



SEP
SECRETARÍA
DE EDUCACIÓN
PÚBLICA



Instituto Politécnico Nacional
"La Técnica al Servicio de la Patria"



Inteligencia Artificial

UNIDAD I

INTRODUCCIÓN A LA INTELIGENCIA ARTIFICIAL



Unidad de competencia

- ❖ Diferencia el tipo de problemas de la Inteligencia Artificial a partir de sus aplicaciones y los agentes inteligentes.

Contenido

1.1 Inteligencia

1.1.1 Tipos de inteligencia .

1.2 Historia de la Inteligencia Artificial .

1.3 Problemas, áreas y aplicaciones de la Inteligencia Artificial .

1.4 Agentes inteligentes.

1.4.1 Estructura .

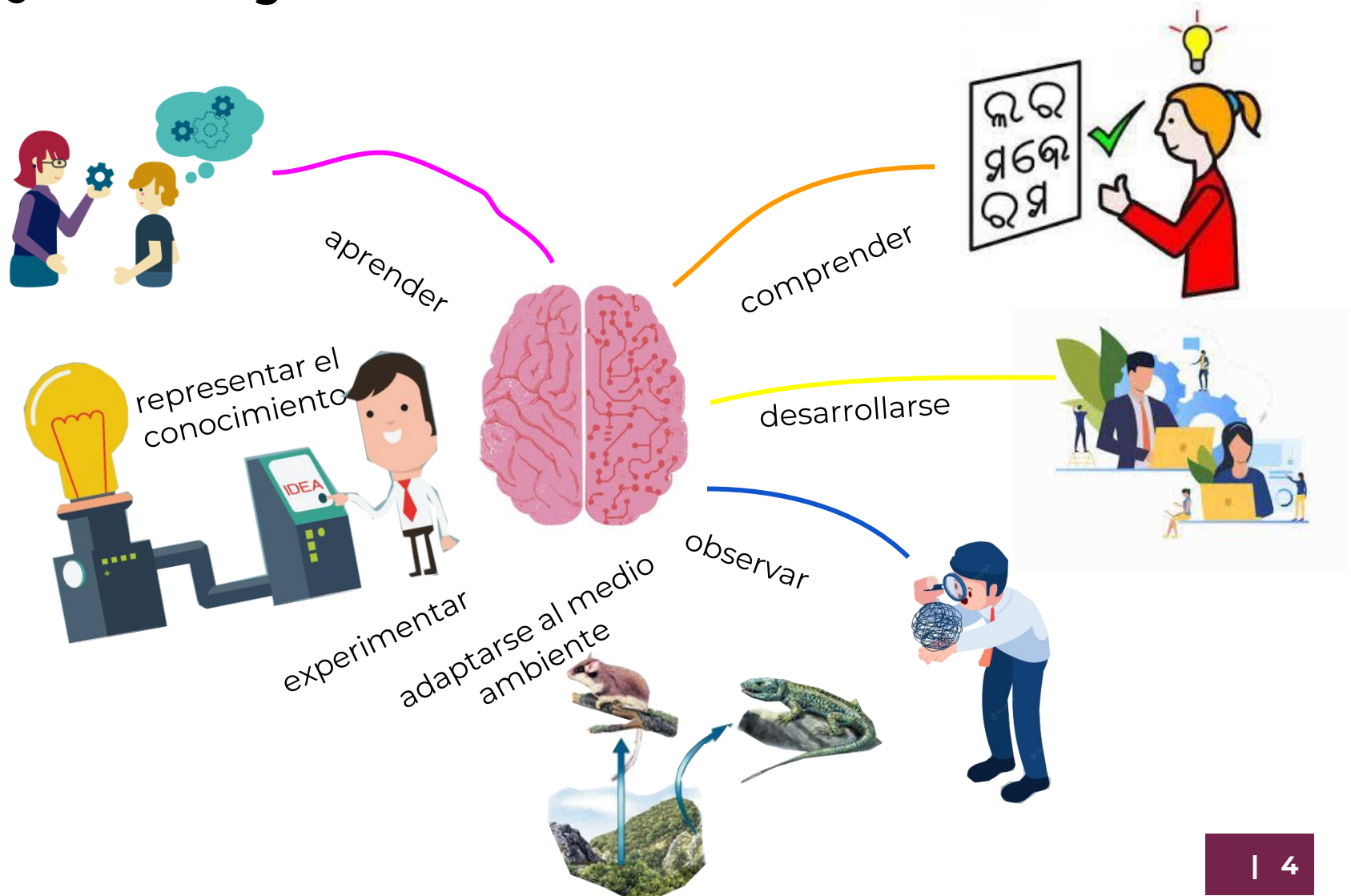
1.4.2 Medio ambiente de trabajo .

1.4.3 Clasificación .

1.4.4 Aplicaciones.



¿Qué es inteligencia?



Tipos de Inteligencia



Tipos de inteligencia

Inteligencia lingüística

Es considerada una de las más importantes. En general se utilizan ambos hemisferios del cerebro y es la que caracteriza a los escritores. El uso amplio del lenguaje ha sido parte esencial para el desarrollo de este tipo de inteligencia.

- Aspectos biológicos: un área específica del cerebro llamada "área de Broca" es la responsable de la producción de oraciones gramaticales. Una persona con esa área lesionada puede comprender palabras y frases sin problemas, pero tiene dificultades para construir frases más sencillas. Al mismo tiempo, otros procesos mentales pueden quedar completamente ilesos.
- Capacidades implicadas: capacidad para comprender el orden y el significado de las palabras en la lectura, la escritura y, también, al hablar y escuchar.
- Habilidades relacionadas: hablar y escribir eficazmente.
- Perfiles profesionales: líderes políticos o religiosos, poetas, vendedores, escritores, etc.

Tipos de inteligencia

Inteligencia musical

También conocida como “buen oído”, es el talento que tienen los grandes músicos, cantantes y bailarines. La fuerza de esta inteligencia radica desde el mismo nacimiento y varía de igual manera de una persona a otra. Un punto importante en este tipo de inteligencia es que por fuerte que sea, necesita ser estimulada para desarrollar todo su potencial, ya sea para tocar un instrumento o para escuchar una melodía con sensibilidad.

- Aspectos biológicos: ciertas áreas del cerebro desempeñan papeles importantes en la percepción y la producción musical. Éstas, situadas por lo general en el hemisferio derecho, no están localizadas con claridad como sucede con el lenguaje. Sin embargo, pese a la falta de susceptibilidad concreta respecto a la habilidad musical en caso de lesiones cerebrales, existe evidencia de “amusia” (pérdida de habilidad musical).
- Capacidades implicadas: capacidad para escuchar, cantar, tocar instrumentos.
- Habilidades relacionadas: crear y analizar música.
- Perfiles profesionales: músicos, compositores, críticos musicales, etc.

Tipos de inteligencia

Inteligencia lógico-matemática

Quienes pertenecen a este grupo, hacen uso del hemisferio lógico del cerebro y pueden dedicarse a las ciencias exactas. De los diversos tipos de inteligencia, éste es el más cercano al concepto tradicional de inteligencia. En las culturas antiguas se utilizaba éste tipo de inteligencia para formular calendarios, medir el tiempo y estimar con exactitud cantidades y distancias.

- Capacidades implicadas: capacidad para identificar modelos, calcular, formular y verificar hipótesis, utilizar el método científico y los razonamientos inductivo y deductivo.
- Habilidades relacionadas: capacidad para identificar modelos, calcular, formular y verificar hipótesis, utilizar el método científico y los razonamientos inductivo y deductivo.
- Perfiles profesionales: economistas, ingenieros, científicos, etc.

Tipos de inteligencia

Inteligencia espacial

Esta inteligencia la tienen las personas que puede hacer un modelo mental en tres dimensiones del mundo o en su defecto extraer un fragmento de él. Esta inteligencia la tienen profesiones tan diversas como la ingeniería, la cirugía, la escultura, la marina, la arquitectura, el diseño y la decoración. Por ejemplo, algunos científicos utilizaron bocetos y modelos para poder visualizar y decodificar la espiral de una molécula de ADN.

- Aspectos biológicos: el hemisferio derecho (en las personas diestras) demuestra ser la sede más importante del cálculo espacial. Las lesiones en la región posterior derecha provocan daños en la habilidad para orientarse en un lugar, para reconocer caras o escenas o para apreciar pequeños detalles.
- Capacidades implicadas: capacidad para presentar ideas visualmente, crear imágenes mentales, percibir detalles visuales, dibujar y confeccionar bocetos.
- Habilidades relacionadas: realizar creaciones visuales y visualizar con precisión.
- Perfiles profesionales: artistas, fotógrafos, arquitectos, diseñadores, publicistas, etc

Tipos de inteligencia

Inteligencia corporal – kinestésica

Los kinestésicos tienen la capacidad de utilizar su cuerpo para resolver problemas o realizar actividades. Dentro de este tipo de inteligencia están los deportistas, cirujanos y bailarines. Una aptitud natural de este tipo de inteligencia se manifiesta a menudo desde niño.

- Aspectos biológicos: el control del movimiento corporal se localiza en la corteza motora y cada hemisferio domina o controla los movimientos corporales correspondientes al lado opuesto. En los diestros, el dominio de este movimiento se suele situar en el hemisferio izquierdo. La habilidad para realizar movimientos voluntarios puede resultar dañada, incluso en individuos que puedan ejecutar los mismos movimientos de forma refleja o involuntaria. La existencia de apraxia específica constituye una línea de evidencia a favor de una inteligencia cinética corporal.
- Capacidades implicadas: capacidad para realizar actividades que requieren fuerza, rapidez, flexibilidad, coordinación óculo-manual y equilibrio.
- Habilidades relacionadas: utilizar las manos para crear o hacer reparaciones, expresarse a través del cuerpo.
- Perfiles profesionales: escultores, cirujanos, actores, modelos, bailarines, etc.

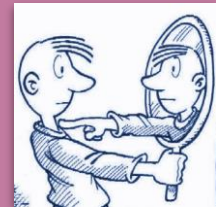
Tipos de Inteligencia artificial

**Sistemas
reactivos**

**Memoria
limitada**

**Teoría de
la mente**

Autoconciencia

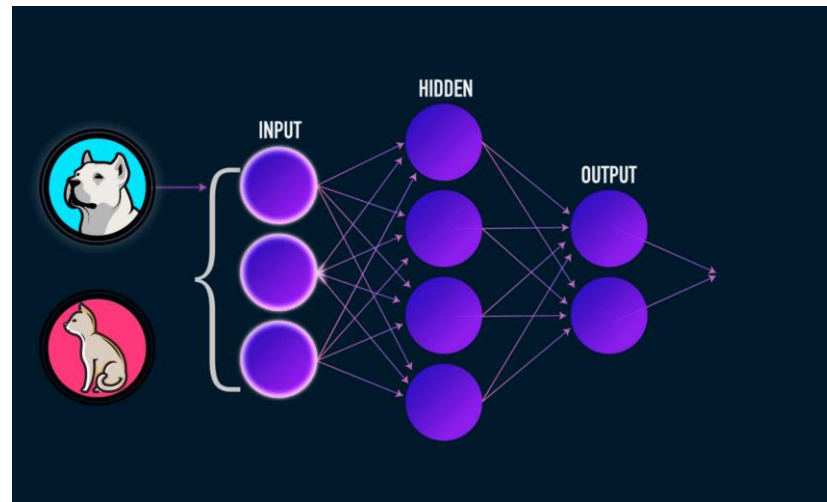


Tipos de inteligencia artificial

Sistemas reactivos.

Las máquinas reactivas constituyen el tipo más básico de IA puesto a que se basan en la toma de decisiones sobre el presente. Al carecer de memoria es imposible que aprendan de experiencias y por tanto no existe una evolución de ninguna índole.

Como ejemplo se tiene a las Redes Neuronales Artificiales calificadoras las cuales reaccionan a un vector de entrada para determinar o clasificar en una clase de salida.



Tipos de inteligencia artificial

Memoria limitada

El segundo tipo de Inteligencia Artificial se refiere a máquinas que sí son capaces de mirar al pasado cercano. Los vehículos autónomos son un ejemplo de ello, puesto que pueden observar la velocidad y dirección de otros automóviles y monitorearlos a través del tiempo. En este sentido el sistema es capaz de tomar decisiones como cambiar de carril, aplicar frenos y acelerar basado en la información recopilada de los últimos segundos.

Pese a lo anterior, dicha experiencia sobre el pasado es transitoria, ya que no se guarda dentro de la biblioteca de experiencias del automóvil y por ende no puede compilar la información durante años o aprender de situaciones nuevas como lo haría un ser humano.



Tipos de inteligencia artificial

Teoría de la mente

Las máquinas de esta categoría son más avanzadas, ya que no solo forman representaciones sobre el mundo, también sobre otros agentes, a esto se le denomina "teoría de la mente" e implica la comprensión de que toda entidad es capaz de tener emociones que afectan su patrón conductual.

Lo anterior resulta importante, ya que, si las máquinas van a vivir entre nosotros deberán poseer una comprensión sobre cómo pensamos y sentimos los seres vivientes y en consecuencia, deberán ajustar su comportamiento.

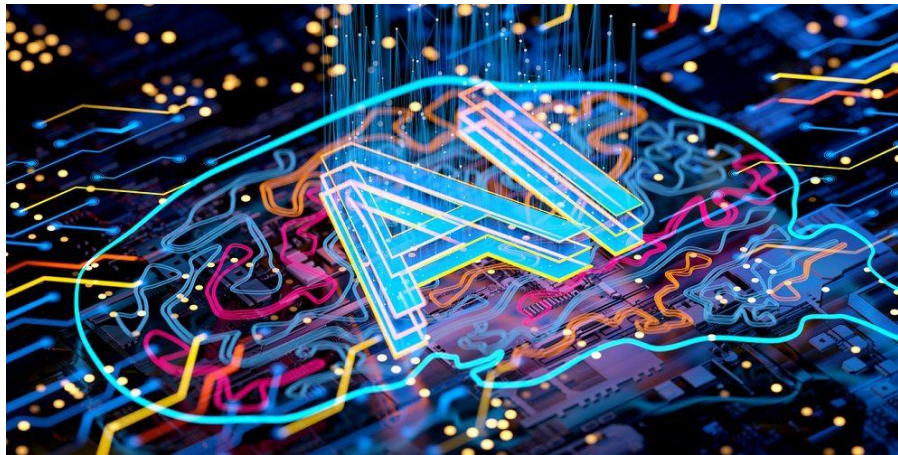
Una aproximación actual a este tipo de inteligencia artificial puede ser encontrada dentro de los modelos de lenguaje interactivos como lo son ChatGPT y LaMDA, los cuales han sido capaces de pasar la prueba de Turing en conversaciones de corta y mediana duración y además cuentan con vastos bancos de información que les permiten "ser conscientes" de otras inteligencias artificiales y además son capaces de modular su lenguaje para adecuarse al de su contraparte humana.

Tipos de inteligencia artificial

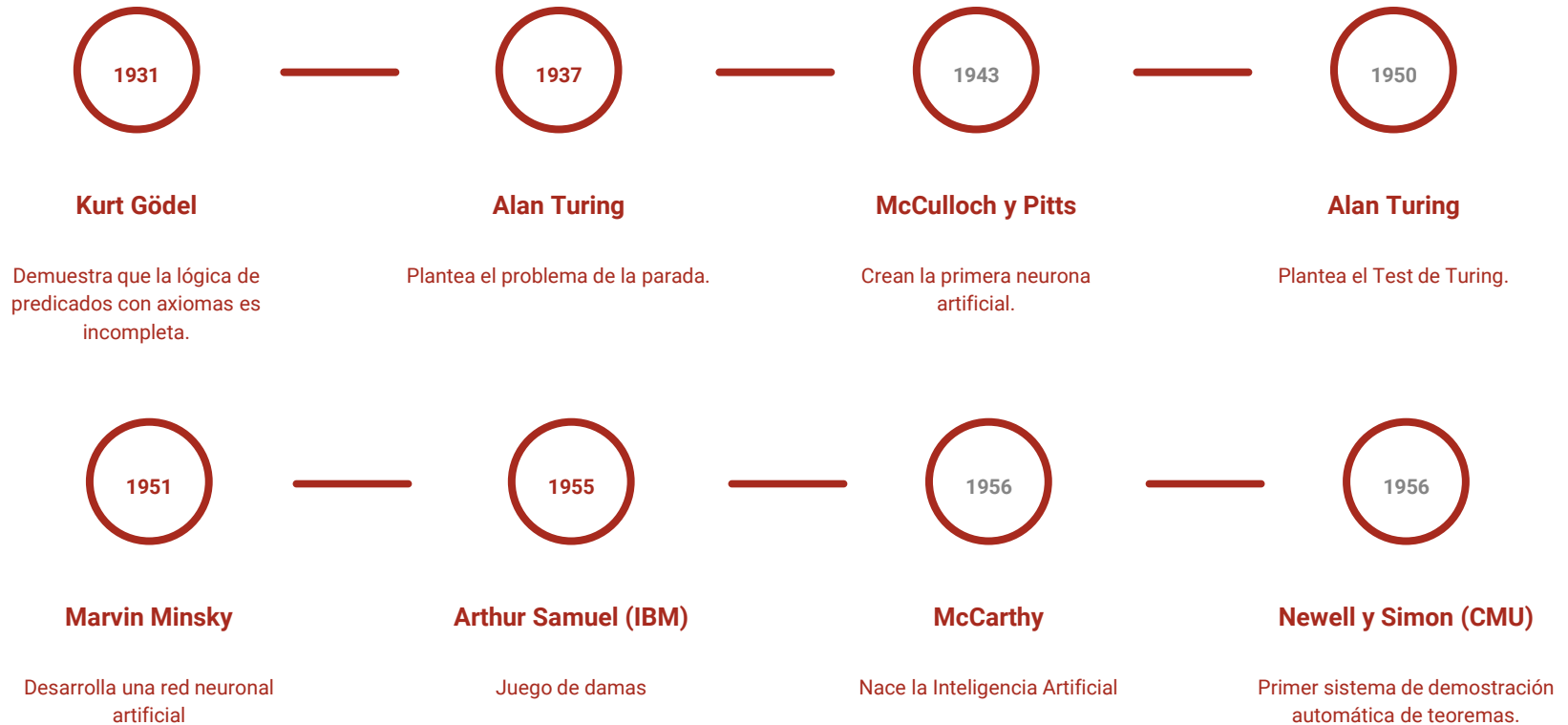
Autoconciencia

Como paso final del desarrollo de la Inteligencia Artificial se espera construir sistemas que puedan formar representaciones no solamente sobre el mundo sino sobre sí mismos, es decir, que tengan conciencia y sean capaces de conocer sus estados internos y comprender los sentimientos de los demás.

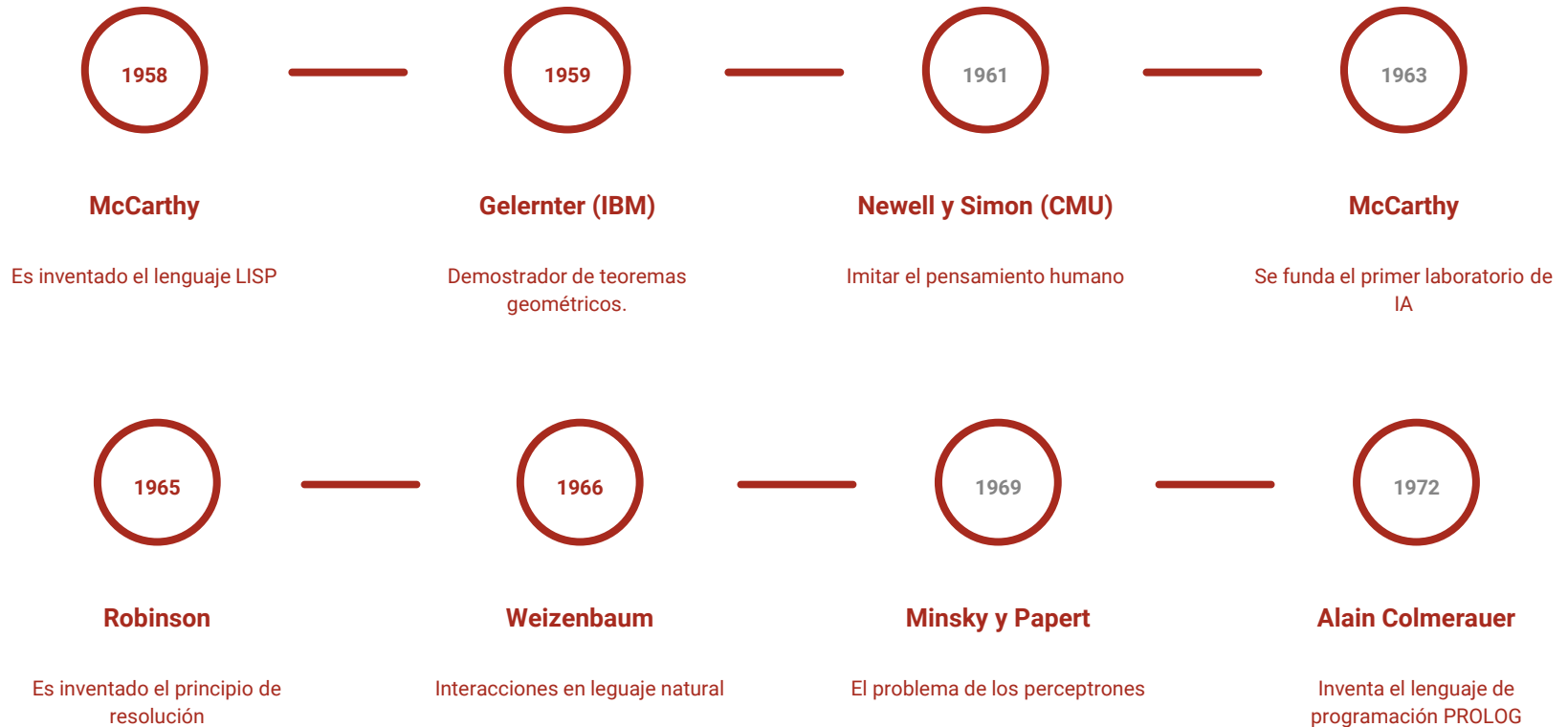
Los cuatro tipos de inteligencia artificial arrojan luz sobre las intenciones que el hombre tiene acerca del futuro de las máquinas. Pese a que aún estamos lejos de la Inteligencia Artificial autoconsciente está claro que eso es lo que se persigue en la última instancia, los esfuerzos se enfocan en el desarrollo de la memoria, el aprendizaje y la capacidad de tomar decisiones basada en experiencias previas.



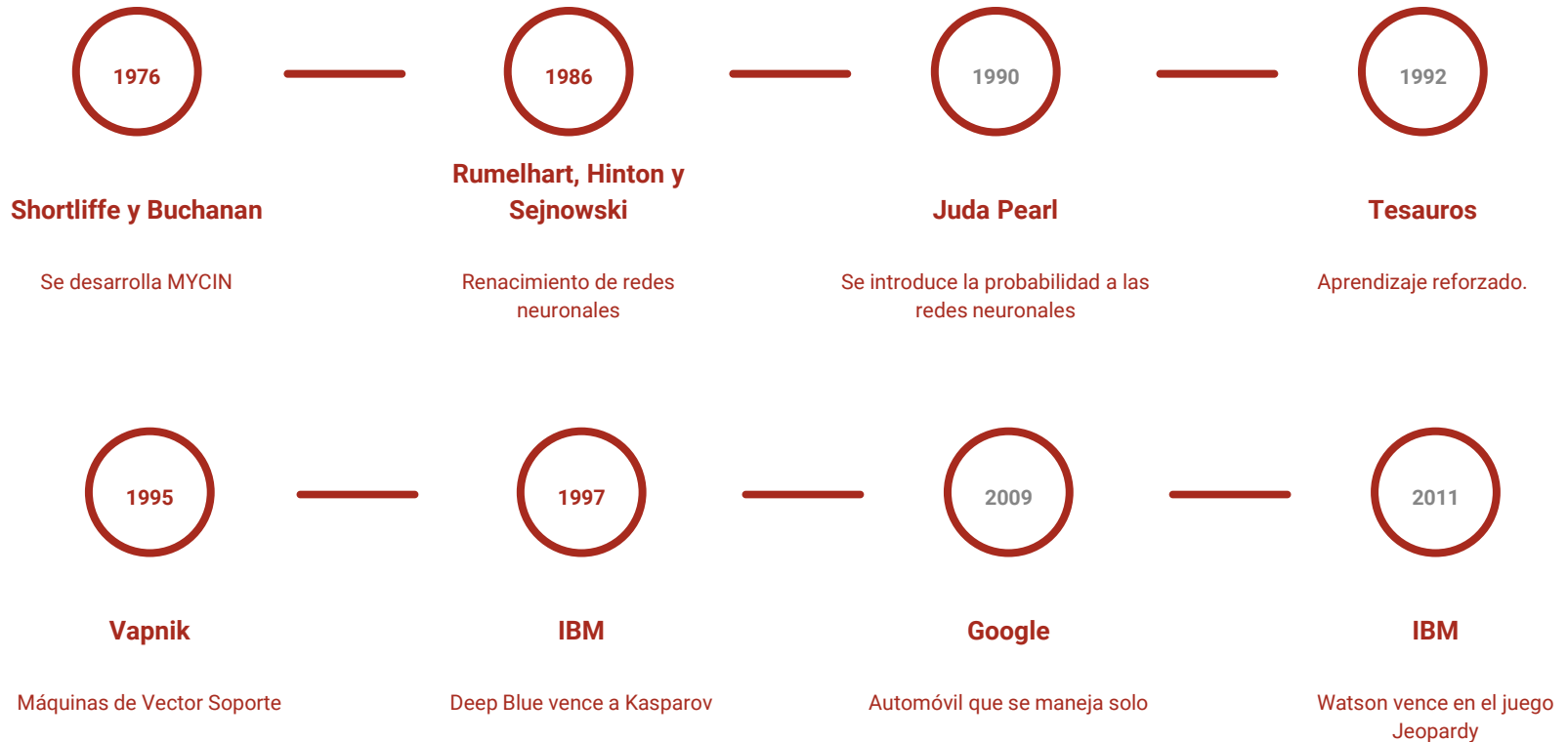
Historia de la Inteligencia Artificial



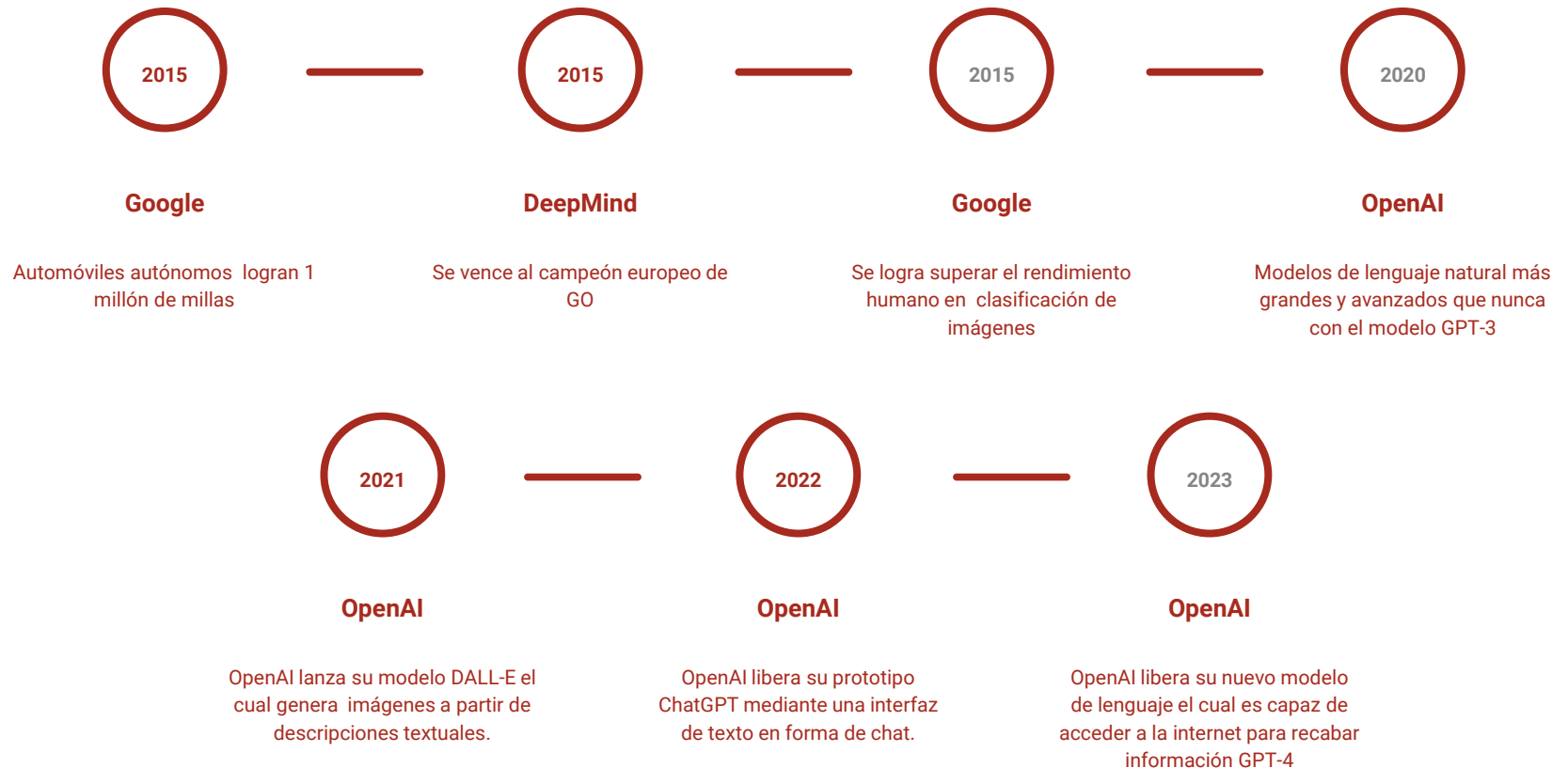
Historia de la Inteligencia Artificial



Historia de la Inteligencia Artificial



Historia de la Inteligencia Artificial



1.3 Problemas, áreas y aplicaciones de la inteligencia artificial

Problemas tecnológicos

Aún existen problemas dentro del desarrollo y la investigación de la Inteligencia Artificial entre estos problemas de carácter tecnológico se encuentran:

- Creciente complejidad de los modelos: Cuanto más se avanza en el desarrollo de modelos de inteligencia artificial más se hace evidente la necesidad de sistemas con la suficiente capacidad computacional para ser capaces de entrenar y ejecutar a los agentes inteligentes.
- El volumen de datos utilizados: Es bien conocida la necesidad de utilización de datos por parte de la Inteligencia Artificial, para poder aprender sobre su entorno, pero regularmente no se toma en cuenta la cantidad de datos que están involucrados. La Inteligencia Artificial no solamente necesita de más datos que el cerebro humano para poder comprender conceptos y características, necesita miles de veces más que los datos requeridos por un humano.

1.3 Problemas, áreas y aplicaciones de la inteligencia artificial

Problemas tecnológicos

- Inteligencia Artificial multitarea: La inteligencia humana es integral, puede desempeñar muchas tareas, y ese es precisamente un fallo de la Inteligencia Artificial desarrollada hasta ahora, ya que es aplicada a un área del conocimiento o actividad específica. Una vez que un cerebro de Inteligencia Artificial ha sido entrenado, puede ser altamente efectivo para tareas como el reconocimiento de animales o para ser el rival ideal en un videojuego, pero de acuerdo con Raia Hadsell, científico de la división de Google DeepMind, no existe una red neuronal artificial en el mundo capaz de hacer más de una tarea a la vez.
- La comprensión de sus conclusiones: Aún queda pendiente poder determinar con exactitud de qué manera la Inteligencia Artificial llega a algunas conclusiones, ya que por ahora sólo es posible determinar el funcionamiento de las redes neuronales en conjunto, pero no se conoce las razones a las que un cerebro artificial llega a determinada conclusión o resultado.

1.3 Problemas, áreas y aplicaciones de la inteligencia artificial

Problemas legales

Los avances dentro del campo de la Inteligencia Artificial ocurren muy rápido especialmente en la época actual, donde desde el año 2010 se han presentado enormes saltos en diferentes campos de la Inteligencia Artificial como es la visión por computadora o el procesamiento del lenguaje natural y pese a que este ritmo tan acelerado favorece el desarrollo tecnológico el ser humano se encuentra en un momento en el que no se cuenta con leyes o regulaciones bien definidas orientadas a los agentes inteligentes y a las posibles consecuencias del uso de los mismos.

1.3 Problemas, áreas y aplicaciones de la inteligencia artificial

Problemas sociales

Han existido muchos libros científicos así como de ciencia ficción en los que se toca este tema, dado a los grandes avances en la investigación de la Inteligencia Artificial nos hemos acercado cada vez más a la era de robots autónomos e internet de las cosas desde el 2005. En enero del 2016 el foro de economía mundial publicó un estudio donde predecían que la industria 4.0 destruiría más de 5 millones de trabajos en los siguientes 5 años, dado a la automatización en fábricas, oficinas, administración, transporte así como muchas otras áreas, desplazando así a las personas encargadas de esos trabajos.

Y pese a que se plantea que este mismo desarrollo generará muchas oportunidades de trabajo dentro del campo de investigación, estos serán principalmente roles altamente especializados, por lo que la mayoría de las personas desplazadas por la Inteligencia Artificial no tendrían la oportunidad de acceder a estas nuevas oportunidades de trabajo.

1.3 Problemas, áreas y aplicaciones de la Inteligencia Artificial

Problemas sociales

Stephen Hawking compartió en una entrevista realizada en www.reddit.com lo siguiente acerca del desempleo causado por la inteligencia artificial y la automatización:

Si las máquinas producen todo lo que necesitamos, el resultado dependerá de cómo las cosas son distribuidas. **Todos pueden disfrutar de una vida de ocio y lujos si la riqueza generada por las máquinas es distribuida, o la mayoría de la gente terminará en la pobreza y miseria si los dueños de éstas máquinas autónomas consiguen evitar la redistribución de la riqueza.** Hasta ahora la segunda opción parece la más probable, con la tecnología creando más desigualdad.

Áreas de la inteligencia artificial



1.3 Problemas, áreas y aplicaciones de la inteligencia artificial

Procesamiento del lenguaje natural

El procesamiento del lenguaje natural implica que la inteligencia artificial sea capaz de entender la estructura y las reglas de uno o varios lenguajes, así como entender la relación que existe entre las palabras y el sentido general que se tiene dentro de una oración, permitiendo comunicar las máquinas con las personas mediante el uso de lenguas naturales, como el español, el inglés o el chino.

Entre las aplicaciones se encuentran:

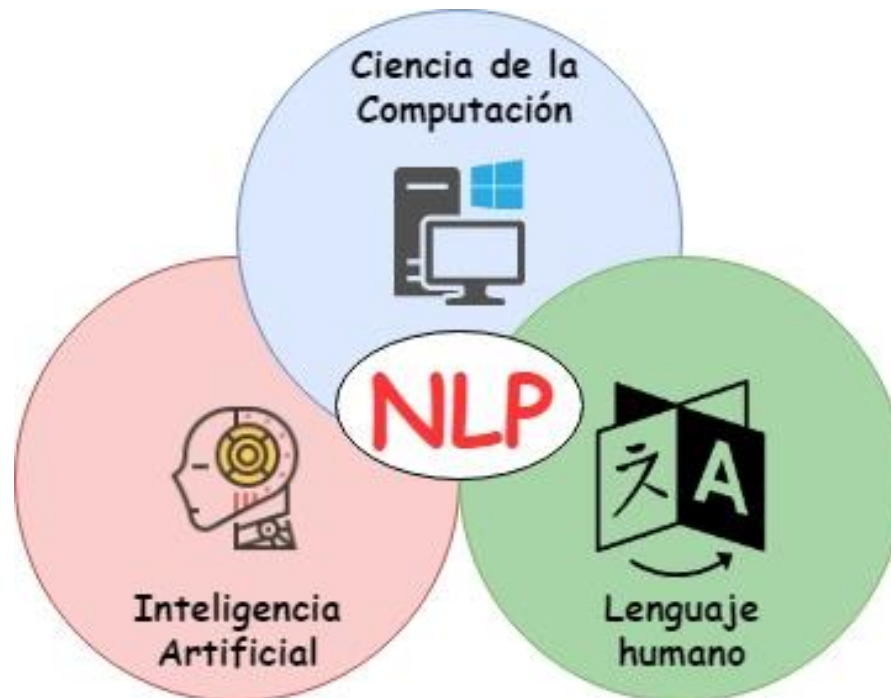
- Revisión ortográfica y gramática
- Extracción de información.
- Clasificación de información.
- Resumen de información.
- Traducciones.



1.3 Problemas, áreas y aplicaciones de la Inteligencia Artificial

Procesamiento del habla

El procesamiento del habla consiste en identificar y entender el habla humana, se busca extraer la información que se encuentra dentro de las ondas de sonido y convertir dicha información a texto y viceversa, entre las aplicaciones de esta área se encuentran: el reconocimiento mediante voz, sistemas de dictado, sistemas de generación de texto a voz, sistemas manejados mediante comandos de voz y sistemas de traducción.



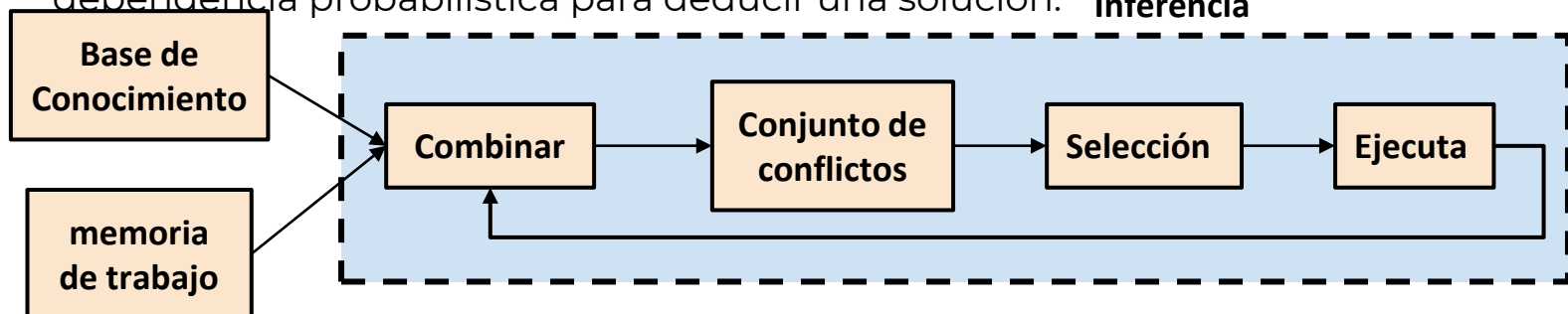
1.3 Problemas, áreas y aplicaciones de la Inteligencia Artificial

Ingeniería y Sistemas Expertos

Los Sistemas Expertos son programas informáticos que tienen el objetivo de solucionar un problema concreto y utilizan la Inteligencia Artificial para simular el razonamiento de un ser humano. Se denominan sistemas expertos porque estos programas imitan la toma de decisiones de un profesional en la materia.

Tipos de Sistemas Expertos:

- **RBO (Rule Based Reasoning):** Están basados en reglas previamente establecidas y abordan las situaciones más complejas a través de reglas deterministas.
- **CBR (Case Based Reasoning):** Basados en casos. Es decir, solucionan problemas utilizando soluciones preexistentes y haciendo una analogía de problemas anteriores.
- **Basados en Redes de Bayes:** Utilizan un conjunto de variables conocidas y su dependencia probabilística para deducir una solución. **Inferencia**



1.3 Problemas, áreas y aplicaciones de la inteligencia artificial

Planeación

Dentro del área de planeación se pretende delegar a sistemas autónomos la toma de decisiones así como el desarrollo de una correcta sucesión de acciones para completar una tarea específica. Para esto se requiere que los sistemas tengan una buena representación del problema.

Modelos del cerebro y evolución

Los modelos del cerebro humano y la evolución corresponden a dos grandes enfoques a la Inteligencia Artificial. El primero, Inteligencia Artificial como modelo del cerebro corresponde a la inteligencia artificial simbólica el cual se enfoca en enseñar a los sistemas autónomos de la misma forma en la que los seres humanos aprenden. Dentro de esta área se manejan abstracciones matemáticas de alto nivel y busca una vista macroscópica de la inteligencia artificial, imitando la psicología humana.

1.3 Problemas, áreas y aplicaciones de la inteligencia artificial

Planeación

Dentro del área de planeación se pretende delegar a sistemas autónomos la toma de decisiones, así como el desarrollo de una correcta sucesión de acciones para completar una tarea específica. Para esto se requiere que los sistemas tengan una buena representación del problema.

Modelos del cerebro y evolución

El segundo enfoque de la inteligencia artificial es el de la evolución, éste enfoque se basa en modelos biológicos y genéticos de seres vivos. La computación basada en neuronas (redes neuronales artificiales) y los algoritmos genéticos son derivaciones de estos conceptos de vida. Estos modelos biológicos de inteligencia artificial no necesariamente son parecidos a sus contrapartes vivas, pero están basados en algoritmos genéticos los cuales evolucionan hasta llegar a una solución como lo hacen los humanos y otras formas de vida.

Aplicaciones de la Inteligencia Artificial

**Planificación
autónoma**

Juegos

**Control
autónomo**

Diagnosis

¿Qué es un Agente Inteligente?

Aunque el término agentes inteligentes no es nuevo para la IA, en los últimos años ha ganado mayor protagonismo. Un **Agente** en rasgos generales es un sistema que procesa información y produce una salida a partir de una entrada.

- En ciencias de la computación un agente es cualquier software o programa que calcula un resultado a partir de una entrada.
- En robótica existen los agentes de hardware (también llamados robots autónomos) los cuales cuentan con sensores con los que pueden percibir su entorno y actuadores con los que pueden interactuar con el entorno en respuesta a las mediciones.



Estructura de los agentes

Todos los agentes tienen una estructura básica o mínima, la cual debe estar acorde con la función que realiza el agente. Esta estructura está compuesta por dos partes principales, la primera es el programa del agente el cual implementa la función del agente y se asume que el programa se ejecutará en algún tipo de hardware con sensores físicos y actuadores, lo cual se conoce como arquitectura.

Agente = arquitectura + programa



Estructura de los agentes

Programas de los agentes

Los programas de los agentes reciben percepciones de su entorno en forma de entradas mediante el uso de los sensores y devuelven una acción mediante el uso de actuadores. Existe una diferencia entre los programas de los agentes, que toman la percepción actual como entrada, y la función del agente que recibe la percepción histórica completa. Los programas de los agentes reciben sólo la percepción actual del agente como entrada porque no hay nada más disponible en el entorno; si las acciones del agente dependen de la secuencia completa de interacciones con el entorno, el agente tendría que recordar las percepciones.

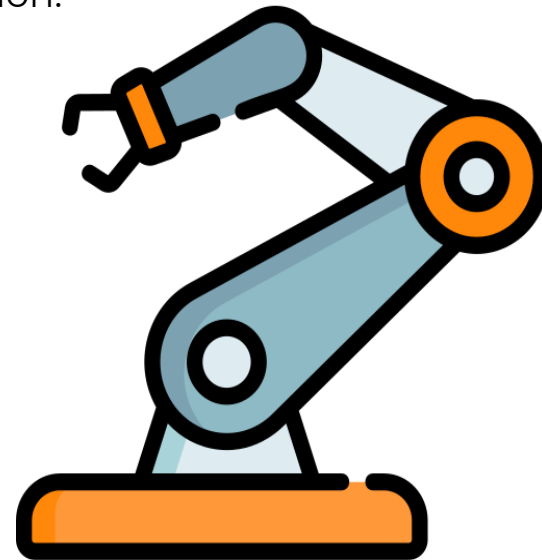
Arquitectura de los agentes

Se trata de un conjunto de elementos físicos, la arquitectura hace que las percepciones de los sensores estén disponibles para el programa, ejecutar el programa y se encarga de que los actuadores pongan en marcha las acciones generadas por los agentes.

Medio ambiente de trabajo

El medio ambiente de trabajo o entornos de trabajo, son esencialmente los problemas para los que los agentes inteligentes son las soluciones. En el diseño de un agente, el primer paso debe ser siempre especificar el entorno de trabajo de la forma más completa posible.

El rango de los entornos de trabajo en los que se utilizan técnicas de **IA** es muy grande. Sin embargo existen características que nos pueden ayudar y categorizar los entornos. Estas características determinan hasta cierto punto el diseño más adecuado para el agente y las técnicas más apropiadas para su creación.



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 puede ser parcialmente observable debido al ruido y la existencia de sensores poco exactos o porque los sensores no reciben información de parte del sistema.

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 los agentes no se tienen que preocupar por la incertidumbre en un entorno observable y determinista.

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 anteriores.

Propiedades de los entornos de trabajo

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 entornos estáticos son más sencillos de trabajar ya que el agente no necesita estar pendiente del mundo mientras está tomando una decisión sobre una acción, no necesita preocuparse por el paso del tiempo.

Discreto vs continuo

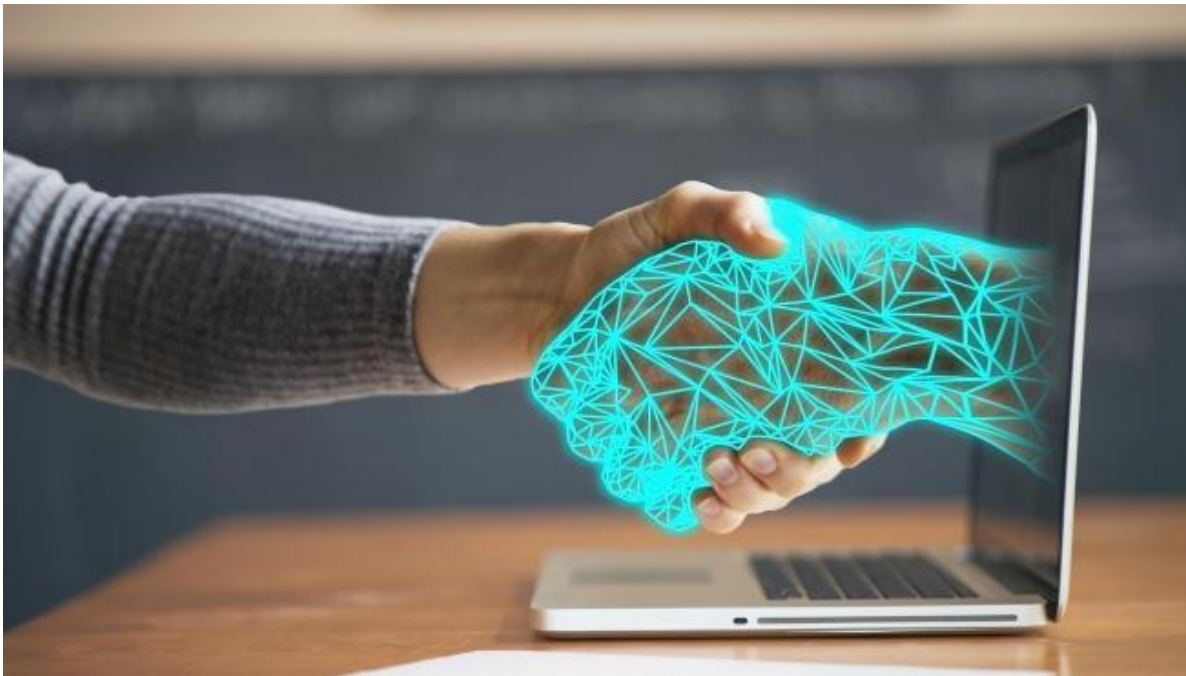
La distinción entre discreto y continuo se puede aplicar al estado del medio, a la forma en la que se maneja el tiempo y a las percepciones y acciones del agente. Por ejemplo, un medio con estados discretos como el del juego del ajedrez tiene un número finito de estados distintos.

Agente individual vs multiagente

La distinción entre el entorno de un agente individual y el de un sistema multiagente es la cantidad de agentes que interaccionan dentro de un entorno. Por ejemplo, un agente resolviendo un crucigrama por sí mismo está en un entorno individual, mientras que un agente que juega ajedrez está en un entorno con dos agentes.

Clasificación de los agentes

La inteligencia de un agente se puede distinguir entre **agentes reactivos** (reflex agents en inglés) los cuales únicamente reaccionan a entradas y **agentes con memoria** (agents with memory en inglés) los cuales también incluyen sus decisiones pasadas.

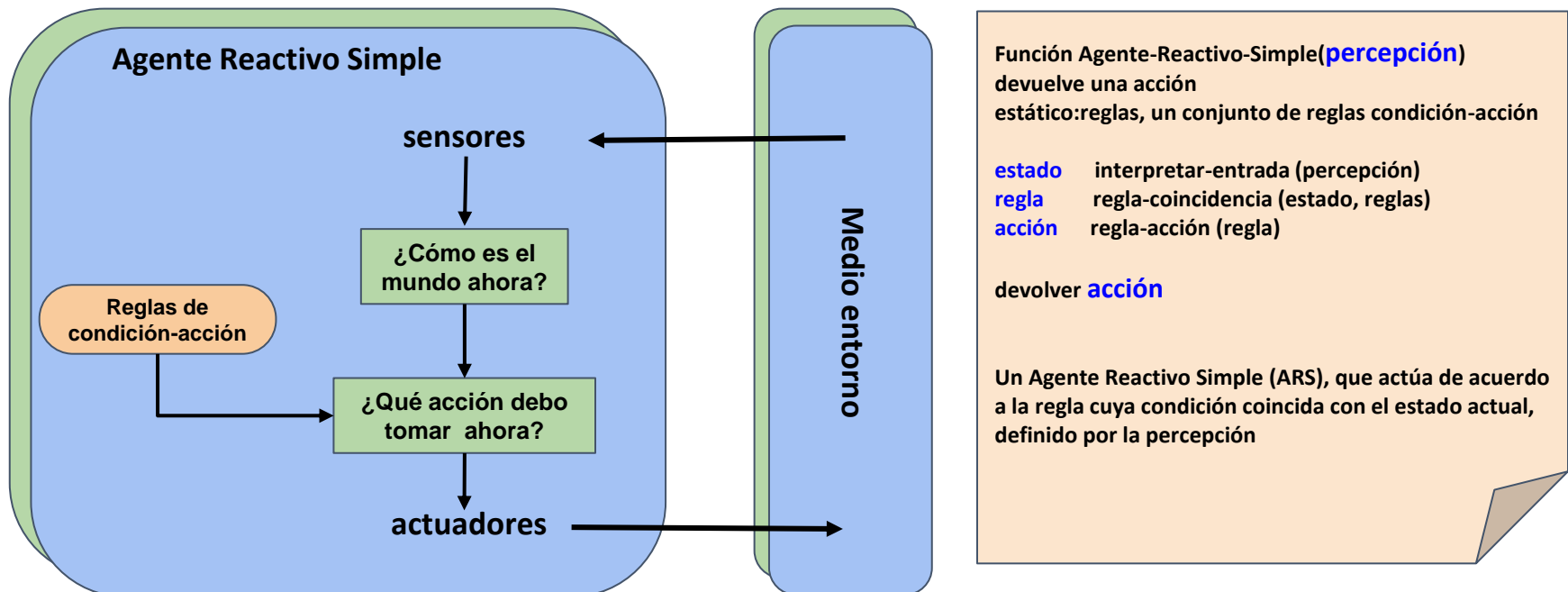


Clasificación de los agentes

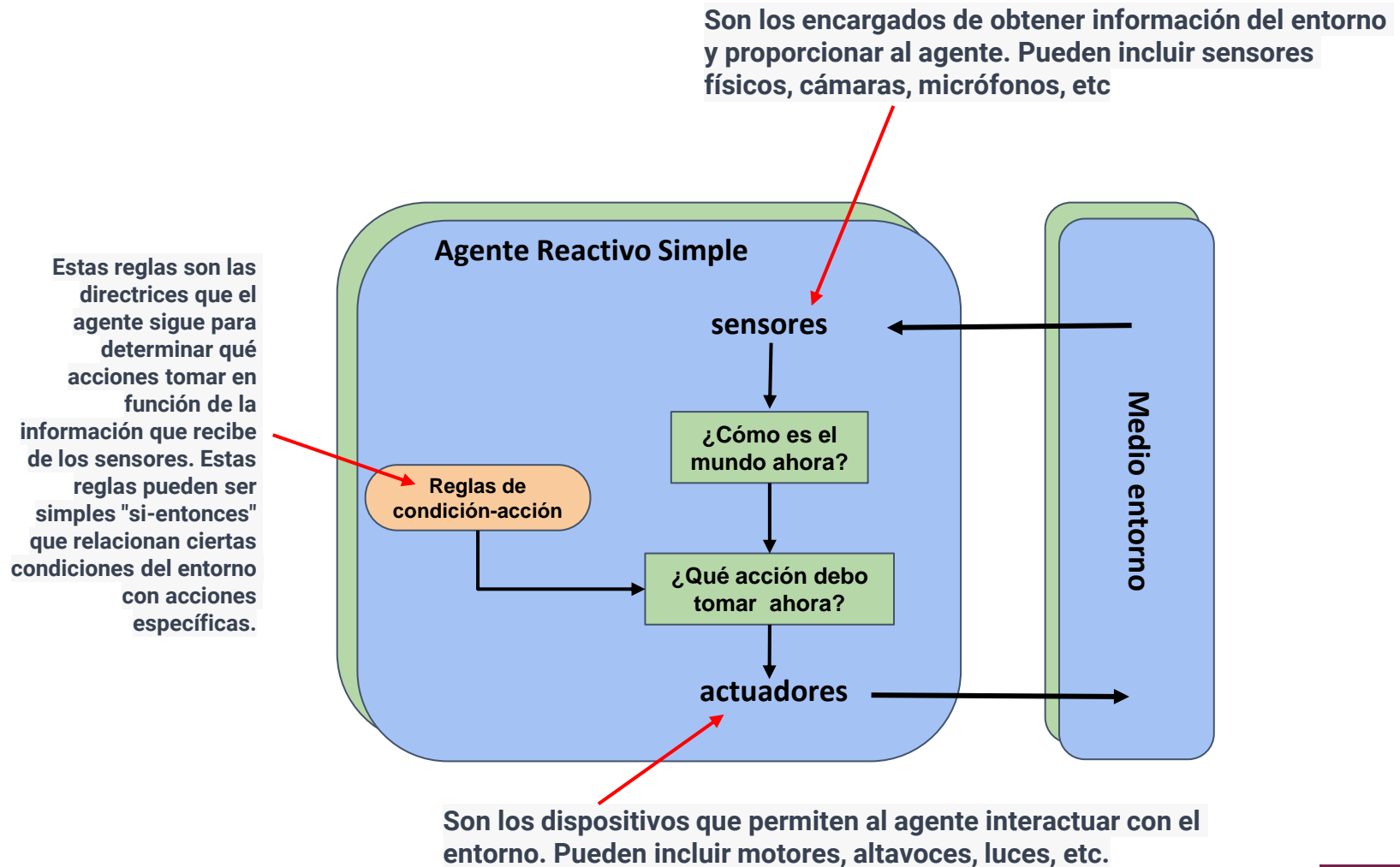
Agentes Reactivos Simples

Estos agentes no tienen una representación interna del estado del entorno ni una planificación a largo plazo. El agente toma decisiones basadas únicamente en la información actual que recibe de los sensores y aplica las reglas predefinidas para seleccionar la acción correspondiente.

Diagrama esquemático:



Agentes Reactivos Simples

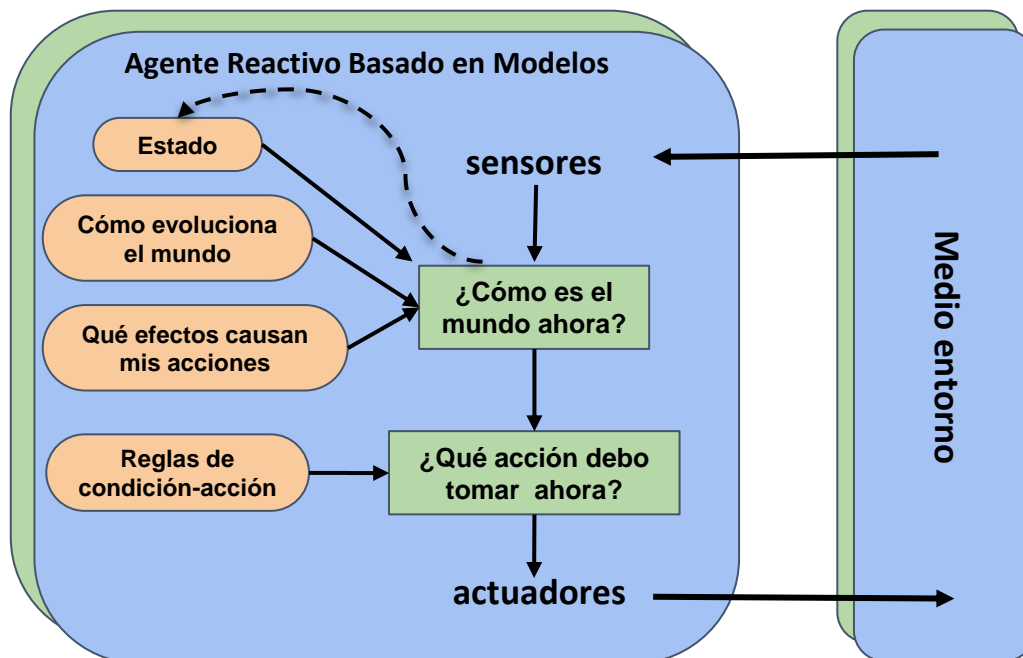


Clasificación de los agentes

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.

Diagrama esquemático



Función Agente-Reactivo-Con-Estado(percepción)
devuelve una acción
estático: *estado*. una descripción actual del mundo
reglas. un conjunto de reglas condición-acción
acción. la acción más reciente, inicialmente ninguna

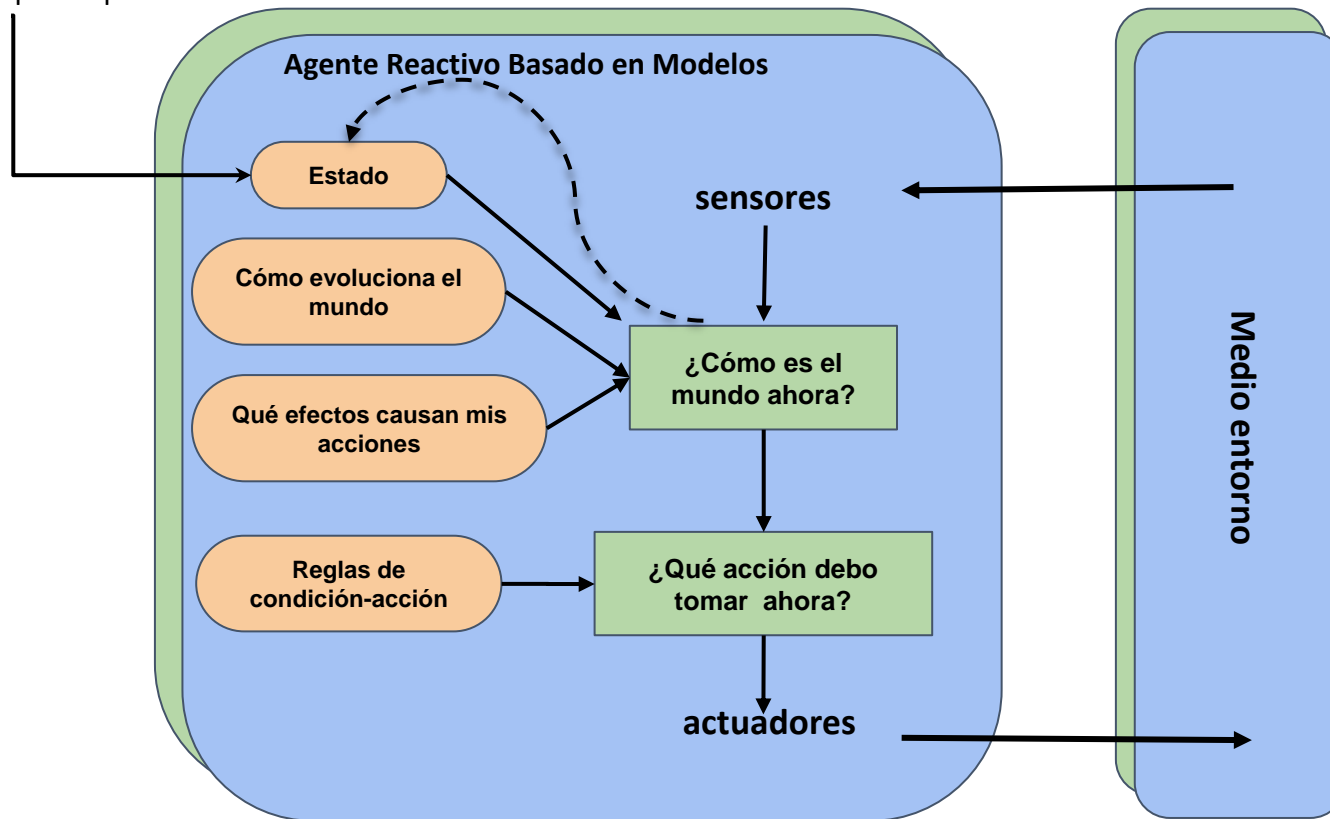
estado Actualizar-Estado(estado, acción, percepción)
regla Regla-Coincidencia(estado, reglas)
acción Regla-Acción[regla]

devolver **acción**

Un agente reactivo basado en modelos, que almacena información sobre el estado actual del mundo utilizando un modelo interno. Después selecciona una acción de la misma forma que un agente reactivo.

Agentes reactivos basados en modelos

Los estados internos almacenan información de las partes del mundo que no pueden ver.

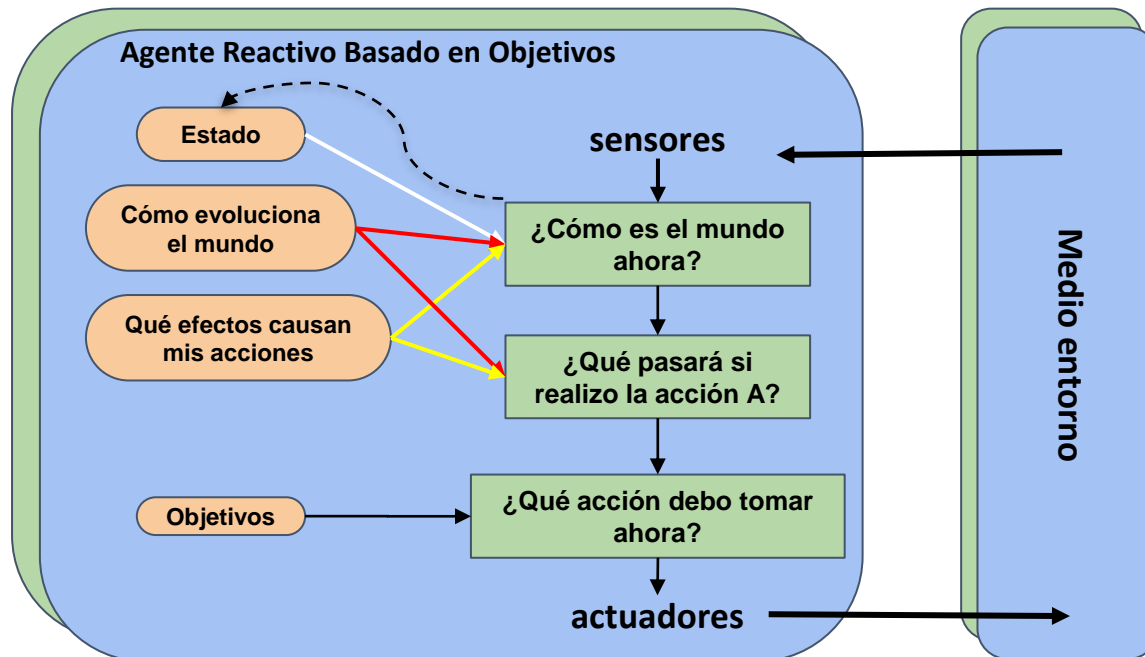


Clasificación de los agentes

Agentes basados en objetivos

El conocimiento sobre el estado actual del mundo no es siempre suficiente para decidir qué hacer. Por ejemplo, en un cruce de carreteras, un auto autónomo puede girar a la izquierda, a la derecha o seguir hacia adelante. La decisión correcta depende de a dónde quiere ir el auto. En otras palabras, además de la descripción del estado actual, el agente necesita algún tipo de información sobre la **meta** que describa las situaciones que son deseables, por ejemplo llegar al destino propuesto por el pasajero.

Diagrama esquemático:



Función Agente-Reactivo-Basado-en-Objetivos (percepción)
devuelve una acción

estático: *estado*. una descripción actual del mundo
objetivos. información sobre su meta donde se describen las situaciones deseables
acción. la acción más reciente, inicialmente ninguna

estado Actualizar-Estado(estado, acción, percepción)

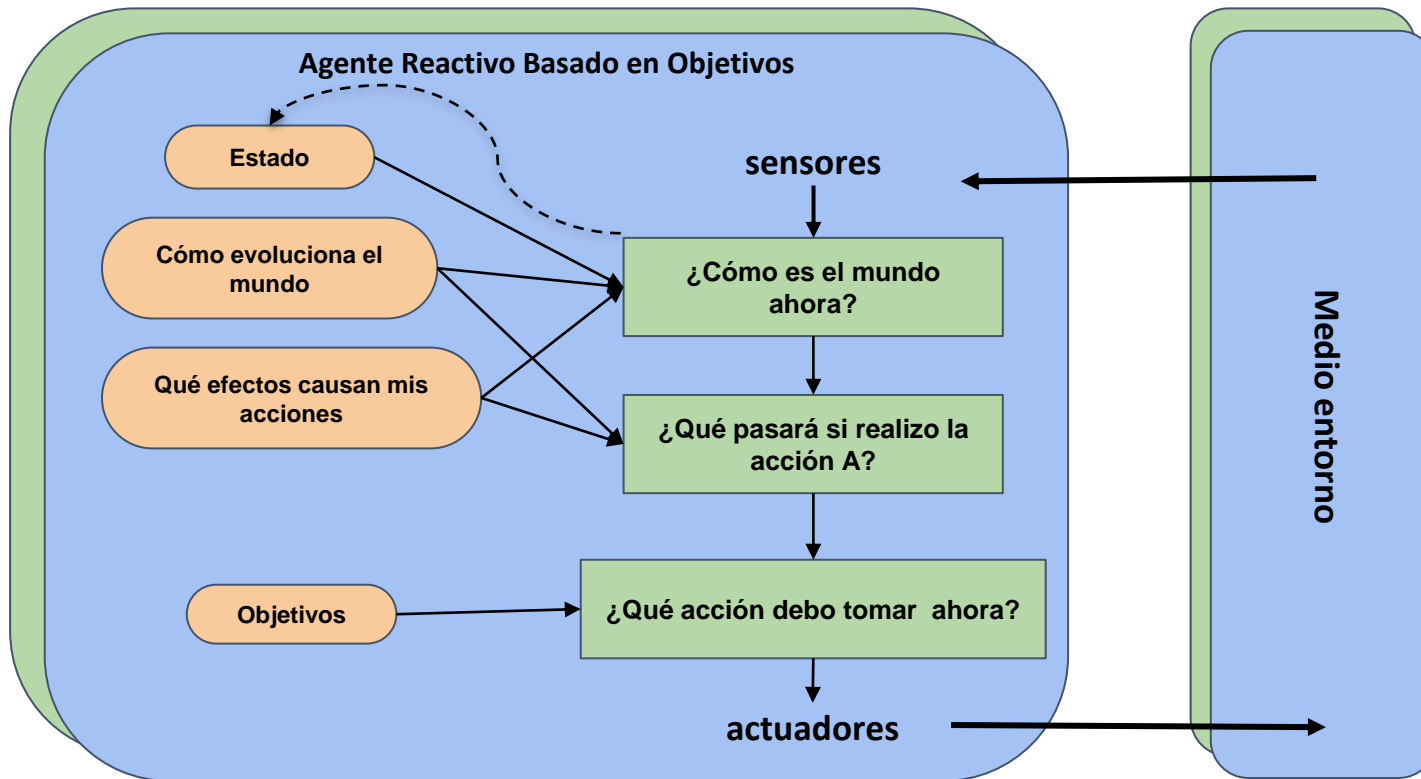
objetivos Objetivo-Coincidencia(estado, objetivos)

acción Objetivo-Acción[objetivo]

devolver **acción**

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 basados en objetivos



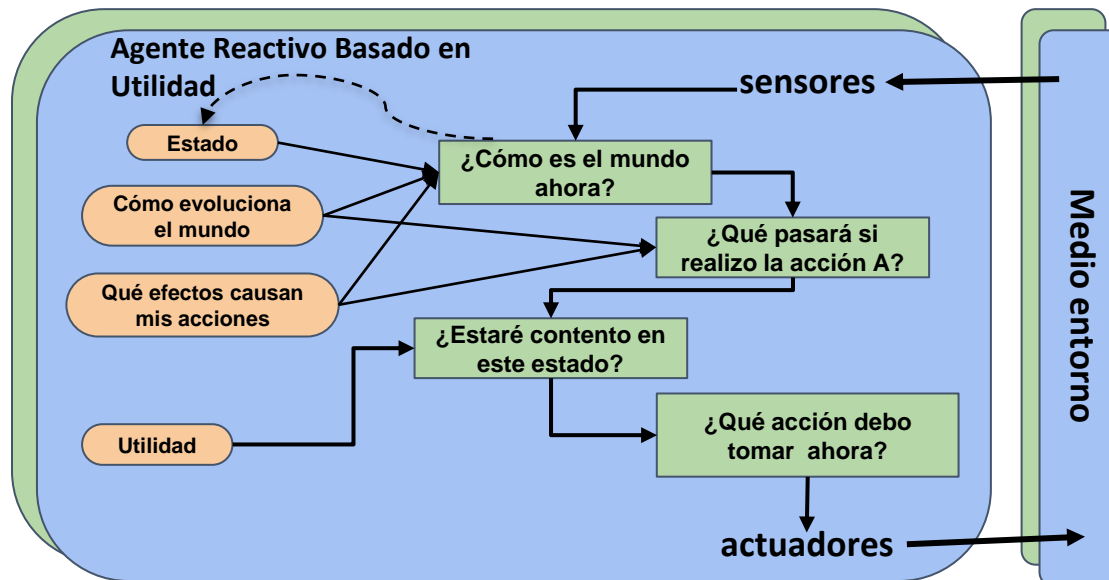
Un agente basado en objetivos y basado en modelos, que almacena información del estado del mundo así como del conjunto de objetivos que intenta alcanzar, y que es capaz de seleccionar la acción que eventualmente lo guiará hacia la consecución de sus objetivos

Clasificación de los agentes

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 a un auto autónomo a su destino (y por tanto alcanzar su objetivo), pero algunas son más rápidas, más seguras, más fiables, o más baratas que otras. Las metas sólo proporcionan una cruda distinción binaria entre estados, mientras que una medida de eficiencia más general debería permitir una comparación entre estados diferentes y decidir qué estado tiene mayor **utilidad** para el agente.

Diagrama esquemático:



Función Agente-Reactivo-Basado-en-Utilidad (percepción)
devuelve una acción

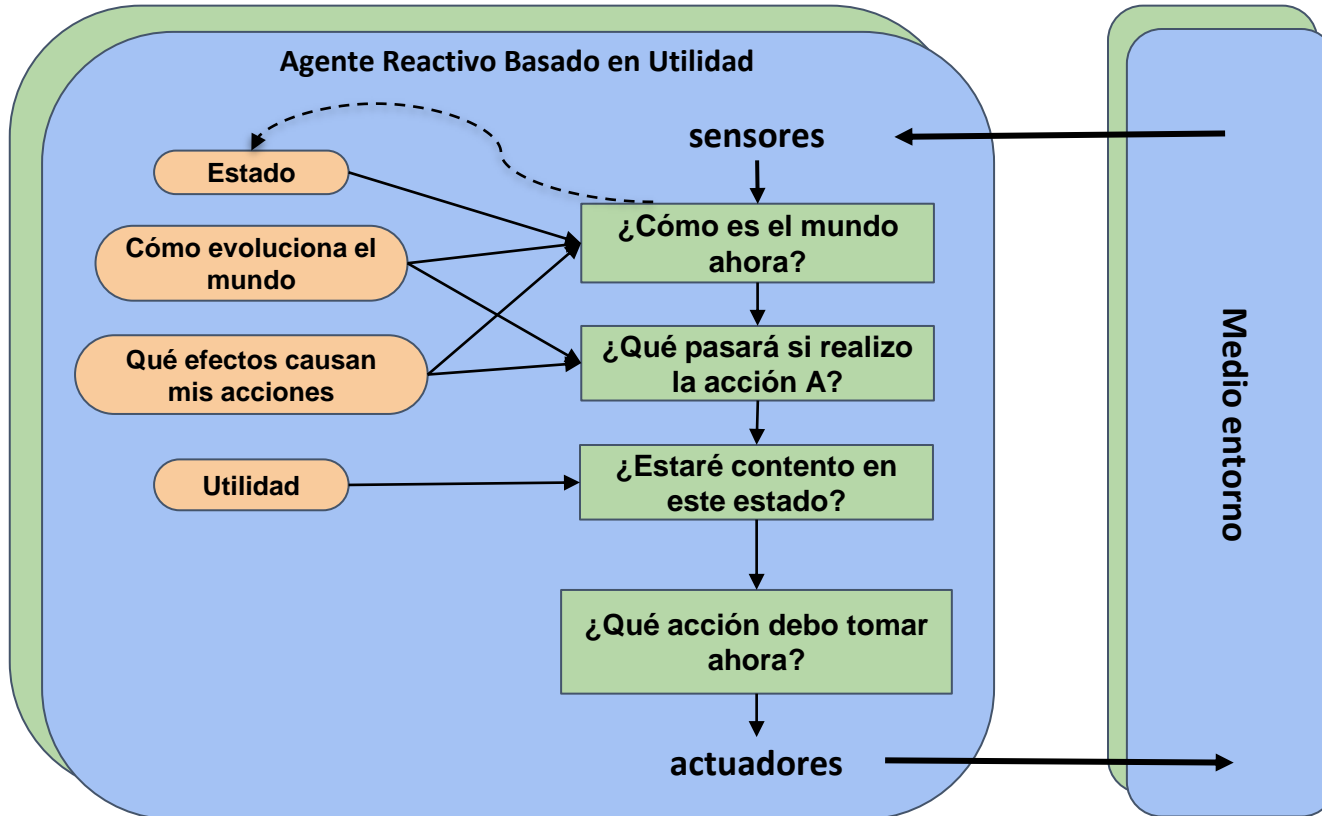
estático: *estado*. una descripción actual del mundo
utilidad. representa un nivel de preferencia respecto al estado actual del agente.
acción. la acción más reciente, inicialmente ninguna

estado Actualizar-Estado(estado, acción, percepción)
utilidad Objetivo-Preferencia(estado, preferencia)
acción Objetivo-Acción[objetivo]

devolver **acción**

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 lleve a alcanzar la mayor utilidad esperada.

Agentes basados en utilidad

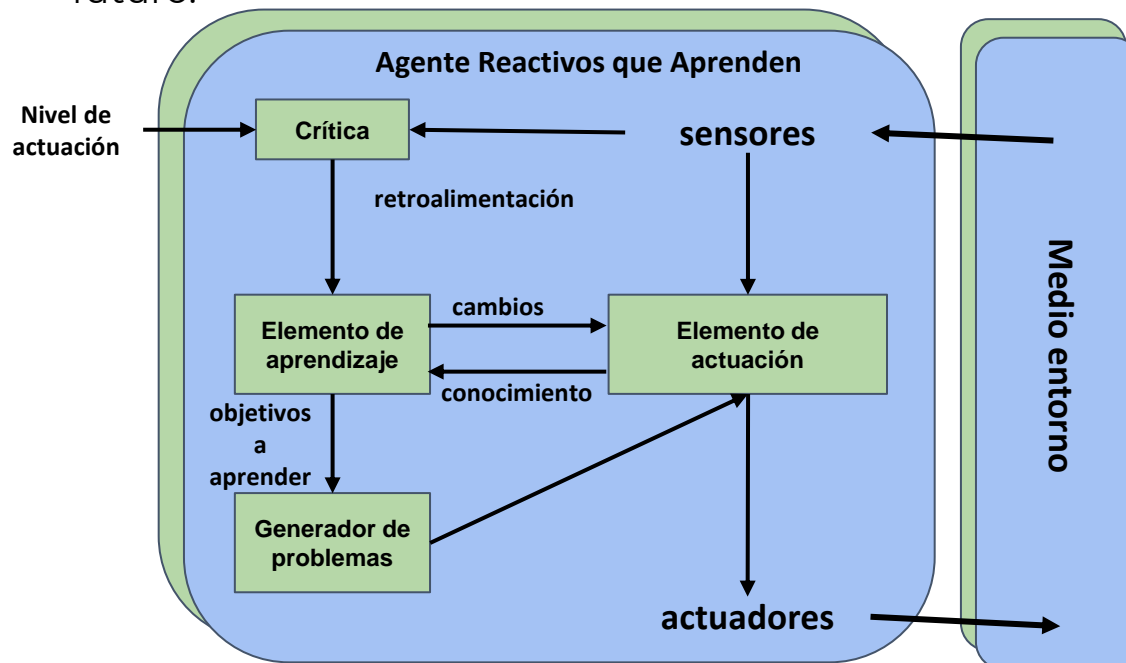


Un agente basado en utilidad y basado en modelos. Utiliza un modelo del mundo. junto con una función de utilidad que calcula sus preferencias entre los estados del mundo. Después selecciona la acción que le lleve a alcanzar la mayor utilidad esperada, que se calcula haciendo la media de todos los estados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

Clasificación de los agentes

Agentes que aprenden

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 está responsabilizado de hacer mejoras y el segundo se responsabiliza de la selección de acciones externas. El elemento de actuación es lo que se considera un agente reactivo, recibe estímulos y determina las acciones a realizar. El elemento de aprendizaje se realimenta con las críticas sobre la actuación del agente y determina cómo se debe modificar el elemento de actuación para proporcionar mejores resultados en el futuro.



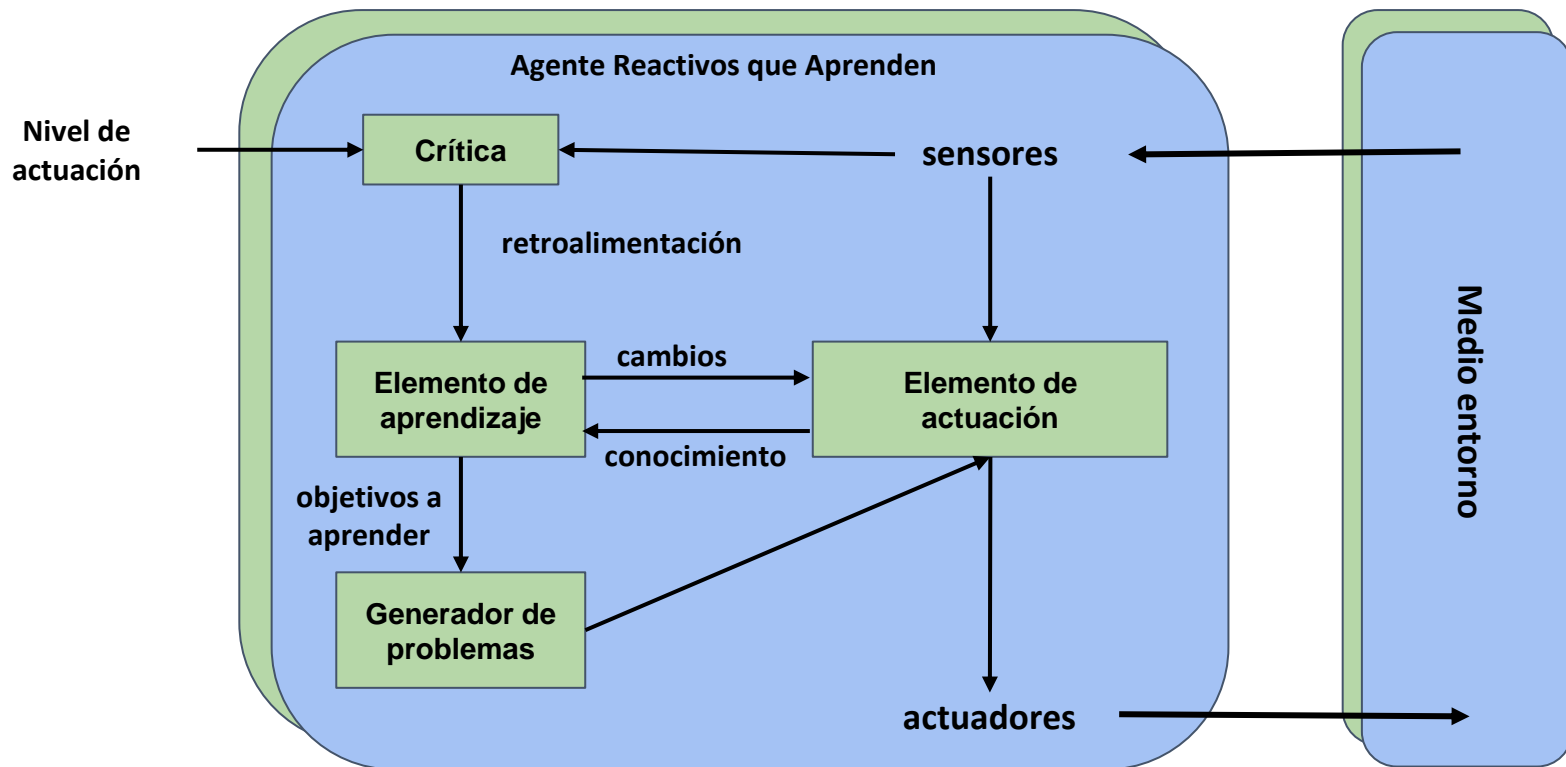
Función Agente-Reactivo-que-Aprende(percepción)
devuelve una acción
estático: *crítica*. retroalimentación de las acciones del agente
elemento de aprendizaje (EA). realiza mejoras al comportamiento del agente.
generador de problemas (GP). sugiere acciones que generen nueva información sobre el entorno
acción. la acción más reciente, inicialmente ninguna

crítica Actualizar-Agente(estado, acción, percepción)
EA Objetivo-Mejoras(estado, preferencia)
GP Mejoras-Crítica(estado, preferencia)
acción Objetivo-Acción[objetivo]

devolver **acción**

Emplea un modelo de aprendizaje que le permite operar en entornos desconocidos y presentar mejoras para adaptarse al mismo..

Agentes que aprenden



Aplicaciones de los agentes inteligentes

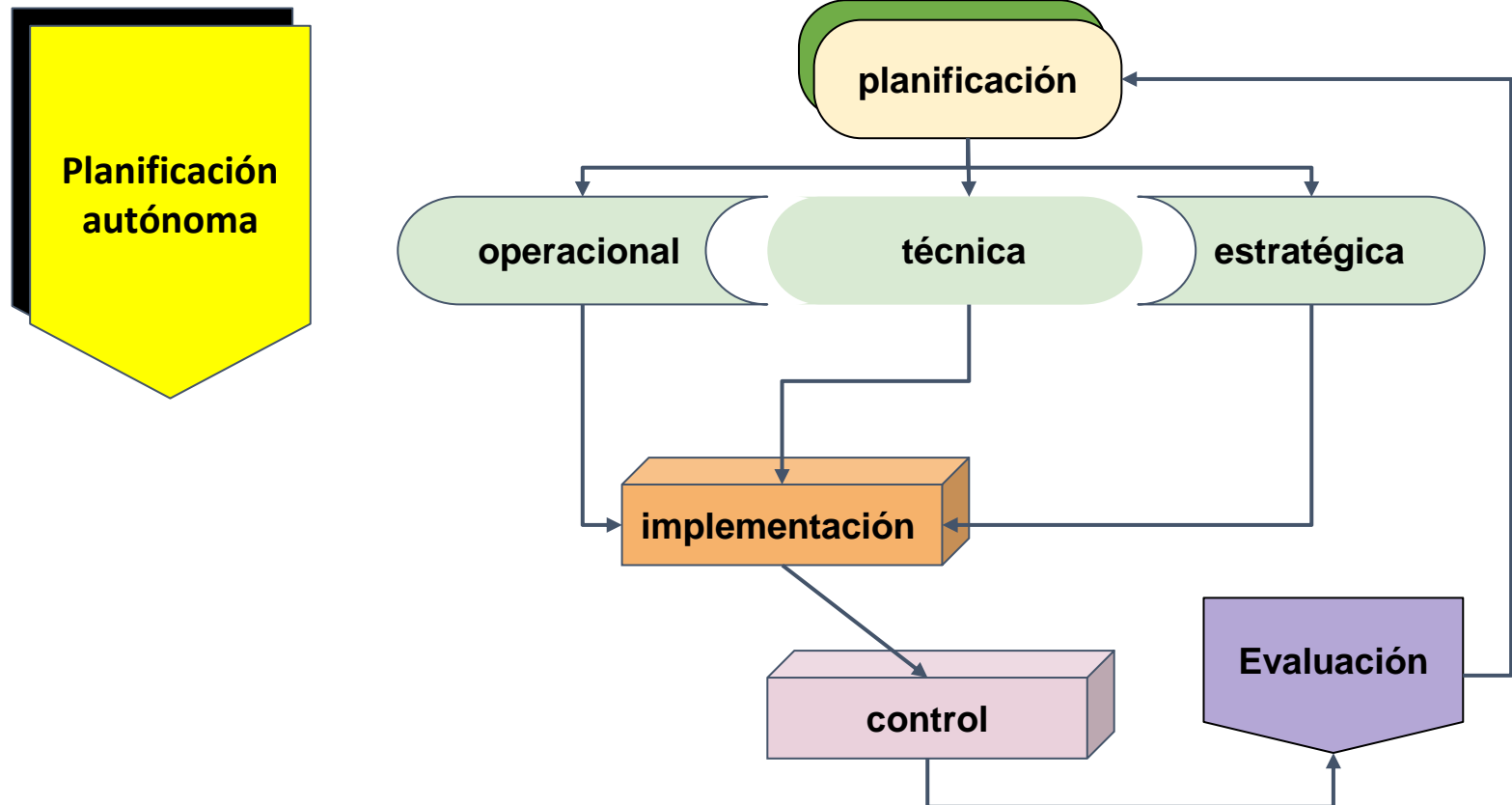
**Planificación
autónoma**

Juegos

**Control
autónomo**

Diagnosis

Aplicaciones de los Agentes Inteligentes



Aplicaciones de los agentes inteligentes

Juegos



Aplicaciones de los agentes inteligentes

Control
autónomo



Aplicaciones de los agentes inteligentes

Diagnosis

