

Temario Inteligencia Artificial

Ismael Sallami Moreno

`ism350zsallami@correo.ugr.es`

`https://ismael-sallami.github.io/`

`https://elblogdeismael.github.io/`

Universidad de Granada

2025

Índice general

1. Tema 1	5
1.1. Tema 1	5
1.1.1. Introducción: Inteligente	5
1.1.2. Definición de IA	5
2. Tema 2	7
2.1. Tema 2: Concepto de Agente Inteligente	7
2.1.1. Agentes Inteligentes	7
2.1.2. Arquitecturas de Agentes	7

Capítulo 1

Tema 1

1.1. Tema 1

1.1.1. Introducción: Inteligente

Debemos de definir la inteligencia artificial usando la palabra inteligente, la cual es difícil de definir. Se dan diferentes definiciones de Harvard y otras universidades/entidades reconocidas.

1.1.2. Definición de IA

Podemos tener varias definiciones de IA, como pueden ser sistemas que actúan/-piensan como humanos, o bien sistemas que actúan/piensan racionalmente. Una de ellas puede ser: "Campo de estudio que busca explicar y emular el comportamiento inteligente en términos de procesos computacionales"

Un ejemplo de IA que menciona el profesor reiteradamente es el caso de los *cajeros automáticos*.

Se garantiza que una pregunta sobre las definiciones de IA cae.

Asociamos el concepto de IA a racional debido a que se quiere desarrollar con la capacidad de poder tomar decisiones como los humanos.

Hay distintos softwares en los dispositivos que podemos relacionarlos con la acción de que *piensan* debido a que usan inteligencia artificial, como es el caso de programas de ajedrez. ¿Es esto correcto? No podemos afirmar esto exactamente, ya que según autores pensar es emular la mente humana, cosa que no es capaz de hacer la IA debido a que solo está entrenado a hacer las mejores jugadas para ganarte, mientras que tú puedes pensar sobre otras temáticas y tener cierto *pensamiento libre*.

Si afirmamos que piensan como humanos, debemos de tener en cuenta el *pensamiento cognitivo*.

La IA según el libro del *Porque* (recibió el premio Turing), tiene 3 niveles:

1. Aprendizaje automático.
2. Simular al humano.
3. Capacidad de imaginar, lo cual, a día de hoy es inalcanzable.

En cuanto a la creatividad, hay cierta *creatividad cognitiva* en temas específicos.

El llamado 'no alineamiento con los humanos' es lo que ocurre cuando una IA no tiene información sobre ese tema o bien no está entrenado, pues en este caso se lo inventa.

El Test de Turing, es aquel en el que si una IA consigue pasar desapercibido como humano, pues efectivamente pasa este Test.

En esta asignatura vamos a seguir la Teoría del Agente, que es la idea más aceptada hoy día y más usada. Este percibe el entorno y actúa de una manera específica.

Se usa la palabra *racional* aludiendo a inteligente.

Un agente racional actúa de manera correcta según la información que posee.

El origen del término de IA recae en John McCarthy. Los pioneros de la IA son varios, entre los que podemos encontrar a Alan Turing, Marvin Minsky,...

En el debate de la IA se menciona una IA fuerte y una débil, que son sistemas que actúan como humanos y racionalmente vs las que piensan como IA.

En cuanto al recorrido de la IA a lo largo de la historia, podemos destacar distintas épocas como es la Edad de Oro, Invierno de IA,...

Von Neumann propone el algoritmo MiniMax, el que afirma que el juego es una situación conflictiva en la que tomamos decisiones sabiendo que los demás también las toman, determinando el resultado en base a las decisiones realizadas. Un hecho histórico destacable, es que Napoleón murió pensando que había perdido jugando al ajedrez pensando que perdió frente al primer sistema inteligente, aunque puede que haya sido un humano. Otro de los avances importantes es el caso de los cohes autónomos.

En cuanto a los tipos de IA podemos mencionar:

- IA débil: una máquina que compite con un humano en una actividad específica.
- IA fuerte: IA que aplica inteligencia para cualquier problema.
- IA general: alcanza el nivel cognitivo humano.
- Super-IA: capacidad que es mucho mayor que cualquier cerebro humano en cualquier campo.

Capítulo 2

Tema 2

2.1. Tema 2: Concepto de Agente Inteligente

La Teoría de Agentes(denotado a continuación como **TA**) es un sistema de ordenador, situado en un entorno, capaz de realizar acciones de manera autónoma y que es flexible en base a las situaciones que se le presentan.

2.1.1. Agentes Inteligentes

Desarrolla un enfoque basado en agentes y lleva a cabo percepciones procesada mediante algoritmos de análisis y toma de decisiones. La TA se usa actualmente en la mayoría de ámbitos y es lo que más se usa hoy en día.

En cuanto a los tipos de agentes encontramos:

- Reactivo: reacciona en base al entorno.
- Pro-activo: no solo actúan en respuesta al entorno, sino que puedan analizar acciones a llevar a cabo.
- Social: son además capaces de interactuar.
- Multiagente: esta implementado como varios agentes interactuando.

La interacción entre agentes se lleva a cabo mediante: cooperación, coordinación y la negociación.

2.1.2. Arquitecturas de Agentes

Podemos distinguir entre:

- Reactivo: reacciona en base a la situación en la que se encuentra, eligiendo la más adecuada en base a lo que sabe.
- Deliberativo: toma decisiones basadas en razonamiento, planificación y modelos internos del mundo.

Bibliografía

- [1] Autor(es), *Título del artículo*, Nombre de la Revista, volumen(número), páginas, año.
- [2] Autor(es), *Título del libro*, Editorial, año.
- [3] Autor(es), *Título del documento*, Nombre de la Conferencia, páginas, año.