

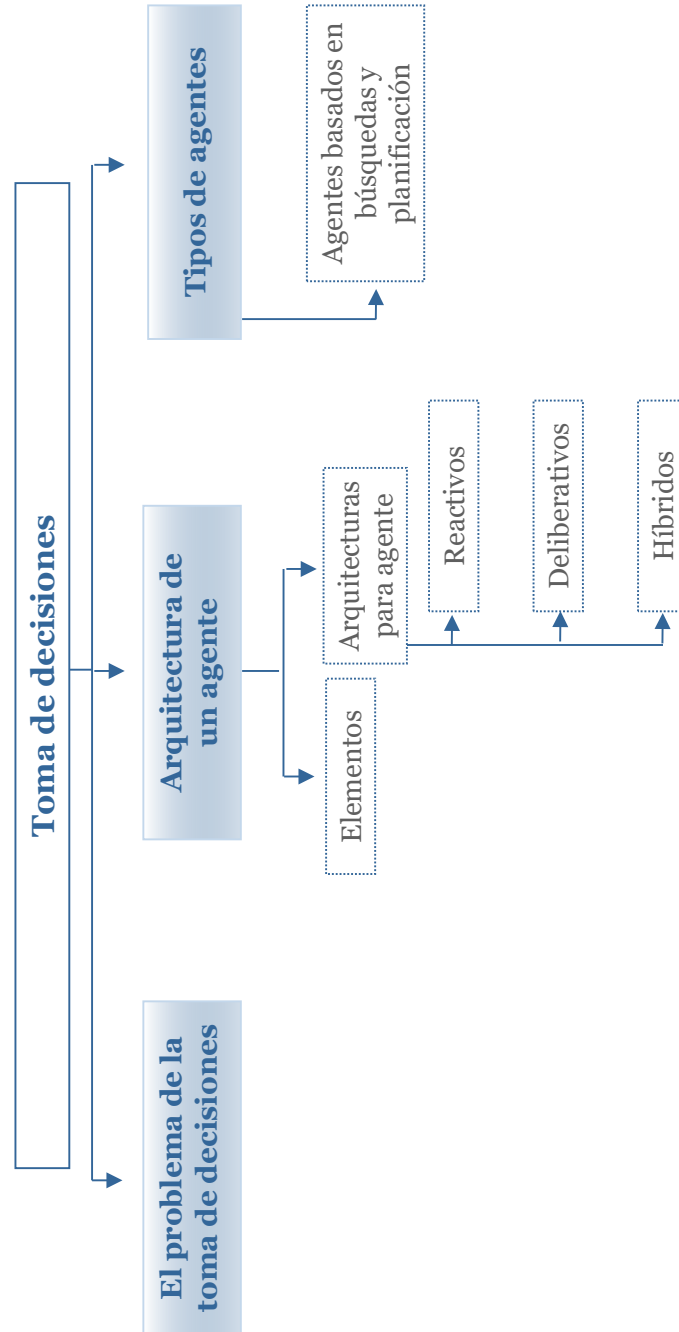
Razonamiento y planificación automática

Introducción a la toma de decisiones

Índice

| | |
|--|----|
| Esquema | 3 |
| Ideas clave | 4 |
| 1.1. ¿Cómo estudiar este tema? | 4 |
| 1.2. Problemas de toma de decisiones | 4 |
| 1.3. Arquitectura de un agente inteligente | 16 |
| 1.4. Tipos de agentes inteligentes | 30 |
| 1.5. Referencias bibliográficas | 32 |
| Lo + recomendado | 34 |
| + Información | 36 |
| Test | 39 |

Esquema



1.1. ¿Cómo estudiar este tema?

Para estudiar este tema lee las **Ideas clave** que encontrarás a continuación.

En este tema estudiaremos los principios necesarios para entender el problema de la toma de decisiones para los sistemas inteligentes. Describiremos la problemática humana de la toma de decisiones y su influencia al construir agentes inteligentes.

Partiremos de las estrategias mentales que pueden ayudar a estructurar la toma de decisiones y plantearemos los mecanismos tradicionales para estructurar los problemas, dividiendo en etapas el proceso de definición del problema.

Completaremos los conceptos iniciales con la definición de **agente inteligente** y de los tipos de agentes inteligentes que nos podemos encontrar según su definición arquitectónica. Presentaremos los distintos elementos que componen estas arquitecturas y su utilidad.

1.2. Problemas de toma de decisiones

La complejidad en el proceso de toma de decisiones es evidente, ya que es necesario evaluar y conocer las condiciones del entorno. Como guías podemos emplear la intuición y las reglas, pero en ocasiones es necesario realizar un análisis exhaustivo de la totalidad de los posibles escenarios. Es en este

contexto cuando las matemáticas aplicadas a la toma de decisiones pueden servir de ayuda.

Todos tenemos dificultades para tomar decisiones, especialmente cuando la decisión debe tomarse frente a una situación de riesgo. En este tipo de casos, la intuición no suele ser una herramienta y las matemáticas son un buen recurso para ayudar a tomar decisiones.

En la actualidad, existen varios desafíos que generan situaciones de riesgo (terrorismo, crisis económico-financieras, ciberseguridad, cambio climático...) y la toma de decisiones en estos casos es muy importante.

Cuando una situación necesita ser analizada y una decisión tomada, es necesario asumir un riesgo; por tanto, es necesario evaluar e identificar el riesgo involucrado y decidir cuáles son las medidas a adoptar para que este sea mínimo. La capacidad que tenemos para tomar decisiones está relacionada con la asunción de riesgo, la creatividad y la búsqueda de alternativas a los retos que aún no existen.

Los elementos a tener en cuenta en la toma de decisiones son:

- ▶ Efecto futuro: cómo afectará la decisión en el futuro. Cualquier decisión influye: algunas lo hacen a corto plazo (decisiones de bajo nivel), pero otras influyen a largo plazo (decisiones de alto nivel).
- ▶ Reversibilidad: es el esfuerzo necesario para revertir una decisión y la velocidad con la que una decisión puede revertirse. Cuando es difícil revertir, es mejor realizar decisiones a alto nivel. Si es simple, será necesario tomar decisiones a bajo nivel.
- ▶ Impacto: está relacionado con otras actividades y áreas que se ven afectadas por la decisión tomada. El impacto puede ser: extenso (abarca varias áreas o actividades y es recomendable aplicar decisiones a alto nivel) o localizado, único

(solo afecta a una o pocas áreas o actividades y se asocia con una decisión de bajo nivel).

- ▶ **Calidad:** la ética, la legalidad, los principios que gobiernan el comportamiento, la imagen y las relaciones laborales, entre otras, entran en juego en este punto. Cuando muchos de estos factores están involucrados, es necesario tomar decisiones a alto nivel. En cambio, cuando hay pocos implicados, es recomendable tomar la decisión a bajo nivel.
- ▶ **Periodicidad:** este elemento se refiere a la frecuencia o excepcionalidad de una decisión. Las decisiones excepcionales son de alto nivel, mientras que las decisiones que deben ser tomadas con mayor frecuencia son decisiones de bajo nivel.

La toma de decisiones a **alto nivel** requiere seguir un proceso de **análisis serio, buscar alternativas, planificar, ejecutar y evaluar**. La toma de decisiones a **bajo nivel** implica **poco esfuerzo** y es posible tomar este tipo de decisiones en **poco tiempo**.

Resumen de elementos de toma de decisiones de alto nivel y decisiones de bajo nivel:

| Alto Nivel | Bajo Nivel |
|---|--|
| Afectan al futuro. | No afectan al futuro. |
| Reversibilidad difícil. | Reversibles. |
| Impacto amplio. | Poco impacto. |
| Afectan a muchos factores importantes de calidad. | Afectan a pocos factores importantes de calidad. |
| Excepcionales. | Frecuentes. |

Tabla 1. Comparativa entre las decisiones de alto y bajo nivel.

Por otra parte, es posible clasificar las decisiones en dos tipos: decisiones **programadas** y decisiones **no programadas**.

Las decisiones **programadas** son de rutina y se repiten periódicamente. Tienen que ver con problemas bien definidos y no requieren un proceso de decisión complejo. Normalmente se administran con una secuencia de pasos, hay un método que es capaz de manejar este tipo de decisiones y existen reglas que permiten dirigir el pensamiento en una dirección específica definida. Por ejemplo: en una emergencia en el hogar, llame al número de emergencia.

Las decisiones no **programadas** están relacionadas con condiciones o entornos desconocidos, situaciones nuevas y no existen reglas o métodos establecidos que puedan servir como guía. Este tipo de decisiones son exclusivas de la situación, también única.

Las decisiones no programadas, en general, necesitan más tiempo para tomarse, ya que hay múltiples variables que deben valorarse y ponderarse, y también la información disponible es incompleta, ya que no es posible anticipar cuál será el impacto en el resultado de la decisión tomada. Por ejemplo: frente al menú de un restaurante que es la primera vez que visitamos.

En cuanto al problema, podemos diferenciar el **problema estructurado** del **no estructurado**.

- ▶ Problema estructurado: el enunciado contiene toda la información necesaria para poder resolverlo.
- ▶ Problema no estructurado: el enunciado no contiene toda la información que sería necesaria para resolverlo. Es necesario que la persona que se enfrenta a él busque información adicional para añadirla.

Etapas necesarias para la resolución de problemas

Primera etapa

Comprender la complejidad del problema es el objetivo de esta etapa. Analizaremos el problema, sintetizándolo por medio del pensamiento hipotético-deductivo, obteniendo una visión general del mismo.

► Identificar el problema

Formularemos e intentaremos verificar la veracidad de una hipótesis que ilustre el problema. La mayoría de las veces, en nuestro día a día, los problemas se presentan de un modo claro y estructurado (en general, alrededor nuestra esperando ser descubiertos). Es importante tener una actitud activa y tener la intención de hacer x o querer hacer y.

La visión correcta del problema es observar las tendencias, analizar la evolución de los hechos y usar la creatividad para anticipar e imaginar qué es posible que suceda. Es necesario tomar distancia al analizar los hechos, realizar un análisis con frialdad y confiar en las experiencias pasadas para predecir lo que sucederá en el futuro, pero sin que estas o los datos entren en conflicto con decisiones futuras. La materia prima para tomar decisiones es la información y, cuanto mejor sea, mejor será la calidad de la decisión.

La **información** es primordial en la toma de decisiones: a mayor calidad de esta, mejor es la calidad de la toma de decisiones. La **creatividad**, la **experiencia** y la **intuición**, junto con el conocimiento de las **experiencias pasadas**, ayudan a la hora de elaborar las predicciones de lo que puede suceder en el futuro.

Todas estas claves deberán considerarse necesarias a la hora de crear mecanismos automáticos para la toma de decisiones y resolución inteligente de problemas, ya sea de forma natural o artificial.

¿Cómo identificar el problema?

En la siguiente tabla se muestran preguntas que sirven para ayudar a delimitar un problema.

| | ¿Dónde se encuentra el problema? | ¿Dónde no está el problema? |
|-----------------|--|--|
| Origen | Determinar los aspectos donde es visible la situación que está provocando el problema. | En qué aspectos no se manifiesta el problema. |
| Magnitud | Número de personas afectadas y sus características (edad, género, contexto socioeconómico...). | En qué áreas no aparece el problema. |
| Lugar | Determinar el área geográfica (oficina, provincia, escuela...). | A qué personas no afecta. |
| Foco | Determinar, del mismo modo, grupos de afectados y el número de integrantes de los grupos. | Dónde no aparece el problema. |
| Historia | Determinar si es un hecho puntual y reciente, si perdura y se alarga en el tiempo... | En qué momento o momentos no se produce la situación problemática. |

Tabla 2. Preguntas para delimitar un problema.

El **análisis DAFO** es otra técnica empleada para diseñar una estrategia para determinar problemas. Se trata de realizar una valoración en grupo sobre el presente y el futuro, valorando lo positivo y lo negativo.

| | |
|--------------|----------------|
| Debilidades. | Amenazas. |
| Fortalezas. | Oportunidades. |

Tabla 3. Análisis DAFO.

Por último, dentro de las técnicas para determinar problemas, vamos a describir la **reunión de discusión**. Consiste en compartir ideas cuando se va a comenzar un proyecto. Requiere una preparación previa concretando los temas que preocupan al equipo, disponer a los asistentes en círculo para favorecer la comunicación y que todos se vean las caras, no superar las dos horas de reunión y no complicarse demasiado realizando los preparativos. Este tipo de técnica se emplea con mucha frecuencia en la metodología Scrum.

Puedes ver más sobre Scrum en el siguiente enlace:

<https://proyectosagiles.org/que-es-scrum/>

Existen varias técnicas para delimitar problemas: preguntarse por el origen, lugar, magnitud, foco o historia; ponerse en el lugar del cliente; análisis DAFO y reunión de discusión.

► Explicar el problema

Es necesario explicar el problema para definir dónde se origina, dónde ocurre, cómo se produce y a quién afecta. Es hora de realizar tareas para profundizar en el problema, comprenderlo y desarrollar una explicación satisfactoria basada en los elementos que lo causan (las causas). De esta manera, será posible comenzar a crear una estrategia para resolverlo.

Para **explicar el problema**, hay que ir más allá de los síntomas y poner el foco en los aspectos que lo provocan: **indagar las causas**. Para lograr explicarlo se pueden seguir los siguientes pasos:

- Valorar la importancia del problema: es clave determinar la importancia de un problema con los criterios objetivos, haciendo un juicio de la importancia del problema en nuestro objetivo.

- Definir límite: encontrar el límite concreto del problema. Los problemas se originan de una manera difusa, mal definida. Antes de encontrar la solución, debemos conocer bien el problema y definir cuál es y cuál no, y buscar el objeto relacionado con el problema.
- Detectar las causas y las consecuencias: al reducir el problema (definir el límite), también lo describimos. Pero es necesario detectar y analizar las causas, ir al origen del problema, determinar y ayudar a predecir las consecuencias.

Existen varias **técnicas para dar explicación a los problemas**: clasificación ERIM de problemas, método de los seis interrogantes, las veinte causas, diagrama de espina de pez y mapas mentales.

- ERIM: se basa en clasificar los problemas según dos criterios (urgencia e importancia), elaborando una matriz del tipo:

| | | | |
|------------|---|---------|---|
| | | Urgente | |
| | | + | - |
| Importante | + | | |
| | - | | |

Tabla 4. Clasificación ERIM.

- Seis interrogantes sobre el problema: ¿cuál es?, ¿cómo sucede?, ¿por qué ocurre?, ¿a quién afecta?, ¿dónde sucede? y ¿cuándo pasa?.
- Las veinte causas: hay que crear una lista de los posibles elementos que causan el problema, posteriormente se van eliminando causas para quedarse con las veinte más importantes.
- Diagrama de espina de pez: también conocido como diagrama Ishikawa, permite mostrar las causas de un problema, organizarlas y asignarles unos pesos o valores. Consiste en una técnica llevada a cabo en varias sesiones divididas en fases.

Primero, se debe dibujar un diagrama en blanco y luego escribir brevemente el problema a analizar. Se identifican las categorías del problema que se consideran apropiadas. Se determina una causa principal y se pregunta sobre el origen de esa causa, esto ayudará a establecer causas secundarias cuya consecuencia es la causa principal (análisis de dispersión). Se analizan en detalle todas las causas, tanto principales como secundarias.

Ejemplo de un diagrama de pez (causa-efecto) o diagrama Ishikawa:

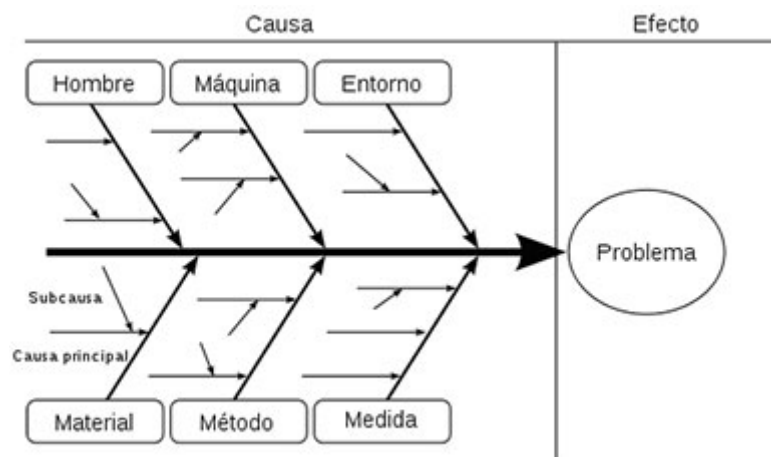


Figura 1. Diagrama espina de pez. Fuente: www.wikipedia.org

- Mapas mentales: sirven fundamentalmente para crear ideas a partir del análisis sistemático y detallado del problema. Los mapas mentales son una forma de expresar un razonamiento radiante (desde el interior hasta el final). La siguiente imagen muestra el mapa mental para hacer un DAFO.



Figura 2. Ejemplo de mapa mental para especificar un análisis DAFO. Fuente: <https://www.leadersummaries.com/ver-resumen/mapas-mentales-para-la-empresa>

Un problema bien definido: la capacidad de explicar el problema nos llevará a construir definiciones consistentes y bien definidas, de tal manera que los agentes inteligentes puedan diseñarse con un objetivo claro, entendiendo correctamente cuál es el objetivo final de su desempeño y su entorno real de desarrollo.

Deberemos, por tanto, crear definiciones de problemas que presenten características que nos permitan trabajar con ellos de modo eficiente. Algunas características de entornos simples (problemas bien definidos) son:

- Discreto:
 - Se puede concebir el mundo en estados.
 - En cada estado hay un conjunto finito de percepciones y acciones.
- Accesible: el agente puede acceder a las características relevantes del entorno.
 - Puede determinar el estado actual del mundo.
 - Puede determinar el estado del mundo que le gustaría alcanzar.
- Estático y determinista: no hay presión temporal ni incertidumbre.
 - El mundo cambia solo cuando el agente actúa.
 - El resultado de cada acción está totalmente definido y es previsible.

Segunda etapa

Crear una estrategia para resolver el problema que reduzca al mínimo los efectos negativos y haga que sea posible alcanzar los logros.

Con el problema bien definido y delimitado, ingresamos en la fase que permite encontrar las soluciones.

► Idear estrategias

Existen muchas técnicas para encontrar estrategias y soluciones. Algunas de las empleadas en la toma de decisiones estratégica pueden ser: *brainstorming*, 4x4x4, la pecera y seis sombreros para pensar. Son interesantes para reflexionar acerca de la complejidad de la toma de decisiones.

► Elegir estrategia y tomar una decisión

Para tomar una decisión, es necesario elegir de entre una serie de posibilidades aquella que sea la más adecuada para resolver el problema. En este punto, es necesario valorar las estrategias que han surgido, usando aquella metodología elegida en la fase de idear estrategias y tomar la decisión. Para elegir hay que valorar:

- Beneficios.
- Probabilidad de éxito.
- Dependencias internas y externas (¿depende de nosotros llevarlo a cabo o depende de más personas?).
- Medios, materiales y personas que son necesarios para llevar a cabo la estrategia.
- Estimación de tiempo necesario para poner en marcha la estrategia.
- Costes.

Desde el punto de vista de los agentes inteligentes, veremos que la estrategia a utilizar dependerá mucho del problema al que nos enfrentemos, de ahí la importancia de definir correctamente este.

Muchas técnicas son generalmente aplicables a muchos problemas, pero existen técnicas que mejoran las posibilidades de éxito frente a ciertos problemas. Por ejemplo, si tenemos que resolver un problema en el que el orden en el que se toman las decisiones no es importante, pero sí el resultado final de ciertas variables, son más eficientes las técnicas de CSP (restricciones) o los algoritmos genéticos. Si encontramos un conocimiento *a priori* alto sobre el problema, pero una definición vaga de los predicados y términos utilizados, deberíamos buscar modelos de inferencia difusos, etc.

Por tanto, es cometido del especialista en planificar la solución el conocer bien el problema antes de plantear estrategias y saber también el mayor número de técnicas posibles para emplear la más conveniente a cada problema.

► Diseñar intervención

A continuación, deberemos planificar las acciones que se van a realizar, determinando una hoja de ruta. Deberemos emplear el mejor diseño arquitectónico posible para poder garantizar la mejor toma de decisiones, de modo eficiente y organizado.

Tercera etapa

Consiste fundamentalmente en conseguir resolver el problema. Es la etapa en la cual se implementa la estrategia óptima ejecutando la intervención. Por último, el proceso de toma de decisiones se cerrará con una evaluación de logros que servirá para mejorar y como aprendizaje.

1.3. Arquitectura de un agente inteligente

Definición de agente inteligente

De las múltiples interpretaciones y técnicas empleadas para construir sistemas que resuelvan problemas, la **inteligencia artificial** es la disciplina derivada de los conocimientos científicos e ingenieriles que estructura las metodologías, técnicas y algoritmos que apoyan la resolución de estos problemas, intentando emplear mecanismos sofisticados de razonamiento y usando de modo eficiente recursos y conocimientos.

Dentro de la inteligencia artificial, la **teoría de agentes inteligentes** diseña los sistemas de toma de decisiones por medio de la filosofía de inmersión de un elemento (agente) en un entorno, con el cual debe interactuar por medio de sus sistemas de percepción (que le proporcionan información de estado de dicho entorno) y sistemas actuadores (que le permiten modificar el estado para conseguir un objetivo).

Es importante resaltar la **definición de agente inteligente** que emplearemos a lo largo de este curso. Para nosotros, un sistema se considerará inteligente siempre que intente obtener una meta por medio de un comportamiento racional. Evitaremos las definiciones de inteligencia artificial «fuerte» que consideran factores más allá de la racionalidad para clasificar un agente como inteligente, dado que necesitaríamos de definiciones reflejo del comportamiento humano para considerar estos agentes «fuertes».

Por nuestra parte, un agente será más inteligente cuanto más autónomo sea, es decir, que necesite menos conocimiento *a priori* para encontrar una solución racional a su problema. El concepto de conocimiento *a priori* surge de cualquier información previa sobre el comportamiento del entorno que se le tenga que proporcionar al

agente para tomar decisiones. Todo conocimiento de este tipo que le ofrezcamos a un agente le resta autonomía y, por tanto, capacidad de tomar decisiones.

La **racionalidad del agente** se medirá a través de una medida de desempeño (normalmente nos referiremos a ella como desempeño, utilidad o idoneidad).



Figura 3. Representación gráfica de agente inteligente.

Elementos de un agente inteligente

En términos generales, un agente inteligente debe organizar su proceso de toma de decisiones en tres fases principales:

- ▶ Sentir: le proporciona al agente la percepción del entorno. Esta información tiene que ser procesada por el sistema de percepción que debe:
 - Extraer de las estructuras de datos que representan la información percibida, datos en estructuras «amigables» para el razonamiento.
 - Procesar dichas estructuras extraídas para generar una representación reducida de elementos relevantes para el razonamiento, conceptos de alto nivel relacionados con el proceso de razonamiento del agente.

- ▶ **Pensar:** el proceso de toma de decisiones racionales del agente. Empleando la información disponible del sistema de percepción y la **memoria** interna del agente, inicia un proceso deliberativo en el cual debe:
 - Razonar, extrayendo de los hechos relevantes observados y recordados conclusiones nuevas, por ejemplo, de la posición de un nuevo agente respecto a la antigua: su trayectoria.
 - Decidir: empleando mecanismos para la toma de decisión de acciones (los algoritmos tradicionales: búsqueda, planificación...).
- ▶ **Actuar:** por medio de sus sistemas de actuadores, el agente interactuará con el entorno para producir aquellos cambios que le permitan obtener un estado que mejore su valoración del desempeño para alcanzar sus metas. Para ello deberá:
 - Generar, transformando sus acciones de alto nivel en estructuras de datos que puedan ser manejadas en el futuro, empleando árboles, grafos, etc.
 - Convertir las acciones empleadas por el agente como resultado del razonamiento en elementos entendibles por el sistema de actuadores (curvas de animación, trazas de información del sistema, variable de estado...).



Figura 4. Fases en las que un agente inteligente debe organizar su proceso de toma de decisiones.

Arquitectura de un agente Inteligente

Una arquitectura de agentes es una metodología empleada para construir agentes. Las arquitecturas de agente tienen un importante papel en el razonamiento práctico.

En ellas se especifica cómo el agente puede ser descompuesto en un conjunto de módulos componentes y cómo estos módulos pueden interactuar. El conjunto total de módulos y sus interacciones determinan las acciones y estados internos futuros. Wooldridge y Jennings proponen, en 1994, una clasificación de arquitectura de agentes en tres categorías principales:

- ▶ Arquitecturas deliberativas.
- ▶ Arquitecturas reactivas.
- ▶ Arquitecturas híbridas.

Arquitecturas de agentes deliberativos

Contienen una representación explícita y lógica del mundo, con lo cual las decisiones (por ejemplo, acerca de las acciones a realizar) son hechas por medio de un razonamiento lógico. Se basan en un razonamiento práctico, decidir en cada momento la acción a realizar facilita la consecución de objetivos.

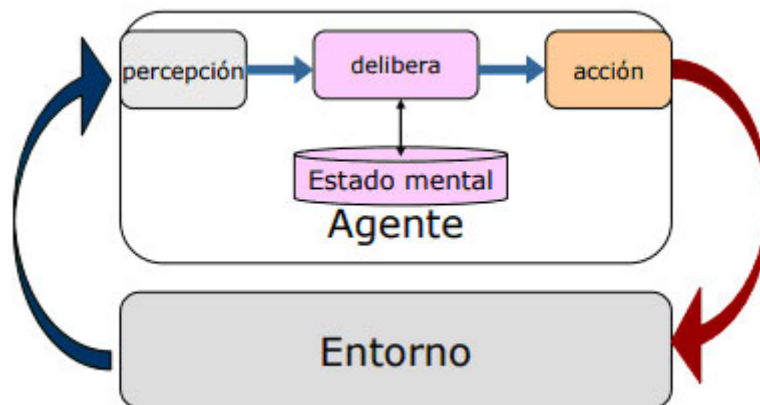


Figura 5. Agente deliberativo. Fuente: Pavón (2006).

- ▶ Ejemplos de arquitecturas deliberativas
 - Agentes basados en planificación (*planning agents*): la comunidad de inteligencia artificial dedicada a la planificación, desde principios de los años

setenta, ha estado estrechamente relacionada con el diseño de agentes. Parece razonable, entonces, que muchas de las innovaciones en el diseño de agentes provengan de esa comunidad. El objetivo es describir a los agentes que pueden encontrar un curso de acciones que, cuando se ejecutan, logran algún objetivo deseado.

- BDI (*Belief, Desire & Intention*) (Georgeff, Pell, Pollack, Tambe y Wooldridge, 1999): la descripción del estado de un agente de estas características se realiza por medio de:
 - Creencias: que representan el conocimiento que el agente tiene de sí mismo y de su entorno. Estas creencias pueden ser incompletas y/o incorrectas.
 - Deseos: objetivos que el agente desea cumplir a largo plazo.
 - Intenciones: subconjunto de las metas que el agente intenta conseguir.

Además de estos elementos básicos, los agentes BDI generan planes por medio de explorar metas intermedias.

- Planes: combinación de intenciones. Las intenciones crean subplanes del plan global del agente, de manera que el conjunto de todos los planes refleja las intenciones del agente.
- Metas: subconjunto de los deseos que el agente puede conseguir. Han de ser realistas y no tener conflictos entre ellas.

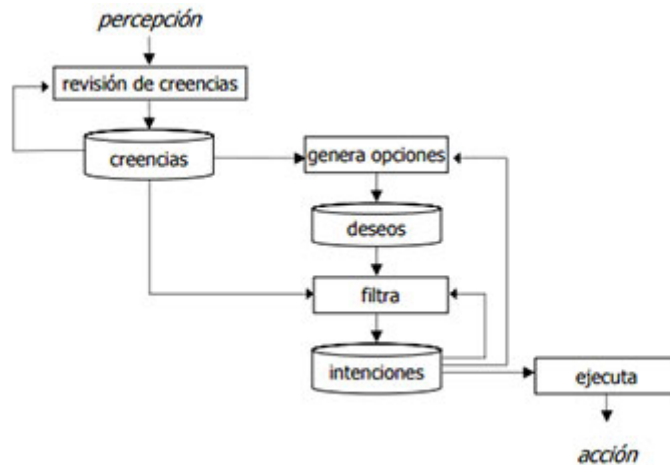


Figura 6. Modelo BDI de agentes. Fuente: Pavón (2006).

Arquitecturas de agentes reactivos

Una arquitectura para agente reactivo es aquella que no incluye ningún tipo de modelo simbólico central del mundo y no utiliza razonamiento simbólico complejo. Pretenden resolver el problema puntual que se encuentran en cada instante por medio de una toma de decisiones rápida.

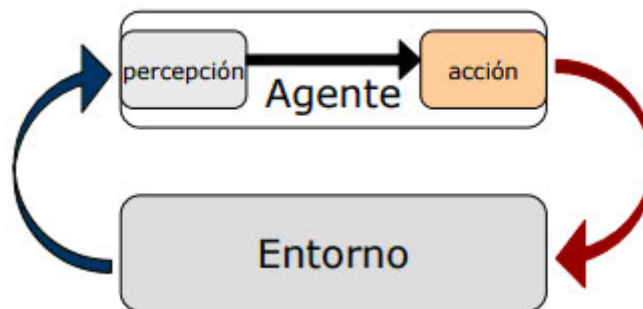


Figura 7. Agente reactivo. Fuente: Pavón (2006).

► Ejemplos:

- *Sumbsumption architecture*: consiste en una jerarquía de comportamientos para la consecución de tareas. Cada comportamiento intenta tomar el control del agente en cada instante. Capas menores representan comportamientos de

tipo más primitivo (tal como evitar obstáculos, por ejemplo) y tienen precedencia sobre las capas superiores de la jerarquía.

- Arquitectura de red de agentes: Pattie Maes (Maes, 1990) ha desarrollado una arquitectura de agentes en la cual un agente se define como un conjunto de módulos de competencia. Estos módulos asemejan ligeramente el comportamiento de la arquitectura subsunción. Cada módulo es especificado por el diseñador en términos de pre- y poscondiciones, y un nivel de activación, que da una indicación real de la relevancia del módulo en una situación particular.

Arquitecturas para agentes híbridos

Un enfoque totalmente deliberativo o totalmente reactivo no es adecuado para diseñar agentes en entornos complejos. Por lo tanto, existen sistemas híbridos que intentan unir estos enfoques. Un enfoque obvio es construir un agente compuesto por dos subsistemas:

- ▶ Deliberativo, que contiene un módulo simbólico del mundo, que desarrolla planes y efectúa decisiones de la manera propuesta por la inteligencia artificial simbólica.
- ▶ Reactivo, que es capaz de reaccionar a eventos que ocurren en el ambiente sin necesitar un razonamiento complejo. A menudo, al componente reactivo se le da cierto grado de precedencia sobre el deliberativo, de tal manera que pueda proveer una pronta respuesta a eventos ambientales importantes.

Este tipo de arquitecturas presenta controladores reactivos en los niveles de toma de decisiones locales o inmediatas, delegando conductas planificadas a las capas superiores con controladores de deliberación. Existen mecanismos para enviar información entre capas, en muchos casos como entradas a capas inferiores que se combinan con las percepciones recibidas en los sensores.

Ejemplos

- ▶ PRS: es una arquitectura basada en creencias, deseos e intenciones (BDI) que incluye una librería de planes, así como una explícita representación simbólica de las creencias, deseos e intenciones.

Las creencias son hechos tanto derivados de la percepción del entorno como del estado interno del agente. Estos hechos son expresados mediante lógica de primer orden. Los deseos son representados como comportamientos del sistema. La librería de planes del PRS contiene un conjunto de planes parcialmente elaborados, llamados áreas de conocimiento (*knowledge areas KAs*), cada uno de los cuales se asocia con una condición para ser invocado.

Estas condiciones determinan cuando el KA debe ser activado. Los KAs pueden ser activados por objetivos o por datos. Pueden ser también reactivos, permitiendo que el PRS responda rápidamente a cambios en su ambiente. Las intenciones son el conjunto de KAs actualmente.

- ▶ COSY: la arquitectura COSY es un BDI (creencias, deseos e intenciones) híbrido que incluye elementos tanto de PRS como de IRMA. Tiene cinco componentes principales:
 - Sensores: reciben entradas perceptibles no comunicativas.
 - Actuadores: permiten al agente realizar acciones no comunicativas.
 - Comunicaciones: permiten al agente enviar mensajes.
 - Cognición: es responsable de mediar entre las intenciones del agente y sus conocimientos acerca del mundo, y de elegir una acción apropiada para realizar.
 - Intención: contiene objetivos a largo plazo, actitudes, responsabilidades y los elementos de control que toman parte en el razonamiento y la toma de decisiones del componente cognición.

Arquitectura de tres capas (*three-layer architecture*)

Es posible agrupar los datos por niveles. Los robots móviles se basan en este tipo de arquitectura.

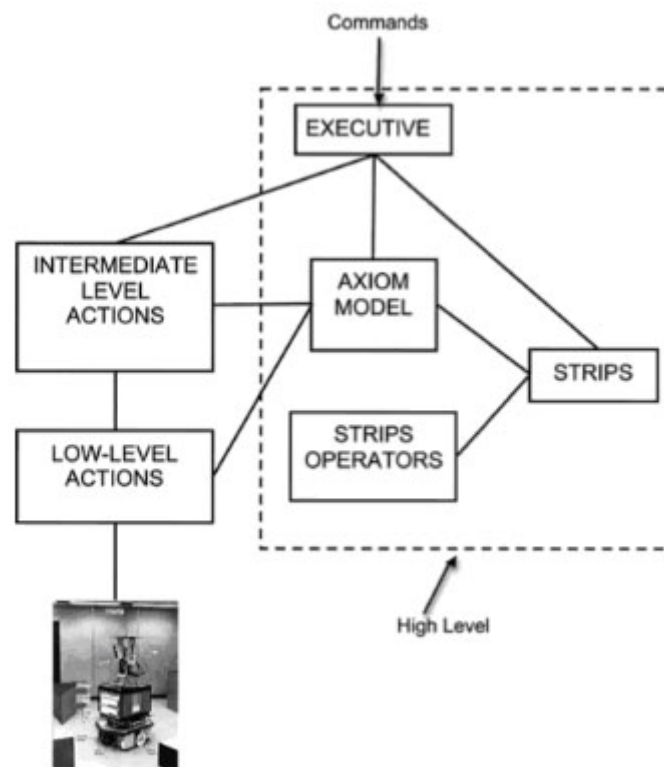


Figura 8. Modelos y arquitecturas de agentes. Fuente: Pavón (2006).

Niveles:

- ▶ De control reactivos.
- ▶ Gobernar secuencias.
- ▶ Consumidores de tiempo.

Arquitecturas multicapa (*multilayer architectures*) (Albus y Barbera, 2005).

Arquitectura de modelo de referencia, múltiples capas de sistemas de control en tiempo real (RSC). Generación de comportamiento (TD), modelado del mundo (WM), procesamiento sensorial (SP), juicio de valor.

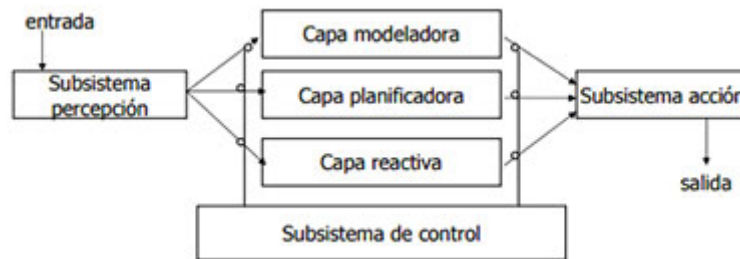


Figura 9. Modelos y arquitecturas de agentes. Fuente: Pavón (2006).

Arquitectura triple torre:

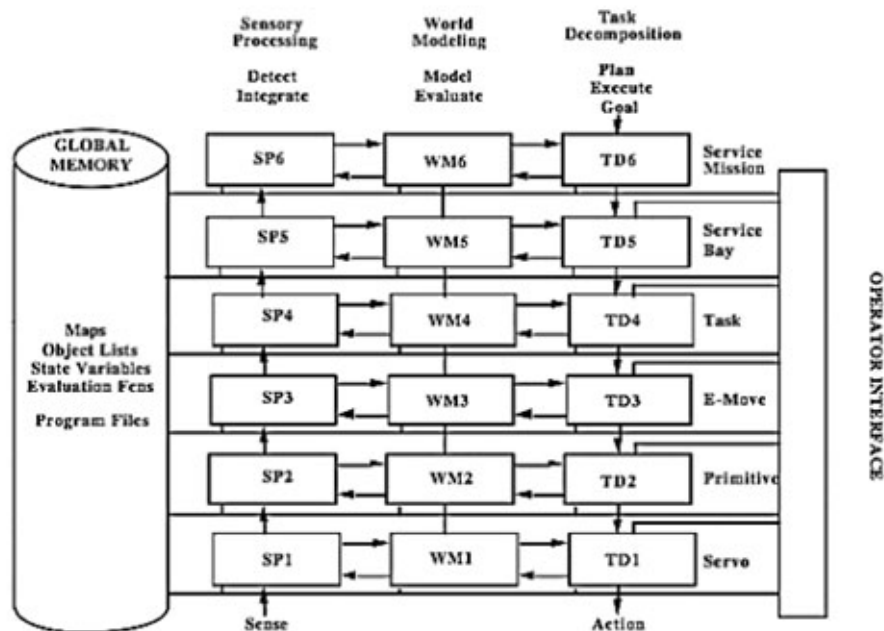


Figura 10. Arquitectura triple torre. Fuente: Pavón (2006).

Arquitecturas cognitivas

Su objetivo es definir las operaciones cognitivas y perceptivas básicas que posee la mente humana. En teoría, cada tarea que los humanos pueden realizar debe consistir en una serie de estas operaciones discretas. Los modelos formalizados pueden usarse para refinar aún más una teoría integral de la cognición y, más inmediatamente, como un modelo comercialmente utilizable.

El Institute of Creative Technologies (s.f.) define la arquitectura cognitiva como:

«hipótesis sobre las estructuras fijas que proporcionan una mente, ya sea en sistemas naturales o artificiales, y cómo funcionan juntas —junto con el conocimiento y las habilidades incorporadas dentro de la arquitectura— para producir un comportamiento inteligente en una diversidad de entornos complejos».

Ejemplos:

- ▶ ACT-R: es una arquitectura cognitiva desarrollada principalmente por John Robert Anderson en la Universidad Carnegie Mellon.

La suposición más importante de ACT-R es que el conocimiento humano se puede dividir en dos tipos de representaciones irreducibles: **declarativo** y de **procedimiento**.

Dentro del código ACT-R, el conocimiento declarativo se representa en forma de fragmentos, es decir, representaciones de vectores de propiedades individuales, cada una de ellas accesible desde una ranura etiquetada.

Los fragmentos se mantienen y se hacen accesibles a través de búferes, que son el *front-end* de lo que son módulos, es decir, estructuras cerebrales especializadas y en gran medida independientes.

Hay dos tipos de módulos:

- Módulo perceptivo-motor: encargado de la interacción con el entorno. Gestiona el flujo de percepción y acción necesario para conectar al agente con el mundo.
- Módulo de memoria: la memoria está dividida en dos tipos. Por un lado, la **memoria a largo plazo** o memoria productiva, que contiene las reglas de producción y, por otro, la **memoria a corto plazo** (memoria de trabajo o memoria declarativa), que contiene los hechos actuales del mundo.

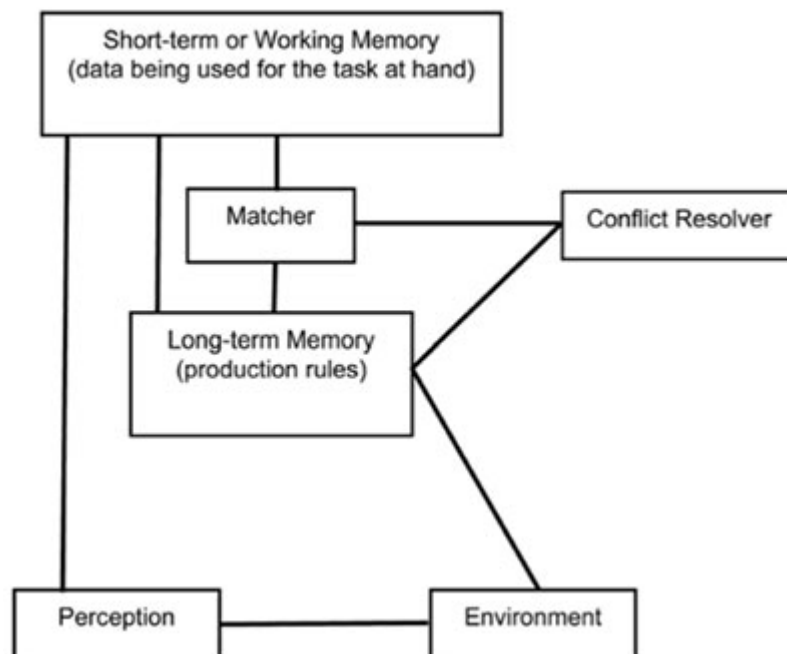


Figura 11. Esquema simplificado de arquitectura cognitiva. Fuente: Institute of Creative Technologies (s.f.)

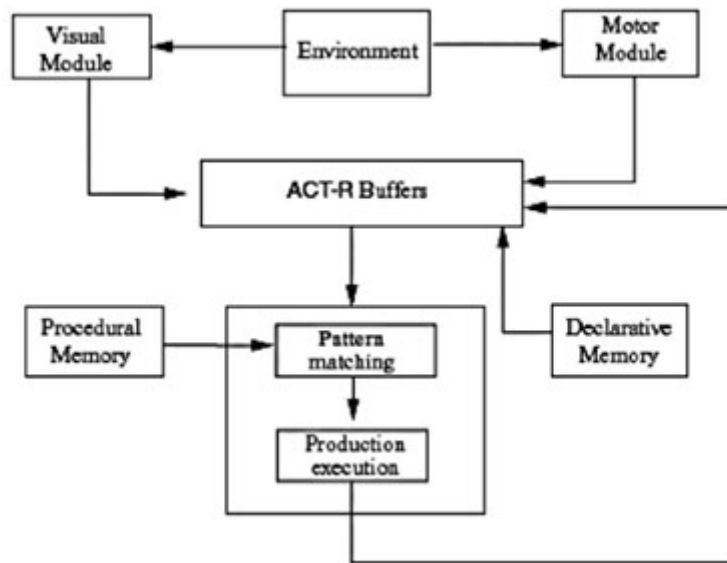


Figura 12. Esquema de un sistema ACT-R. Fuente: <http://act-r.psy.cmu.edu/about/>

- SOAR: también creada en la Universidad de Carnegie Mellon por Laird y Newell (Laird, Newell y Rosenbloom, 1987). El objetivo final de una arquitectura cognitiva es proporcionar una base para un sistema capaz de un comportamiento inteligente general. Es decir, el objetivo es proporcionar la estructura subyacente que permitiría a un sistema realizar toda la gama de tareas cognitivas, emplear toda la gama de métodos de resolución de problemas y representaciones apropiadas para las tareas, y conocer todos los aspectos de la tarea y su rendimiento en ellos.

Para alcanzar este objetivo se desarrollan los bloques de construcción computacionales fijos necesarios para los agentes inteligentes generales: agentes que pueden realizar una amplia gama de tareas y codificar, utilizar y aprender todo tipo de conocimientos para comprender la gama completa de capacidades cognitivas que se encuentran en humanos, como la toma de decisiones, la resolución de problemas, la planificación y la comprensión del lenguaje natural. Es, a la vez, una teoría de lo que es la cognición y una implementación computacional de esa teoría.

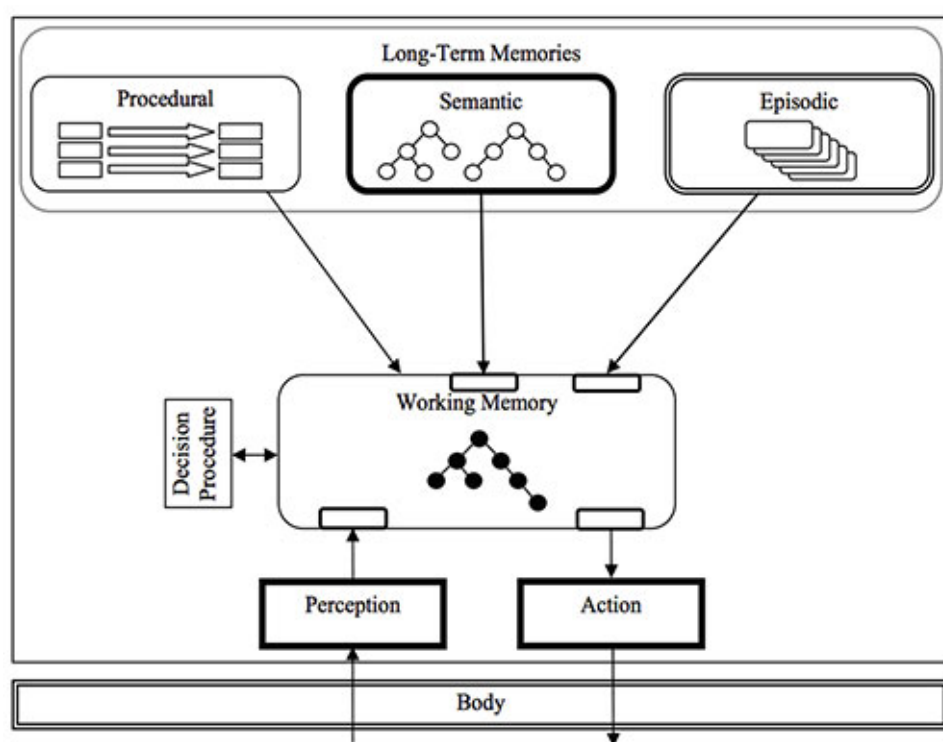


Figura 13. Arquitectura SOAR. Fuente: Laird, Newell y Rosenbloom (1987)

El diseño de SOAR se basa en la hipótesis de que todo comportamiento deliberado orientado a **objetivos** puede ser entendido como la selección y aplicación de **operadores** a un **estado**. Un estado es una representación de la situación actual del problema a resolver, un operador transforma un estado (realiza cambios en la representación) y un objetivo es un resultado deseado para el problema.

SOAR se ejecuta continuamente intentando aplicar el operador actual y seleccionar el siguiente operador (un estado puede tener solo un operador a la vez), hasta que se logre el objetivo.

SOAR tiene memorias separadas (con diferentes modos de representación) para la descripción de su estado actual y de su conocimiento *procedural* a largo plazo. La **memoria a corto plazo** almacena los datos producidos por los sensores, los resultados de inferencias intermedias sobre los datos más actuales, las metas actualmente activas para el agente y los operadores (acciones y planes) que están también activos.

Mientras que la **memoria a largo plazo o memoria de producción** mantiene el conocimiento de responder a situaciones por medio de procedimientos que almacenan el conocimiento específico para la resolución del problema. Esto lo lleva a cabo por medio de reglas de inferencia y conocimiento para la selección y aplicación de operadores frente a estados específicos.

1.4. Tipos de agentes inteligentes

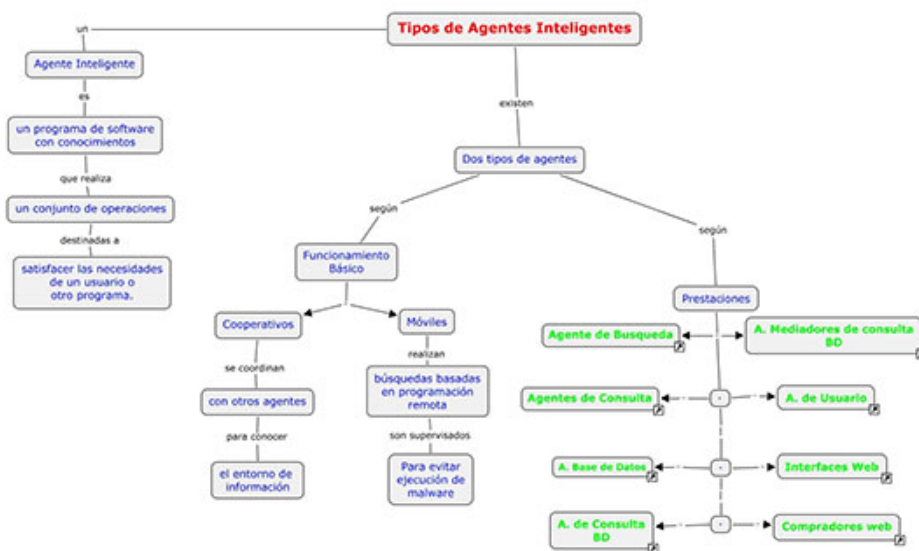


Figura 14. Tipos de agentes inteligentes. Fuente: (Heredia, s.f.).[BMD1]

Puedes consultar el mapa conceptual completo y navegar a través de él para ver en detalle los tipos de agentes inteligentes en:

<http://cmapspublic.ihmc.us/rid=1LN08Q062-151SR17-2805/tipos%20de%20agentes%20inteligentes.cmap>

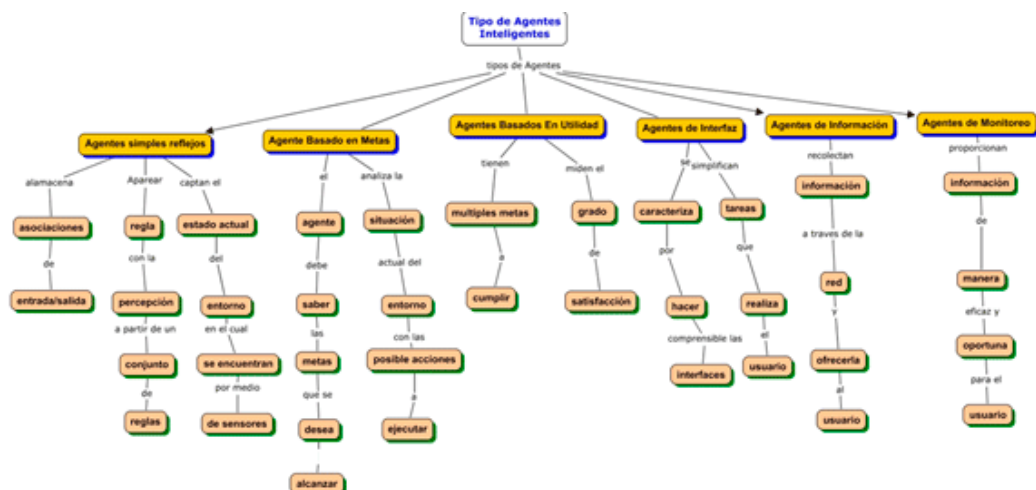


Figura 15. Tipos de agentes inteligentes. Fuente: (Heredia, s.f.).

Agentes basados en búsquedas y planificación

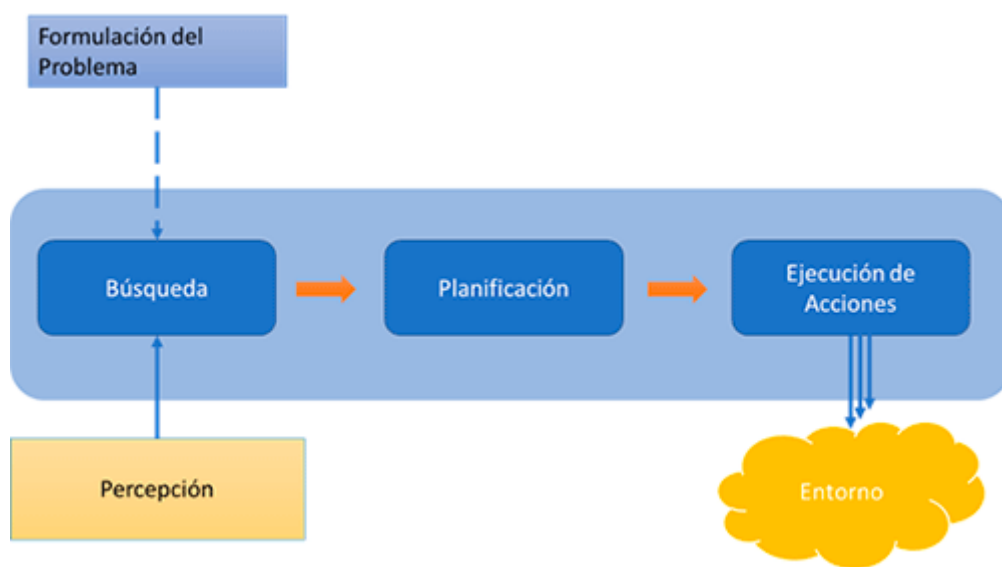


Figura 16. Agentes basados en búsquedas y planificación.

Estos agentes resuelven el problema de alcanzar una meta u objetivo deseado en el entorno en el que se encuentran por medio de la exploración del espacio de estados. Para ello realizan una búsqueda de las posibles alternativas (acciones disponibles) dado el estado en el que están. Con esta decisión simulan o estiman el estado resultante del mundo al operar de esa manera e intentan encontrar qué secuencia de acciones es la que les permite alcanzar la meta que desean.

Los algoritmos de búsqueda emplean la descripción del estado como punto de partida para toda exploración, pero pueden tomar o no en consideración el hecho de que los estados del mundo presentan estructura.

Empezaremos no teniendo en cuenta dicha estructura y presentaremos los algoritmos generales de búsqueda en espacios de estados (amplitud, profundidad, A*, multiagentes...), y luego introduciremos aquellos modelos de razonamiento en los cuales la estructura del estado nos permita alcanzar unos mejores modelos de exploración del espacio de estados, acelerando el proceso de toma de decisiones.

1.5. Referencias bibliográficas

Albus, J. y Barbera, A. (2005). RCS: A cognitive architecture for intelligent multi agent systems. *Annual Reviews in Control*, 29(1), 87-99.

Georgeff, M., Pell, B., Pollack, M., Tambe, M. y Wooldridge, M. (1999). *The Belief-Desire-Intention Model of Agency*. Conferencia presentada en Intelligent Agents V, Agent Theories, Architectures, and Languages, París.

Heredia, C. (s.f.) *Tipos de agentes inteligentes*. Recuperado de <http://cmapspublic.ihmc.us/rid=1LN08Q062-151SR17-2805/tipos%20de%20agentes%20inteligentes.cmap>

Laird, J., Newell, A. y Rosenbloom, P. (1987). *SOAR: an Architecture for General Intelligence*. Michigan: Carnegie-Mellon University. Recuperado de <https://pdfs.semanticscholar.org/c86e/4e9f058f06e2e452ff34c08b2f29b957acb9.pdf>

Maes, P. (1990). *Designing Autonomous Agents*. Cambridge, MA: The MIT Press.

Müller, J. (1996). *The design of intelligent agents: a layered approach. Lecture Notes in Computer Science, Vol. 1177*. Berlín: Springer-Verlag.

Pavón, J. (2006). *Agentes inteligentes. Modelos y arquitecturas de agentes*. Recuperado de <http://www.fdi.ucm.es/profesor/jpavon/doctorado/arquitecturas.pdf>

USC Institute for Creative Technologies. (s. f.) Cognitive Architecture. Recuperado de <http://cogarch.ict.usc.edu/>

Wooldridge, M. y Jennings, N. (1994). *Agent Theories, Architectures and Languages: A Survey*. Conferencia presentada en Intelligent Agents ECAI-94, Ámsterdam. Recuperado de <https://www.csee.umbc.edu/~finin/papers/atal.pdf>

Lo + recomendado

No dejes de leer

Agentes de la información

Hípola, P. y Vargas-Quesada, B. (1999). Agentes inteligentes: definición y tipología: los agentes de la información. *El profesional de la información*, 8(4), 13-21.

Artículo que define, identifica y clasifica los agentes inteligentes. Se centra, sobre todo, en los encargados de la recuperación de información.

Accede al documento a través del aula virtual o desde la siguiente dirección web

<http://eprints.rclis.org/18300/1/epi1999e%5B1%5D.pdf>

Agentes inteligentes

Flores, E. (1997). Agentes inteligentes: el siguiente paso en la inteligencia artificial. *Revista de Información, Tecnología y Sociedad*, 1.

En este artículo puedes profundizar en los agentes inteligentes (qué son y qué características tienen), además de ver algunas de sus aplicaciones.

Accede al documento a través del aula virtual o desde la siguiente dirección web

http://www.revistasbolivianas.org.bo/scielo.php?pid=S1997-40442008000200026&script=sci_arttext

No dejes de ver

It's a Risky Life

David Ríos, titular de la cátedra AXA del ICMAT, presenta los vídeos *It's a risky life*, donde, en clave de humor, aborda conceptos matemáticos esenciales sobre la toma de decisiones, el riesgo y la influencia que tienen en la sociedad actual.



Accede al vídeo a través del aula virtual o desde la siguiente dirección web:

https://www.youtube.com/watch?v=2OAaU5Vtjug&list=PLH_XnVAPg2hw5P0K5GXW5yvH60xxl1ckn

A fondo

Técnicas para definir estrategias de resolución de problemas

Crea Business Idea (2010). *Manual de la creatividad empresarial*. España: Sudoe

Ejemplos y técnicas de resolución de problemas en el ámbito de la toma de decisiones empresariales.

Accede al documento a través del aula virtual o desde la siguiente dirección web

http://www.creabusinessidea.com/creativo_noticias_detalle.php?id=172

Ejemplos adicionales de tipos de agentes

Universidad Politécnica de Madrid (2000). *Informe técnico: Arquitecturas de agente*.

Puedes encontrar ejemplos de arquitecturas de agentes en el informe técnico del Proyecto DAMMAD de 2000.

Accede al documento a través del aula virtual o desde la siguiente dirección web

<http://absta.info/u-niversidad-politcnica-de-madrid-anexo-i-informe-tnico-arqui.html?page=5>

Webgrafía

Instituto de Tecnologías Creativas

Página web del Instituto de Tecnologías Creativas de la University of Southern California.



Accede a la página web a través del aula virtual o desde la siguiente dirección web:

<http://cogarch.ict.usc.edu/>

Bibliografía

Bratman, M., Israel, D. y Pollack, M. (1988). Plans and resource-bounded practical reasoning. *Computational Intelligence*, 4, 349-355.

Brenner, W., Zarnekow, R. y Wittig, H. (1998). *Intelligent Software Agents*. Berlín: Springer.

Brooks, R. (1991). Intelligence without representation. *Artificial Intelligence*, 47(1-3), 139-159.

Carver, N. y Lesser, V. R. (1992). *The Evolution of Blackboard Control Architectures* (informe). Estados Unidos: Computer Science Department, University Massachusetts.

Chapman, D. y Agre, P. (1986). Abstract reasoning as emergent from concrete activity. En Georgeff, M. y Lansky, A. (Eds.), *Reasoning About Actions and Plans* (pp. 411-424). San Mateo, CA: Morgan Kaufmann Publishers.

Ferber, J. (1999). *Multi-agent systems: an introduction to distributed artificial intelligence*. Harlow: Addison-Wesley.

Genesereth, M. y Ketchpel, (1994). Software Agents. *Communications of the Association for Computing Machinery*, 48-53.

Shoham, Y. (1993). Agent-oriented programming. *Artificial Intelligence*, 60, 51-92.

Weiss, G. (1999). *Multiagent Systems*. Massachusetts: The MIT Press.

Wooldridge, M. y Jennings, N. R. (1995). Modelos y arquitecturas de agente. *Intelligent Agents: Theory and Practice. The Knowledge Engineering Review*, 10(2), 115-152.

1. ¿Todas las técnicas de inteligencia artificial son aplicables a todos los problemas?
 - A. Sí, existen algoritmos que resuelven todos los problemas.
 - B. No, debemos buscar la técnica que mejor se adapta a un problema de entre las distintas que son aplicables.
 - C. Sí, los algoritmos genéticos resuelven todos los problemas.
 - D. No, cada problema es solo resuelto por un único algoritmo.

2. ¿Se dice que en un problema bien definido un estado es discreto?
 - A. Sí, se puede concebir el mundo en estados.
 - B. En cada estado hay un conjunto finito de percepciones y acciones.
 - C. Cada estado presenta una cardinalidad inferior a diez elementos.
 - D. A y B son correctas.

3. La percepción de un agente...
 - A. Proviene de los sensores.
 - B. Es creada exclusivamente por el agente.
 - C. Contiene información no alcanzable por los sensores.
 - D. Se descompone en acciones.

4. ¿En la fase de razonamiento de un agente?
 - A. Se deciden las acciones a realizar.
 - B. Se crean juicios e inferencias derivadas de las percepciones actuales y de los hechos recordados.
 - C. Se filtran los valores percibidos por los sensores.
 - D. La A y la C son correctas.

5. ¿Un agente reactivo?

- A. No incluye ningún tipo de modelo simbólico central del mundo.
- B. Tiene procesos cognitivos con semejanzas biológicas.
- C. No emplea la percepción dentro del modelo de razonamiento.
- D. Funciona de modo asistido por el usuario.

6. La arquitectura SOAR:

- A. Ejecuta continuamente intentando aplicar el operador actual y seleccionar el siguiente operador.
- B. Es una arquitectura reactiva.
- C. Emplea un sistema de creencias para representar sus intenciones.
- D. La B y C son correctas.

7. Las arquitecturas multicapa:

- A. Tienen dos capas.
- B. Tienen un sistema de percepción reactivo que toma los componentes de actuación.
- C. Suele tener una capa reactiva para gestionar los eventos de corto plazo.
- D. Son cognitivas.

8. Los algoritmos de búsqueda...

- A. Emplean la descripción del estado como punto de partida para toda exploración.
- B. Crean sistemas de restricciones.
- C. Son reactivos.
- D. La B y C son correctas.

9. ¿Qué arquitecturas se basan en modelos de la mente humana?

- A. Arquitecturas reactivas.
- B. Arquitecturas cognitivas.
- C. Arquitecturas deliberativas.
- D. Ninguna de las anteriores.

10. ¿La memoria interna del agente?

- A. Almacena solo las percepciones de un instante determinado.
- B. Guarda una secuencia episódica.
- C. Almacena las acciones.
- D. Ninguna de las anteriores.