

# Módulo I: Fundamentos de la Inteligencia Artificial

## Tema 1. Introducción a la Inteligencia Artificial

**José Manuel Fuertes García**

`jmf@ujaen.es`

Departamento de Informática  
Universidad de Jaén

30 de enero de 2018

# Objetivos

- ▶ Entender los diferentes enfoques para definir la Inteligencia Artificial, características, ventajas e inconvenientes
- ▶ Conocer los fundamentos de la Inteligencia Artificial y áreas de conocimiento implicadas
- ▶ Conocer elementos significativos de la historia y estado del arte de la Inteligencia Artificial

# Índice

¿Qué es la Inteligencia Artificial?

Fundamentos de la Inteligencia Artificial

Historia de la Inteligencia Artificial

Estado del Arte

# Inteligencia

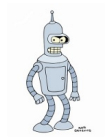
- ▶ Las capacidades mentales son muy importantes para el ser humano.
- ▶ Durante miles de años uno de los principales objetivos de la ciencia ha sido entender cómo pensamos.

*Cómo podemos percibir, entender, predecir y manipular un mundo mucho más grande y complicado que nosotros mismos.*

- ▶ La I.A. además de intentar comprender, se esfuerza en construir entidades *inteligentes*.

# Entidades Inteligentes

- ▶ ¿Qué es una *entidad inteligente*?
  - ▶ Hoy día tenemos edificios inteligentes, aspiradoras inteligentes, sillas inteligentes...



¿Entidades inteligentes?

# La Inteligencia Artificial

## Características:

- ▶ Abarca una gran variedad de subcampos.
- ▶ Sintetiza y automatiza tareas intelectuales.
- ▶ Es potencialmente relevante para cualquier ámbito de la actividad intelectual humana.
- ▶ Es, en este sentido, un campo genuinamente universal.

## Objetivo de la I.A.

*Estudiar los entes inteligentes.*

- ▶ Ambito **científico**: entender los entes inteligentes.
- ▶ Ambito **ingenieril**: *construir* entes inteligentes.

# Definiciones de Inteligencia Artificial

- ▶ Sistemas que piensan como humanos

“El nuevo y excitante esfuerzo de hacer que los computadores piensen... máquinas con mentes, en el sentido más literal” (Haugeland, 1985)

- ▶ Sistemas que piensan racionalmente

“El estudio de las facultades mentales mediante el uso de modelos computacionales” (Charniak y McDermott, 1985)

- ▶ Sistemas que actúan como humanos

“El arte de desarrollar máquinas con capacidad para realizar funciones que cuando son realizadas por personas requieren de inteligencia” (Kurzweil, 1990)

- ▶ Sistemas que actúan de forma racional

“La inteligencia computacional es el estudio del diseño de agentes inteligentes” (Poole et al., 1998)

“IA... está relacionada con conductas inteligentes en artefactos” (Nilsson, 1998)

I.A. *fuerte*



I.A. *débil*

## Sistemas que *piensan* como humanos

Para decir que un programa piensa como un humano es necesario un mecanismo para determinar cómo piensan los humanos.



Elaborar un **modelo cognitivo**.

- ▶ Abrir la *caja negra* de la mente humana.
- ▶ Analizar los procesos mentales (introspección, experimentos).
- ▶ Desarrollar una teoría sobre estos procesos mentales.
- ▶ Aplicar esta teoría en la simulación de estos procesos en un ordenador.



## Ejemplo

Sistema de resolución general de problemas (*General Problem Solver*, GPS, Newell y Simon, 1961).

- ▶ **Objetivo:** Más que resolver correctamente problemas, seguir la pista de las etapas del proceso de razonamiento.
  - ▶ Resuelve problemas mediante la descomposición en problemas más simples.
  - ▶ Se centra en la comparación de los pasos de razonamiento del GPS con los pasos seguidos por una persona al resolver el mismo problema.

# La ciencia cognitiva

## Campo interdisciplinar:

Modelos computacionales (IA) + Técnicas experimentales (psicología).

- ▶ Elabora teorías precisas y verificables sobre el funcionamiento de la mente humana.
- ▶ IA y ciencia cognitiva actualmente se desarrollan con rapidez apoyándose mutuamente.

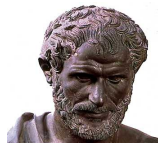
# Las *leyes del pensamiento*

- ▶ Pretenden representar la manera *correcta* de pensar.
  - ▶ Aristóteles fue uno de los primeros en representar un proceso de razonamiento irrefutable mediante

**silogismos:**

*Estructuras de argumentación mediante las que siempre se llega a una conclusión correcta si se parte de premisas correctas.*

- ▶ Estas leyes de pensamiento supuestamente gobiernan la forma de operar de la mente.
  - ▶ Su estudio fue el inicio de la **Lógica**.
- ▶ *En principio*, deberían permitir superar a la inteligencia humana.



Aristóteles

- ▶ La tradición *logista* dentro de la IA intenta construir sistemas inteligentes a partir de programas que resuelvan cualquier problema *resoluble* descrito en notación lógica.

### Inconvenientes de este enfoque:

- ▶ No es fácil expresar cualquier tipo de conocimiento en notación lógica.
- ▶ Es muy diferente resolver un problema *en principio* a resolverlo en la práctica.
- ▶ Estas dificultades están presentes en todo intento de construir sistemas de razonamiento computacional.

# Agentes racionales

## Comportamiento racional:

Hacer *lo correcto*.

- ▶ **Lo correcto**: aquello con lo que se espera alcanzar el mejor resultado, dada la información disponible.
- ▶ No implica el pensamiento obligatoriamente (p.e. el acto reflejo de parpadear) pero el pensamiento debería estar al servicio de la acción racional.

# ¿Qué es un agente racional?

Agente:

Entidad que percibe y actúa.

Agente racional:

Aquel que percibe, razona y actúa con la intención de alcanzar el mejor resultado o, cuando hay incertidumbre, el mejor resultado esperado.

- Función desde la historia de percepciones hasta las acciones:

$$f : P^* \longrightarrow A$$

- Para un entorno y tarea dados, buscamos el agente (o clase de agentes) con el mejor resultado.
- Las limitaciones computacionales hacen que la racionalidad perfecta sea inalcanzable  $\implies$  diseño del mejor programa para unos recursos de máquina determinados.

# Características de los agentes racionales

- ▶ Enfoque relativo al contexto:
  - ▶ Actuar de forma **correcta** en un **entorno determinado**.
- ▶ No se limitan a la inferencia racional (lógica):
  - ▶ A veces es imposible determinar formalmente cuál es la mejor acción.
  - ▶ En algunas situaciones es *racional* emprender una acción *buena* inmediatamente en lugar de esperar hasta determinar la alternativa óptima.
  - ▶ Se pueden determinar acciones racionales mediante inferencias no lógicas.

# Ventajas de los agentes racionales

- ▶ Pone énfasis en una perspectiva ingenieril.
- ▶ Destaca la relación entre comportamientos inteligentes y el entorno en el que se desarrollan.
- ▶ Proporciona criterios transparentes para evaluar conducta inteligente.
- ▶ Permite una concepción integrada de las distintas técnicas y subáreas de la Inteligencia Artificial.

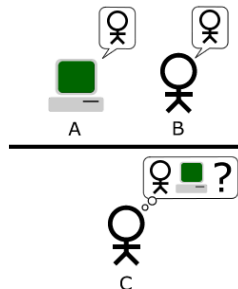


# La prueba de Turing

- Propuesta en 1950 por Turing:



Alan Turing



- **Objetivo:** proporcionar una definición operacional y satisfactoria de inteligencia.
  - Un evaluador humano formula preguntas y el interlocutor (humano o sistema de IA) responde.
  - El sistema de IA supera la prueba si el evaluador no es capaz de distinguir entre él y el humano.

# Utilidad del test de Turing

## Capacidades requeridas:

- ▶ Procesamiento del lenguaje natural.
- ▶ Representación del conocimiento.
- ▶ Razonamiento automático.
- ▶ Aprendizaje automático.

## Prueba global de Turing:

- ▶ Incluye señales de vídeo y objetos físicos.
- ▶ Requiere visión computacional y robótica.
- ▶ Problema de *la habitación china*.
- ▶ La máquina ELIZA (Weizenbaum, 1966) superó una versión simplificada del test.

# Índice

¿Qué es la Inteligencia Artificial?

Fundamentos de la Inteligencia Artificial

Historia de la Inteligencia Artificial

Estado del Arte

# Fundamentos de la I.A.

## Filosofía

- ▶ ¿Se pueden utilizar reglas formales para extraer conclusiones válidas?
- ▶ ¿Cómo se genera la inteligencia mental a partir de un cerebro físico?
- ▶ ¿De dónde viene el conocimiento?
- ▶ ¿Cómo se pasa del conocimiento a la acción?
- ▶ Aspectos:
  - ▶ Lógica, métodos de razonamiento.
  - ▶ La mente como sistema físico.
  - ▶ Fundamentos de aprendizaje, lenguaje, racionalidad.

# Fundamentos de la I.A.

## Matemáticas

- ▶ ¿Qué reglas formales son las adecuadas para obtener conclusiones válidas?
- ▶ ¿Qué se puede computar?
- ▶ ¿Cómo razonamos con información incierta?
- ▶ Aspectos:
  - ▶ Representación formal y pruebas.
  - ▶ Algoritmos, computación, (in)decisión, (in)flexibilidad.
  - ▶ Probabilidad.

# Fundamentos de la I.A.

## Psicología

- ▶ ¿Cómo piensan y actúan los humanos y los animales?
- ▶ Aspectos:
  - ▶ Adaptación.
  - ▶ Fenómenos de percepción y control motor.
  - ▶ Técnicas experimentales (psicofísica, etc.).

## Economía

- ▶ ¿Cómo se deben tomar las decisiones para maximizar el rendimiento?
- ▶ ¿Cómo se deben realizar las acciones cuando los resultados se obtienen en un futuro lejano?
- ▶ Aspectos:
  - ▶ Teoría formal de las decisiones racionales.

# Fundamentos de la I.A.

## Lingüística

- ▶ ¿Cómo está relacionado el lenguaje con el pensamiento?
- ▶ Aspectos:
  - ▶ Representación del conocimiento.
  - ▶ Gramática.

## Neurociencia

- ▶ ¿Cómo procesa la información el cerebro?
- ▶ Aspectos:
  - ▶ Sustrato físico plástico para la actividad mental.

# Fundamentos de la I.A.

## Teoría de Control

- ▶ ¿Cómo pueden los artefactos funcionar bajo su propio control?
- ▶ Aspectos:
  - ▶ Sistemas homeostáticos, estabilidad.
  - ▶ Diseño de agentes óptimos simples.



# Índice

¿Qué es la Inteligencia Artificial?

Fundamentos de la Inteligencia Artificial

Historia de la Inteligencia Artificial

Estado del Arte

# Historia

## Definiciones distintas en tiempos distintos.

- ▶ 1956: La ciencia e ingeniería dedicadas a la creación de máquinas inteligentes.
- ▶ 2000-: Estudio y diseño de agentes inteligentes.

# Historia

- ▶ **1943** McCulloch y Pitts: modelo del circuito Booleano del cerebro.
- ▶ **1950** *Computing Machinery and Intelligence* de Turing.
- ▶ **1952-69** Entusiasmo inicial, grandes esperanzas.
  - ▶ Primeros programas de la IA, incluidos el programa para el juego de las damas de Samuel, el Teórico Lógico de Newell y Simon y el demostrador de teoremas de geometría de Gelernter
- ▶ **1956** Dartmouth: adopción de la inteligencia artificial.
- ▶ **1965** Algoritmo completo para la demostración de teoremas para la lógica de primer orden de Robinson.

# Historia

Afirmaciones de algunos participantes en la conferencia de Dartmouth

- ▶ Herbert Simon y Allen Newell, 1958: *"Dentro de diez años un ordenador será campeón de ajedrez del mundo"* *"dentro de diez años un ordenador descubrirá y demostrará un nuevo e importante teorema matemático"*
- ▶ Herbert Simon, 1965: *"las máquinas serán capaces, en veinte años, de hacer cualquier trabajo que una persona puede hacer"*
- ▶ Marvin Minsky, 1967: *"Dentro de una generación ... el problema de la creación de la inteligencia artificial se resolverá sustancialmente"*

# Historia

- ▶ **1966-74** La IA descubre la complejidad computacional. La investigación de las redes neuronales casi a punto de desaparecer.
- ▶ **1969-79** Primeros avances de los sistemas basados en el conocimiento.
- ▶ **1980-88** Florece la industria de los sistemas expertos.
- ▶ **1988-93** Quiebra de la industria de los sistemas expertos: *el invierno de a IA*.
- ▶ **1985-95** Regreso de las redes neuronales.
- ▶ **1988-** Resurgimiento de la probabilidad; aumento general del enfoque técnico.
- ▶ **1995-** Agentes por todas partes...

# Índice

¿Qué es la Inteligencia Artificial?

Fundamentos de la Inteligencia Artificial

Historia de la Inteligencia Artificial

Estado del Arte

# ¿Qué puede hacer actualmente la I.A.?

- ▶ Planificación autónoma:
  - ▶ Programa de la NASA Agente Remoto (planificación de las operaciones de una nave espacial).
- ▶ Juegos:
  - ▶ Deep Fritz gana a Vladimir Kramnik en 2006.
  - ▶ Jonathan Schaeffer *resuelve* el juego de las Damas (2007).
- ▶ Control autónomo:
  - ▶ Sistema de visión por computador ALVINN.
- ▶ Diagnóstico:
  - ▶ Programas de diagnóstico médico.



# ¿Qué puede hacer actualmente la I.A.?

- ▶ Planificación logística:
  - ▶ DART (Dynamic Analysis and Replanning Tool) para la planificación y organización logística del transporte.
- ▶ Robótica:
  - ▶ HipNav, asistente robot para operaciones de microcirugía.
- ▶ Procesamiento del lenguaje y resolución de problemas:
  - ▶ PROVER B, resuelve crucigramas.
  - ▶ IBM Watson empata en el juego Jeopardy! (febrero de 2011).

