



Inteligencia Artificial

Universidad Tecnológica de la Mixteca

Instituto de Física y Matemáticas

M.C. Alejandro Hernández T.

Programa de estudios 1/4

1.- Introducción

- i. ¿Qué es la inteligencia artificial?
- ii. Los fundamentos de la inteligencia artificial
- iii. Historia de la inteligencia artificial
- iv. Los últimos avances

2.- Agentes Inteligentes

- i. Introducción
- ii. Cómo debe proceder un agente
- iii. Estructura de los agentes inteligentes
- iv. Ambientes

3.- Solución de problemas mediante la búsqueda

- i. Agentes que resuelven problemas
- ii. Formulación de problemas
- iii. Problemas de ejemplo
- iv. Búsqueda de soluciones
- v. Estrategias de búsqueda
- vi. Cómo evitar estados repetidos
- vii. Búsqueda mediante la satisfacción de restricciones

Programa de estudios 2/4

4.- Métodos de búsqueda respaldados con información

- i. Búsqueda preferente por lo mejor
- ii. Funciones Heurísticas
- iii. Búsqueda limitada por la capacidad de la memoria
- iv. Algoritmos de mejoramiento iterativo

5.- Agentes que razonan de manera lógica

- i. Un agente basado en el conocimiento
- ii. El ambiente del mundo de wumpus
- iii. Representación, razonamiento y lógica
- iv. Lógica propositiva: un tipo de lógica muy sencillo
- v. Un agente para el mundo de wumpus

6.- Lógica de primer orden

- i. Sintaxis y semántica
- ii. Extensiones y variaciones en la notación
- iii. Uso de una lógica de primer orden
- iv. Agentes lógicos para el mundo de wumpus
- v. Un agente reflejo simple
- vi. Como representar los cambios en el mundo
- vii. Cómo deducir propiedades ocultas del mundo
- viii. Preferencia de una acción en vez de otra
- ix. Hacia un agente basado en metas

Programa de estudios 3/4

7.- La inferencia en la lógica de primer orden

- i. Reglas de inferencia relacionadas con cuantificadores
- ii. Ejemplo de una demostración
- iii. *Modus ponens: generalizado*
- iv. Encadenamiento hacia adelante y hacia atrás
- v. Completez
- vi. Resolución: un procedimiento completo de inferencia
- vii. Completez de una resolución

8.- Incertidumbre

- i. Cómo actuar ante la incertidumbre
- ii. Notación básica en probabilidad
- iii. Los axiomas de probabilidad
- iv. La regla de bayes y cómo se emplea
- v. ¿Cuál es el origen de las probabilidades?

9.- Sistemas de razonamiento probabilístico

- i. Representación del conocimiento en un dominio incierto
- ii. La inferencia en las redes de creencia
- iii. La inferencia en redes de creencia con múltiples conexiones
- iv. La ingeniería del conocimiento en el razonamiento incierto
- v. Otros enfoques para el razonamiento incierto

Programa de estudios 4/4

10.- Aprendizaje

- i. Un modelo general de agentes con capacidad de aprender
- ii. Aprendizaje inductivo
- iii. Aprendizaje con árboles de decisión
- iv. Empleo de la teoría de la información
- v. Aprendizaje de las descripciones lógicas generales
- vi. Por qué funciona el aprendizaje: teoría del aprendizaje computacional

11.- El aprendizaje en las redes neuronales y de creencia

- i. Cómo funciona el cerebro
- ii. Redes neuronales
- iii. Perceptrones
- iv. Redes de prealimentación con varios niveles
- v. Aplicaciones de las redes neuronales
- vi. Métodos bayesianos empleados en las redes de creencias para aprendizaje

12.- Fundamentos filosóficos

- i. Las grandes interrogantes
- ii. Los fundamentos del razonamiento y de la percepción
- iii. Acerca de la posibilidad de lograr un comportamiento inteligente
- iv. Intencionalidad y conciencia

13.- Inteligencia Artificial: presente y futuro

- i. ¿Hemos triunfado?
- ii. ¿Qué es lo que estamos tratando de hacer?
- iii. ¿Y si triunfamos?

Método de evaluación

- Examen semanal cada lunes que contará como tarea.
- Exámenes parciales en las fechas que marca el calendario oficial.
- **Libro de texto:**
Inteligencia Artificial: un Enfoque Moderno. Russel, Stuart j. Norvig, Peter
México: Prentice Hall Interamericana Pearson Educación , 2000.
- Capítulos:1,2,3,4,6,7,9,14,15,18,19,26,27.

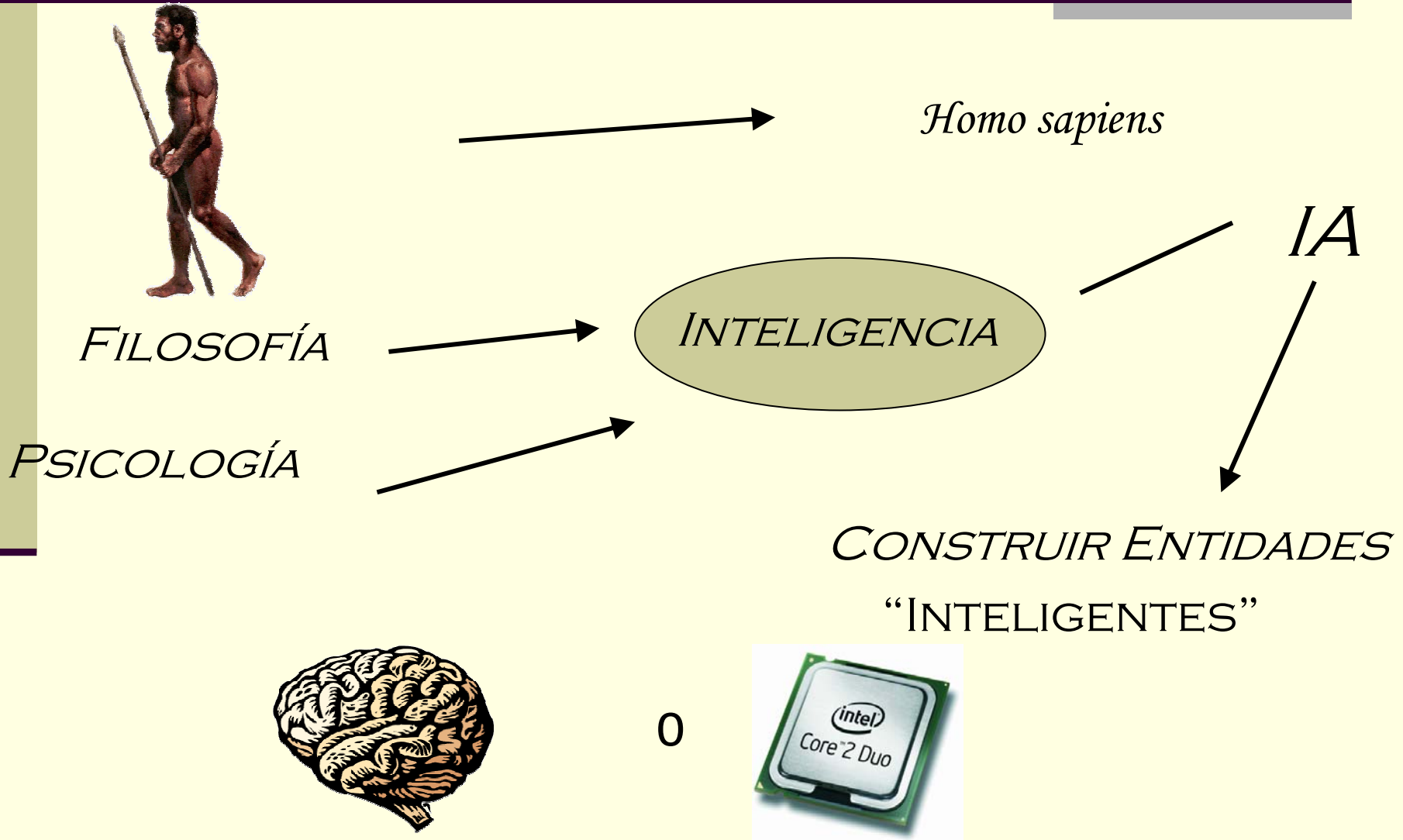
Capítulo 1

Introducción

Donde se intentará explicar por qué se considera a la inteligencia artificial un tema digno de estudio y donde se intentará definirla con exactitud; es esta tarea muy recomendable antes de emprender de lleno su estudio.

Introducción

¿Qué es la inteligencia artificial?



Visión general

- Agentes inteligentes.
- Investigación y juegos.
- Sistemas lógicos.
- Sistemas de planificación.

- Incertidumbre: probabilidad y teoría de la decisión.
- Aprendizaje.

- Lenguaje.
- Percepción.
- Robótica.
- Aspectos filosóficos.

¿Qué es la inteligencia artificial?

- La **meta** de la inteligencia artificial (IA) es lograr la **construcción y comprensión de entidades inteligentes**, que permitan resolver los mismos problemas que resuelven los humanos. La IA pretende alcanzar **metas científicas y metas de ingeniería**. En IA subyace una cuestión básica - **¿pueden pensar las máquinas?** - que ha constituido el interés de filósofos, científicos e ingenieros.
- Las **definiciones** de “Inteligencia Artificial” son muy diversas, por lo que usaremos la clasificación propuesta en (Russell y Norvig, 1995). En ella se analizan diversas definiciones de investigadores y se ordenan extrayendo variables comunes. Así resulta una clasificación según dos ejes coordinados: pensamiento y comportamiento.

¿Qué es la inteligencia artificial?

- **Inteligencia artificial:** Inteligencia exhibida por artefactos creados por humanos (es decir, artificial).

Pensamiento Humano	Pensamiento Racional
Actuación Humana	Actuación Racional

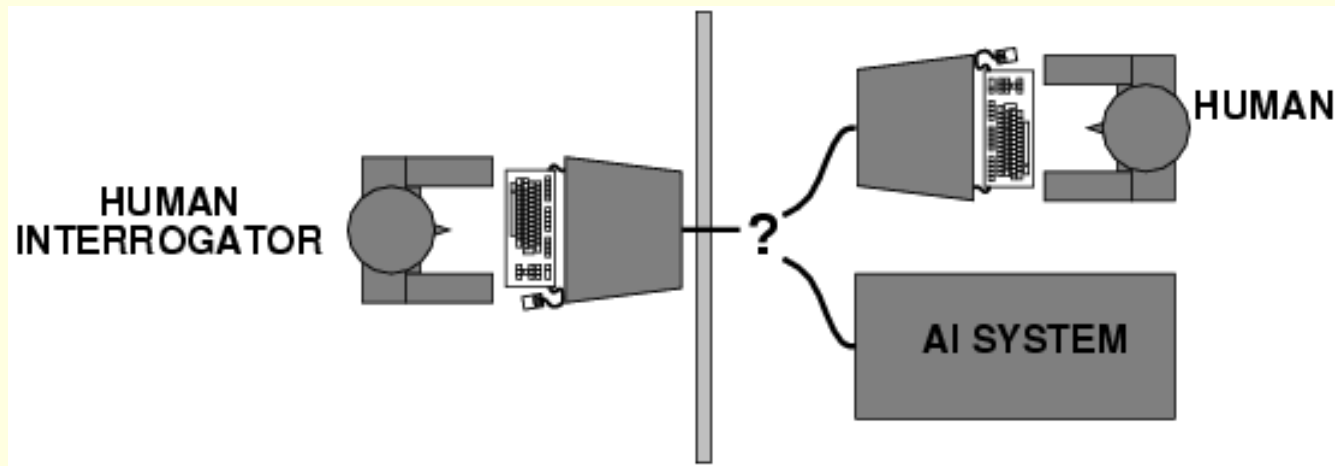


Existirá Inteligencia Artificial cuando no seamos capaces de distinguir entre un ser humano y un programa de computadora en una conversación a ciegas.

Pensamiento y Actuación	Pensamiento	“La interesante tarea de lograr que las computadoras piensen....máquinas con mente en su amplio sentido literal” (Haugeland, 1985) “La automatización de actividades que vinculamos con procesos de pensamiento humano, actividades tales como toma de decisiones, resolución de problemas, aprendizaje,...” (Bellman, 1978) “La IA está relacionada con el intento por desarrollar programas complejos de computador que sean capaces de realizar tareas cognitivas difíciles” (Eysenck, 1990)	“El estudio de las facultades mentales mediante el uso de modelos computacionales” (Charniack y McDermott, 1985) “El estudio de los cómputos que permiten percibir, razonar y actuar” (Winston, 1992)
	Actuación	“El arte de crear máquinas con capacidad de realizar tareas que normalmente requieren de inteligencia” (Nilsson, 1971) “El estudio de cómo lograr que las computadoras realicen tareas que, por el momento, los humanos hacen mejor” (Rich y Knight, 1991) “La IA es la parte de la ciencia de los computado-res, concerniente con el diseño de sistemas de computadores inteligentes, esto es, sistemas que exhiben las características que asociamos con la inteligencia en la conducta humana” (Barr y otros, 1982)	“Un campo de estudio que se enfoca a la explicación y emulación de la conducta inteligente en función de procesos computacionales” (Schalkoff, 1990) “La rama de la ciencia de la computación que se ocupa de la automatización de la conducta inteligente” (Luger y Stubblefield, 1997)
		Humana	Racional
		Eficiencia	

¿Qué es la inteligencia artificial?

Actuar como humano: *El enfoque de la prueba de Turing.*



¿Qué es la inteligencia artificial?

Actuar como humano: *El enfoque de la prueba de Turing.*

Se pretende dar una definición de **inteligencia** mediante la prueba de Turing (Alan Turing, 1950). Se sostuvo que una conducta inteligente consistía en la capacidad para lograr eficiencia a nivel humano como para engañar a un evaluador. Esta prueba consistía en que un humano evaluara a un computador por medio de un teletipo, considerándose prueba aprobada si el humano era incapaz de determinar si se trataba de un humano o de una máquina. En la actualidad, para que una computadora pueda pasar una prueba de este tipo, debería procesar un lenguaje natural, representar el conocimiento, razonar automáticamente y poseer autoaprendizaje. Para aprobar la **prueba** total de **Turing**, es necesario que la máquina disponga de vista (que le permita percibir objetos) y robótica (para desplazar estos objetos).

¿Qué es la inteligencia artificial?

Pensar como humano: *El modelo del enfoque cognoscitivo.*

Se basa en el **desarrollo de programas que emulen el modo de pensamiento y razonamiento humano**. Es necesario definir modelo de cómo piensa un humano si queremos afirmar que un determinado programa utiliza razonamiento humano. Hay dos formas de fijar estos modelos: por introspección o mediante experimentación psicológica. Esta aproximación es la utilizada en el campo interdisciplinario de la ciencia cognitiva. Un ejemplo de trabajo bajo este enfoque lo constituye el Solucionador General de Problemas ("**General Problem Solver**") (Newell y Simon, 1961). Este programa intenta la resolución de problemas mediante los mismos pasos de razonamiento que los humanos. Los problemas se resuelven mediante la descomposición de un problema en subproblemas y en la generación de nuevos problemas basados en la diferencia entre la solución actual y la deseada. Estos autores pretendían detectar los pasos de razonamiento que siguen los humanos.

¿Qué es la inteligencia artificial?

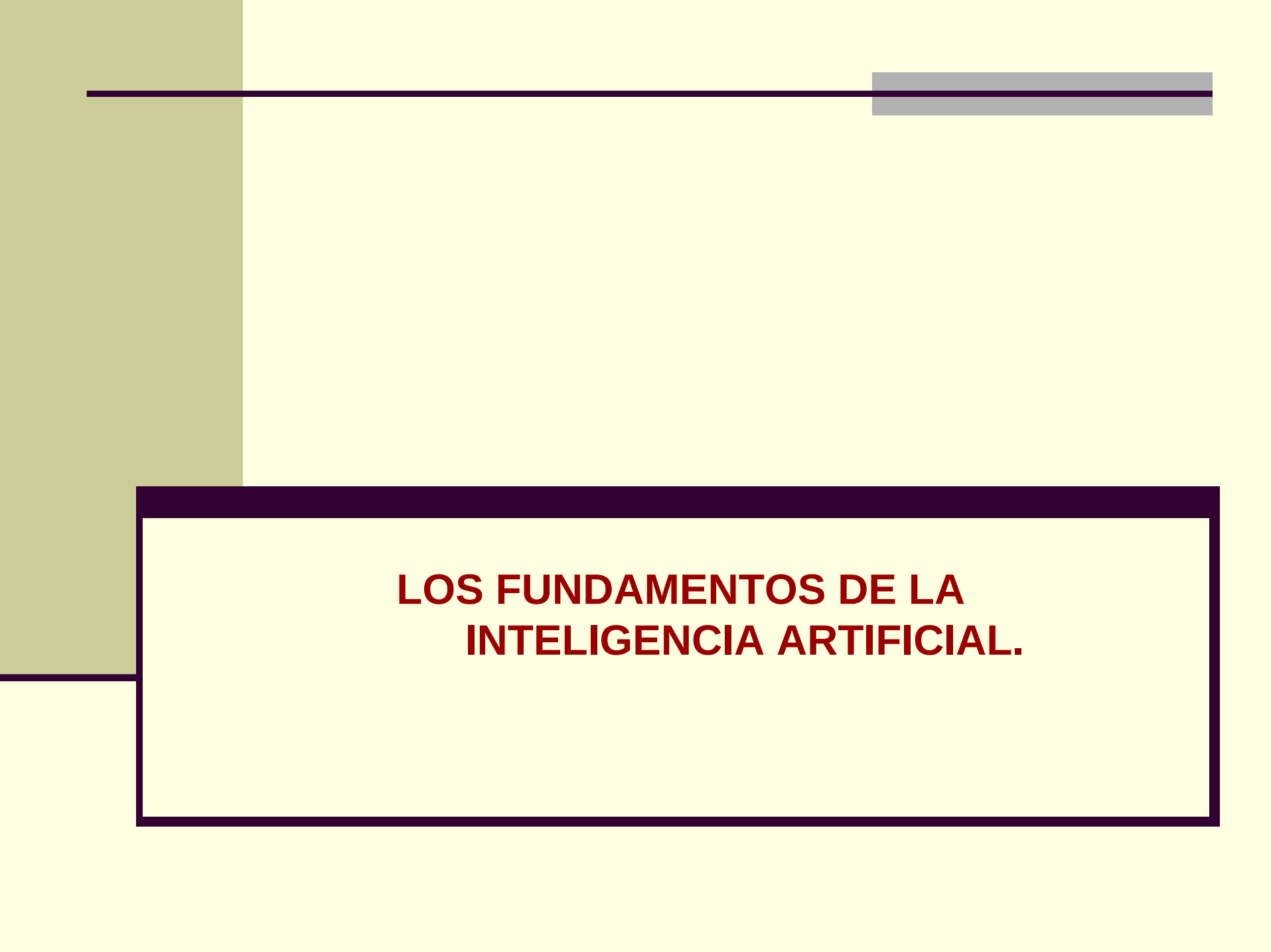
Pensar racionalmente: *El enfoque de las leyes del pensamiento.*

Es una **aproximación basada en las herramientas de la lógica formal, como mecanismo, según Aristóteles, de codificar “la manera correcta de pensar”**. Durante finales del siglo XIX y principios del XX se produjo un desarrollo de la lógica que permitió describir problemas en forma lógica para su resolución. Un ejemplo de trabajo bajo este enfoque lo constituye el **motor de inferencia** y la inmensa **base de conocimiento denominada CYC** (Lennat, 1984). La intención de CYC es construir un almacén del conocimiento utilizado habitualmente por los humanos (conocimiento de "sentido común"). Todo el conocimiento en CYC se encuentra representado de forma declarativa y el motor de inferencia utiliza el razonamiento deductivo. No se produce ningún intento de comprensión del funcionamiento de la capacidad humana de razonamiento en CYC. Para tomar una decisión, CYC genera todos los hechos consistentes con su base de conocimiento y los examina de forma lógica (utilizando la deducción) para alcanzar sus conclusiones.

¿Qué es la inteligencia artificial?

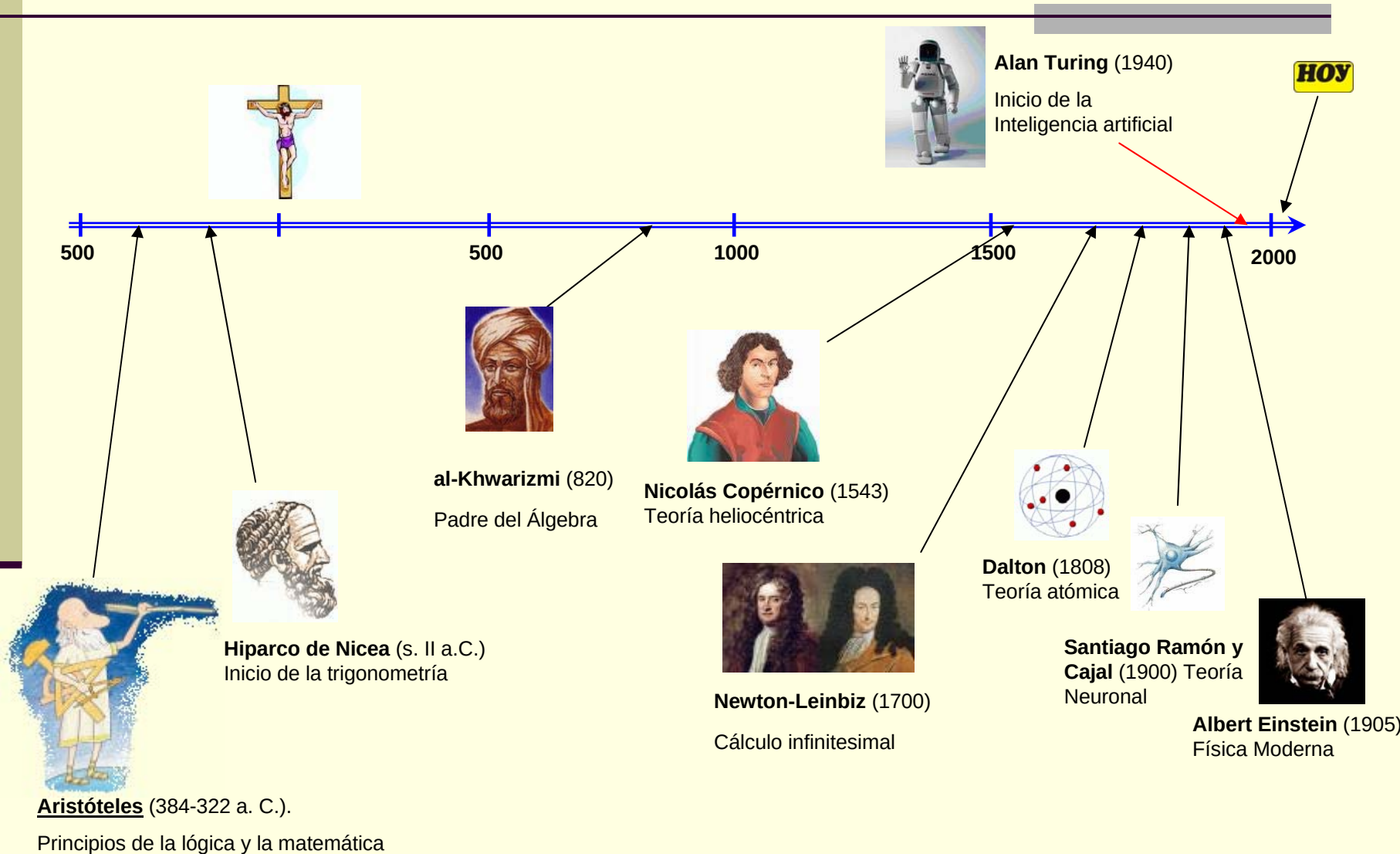
Actuar en forma racional: *El enfoque del agente racional.*

Se basa en **generar** acciones racionales o "**acciones correctas**" para lograr los objetivos deseados. Consiste en construir y estudiar agentes racionales. Un ejemplo de este tipo de enfoque lo constituye el ordenador **Deep Blue** que en 1997 derrotó al campeón del mundo de ajedrez Gary Kasparov. Los jugadores humanos de ajedrez no pueden examinar todos los movimientos en cada posición, por tanto se centran en explorar movimientos prometedores. En comparación Deep Blue era capaz de examinar 100 millones de posiciones de ajedrez por segundo (Newborn, 1996).



LOS FUNDAMENTOS DE LA INTELIGENCIA ARTIFICIAL.

¿Qué tan vieja es la IA?



Parentela de la IA

Esta sección presenta una breve historia de las disciplinas que han contribuido con ideas, puntos de vista, y técnicas al desarrollo de la Inteligencia Artificial.

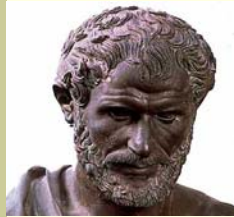
- **Filosofía** (desde el año 428 a.C hasta el presente).
- **Matemáticas** (aproximadamente desde el año 800 al presente).
- **Economía** (desde el año 1776 hasta el presente).
- **Neurociencia** (desde el año 1861 hasta el presente).
- **Psicología** (desde el año 1879 hasta el presente).
- **Ingeniería computacional** (desde el año 1940 hasta el presente).
- **Teoría de Control y cibernética** (desde el año 1948 hasta el presente).
- **Lingüística** (desde el año 1957 hasta el presente).

Filosofía

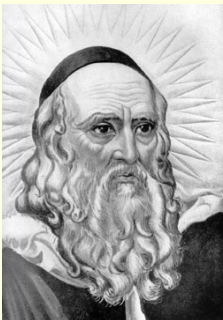
- ¿Se pueden utilizar reglas formales para extraer conclusiones válidas?
- ¿Cómo se genera la inteligencia mental a partir de un cerebro físico?
- ¿De dónde viene el conocimiento?
- ¿Cómo se pasa del conocimiento a la acción?

Síntesis: Lógica, métodos de razonamiento, la mente como un sistema físico, fundamentos del aprendizaje, lenguaje, racionalidad.

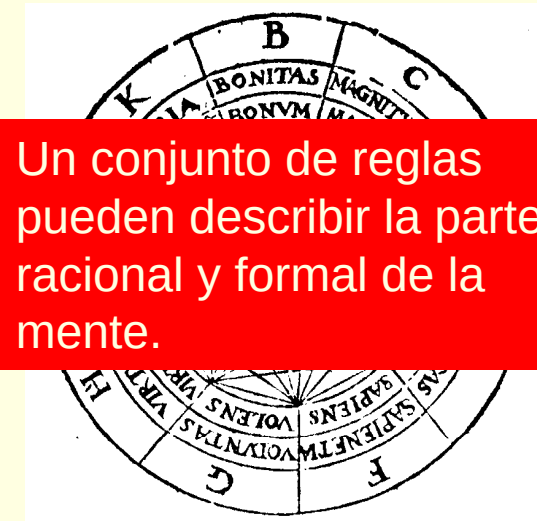
Filosofía



➤ **Aristóteles** (384-322 a. C.). Primero en formular un conjunto preciso de leyes que gobernaban la parte racional de la inteligencia. Desarrolló un sistema que permitía, a partir de premisas iniciales, extraer conclusiones mecánicamente razonando mediante silogismos.



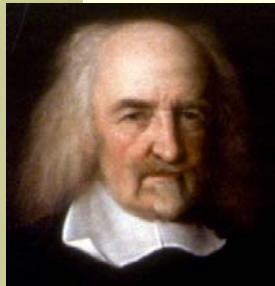
➤ **Ramón Lull** (d. 1315). Tuvo la idea de que el razonamiento útil se podía obtener por medios artificiales.



Un conjunto de reglas pueden describir la parte racional y formal de la mente.

Diagrama de un generador de conceptos por Lull.

Filosofía



➤ **Thomas Hobbes** (1588-1679). Propuso que el razonamiento era como la computación numérica, de forma que “*Nosotros sumamos y restamos silenciosamente en nuestros pensamientos*”.



➤ **René Descartes** (1596-1650). “*La mente como un sistema lógico*”

Proporciona la primera discusión clara sobre la distinción entre la mente y la materia y los problemas que surgen. A pesar de ser un gran defensor de la capacidad de razonamiento, Descartes fue también defensor del **dualismo** (existe una parte de la mente que está al margen de la naturaleza). El **materialismo**, al contrario que el dualismo, considera que las operaciones del cerebro realizadas de acuerdo a las leyes de la física constituyen la mente.

Filosofía

“La Mente como un sistema físico”

Dada una mente física que gestiona el conocimiento, el siguiente problema es establecer las fuentes de este conocimiento:

- Movimiento **empírico**: *“Nada existe en la mente que no haya pasado antes por los sentidos”*.
- **Positivismo lógico**: sostiene que todo el conocimiento se puede caracterizar mediante teorías lógicas relacionadas, en última instancia, con sentencias de observación que corresponden a estímulos sensoriales.
- **Teoría de la confirmación** (Carnap y Carl Hempel (1905-1997)): explica cómo el conocimiento se obtiene de la experiencia. Primera teoría en mostrar la mente como un proceso computacional.

Filosofía

“Relación entre Conocimiento y Acción”

- Aristóteles argumenta que las acciones se pueden justificar por la conexión lógica entre los objetos y el conocimiento de los efectos de las acciones.

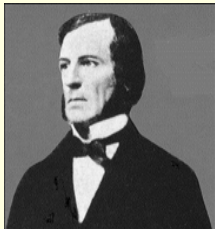
Matemáticas

- ¿Qué reglas formales son las adecuadas para obtener conclusiones válidas?
- ¿Qué se puede computar?
- ¿Cómo razonamos con información incierta?

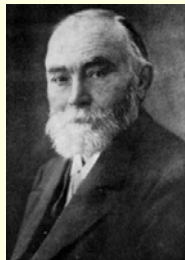
Síntesis: Representación formal y demostración de algoritmos, computación, (in)decibilidad, (in)tratabilidad, probabilidad.

Matemáticas

■ Los filósofos delimitaron las ideas más importantes de la IA, pero para pasar de ahí a una ciencia formal, es necesario contar con una formulación matemática en tres áreas fundamentales: **lógica**, **computación** y **probabilidad**.



■ El desarrollo de la Lógica Formal comienza con el trabajo de George Boole (185-1864) que definió la **lógica proposicional** o Booleana.



■ Gottlob Frege (1848-1925), en 1879, extendió la lógica de Boole para incluir objetos y relaciones, creando así la **lógica de primer orden**, que en la actualidad es el sistema más básico de representación de conocimiento.

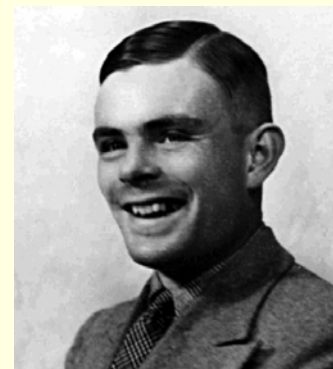
Matemáticas



- **Boole**, entre otros, presentó algoritmos para llevar a cabo deducciones lógicas.
- **David Hilbert** (1862-1943) presentó una lista de **23 problemas** que, acertadamente predijo, mantendrían ocupados a los matemáticos todo ese siglo. En el problema 23 se preguntaba si existe un algoritmo que permita determinar la validez de cualquier proposición lógica en la que aparezcan números naturales (*Entscheidungsproblem*, o problema de decisión).
- **Kurts Gödel** (1906-1978), mediante su **teorema de incompletitud**, demostró que en cualquier lenguaje que tuviera la capacidad para expresar las propiedades de los números naturales, existen aseveraciones verdaderas no decidibles en el sentido de que no es posible decidir su validez mediante ningún algoritmo.

Matemáticas

■ La afirmación anterior llevó a Alan Turing a tratar de caracterizar exactamente aquellas funciones que sí eran susceptibles de ser caracterizadas. Aunque no es posible dar una definición formal a la noción de cálculo o procedimiento efectivo, la tesis de Church-Turing, que afirma que la máquina de Turing es capaz de calcular cualquier función computable, es ampliamente aceptada ya que proporciona una definición suficiente.



Matemáticas

■ **Intratabilidad.** Se dice que un problema es intratable si el tiempo necesario para la resolución de casos particulares de dicho problema crece exponencialmente con el tamaño de dichos casos. La diferencia entre crecimiento polinomial y exponencial de la complejidad se consideró por primera vez a mediados de los años 60 (Cobham, 1964; Edmonds, 1965).

■ ¿Cómo se puede conocer un problema intratable? La teoría de la **NP-Complejidad** propuesta por primera vez por Steven Cook (1971) y Richard Karp (1972) propone un método. Demostraron la existencia de grandes clases de problemas de optimización combinatoria NP-completos. Toda clase de problemas a la que la clase de problemas NP completos se pueda reducir será seguramente intratable.

Matemáticas

■ Además de la lógica y el cálculo, la tercera gran aportación de las matemáticas a la IA es la teoría de la **probabilidad**.



- Gerolamo Gardano (1501-1576) fue el primero en proponer la idea de probabilidad, presentándola en términos de los resultados de juegos de apuesta.



- Thomas Bayes (1702-1761) propuso una regla para la actualización de probabilidades subjetivas a la luz de nuevas evidencias. El análisis Bayesiano conforma la base de las propuestas más modernas que abordan el razonamiento incierto en sistemas de IA.

Economía

- ¿Cómo se debe llevar a cabo el proceso de toma de decisiones para maximizar el rendimiento?
- ¿Cómo se deben llevar a cabo acciones cuando otros no colaboren?
- ¿Cómo se deben llevar a cabo acciones cuando los resultados se obtienen en un futuro lejano?

Síntesis: Teoría formal de las decisiones racionales.

Economía



■ La ciencia de la economía comenzó el 1776, cuando el filósofo escocés Adam Smith (1723-1790) publicó *An Inquiry into Nature and Causes of Wealth of Nations*.

■ La mayor parte de la gente cree que la economía sólo se trata de dinero, pero los economistas dicen que ellos realmente estudian cómo la gente toma decisiones para maximizar los beneficios esperados.



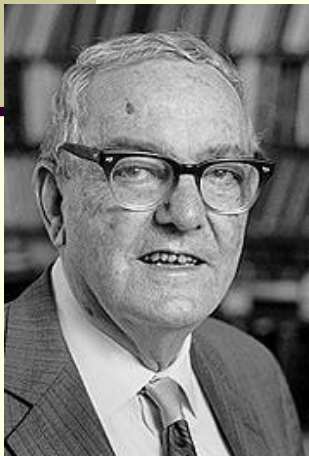
■ Leon Walras (1834-1910) formalizó el tratamiento matemático del beneficio deseado o **utilidad**, y fue posteriormente mejorado por Frank Ramsey (1932) y después por John Von Neumann y Oskar Morgenstern..

Economía

- La **teoría de la decisión**, que combina la teoría de la probabilidad con la utilidad, proporciona un marco completo y formal para la toma de decisiones realizadas bajo incertidumbre.
- **Teoría de juegos**. Cuando se trata de pequeñas economías, la situación se asemeja más a la de un juego.
- **Investigación operativa**. Este campo persigue cómo tomar decisiones racionales cuando los resultados de las acciones no son inmediatos y por el contrario, se obtienen de forma secuencial.
- El trabajo en la economía y la investigación operativa ha contribuido en gran medida a la noción de **agente racional** que veremos aquí.

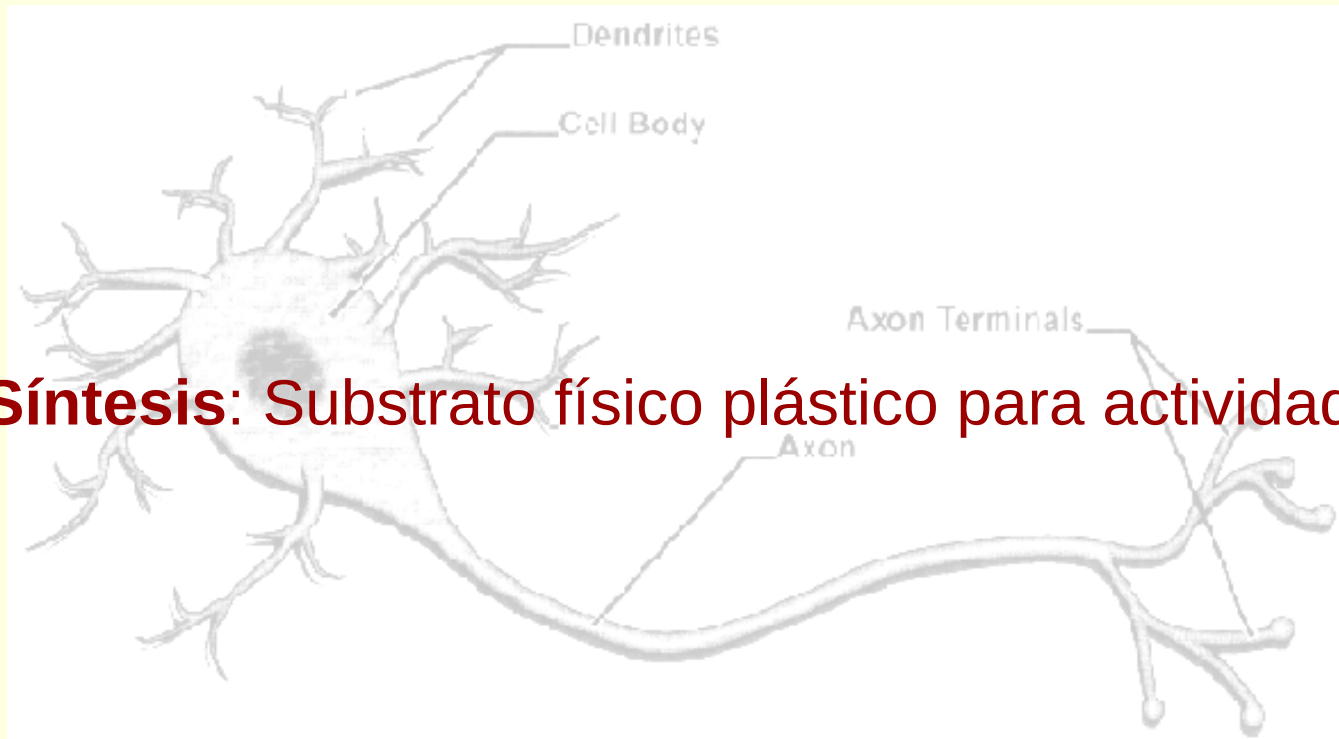
Economía

■ Herbert Simon (1916-2001), uno de los primeros investigadores de la IA, ganó el premio Nobel en Economía en 1978 por su trabajo en el que mostró que los **modelos basados en satisfacción** (que toman decisiones que son suficientemente buenas, en vez de hacer laboriosos cálculos para alcanzar decisiones óptimas) proporcionaban una descripción mejor del comportamiento humano real.



Neurociencia

¿Cómo procesa información el cerebro?



Síntesis: Substrato físico plástico para actividad mental.

Neurociencia

- ⌘ La **neurociencia** es el estudio del sistema neurológico, y en especial, del cerebro.
- ⌘ Uno de los grandes misterios de la ciencia es la forma en la que se genera el pensamiento en un cerebro.
- ⌘ La conclusión verdaderamente increíble de los estudios del cerebro es que *una colección de simples células puede llegar a generar razonamiento, acción y conciencia o, dicho en otras palabras, los cerebros generan las inteligencias* (Searle, 1992)

Psicología

¿Cómo piensan y actúan los animales?

Síntesis: Adaptación, fenómeno de percepción, técnicas experimentales.



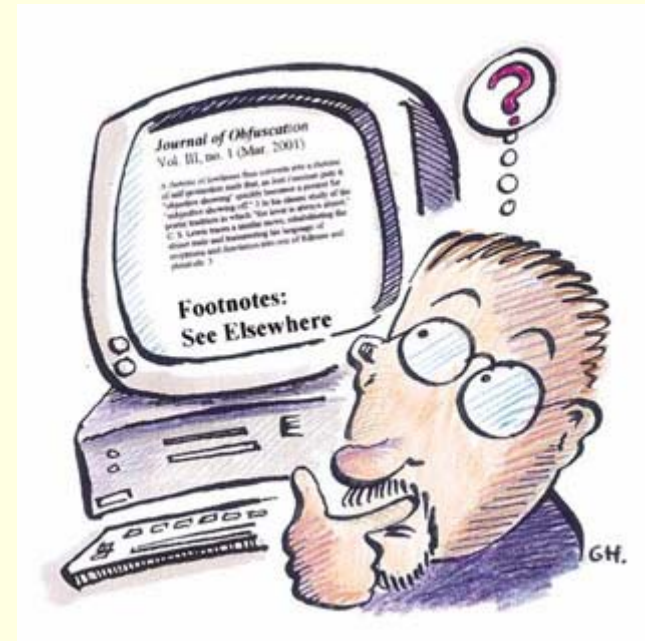
Introducción

¿Qué es la inteligencia artificial?

Los fundamentos de la IA

Ingeniería Computacional

¿Cómo se puede construir un computador eficiente?



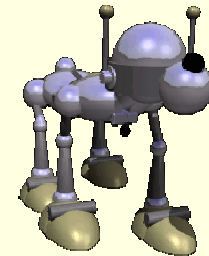
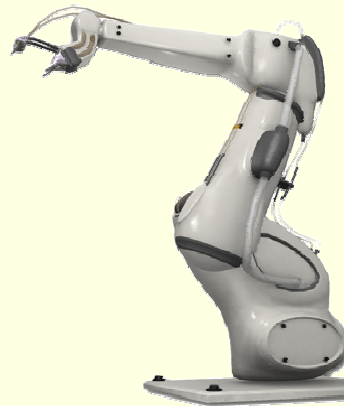
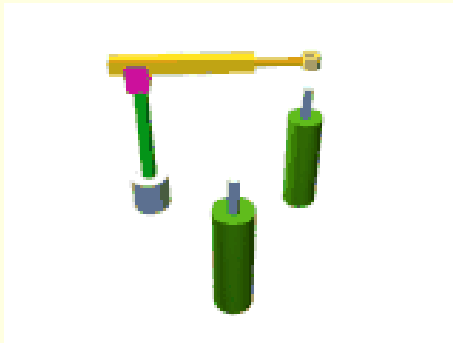
Ingeniería Computacional

- Para que la IA pueda llegar a ser una realidad, se necesitan dos cosas: **inteligencia** (**ciencia**) y un **artefacto** (**ingeniería**).
- El computador ha sido el artefacto elegido.



Teoría de Control y Cibernética

¿Cómo pueden los artefactos operar bajo su propio control?



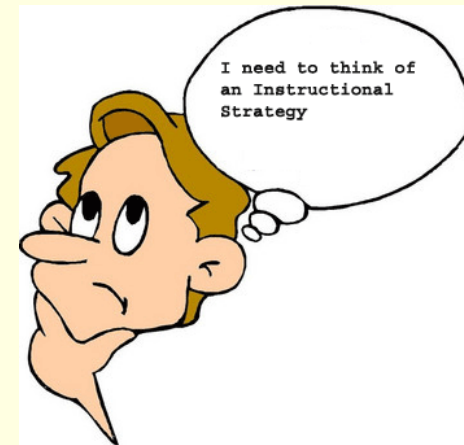
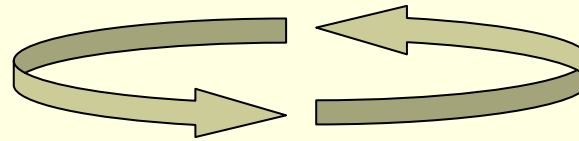
Síntesis: sistemas homeostáticos, estabilidad, diseño de agentes óptimos simples.

Teoría de Control y Cibernética

- La teoría de control moderna tiene como objetivo el diseño de sistemas que maximizan una función objetivo en el tiempo. Esto se asemeja ligeramente a nuestra visión de lo que es la IA.
- ¿Por qué son entonces la IA y la Teoría de Control dos campos distintos?
- La IA se fundó, en parte, para escapar de las limitaciones matemáticas de la teoría de control en los años 50. Las herramientas de inferencia lógica y computación permitieron a los investigadores de la IA afrontar problemas relacionados con el lenguaje, visión y planificación, que estaban completamente fuera del punto de mira de la teoría de control.

Lingüística

¿Cómo está relacionado el lenguaje con el pensamiento?



Síntesis: Representación del conocimiento, gramática.

Lingüística

- La lingüística moderna y la IA nacieron al mismo tiempo y maduraron juntas, solapándose en un campo híbrido llamado lingüística computacional o **procesamiento del lenguaje natural**.
- Gran parte de los primeros trabajos de la investigación en el área de la representación del conocimiento estaban vinculados al lenguaje y a la búsqueda de información en el campo del lenguaje, y su base eran las investigaciones realizadas durante décadas en el análisis filosófico del lenguaje.



HISTORIA DE LA INTELIGENCIA ARTIFICIAL

Resumen de la historia de IA

- 1943 McCulloch & Pitts: modelo de cerebro
- 1950 Turing's "Computing Machinery and Intelligence"
- 1956 Se adopta el término "Artificial Intelligence"
- 1952—69 Look, Ma, no hands!
- 1950s Primeros programs de IA, programa de ajedrez de Samuel, Newell & Simon's Logic Theorist, Gelernter's Geometry Engine
- 1965 algoritmo completo de razonamiento lógico de Robinson
- 1966—73 La IA descubre la complejidad computacional
Casi desaparece la investigación en Redes Neuronales
- 1969—79 Primeros desarrollos de los Sistemas Basados en Conocimiento
- 1980-- La IA se convierte en una industria
- 1986-- Las Redes Neuronales vuelven a la popularidad
- 1987-- La IA se convierte en una ciencia
- 1995-- Emergencia de los agentes inteligentes

Génesis de la IA (1943-1956)

- El primer trabajo reconocido en la IA fue el realizado por Warren McCulloch y Walter Pitts (1943), en el cual se proponía un modelo formado por neuronas artificiales. Estos autores apuntaron la posibilidad del aprendizaje de neuronas definidas para ese propósito. A los trabajos realizados por estos autores siguieron los programas de ajedrez para computadores escritos por Claude Shannon (1950) y Alan Turing (1953). De forma simultánea, Marvin Minsky y Dean Edmons construyeron la primera computadora de red neuronal en la universidad de Princeton.
- En el verano de 1956, John McCarthy, emplazado en Dartmouth College, junto con otros investigadores de la disciplina consiguen congregarse en Dartmouth durante dos meses a diez investigadores interesados en las redes neuronales y el estudio de la inteligencia. A pesar de que durante este tiempo no se produjeron avances notables, se obtuvo un consenso en el nombre asignado a este campo, **Inteligencia Artificial**, que fue propuesto por McCarthy.

Entusiasmo inicial: *High Hopes*

- Tras el nacimiento de la IA siguieron varios años de investigaciones exitosas teniendo en cuenta la disponibilidad y capacidad de computadoras de la época. El solucionador general de problemas (SGP) (**Newell y Simon, 1961**), que resolvía un número reducido de problemas de los humanos, constituyó el primer programa que incorporaba pensamiento de humano.
- En IBM, Nathaniel Rochester y sus colegas desarrollaron algunos de los primeros programas de IA. Se construyó el demostrador de teoremas de geometría (DTG) (Herbert Gelernter, **1959**), que demostraba teoremas utilizando axiomas explícitamente representados. A principios de 1952, Arthur Samuel desarrolló una serie de programas para el juego de damas, que permitió posteriormente jugar a nivel de torneo.
- En el año **1958 McCarthy** hizo tres contribuciones al campo de la IA: la **definición** del lenguaje de programación de alto nivel **Lisp**, la invención del **tiempo compartido**, y finalmente, el desarrollo del **escucha consejos**, programa considerado como el primer sistema de IA completo.

Entusiasmo inicial: *High Hopes*

- Durante años, Marvin Minsky y McCarthy, fueron definiendo el campo de la IA. Estos investigadores separaron sus enfoques pues McCarthy se centró más en la representación y razonamiento en lógica formal, mientras que Minsky estaba más interesado en lograr que los programas funcionaran, desarrollando un punto de vista antilógico.
- En 1963, McCarthy inició el laboratorio de IA en Stanford, donde se ponía énfasis en los métodos de propósito general para el razonamiento lógico. Como aplicaciones de la lógica pueden citarse los sistemas de planificación y respuesta de preguntas de Cordell Green (1969) y el proyecto de robótica de Shakey, que fue el primero en demostrar la total integración del razonamiento lógico y la actividad física.

Entusiasmo inicial: *High Hopes*

- Los **estudiantes** supervisados por **Minsky** se centraron en problemas de dominios limitados, conocidos como **micromundos**, cuya solución parecía requerir de la inteligencia. Se pueden citar el programa **SAINT** de James Slagle (1963), que resolvía problemas de integración de forma cerrada, el programa **ANALOGY** de Tom Evans (1968), que resolvía problemas de analogía simétrica, el programa de Bertram Raphael, **SIR** (Semantic Information Retrieval – recuperación de información semántica) (1968), y el programa **STUDENT** de Daniel Bobrow (1967), que podía resolver problemas de álgebra.
- Destacan las **actividades enfocadas a las redes neuronales** de McCulloch y Pitts. Los métodos de aprendizaje de Hebb fueron reforzados por Bernie Widrow, quien llamó adalinas a sus redes y por Frank Rosenblatt (1962) con sus perceptrones.

Una Dosis de Realidad (1966-1974) ...volando bajo

- Los investigadores de IA con frecuencia citaban el comentario hecho por **Herbert Simon** en **1957**:

“Sin afán de sorprenderlos y dejarlos atónitos, debo informarles lisa y llanamente de que actualmente en el mundo existen máquinas capaces de pensar, aprender y crear. Además, su capacidad para hacer lo anterior aumentará rápidamente hasta que – en un futuro previsible – la magnitud de problemas que tendrán capacidad de manejar irá a la par con la capacidad de la mente humana para hacer lo mismo”.

Una Dosis de Realidad (1966-1974) ...volando bajo

- ❖ En 1958, Simon hizo afirmaciones demasiado optimistas como que en 10 años sería posible realizar mediante la máquina la demostración de un nuevo e importante teorema matemático. El problema radicó en que aquellos métodos que parecían funcionar bien en ejemplos pequeños, fallaron rotundamente cuando se utilizaron problemas más variados y de mayor dificultad.
- ❖ El **primer tipo de obstáculo** se originó en los antiguos programas, que contaban con **poco o ningún conocimiento de la materia objeto de estudio**. Por ejemplo, el programa **ELIZA** de Wizenbaum (1965), mediante el que supuestamente era posible mantener una conversación sobre cualquier tema, de hecho recibía y manipulaba las oraciones que introducía un humano mediante teclado.

Una Dosis de Realidad (1966-1974) ...volando bajo



El **segundo obstáculo** fue el de la **intratabilidad** de muchos de los problemas que se estaban intentando resolver mediante la IA. Los primeros programas de IA funcionaron porque los **micromundos** con los que se trabajaba no contenían muchos objetos. Antes de que surgiera la teoría de la completitud-NP, se pensaba que para pasar de problemas sencillos a complejos bastaba con instalar hardware más veloz y memoria de mayor capacidad. El hecho de que, en principio, un programa sea capaz de encontrar una solución no implica que tal programa encierre todos los mecanismos necesarios para encontrar la solución en la práctica.



La incapacidad para manejar la *explosión combinatoria* fue una de las principales críticas que se le hicieron a la IA en el **informe Lighthill** (Lighthill, **1973**), en el cual se basó el **gobierno británico** para retirar la ayuda a los investigadores de IA.



El **tercer obstáculo** derivó de las **limitaciones** inherentes a las estructuras básicas que se utilizaban en la generación de la conducta inteligente (las redes neuronales casi desaparecen).

Sistemas basados en el conocimiento: ¿clave del poder? (1969-1979)

- Durante la primera década de la IA los problemas eran resueltos siguiendo un mecanismo de búsqueda de propósito general en el que se combinaban pasos de razonamiento elementales para encontrar soluciones completas. A estos procedimientos se les ha denominado métodos débiles, dado que cuentan con una información débil sobre el dominio.
- El programa **DENDRAL** (Buchanan et al, 1969), que trataba de inferir una estructura molecular a partir de la información proporcionada por un espectrómetro de masas, constituye uno de los primeros ejemplos de este enfoque. Este programa resultó inviable para moléculas de tamaño decente. La trascendencia de DENDRAL fue la de ser el **primer sistema de conocimiento intensivo que lograba funcionar**.

Sistemas basados en el conocimiento: ¿clave del poder? (1969-1979)

- Feigenbaum y otros en Stanford dieron inicio al **Proyecto de Programación Heurística (HPP)**, que trataba de determinar el grado con el que la nueva metodología de los **sistemas expertos** podía aplicarse a otras áreas de la actividad humana. Feigenbaum, Buchanan y el doctor Edward Shortliffe diseñaron el programa **MYCIN**, para el diagnóstico de las infecciones sanguíneas. Con el respaldo de 450 reglas, MYCIN era capaz de hacer diagnósticos tan buenos como los realizados por un experto. Tenía varias diferencias con el DENDRAL. En primer lugar, a diferencia de las reglas del DENDRAL, no se contaba con un modelo teórico desde el que se pudiera deducir las reglas del MYCIN. Era necesario obtenerlas a partir de amplias consultas con los expertos, quienes las habían obtenido de su experiencia en diversos casos. En segundo lugar, las reglas deberían reflejar la incertidumbre inherente al conocimiento médico. A este programa siguieron otros procedimientos para la realización de diagnósticos médicos.
- La importancia del conocimiento se demostró también en el área de la comprensión del lenguaje natural.

La IA se convierte en una industria (1980-1988)

- El **primer sistema experto comercial**, **R1**, que se utilizaba en la elaboración de pedidos de nuevos sistemas de computación, inició sus actividades en Digital Equipment Corporation (McDermott, 1982). En 1986 representaba para la compañía ahorros de 40 millones de dólares al año. En 1988, el grupo de IA de la DEC había distribuido ya 40 sistemas expertos, y esperaba tener más. Casi todas las compañías importantes de USA contaban con su propio grupo de IA, en el que se utilizaba o investigaba la tecnología de los sistemas expertos.
- En **1981** los **japoneses** anunciaron el **proyecto de la quinta generación**, un plan de 10 años para construir computadoras inteligentes en las que se corriese Prolog, de manera idéntica a como las computadoras comunes corren código máquina. Este proyecto suscitó el interés en IA, y aprovechando el temor a un posible dominio de los japoneses, investigadores y empresas obtuvieron apoyo para una inversión semejante en Estados Unidos

Acontecimientos recientes (1987 al presente)

- En la actualidad es más común **construir sobre teorías ya existentes** que proponer teorías totalmente novedosas, tomar como base rigurosos teoremas o sólidas evidencias experimentales más que la intuición, y **demostrar la utilidad de las aplicaciones en el mundo real** más que crear ejemplos de juguete.
- La aparición de Probabilistic Reasoning and Intelligent Systems, (Judea Pearl, 1988) señaló la bienvenida al uso de la probabilidad y de la teoría de decisiones en IA. La invención del formalismo de red de creencia obedece a la necesidad de poder razonar eficientemente ante una combinación de evidencias inciertas. Este enfoque supera a los sistemas de razonamiento probabilístico de las décadas de los años 60 y 70, y actualmente prevalece en la investigación que se hace en IA sobre razonamiento incierto y sistemas expertos.

Acontecimientos recientes (1987 al presente)

- También se han producido **revoluciones en robótica, visión por computador, autoaprendizaje** de máquina (incluidas las redes neuronales) y **representación del conocimiento**. Una comprensión de los problemas y de su complejidad, aunada a una mayor capacidad de manejo matemático, han favorecido la factibilidad de agendas de investigación y de sólidos métodos.

Los últimos avances

- El gran maestro internacional **Arnold Denker** estudia las piezas que tiene ante sí en el tablero. Reconoce que no hay esperanza alguna, perderá el juego. Su oponente, **HITECH**, se convierte así en el **primer programa de computadora que logra derrotar a un gran maestro de la talla de Denker en un juego de ajedrez** (Berliner, 1989).
- '**Quiero ir de Boston a San Francisco**', dice el viajero a través de un micrófono. '¿Qué datos tiene para el viaje?', es la respuesta. El viajero explica que desea salir el 20 de octubre, en vuelo directo, con la tarifa más barata y regresar el domingo siguiente. **PEGASUS**, un **programa para comprensión del lenguaje** se hace cargo de todo y el resultado es una reserva confirmada que le permite al viajero ahorrarse \$894 del precio normal del boleto. Dado que de diez palabras, este intérprete entiende mal una, es capaz de recuperarse de estos errores gracias a su conocimiento sobre la forma en la que se estructura un diálogo (Zue et al, 1994).

Los últimos avances

- Uno de los más reconocidos expertos en **patología de nodos linfáticos** plantea un caso extremadamente difícil al sistema experto y evalúa la respuesta dada por éste. La respuesta del sistema le causa risa. Ligeramente preocupados, los creadores del sistema le sugieren que pida a la computadora una explicación del diagnóstico dado. La máquina señala los principales factores que llevaron a esa decisión y explica la sutil interacción de varios de los síntomas en este caso. El experto, después de un tiempo, admite su error (Heckerman, 1991).

Los últimos avances

- Desde una cámara suspendida en un semáforo de un cruce, el monitor de tráfico observa el panorama. Si alguien observara el monitor principal leería: "Citroen AX está dando la vuelta a la Plaza de la Concordia en los Campos Alíseos"; "Camión carguero grande se ha detenido en la Plaza de la Concordia", y así durante toda la noche. Ocasionalmente, leerían: "Grave incidente en la Plaza de la Concordia, una camioneta que avanzaba rápidamente se ha estrellado contra una motocicleta", y se haría una **llamada automática a los servicios de emergencia** (King et al, 1993; Koller et al, 1994).

Se han citado sólo algunos ejemplos del tipo de sistemas de inteligencia artificial que existen en la actualidad. No se trata de magia ni de ciencia ficción: es **ciencia, ingeniería y matemáticas**.

¿De qué es capaz la IA actualmente?

Planificación Autónoma

El programa de la NASA **Agente Remoto** se convirtió en el primer programa de planificación autónoma a bordo que controlaba la planificación de las operaciones de una nave espacial desde abordó (Jonsson et al, 2000).

Este agente generaba planes a partir de objetivos especiales especificados desde tierra.

¿De qué es capaz la IA actualmente?

Juegos

Deep Blue de IBM fue el primer sistema que derrotó a un campeón del mundo en una partida de ajedrez, cuando superó a Gary Kasparov en una partida de exhibición en 1997.

El valor de las acciones de IBM se incrementó en 18 millones de dólares.

¿De qué es capaz la IA actualmente?

Control Automático

El sistema de visión por computador **ALVINN** fue entrenado para dirigir un coche de forma que siguiese una línea. Se instaló en una furgoneta controlada por computador y se utilizó para dirigir al vehículo por EEUU.

Durante 2850 millas controló la dirección del vehículo en el 98% del trayecto. Una persona lo sustituyó en el 2% restante, principalmente en vías de salida.

¿De qué es capaz la IA actualmente?

Diagnosis

Los programas de diagnóstico médico basados en el análisis probabilista han llegado a alcanzar niveles similares a los de los médicos expertos en algunas áreas de la medicina.

Heckerman (1991) describe un caso en el que un destacado experto en la patología de los nodos linfáticos se mofó del diagnóstico generado por un programa en un caso especialmente difícil. Se le preguntó al computador cómo había generado el diagnóstico, éste indicó los factores más importantes en los que hacía basado su decisión y explicó la ligera interacción existente entre varios de los síntomas en este caso. Finalmente, el experto aceptó el diagnóstico del programa.

¿De qué es capaz la IA actualmente?

Planificación Logística

Durante la crisis del Golfo Pérsico de 1991, las fuerzas de los Estados Unidos desarrollaron la herramienta *Dynamic Analysis and Replanning Tool* (**DART**) para automatizar la planificación y organización logística del transporte, lo que incluía hasta 50000 vehículos, carga y personal a la vez.

¿De qué es capaz la IA actualmente?

Robótica

Muchos cirujanos utilizan hoy en día asistentes robot en operaciones de microcirugía. **HipNav** (1996) es un sistema que utiliza técnicas de visión por computador para crear un modelo tridimensional de la anatomía interna del paciente y después utiliza un control robotizado para guiar el implante de prótesis de cadera.

¿De qué es capaz la IA actualmente?

Procesamiento de Lenguaje y resolución de problemas

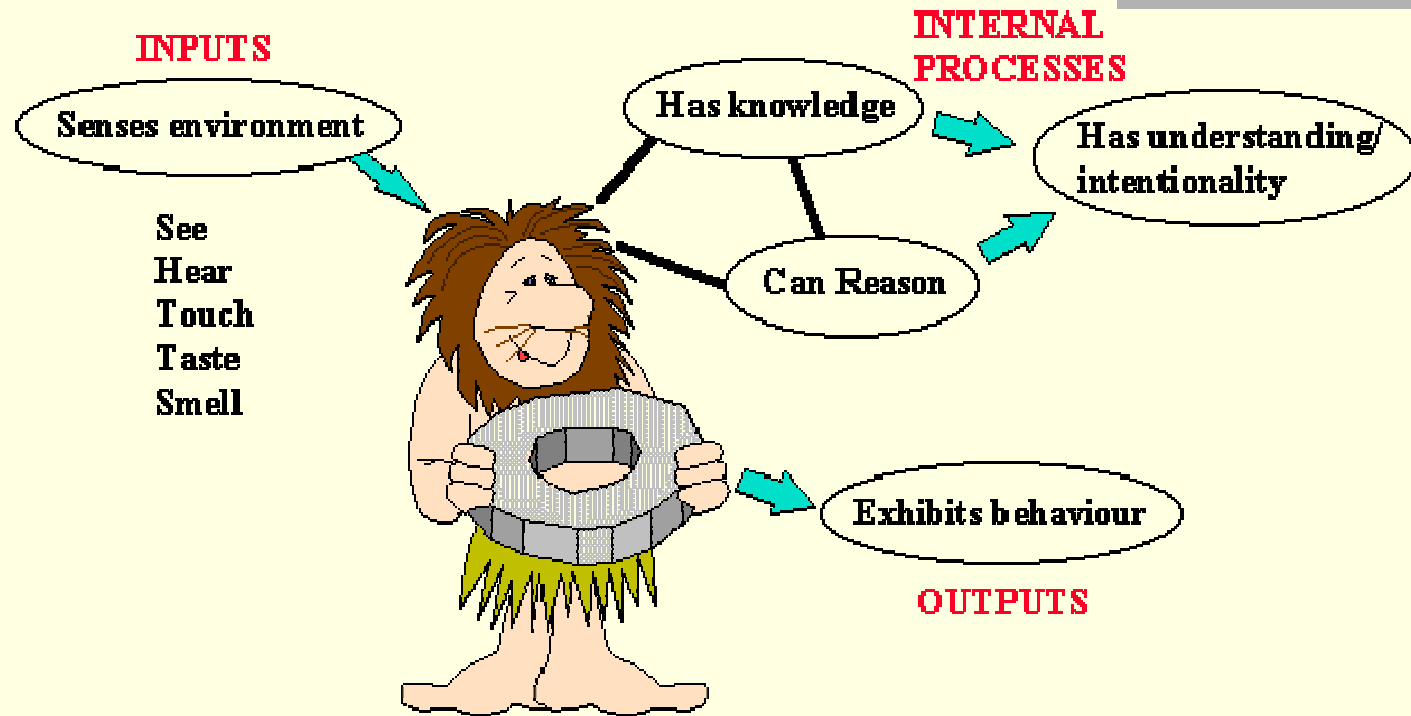
PROVER B (1999) es un programa que resuelve crucigramas mejor que la mayoría de los humanos, utilizando restricciones en programas de relleno de palabras, una gran base de datos de crucigramas, varias fuentes de información como diccionarios y bases de datos online, que incluyen la lista de películas y actores que intervienen en ellas entre otras cosas.

Capítulo 2

Agentes Inteligentes

Donde se explicará que hace un agente inteligente, como se relaciona éste con su medio, cómo se le evalúa y cómo se podría construir uno.

Agentes Inteligentes



- Se define **agente** como todo aquello que **percibe** su ambiente mediante **sensores** y que **responde o actúa** en tal ambiente por medio de **efectores**.

¿Cómo debe proceder un agente?

Un **agente racional** es aquel que hace lo correcto, es decir que obtiene el mejor desempeño. Habrá que decidir *cómo* y *cuándo* evaluar ese buen desempeño del agente. El término **medición del desempeño** se aplica al *cómo*: es el criterio que se utiliza para decidir qué tan exitoso ha sido un agente. El *cuándo* evaluar el desempeño es también importante.

■ Es conveniente hacer una distinción entre **racionalidad** y **omnisciencia**. Un agente omnisciente es aquel que conoce el resultado real de sus acciones, y su conducta es congruente con ello. Sin embargo, *en la realidad no existe la omnisciencia*. No se puede responsabilizar a una agente por no haber tomado en cuenta algo que no podía percibir, o por no emprender una acción de la que es incapaz. *Si se especifica que un agente inteligente siempre debe hacer lo que es realmente correcto, será imposible diseñar un agente.*

Agentes Inteligentes

En resumen, el **carácter de racionalidad** de lo que se hace en un momento dado dependerá de cuatro factores

- De la **medida** con la que se evalúa el grado de éxito logrado.
- De todo lo que hasta ese momento haya percibido el agente. A esta historia perceptual completa se le denomina la **secuencia de percepciones**.
- Del **conocimiento** que posea el agente acerca del medio.
- De las **acciones** que el agente puede emprender.

Agentes Inteligentes

- En todos los casos de posibles secuencias de percepciones, un **agente racional ideal** deberá emprender todas aquellas acciones que favorezcan obtener el máximo de su medida de rendimiento, basándose en las evidencias aportadas por la secuencia de percepciones y en todo el conocimiento incorporado en tal agente.

El mapeo ideal de las secuencias de percepciones

❖ Una vez que nos hemos dado cuenta de que el comportamiento de un agente depende exclusivamente de la secuencia de sus percepciones en un momento dado, sabemos que es posible caracterizar cualquier agente en particular elaborando una tabla de las acciones que éste emprende como respuesta a cualquier secuencia de percepciones posible. Esta lista se denomina **mapeo** de secuencias de percepciones para acciones. *El especificar qué tipo de acción deberá emprender un agente como respuesta a una determinada secuencia de percepciones constituye el diseño de un agente ideal.*

El mapeo ideal de las secuencias de percepciones

- Para **realizar el mapeo** entre secuencias de percepciones y acciones no es necesario enumerar la lista anterior de forma exhaustiva. Bastaría con definir una especificación del mapeo. Por ejemplo, la utilización del método de Newton para obtener la raíz cuadrada de un número positivo.

Autonomía

- En la definición anterior de agente racional ideal hay que prestar atención a la parte de "**conocimiento integrado**". Si las acciones emprendidas por el agente están basadas exclusivamente en un conocimiento integrado y no en percepciones, se dice que el agente **no tiene autonomía**.

Autonomía

La conducta de un agente se basa tanto en su propia experiencia como en el conocimiento integrado que sirve para construir el agente para un ambiente específico en el que va a operar. *Un **sistema** será **autónomo** (no bajo el control inmediato de un humano) en la medida en que su conducta está definida por su propia experiencia.* Si el agente cuenta con poca o ninguna experiencia, tendrá que comportarse de manera más o menos aleatoria a menos que el diseñador le de algún tipo de ayuda.

Ejemplos de diversos tipos de agentes y sus correspondientes descripciones PAMA (Percepciones, Acciones, Metas y ambiente)

<i>Tipo de agente</i>	<i>Percepciones</i>	<i>Acciones</i>	<i>Metas</i>	<i>Ambiente</i>
Sistema para diagnósticos médicos	Síntomas, evidencias y respuestas del paciente	Preguntas, pruebas, tratamientos	Paciente saludable, reducción al mínimo de costos	Paciente, hospital
Sistema para el análisis de imágenes de satélite	Píxeles de intensidad y colores diversos	Imprimir una clasificación de escena	Clasificación correcta	Imágenes enviadas desde un satélite en órbita
Robot clasificador de partes	Píxeles de intensidad variable	Recoger partes y clasificarlas poniéndolas en botes	Poner las partes en el bote que les corresponda	Banda transportadora sobre la que se encuentran las partes
Controlador de una refinería	Lecturas de temperatura y presión	Abrir y cerrar válvulas; ajustar la temperatura	Lograr pureza, rendimiento y seguridad máximos	Refinería
Asesor interactivo de inglés	Palabras escritas a máquina	Ejercicios impresos, sugerencias y correcciones	Que el estudiante obtenga la máxima calificación en una prueba	Grupo de estudiantes

Estructura de los agentes inteligentes

Hasta el momento se ha hablado de los agentes describiendo su conducta, la acción que se realiza después de una secuencia de percepciones dada.

Ahora, se trata de centrarse en el núcleo del problema y hablar sobre cómo trabajan internamente.

El trabajo de la IA es diseñar el **programa del agente**, que implemente la función del agente que proyecta las percepciones en las acciones.

Estructura de los agentes inteligentes

✧ Trabajo de la IA: diseñar el **programa de agente**, es decir, una función que permita implantar el mapeo del agente para pasar de percepciones a acciones. Este programa se ejecutará en algún dispositivo de cómputo, que se llamará **arquitectura**. La relación entre agentes, arquitectura y programas podría resumirse de la siguiente manera:

agente = arquitectura + programa

✧ Antes de proceder al diseño de un programa de agente, es necesario contar con una idea bastante precisa de las posibles percepciones y acciones que intervendrán, qué metas o medidas de desempeño se supone lleve a cabo el agente, así como del tipo de ambiente en que tal agente operará.

Programas de agentes

Hay que tener en cuenta la 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. (Si las acciones del agente dependen de la secuencia completa de percepciones, el agente tendría que recordar las percepciones).

La manera más sencilla de hacer un programa de agente es mediante una tabla de consulta. Se guarda en memoria la totalidad de la secuencia de percepciones y se utiliza ésta para hacer localizaciones en la tabla, en la que se encuentran todas las acciones que corresponden a todas las posibles secuencias de percepción. Este enfoque está destinado al fracaso por varias razones:

Programas de agentes

- ✧ Para elaborar la tabla de algo tan sencillo como un agente cuya única tarea es jugar ajedrez, se necesitarían 35^{100} entradas.
- ✧ La elaboración de la tabla tomaría demasiado tiempo al diseñador.
- ✧ El agente carece totalmente de autonomía, puesto que el cálculo de lo que se considera mejor acción está ya integrado. Si el ambiente se modificara en algún aspecto no previsto, el agente se vería completamente perdido.
- ✧ Incluso se dotara al agente de un mecanismo de aprendizaje, con el fin de que tuviera cierto grado de autonomía, necesitaría una eternidad para aprender el valor correcto de todas las entradas de la tabla.

Programas de agentes

A pesar de lo expuesto, el agente conducido mediante tabla sí es capaz de hacer lo que deseamos: implanta el deseado mapeo del agente. No basta con afirmar: "No puede ser inteligente"; lo importante es comprender por qué un agente que razona puede tener un mejor desempeño al superar las cuatro desventajas mencionadas anteriormente.

Programas de agentes ejemplo

Consideraremos la tarea del diseño de un conductor de taxi automatizado, aunque tal sistema está fuera del alcance de la tecnología actual. Habrá que considerar las PAMA que corresponden a un **taxi**.

- **Percepciones**: Cámaras, velocímetro, sistema satélite global de ubicación, sonar, micrófono.
- **Acciones**: Manejo del volante, acelerar, frenar, hablar con el pasajero.
- **Metas**: Un viaje seguro, rápido, sin infracciones, cómodo, obtención máxima de ganancias.
- **Ambiente**: Caminos, tráfico, peatones, clientes.

Programas de agentes ejemplo

El taxi necesita saber dónde se encuentra, quién más circula por su vía y a qué velocidad circula. Esta información se obtiene de las **percepciones**. Las **acciones** que puede producir este conductor de taxi son más o menos las mismas que las de un conductor humano. ¿Cuál sería la **medida de desempeño** que desearía satisfacer el conductor automatizado? Entre las cualidades deseables están las de llegar al destino correcto, reducir al mínimo el consumo de combustible, desgaste del vehículo, tiempo de recorrido, su costo, violaciones al reglamento de tráfico y las molestias ocasionadas a otros conductores; ofrecer el máximo de seguridad y comodidad al pasajero y producir el máximo de ganancias. Dado que el logro de algunos de estos objetivos está en conflicto con el de otros, habrá que establecer concesiones. Por último, habrá que definir qué tipo de **ambiente** de conducción enfrentará el taxi. Cuanto más restringido sea el ambiente, más fácil será el problema del diseño.

Programas de agentes

tipos de agentes

En lo que resta de esta lección se presentan los cuatro tipos básicos de programas para agentes que encarnan los principios que subyacen en casi todos los sistemas inteligentes.

- **Agentes reactivos simples.**
- **Agentes reactivos basados en modelos.**
- **Agentes basados en objetivos o metas.**
- **Agentes basados en utilidad.**

Agentes reactivos simple

- Éste es el tipo de agente más sencillo.
- Seleccionan las acciones sobre la base de las percepciones actuales, ignorando el resto de las percepciones históricas.
 - Por ejemplo, el agente aspiradora, cuya función de agente aparece en la siguiente transparencia, es un agente reactivo simple porque toma sus decisiones sólo con base en la localización actual y si ésta está sucia.

Agentes reactivos simple

Secuencia de percepciones	Acción
[A,Limpio]	<i>Derecha</i>
[A,Sucio]	<i>Aspirar</i>
[B,Limpio]	<i>Izquierda</i>
[B,Sucio]	<i>Aspirar</i>
[A,Limpio], [A,Limpio]	<i>Derecha</i>
[A,Limpio], [A,Sucio]	<i>Aspirar</i>
--	--
--	--
--	--
[A,Limpio], [A,Limpio], [A,Limpio]	<i>Derecha</i>
[A,Limpio], [A,Limpio], [A,Sucio]	<i>Aspirar</i>
--	--
--	--
--	--

Agentes reactivos simple

■ Hay que tener en cuenta que el programa para el agente aspiradora es muy pequeño comparado con su tabla correspondiente.

- La reducción más clara se obtiene al ignorar la historia de la percepción.
- Otra reducción se basa en el hecho de que cuando la cuadrícula actual está sucia, la acción no depende de la localización.

■ **Regla de condición-acción:** se lleva a cabo algún tipo de procesamiento sobre las señales visuales para establecer la condición.

- Imagínese que es el conductor del taxi automático. Si el coche que circula delante frena (*condición*), y las luces de freno se encienden, entonces lo advertiría y comenzaría a frenar (*acción*).

Agentes reactivos simple

Programa para el agente de la aspiradora

función AGENTE-ASPIRADORA-REACTIVO([localización, estado]) **devuelve** acción

si *estado* = *Sucio* **entonces devolver** *Aspirar*

de otra forma, si *localización* = *A* **entonces devolver** *Derecha*

de otra forma, si *localización* = *B* **entonces devolver** *Izquierda*

Un Agente Reactivo Simple

función AGENTE-REACTIVO-SIMPLE (*percepción*) **devuelve** una acción

estático: *reglas*, un conjunto de reglas condición-acción

estado <- INTERPRETAR-ENTRADA (*percepción*)

regla <- REGLA-COINCIDENCIA (*estado*, *reglas*)

acción <- REGLA-ACCIÓN[*regla*]

devolver *acción*

Agentes reactivos simple

Los agentes reactivos simples tienen la gran propiedad de ser simples, pero poseen una inteligencia muy limitada.

El agente del pseudocódigo anterior funcionará sólo si se puede tomar la decisión correcta sobre la base de la percepción actual, lo cual es posible sólo si el entorno es totalmente observable.

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.*
- 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.

Agentes reactivos basados en modelos

La actualización de la información de estado interno según pasa el tiempo requiere codificar dos tipos de conocimiento en el programa de agente.

- Información de cómo evoluciona el mundo independientemente del agente (por ejemplo, que un coche que está adelantando estará más cerca, detrás, que en un momento inmediatamente anterior.)
- Se necesita más información sobre cómo afectan al mundo las acciones del agente (por ejemplo, que cuando el agente gire hacia la derecha, el coche gira hacia la derecha o que después de conducir durante 5 minutos hacia el norte en la autopista se avanzan 5 millas hacia el norte a partir del punto en el que se estaba 5 minutos antes).

Agentes reactivos basados en modelos

- Este conocimiento acerca de **cómo funciona el mundo**, se denomina **modelo del mundo**.
- Un agente que utilice este modelo es un **agente basado en modelos**.

función AGENTE-REACTIVO-CON-ESTADO (*percepción*) **devuelve** una acción
estático: *estado*, una descripción actual del estado 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
```

Agentes reactivos basados en modelos

- El conocimiento sobre el estado actual del mundo no es siempre suficiente para decidir qué hacer.

Por ejemplo, en un cruce de carreteras, el taxista puede girar a la izquierda, girar a la derecha o seguir hacia delante. La decisión correcta depende de a dónde quiera ir el taxi.

Además de la descripción del estado actual, el agente necesita algún tipo de información sobre su meta que describa las situaciones que son deseables, por ejemplo, llegar al destino propuesto por el pasajero.

Agentes reactivos basados en objetivos

- En algunas ocasiones, la selección de acciones basadas en objetivos es directa; cuando alcanzar los objetivos es el resultado inmediato de una acción individual.
- En otras ocasiones, puede ser más complicado, cuando el agente tiene que considerar secuencias complejas para encontrar el camino que le permite alcanzar el objetivo. Búsqueda y planificación son los subcampos de la IA centrado en encontrar secuencias de acciones que permitan a los agentes alcanzar sus metas.

Agentes reactivos basados en utilidad

- Las metas por sí solas no son suficientes para generar comportamiento de gran calidad en la mayoría de los entornos.
 - Por ejemplo, hay muchas secuencias de acciones que llevarán al taxi a su destino (y por tanto a alcanzar su meta), pero algunas son más rápidas, más seguras, más fiables o más baratas.

*Una **función de utilidad** proyecta un estado (o una secuencia de estados) en un número real, que representa un nivel de felicidad. Permite tomar decisiones racionales:*

- Cuando haya objetivos conflictivos (ej., velocidad y seguridad).
- Cuando haya varios objetivos por los que se pueda guiar el agente y ninguno de ellos se pueda alcanzar con certeza.

Propiedades de los ambientes

Las diferencias básicas entre los distintos ambientes son las siguientes:

- ✿ **Accesibles y no accesibles.**
- ✿ **Deterministas y no deterministas.**
- ✿ **Episódicos y no episódicos.**
- ✿ **Estáticos y Dinámicos.**
- ✿ **Discretos y Continuos.**

Propiedades de los ambientes

- **Accesibles y no accesibles.** Si el aparato sensorial de un agente le permite tener acceso al estado total de un ambiente, se dice que éste es accesible a tal agente. Un agente es realmente accesible si los sensores detectan todos los aspectos relevantes a la elección de una acción.

Propiedades de los ambientes

Deterministas y no deterministas. Si el estado siguiente de un ambiente se determina completamente mediante el estado actual y las acciones escogidas por los agentes, se dice que el ambiente es determinista. Un agente no tiene por qué preocuparse de la incertidumbre en un ambiente accesible y determinista. Sin embargo, si el ambiente es inaccesible, podría parecer que es no determinista. Sería más conveniente clasificar un ambiente como determinista o no determinista considerando el punto de vista del agente.

Propiedades de los ambientes

- **Episódicos y no episódicos.** La experiencia del agente se divide en episodios, cada uno de los cuales consta de un agente que percibe y actúa. Los episodios no dependen de las acciones producidas en episodios anteriores.

Propiedades de los ambientes

- **Estáticos y Dinámicos.** Si el ambiente se puede modificar mientras el agente se encuentra deliberando, se dice que tal ambiente se comporta en forma dinámica en relación con el agente; de lo contrario, se dice que es estático. Si el ambiente no cambia, pero sí se modifica la calificación asignada al desempeño de un agente, se dice que el ambiente es semidinámico.

Propiedades de los ambientes

- **Discretos y continuos.** Si existe una cantidad limitada de percepciones y acciones distintas, se dice que el ambiente es discreto. El ajedrez es discreto pues existe una cantidad fija de posibles jugadas en cada ronda. La conducción de un taxi es continua.

Ejemplos de Ambientes y sus características

Ambiente	Accesible	Determinista	Episódico	Estático	Discreto
Ajedrez con reloj	Sí	Sí	No	Semi	Sí
Ajedrez sin reloj	Sí	Sí	No	Sí	Sí
Conducir un taxi	No	No	No	No	No
Sistema de diag. Médico	No	No	No	No	No
Sistema de análisis de imag.	Sí	Sí	Sí	Semi	No
Robot clasificador de partes	No	No	Sí	No	No
Controlador de refinería	No	No	No	No	No
Asesor de ingles interactivo	No	No	No	No	Sí