acción a realizar.
- Un módulo se puede subsumir en otro. Si el módulo superior del esquema se cumple, se ejecuta este en lugar de los módulos inferiores.

Arquitectura de subsunción

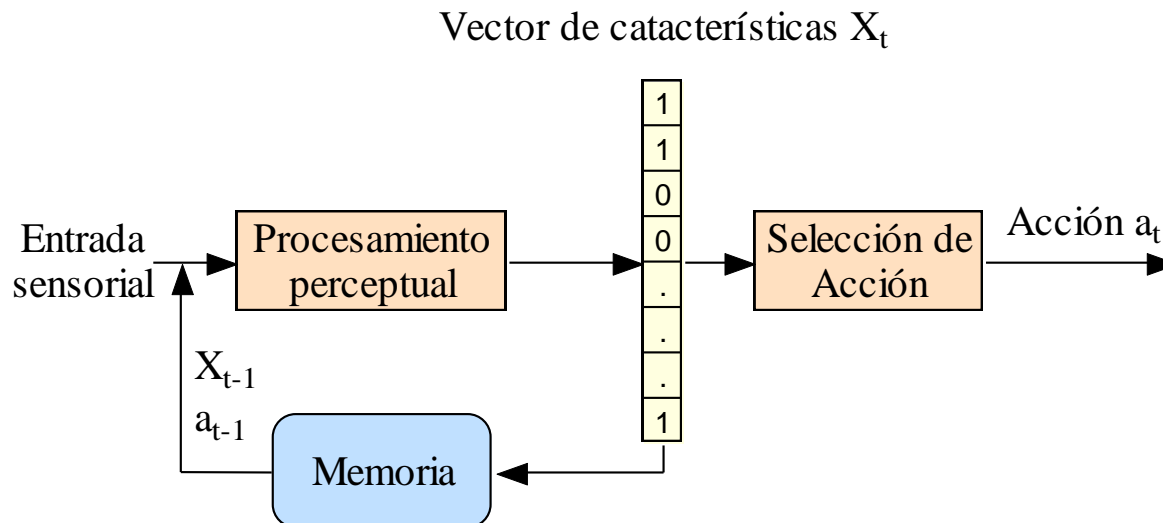


Agentes reactivos con memoria

- Limitaciones del sistema sensorial de un agente.
- Mejorar la precisión teniendo en cuenta la historia sensorial previa: sistemas con memoria

Agentes reactivos con memoria

la representación de un estado en el instante $t+1$ es función de la entradas sensoriales en el instante $t+1$, la representación del estado en el instante anterior t y la acción seleccionada en el instante anterior t .



Ejemplo

- Usaremos las características $w_i = s_i$ $i=2,4,6,8$ y las características restantes del siguiente modo

$w_1=1$ si en el instante anterior $w_2=1$ y el robot se movió al este
 $w_3=1$ si en el instante anterior $w_4=1$ y el robot se movió al sur
 $w_5=1$ si en el instante anterior $w_6=1$ y el robot se movió al oeste
 $w_7=1$ si en el instante anterior $w_8=1$ y el robot se movió al norte

$w_2\overline{w_4} \rightarrow \text{este}$

$w_4\overline{w_6} \rightarrow \text{sur}$

$w_6\overline{w_8} \rightarrow \text{oeste}$

$w_8\overline{w_2} \rightarrow \text{norte}$

$w_1 \rightarrow \text{norte}$

$w_3 \rightarrow \text{este}$

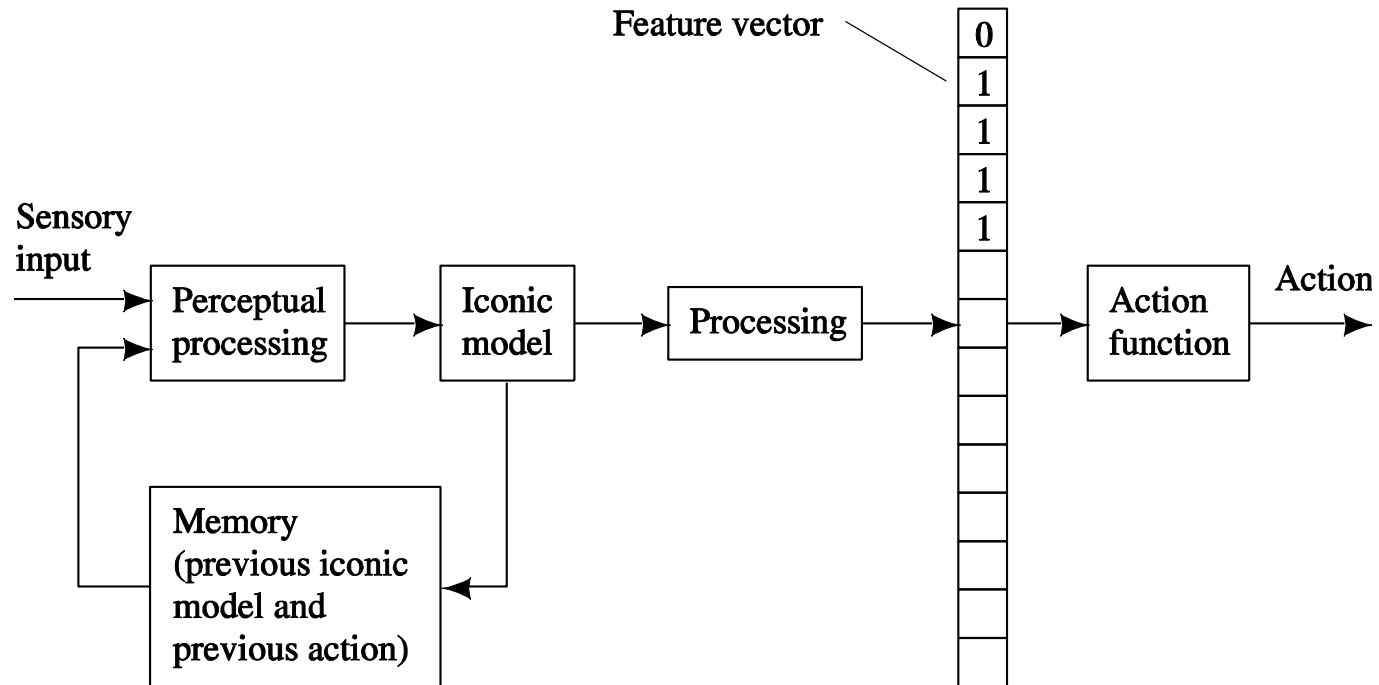
$w_5 \rightarrow \text{sur}$

$w_7 \rightarrow \text{oeste}$

$1 \rightarrow \text{norte}$

Implementación de la memoria con representaciones icónicas

- Adicionalmente el robot podría utilizar otras estructuras de datos: matriz que almacene el mapa con las casillas libres u ocupadas en el momento en el que se percibieron.



Campo de potencial artificial

	1	1	1	1	1	1	1	?
1	0	0	0	0	0	0	0	?
1	0	0	0	0	0	0	0	?
1	0	0	0	0	0	0	0	?
1	0	0	0	0	0	0	0	?
1	0	0	R	0	0	0	0	?
1	0	0	0	0	0	0	0	?
1	0	0	0	0	0	0	0	?
1	0	0	0	0	0	0	0	?
1	?	?	?	?	?	?	?	?
?	?	?	?	?	?	?	?	?

Componente atractiva:

$$p_a(X) = k_1 d(X)^2$$

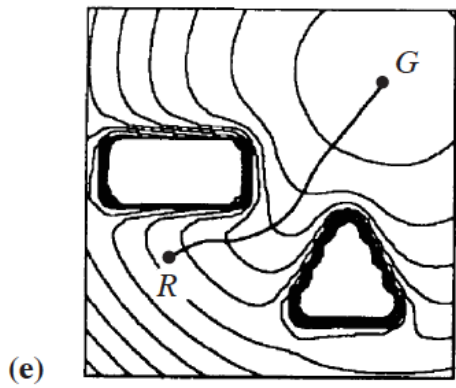
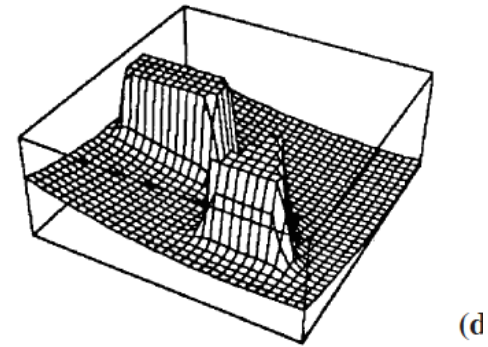
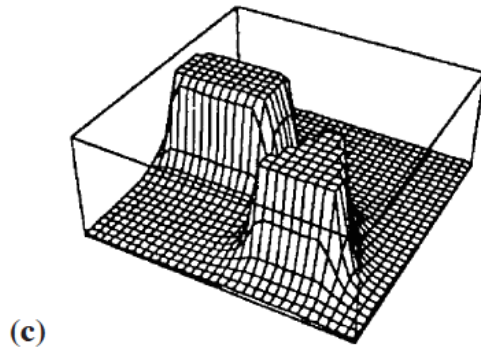
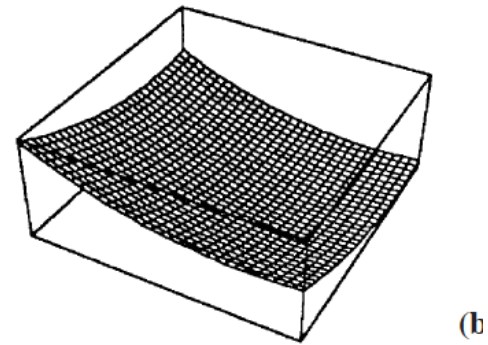
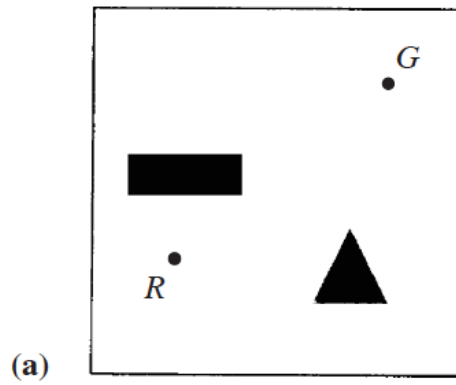
Componente repulsiva:

$$p_r(X) = \frac{k_2}{d_0(X)^2}$$

Potencial:

$$\text{Potencial} = p_a + p_r$$

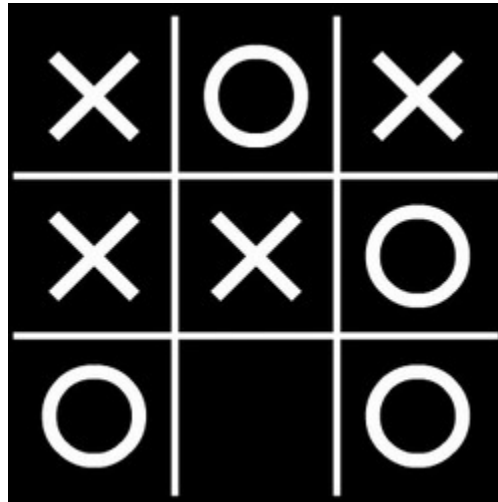
© 1998 Morgan Kaufman Publishers



Ejemplo de agente reactivo: un robot que recorre un pasillo



Ejemplos de agente reactivo: un agente que juega al tres en raya



Características de los agentes reactivos

- Se diseñan completamente y por tanto es necesario anticipar todas las posibles reacciones para todas las situaciones
 - Realizan pocos cálculos
 - Almacenan todo en memoria

¿Qué tipo de agente es ChatGPT?

- ChatGPT es un **agente deliberativo**, ya que utiliza un proceso interno para generar respuestas basadas en entradas del usuario, un vasto modelo de conocimiento, y razonamiento complejo. Evalúa el contexto y decide la mejor "acción" (respuesta) a través de un procesamiento reflexivo y estructurado. Sin embargo, también puede presentar características reactivas, ya que responde rápidamente a estímulos (mensajes) sin requerir planificación a largo plazo. Por lo tanto, combina elementos deliberativos con un alto grado de eficiencia reactiva. (Respuesta de Copilot).