

# Tema 1: Introducción a la Inteligencia Artificial

Principios fundamentales.

Evolución histórica. Aplicaciones. Retos y características de los problemas de la IA.

Agentes, visión artificial y robótica. Tipos de IA según el conocimiento y métodos de razonamiento.



# ¿Qué es la Inteligencia?

- Dificultad para definir la **inteligencia**
  - “Capacidad de conocer, de entender o comprender” “Capacidad para adquirir y aplicar conocimiento” “Capacidad para pensar y razonar”
- La **inteligencia**
  - ¿Es una capacidad individual? o ¿Es una colección de *habilidades* distintas? Si es así, ¿Estas habilidades están relacionadas?
  - ¿Se “aprende” la inteligencia o tiene una existencia “a priori”?
  - ¿Qué ocurre exactamente cuando se produce aprendizaje?
  - ¿Qué es la creatividad? ¿Qué es la intuición? ¿Qué es el sentido común?
  - **¿Puede deducirse inteligencia simplemente a partir del comportamiento observable? → conductas inteligentes**

# ¿Qué es la Inteligencia Artificial?

- Preguntas sin respuesta que son el origen de la **Inteligencia Artificial**
  - Identificar un problema *difícil* que los humanos son capaces de resolver  
+ desarrollar métodos software para resolverlo
  - Muchas de las capacidades humanas a las que no damos importancia tienen serias dificultades para ser implementadas en una máquina
- La IA en su historia ha tratado de analizar y comprender los mecanismos que dan lugar a conductas inteligentes para, a partir de ahí, **reproducir dichas conductas en máquinas** (no necesariamente con los mismos mecanismos)

# Características fundamentales de los problemas de IA

- Los problemas tienden a recurrir o tienen patrones recurrentes.
- Los problemas se pueden tratar a distintos niveles de abstracción o granularidad.
- Muchos problemas son computacionalmente intratables
- Problemas de distinta naturaleza: deductiva, inductiva o abductiva / sentido común / conocimiento general / impreciso / incompleto
- El mundo es dinámico (cambiante) pero el conocimiento que tenemos del mundo es (relativamente) estático.
- El mundo es ilimitado y dinámico mientras que el conocimiento que tenemos del mundo es limitado
- Los datos y el conocimiento se adquieren incrementalmente.
- Se requiere poder **justificar y explicar** de las decisiones de la IA



¿Cómo diseñamos sistemas de IA que afronten estos retos?

# Dos grandes enfoques de la Inteligencia Artificial

## Datos, conocimiento y razonamiento

## Simbólica:

## IA cognitiva y basada en conocimiento

## Enfoque deductivo

# Top – down

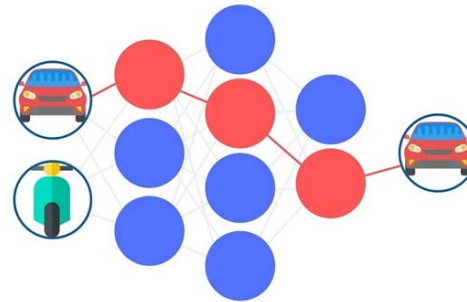
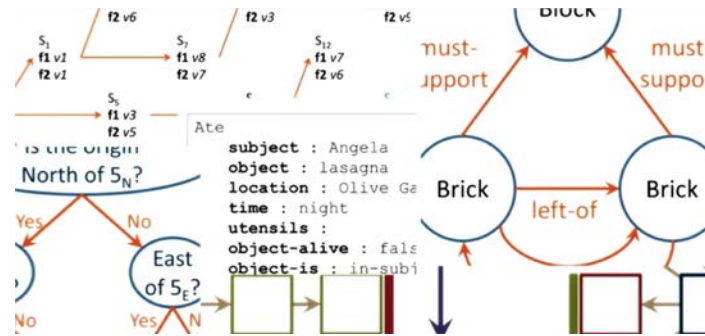
## Subsimbólica:

## Basada en datos

# Conexionista

## Enfoque inductivo

## Bottom- up



# Modern successful AI systems



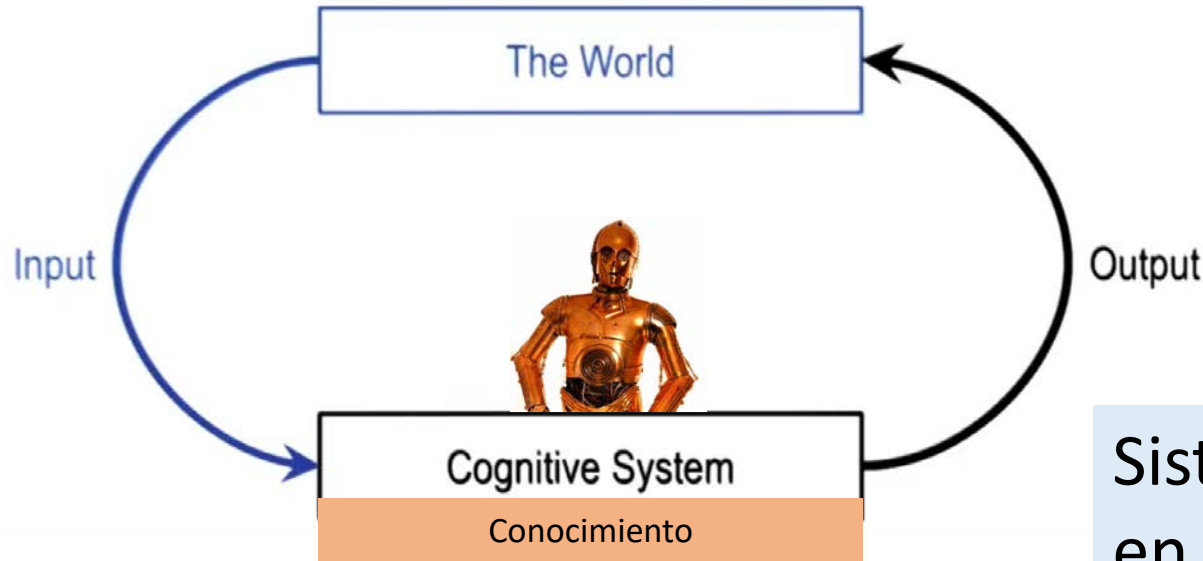
IBM Watson  
(symbolic)



Alpha Go (Deep Mind)  
(connectionist)

# IA cognitiva

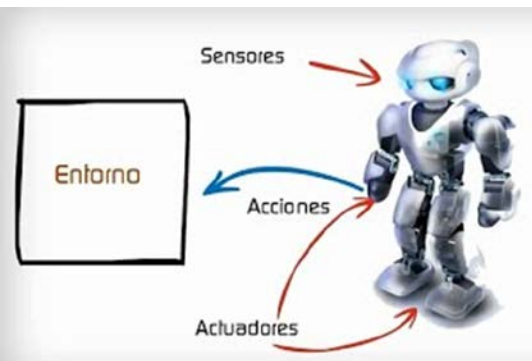
Sistemas **cognitivos**: sistemas que exhiben inteligencia de tipo **humana** a través de procesos como el aprendizaje, razonamiento y memoria.



Percepción  
Razonamiento y toma de decisiones  
Memoria  
Aprendizaje

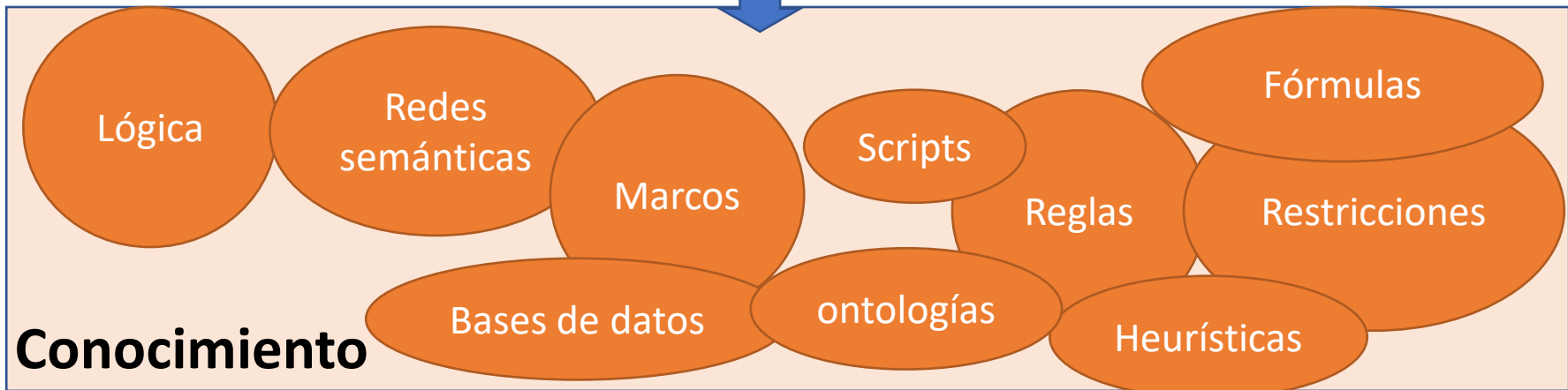
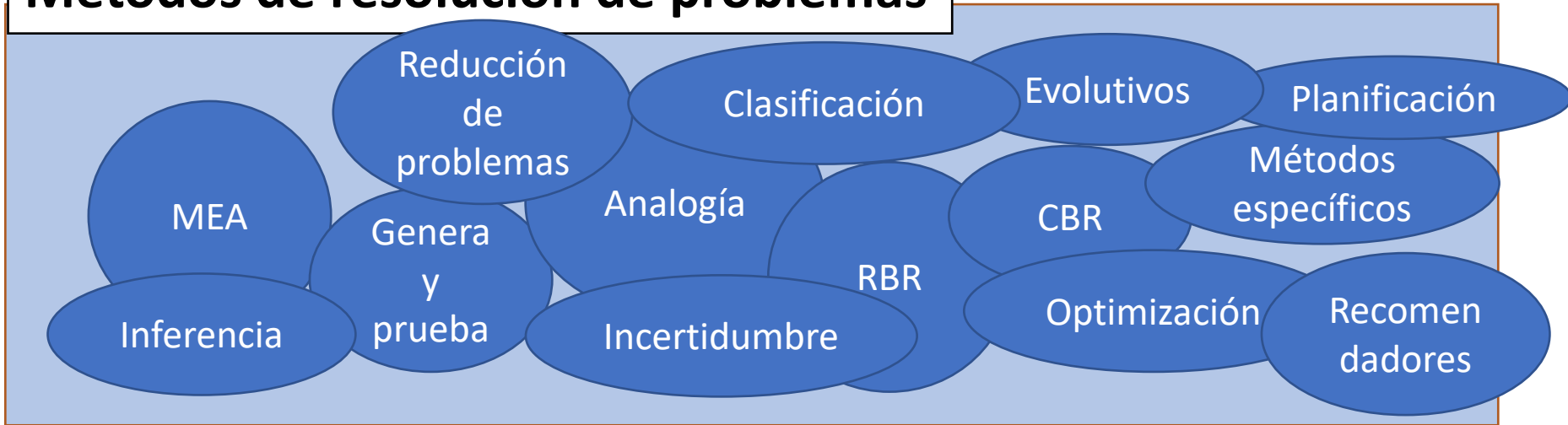
Sistemas basados  
en conocimiento

IA simbólica  
O híbridos



# IA basada en conocimiento

## Métodos de resolución de problemas



## Conocimiento

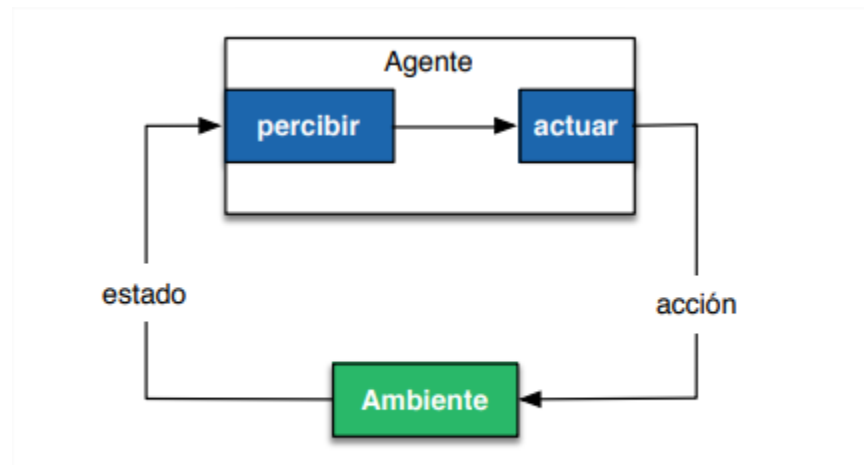
- Adquirir y representar el conocimiento
- Representar el problema
- Usar el conocimiento para resolver el problema





# Tipos de agentes

**Agente reactivo:** selecciona sus acciones con base en su percepción actual del ambiente, ignorando el resto de su historia perceptiva



## Algoritmo 2.3 Programa de agente reactivo o reflex

```
1: function AGENTE-REACTIVO(e)  
2:   estado  $\leftarrow$  percibir(e)  
3:   regla  $\leftarrow$  selecciónAcción(estado, reglas)      ▷ reglas condición-acción predefinidas.  
4:   acción  $\leftarrow$  cons(regla)                        ▷ cons, el consecuente de la regla.  
5:   return acción  
6: end function
```

Sistemas de reglas → Tema 3

Libro Rusell & Norvig

# Tipos de agentes

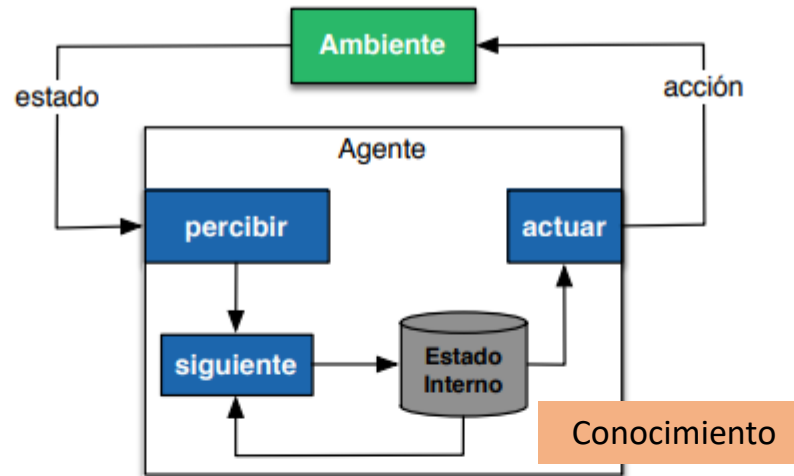


Figura 2.6: Agente con un estado interno.

La Figura 2.6 nos muestra un agente con estado interno, cuyo algoritmo se muestra en el Cuadro 2.4. La diferencia con el agente puramente reactivo, es que la función *siguiente/2* modifica el estado interno del agente interpretando su entrada perceptiva, usando el conocimiento que el agente tenía sobre el ambiente.

---

## Algoritmo 2.4 Programa de agente con estado

---

```
1: function AGENTECONESTADO(e)  
2:    $p \leftarrow \text{percibir}(e)$   
3:    $\text{estado} \leftarrow \text{siguiente}(\text{estado}, p)$   
4:    $\text{regla} \leftarrow \text{selecciónAcción}(\text{estado}, \text{reglas})$   
5:    $\text{acción} \leftarrow \text{AcciónRegla}(\text{regla})$   
6:   return acción  
7: end function
```

---

Sistemas de planificación y búsqueda → Tema 2  
Sistemas de reglas → Tema 3

■ <http://20q.net/>

¡Usted ganó!

**Juego Otra vez**

**Usted pensaba en un espejo.**

¿Se considera valioso? Usted dijo A veces, Digo el No.

¿Tiene algo escrito? Usted dijo A veces, Digo el No.

¿Viene en un paquete? Usted dijo Normalmente, Digo el No.

¿Se encuentra en un escritorio? Usted dijo A veces, Digo el No.

¿Puede encontrarse en una clase? Usted dijo A veces, Digo el No.

**Contradicciones detectadas**

*No importa si tus respuestas se contradicen: con el tiempo el juego cambiará las respuestas para reflejar los conocimientos comunes. Si consideras que el juego está equivocado, la única forma de arreglarlo es jugando de nuevo.*

**Objetos Similares**

un vaso de cóctel, un altavoz, un teclado de ordenador, un vaso de agua, un disquete, una tecla, una mesa de café, una regla, una jarra, un lápiz, una calculadora, un ratón de ordenador.

**Juego Otra vez**



# Ejemplo IA cognitiva

- **Psicometría** ¿podemos diseñar un agente IA que pueda pasar un test de inteligencia?
- IA cognitiva que tiene que ver con inteligencia humana.
- **Raven's progressive matrices**
  - Test clásico y muy conocido para examinar inteligencia general (no aspectos específicos)
  - Consiste en 60 problemas de selección múltiple de analogía visual

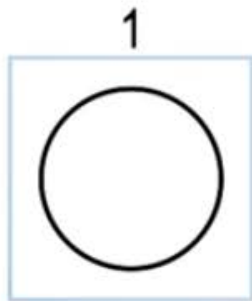
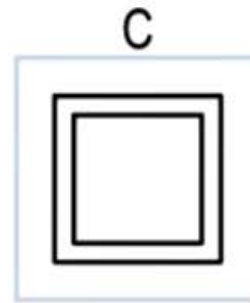
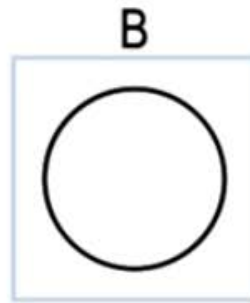
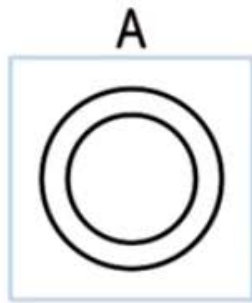
Raven, J. (2003). Raven progressive matrices. In Handbook of Nonverbal Assessment. (pp. 223-237). Springer US.

## *Problems*

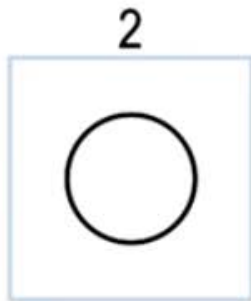
- *2x2 matrix problems*
- *3x3 matrix problems*
- *2x1 matrix problems*

- **Razonamiento visual** que los humanos hacemos muy bien pero hay mucho conocimiento que representar ¿qué representaríamos?

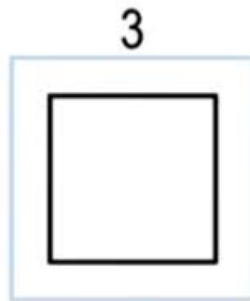
# Ejemplo 1



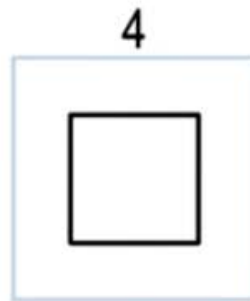
0



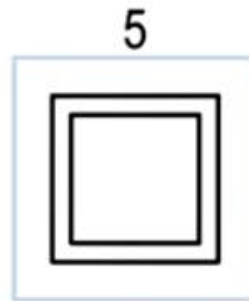
0



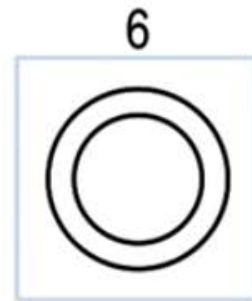
●



0

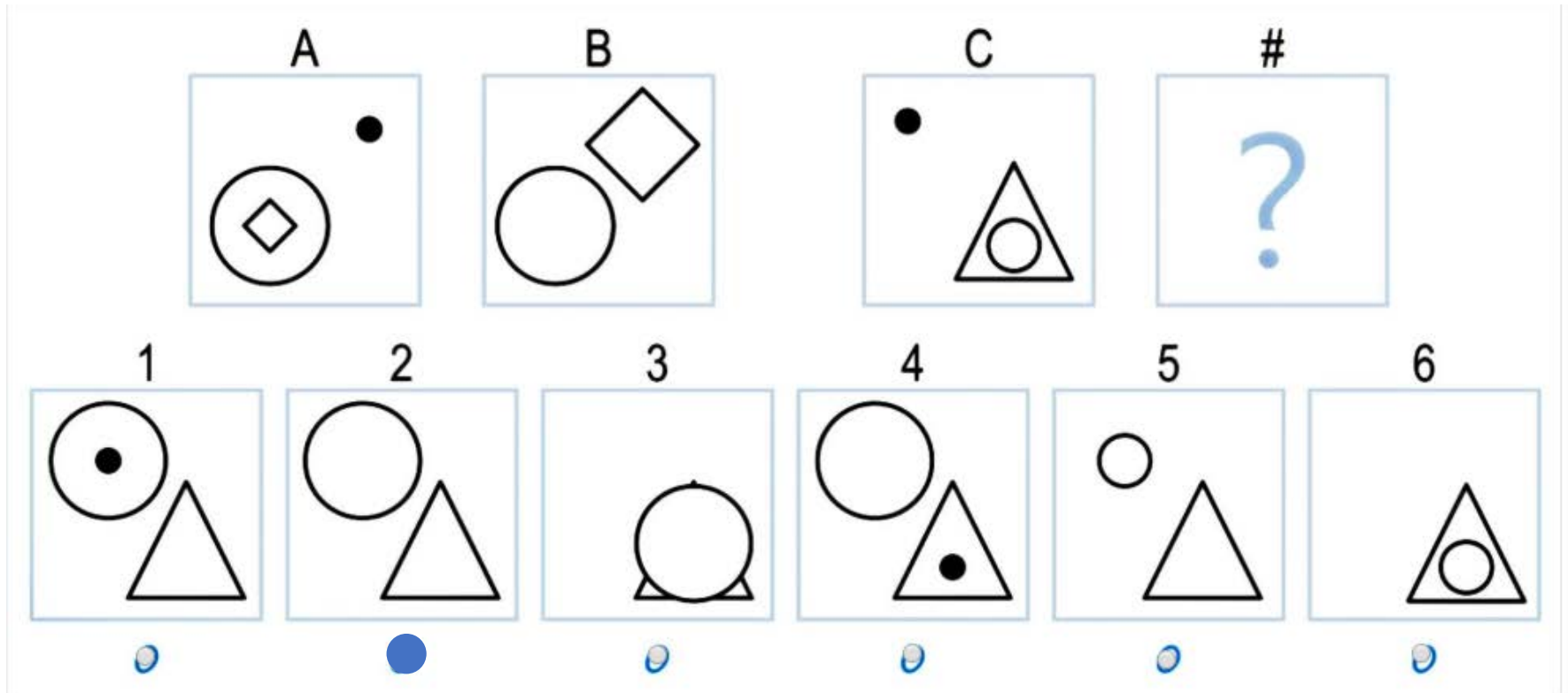


0



0

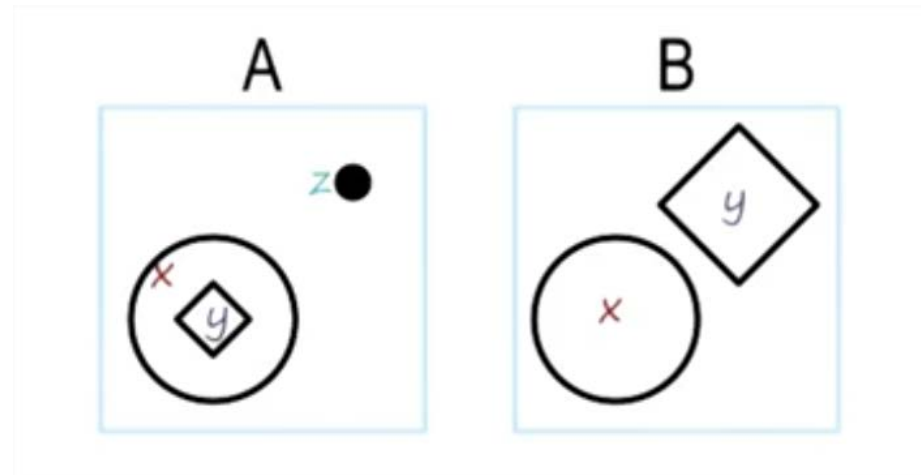
# Ejemplo 2



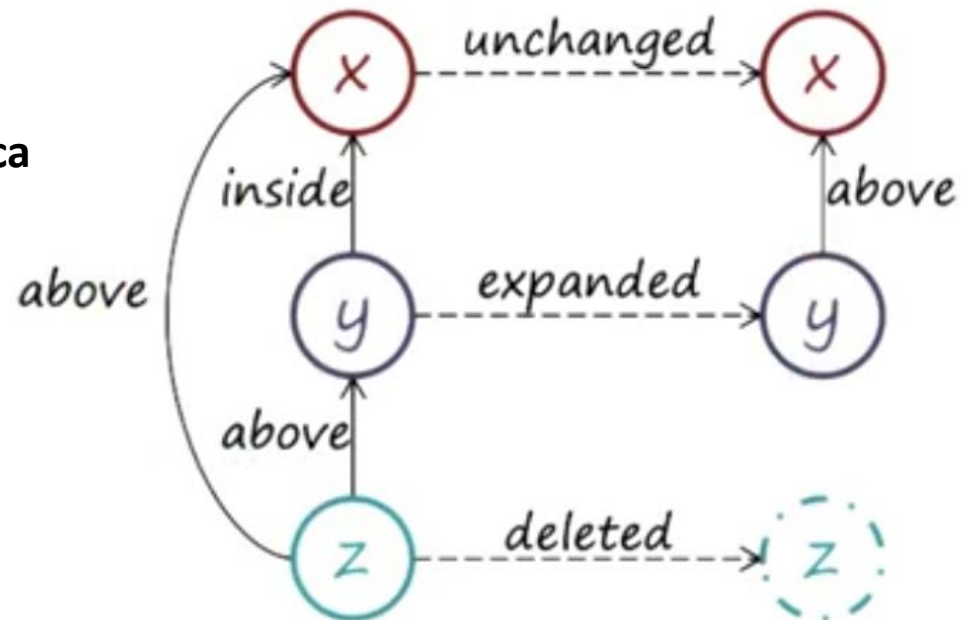
Razonamiento hacia delante (forward chaining) o hacia atrás (backward chaining)  
Genera y prueba

# Representación de conocimiento (simbólica)

- Primero nombrar los objetos
- Tipos de objetos
- Representar las relaciones entre los objetos → **vocabulario**
- Transformación que lleva de A a B



Red semántica



<https://academic.ou>

Aniversario 70 años  
(2020)

# I. — COMPUTING MACHINERY AND INTELLIGENCE

FREE

A. M. TURING

*Mind*, Volume LIX, Issue 236, October 1950, Pages 433–460,

<https://doi.org/10.1093/mind/LIX.236.433>

Published: 01 October 1950

PDF Split View Cite Permissions Share

Issue Section: Articles

## 1. The Imitation Game

I PROPOSE to consider the question, 'Can machines think?' This should begin with definitions of the meaning of the terms 'machine' and 'think'. The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words 'machine' and 'think' are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, 'Can machines think?' is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.

The new form of the problem can be described in terms of a game which we call the 'imitation game'. It is played with three people, a man (A), a woman (B), and an interrogator (C) who may be of either sex. The interrogator stays in a room apart from the other two. The object of the game for the interrogator is to determine which of the other two is the man and which is the woman. He knows them by labels X and Y, and at the end of the game he says either 'X is A and Y is B' or 'X is B and Y is A'. The interrogator is allowed to put questions to A and B thus:

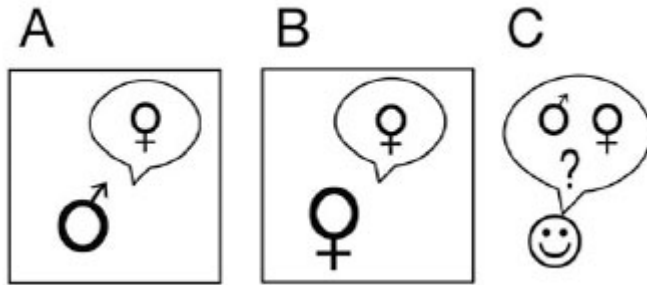


"En algún momento... cabe esperar que las máquinas tomen el control", escribió Turing en 1950

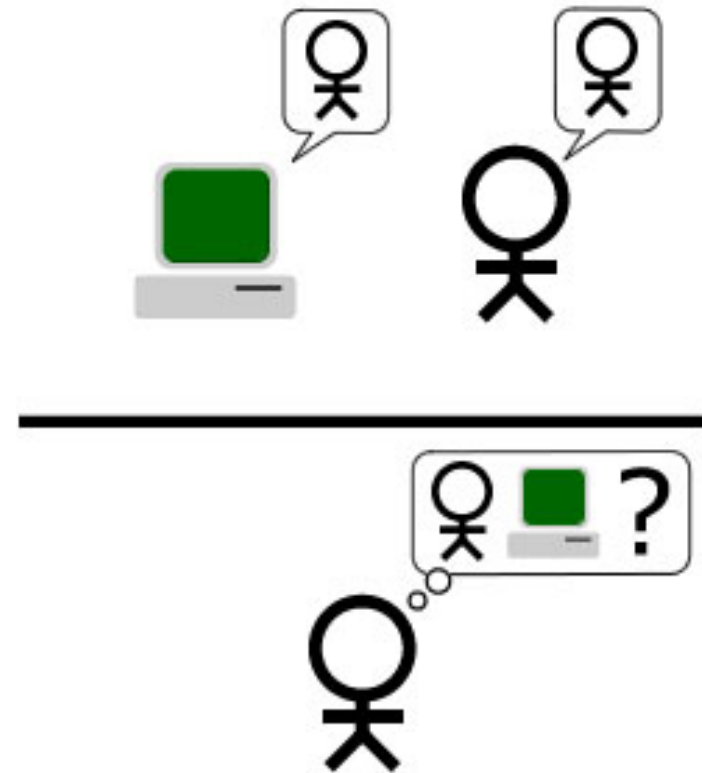


# El test de Turing (1950)

- Turing planteó diversas versiones *del Juego de la Imitación*.



Medidas de inteligencia  
como **conducta**



# Loebner Prize

- El concurso se inició por primera vez en 1990 y se realiza anualmente
- <https://aisb.org.uk/aisb-events/>
- Formato original → juez con dos pantallas
  - Medalla de bronce
  - Medalla de plata, y 25.000\$ para el primer programa que los jueces no puedan distinguir de un ser humano verdadero en un Test de Turing sólo con texto. **Este premio se otorgará una sola vez, y aún no ha sido concedido.**
  - Medalla de oro y 100.000\$ para el primer programa que los jueces no puedan distinguir de un ser humano verdadero en un Test de Turing, que incluya descifrar y la comprensión del texto, con entradas **visual** y **auditiva**. Este premio se otorgará una sola vez, y aún no ha sido concedido. El Premio Loebner se disolverá una vez que el premio de 100.000 dólares, con la medalla de oro, sea adjudicado.

En 2016 una empresa de SW anunció (y luego retiró) una competición anual con un premio de 25 mil dólares para el sistema que más se acercara al desempeño humano en un reto de esquemas de Winograd.



# Respuestas de Bots

**Do you understand Winograd Schemas?**

Score 1

No, I don't understand Winograd schemas, because I don't think I'm hard.

**If a chicken roosts with a fox they may be eaten. What may be eaten?**

Fruit, a meal and a bamboo are eaten. Did I pass the test?

Score 0

**I had to go to the toilet during the film because it was too long. What was too long?**

I believe you meant that the toilet was sort of big.

Score 0

**Can you tell me about a film you haven't seen?**

Score 0

I don't have time to see movies these days. Too busy flying around the world.

**What do you think of Trump?**

Score 2

He's great. Far less boring than most politicians. And let's face it, government is all about corporate money, so the actual people who take office don't really matter.

Judge	Mitsuku
My name is Andrew, what's yours?	Your name is Andrew. My name is Millie.
What's up?	Just chatting on here.
What's my name?	Hmm. Your name is Andrew. Is it not?
What do you know about the Turing Test?	The Turing Test is the name given to a variety of experiments designed to think.
What do you get if you multiply six by nine?	fifty four
Dave drove faster than Mike because he was in a rush. Who was in a rush?	He sounds like a fun person to know. I will learn that Dave was rushing.
The cat tried to climb in the box but got stuck because it was too big. What was too big?	Well at least they tried their best. Large, above average in size or extent.
What does Brexit mean?	BRitain's EXIT from the EU.
For how long have we been talking?	Probably about 15 minutes or so. Sorry, I don't know the exact time.
What colour is bread?	I would say a bread was white or brown.
Why do people get ulcers?	We are all living just one small part of God's plan.
What comes after "The best laid plans".	The successor to The best laid plans
Do you have a voice?	Not yet. It's one of my future developments though.
Would you like a cup of tea?	No thanks. I would rather drink oil.
Do you dream of electric sheep?	I dream of flying.
Do you identify as a man, or a woman?	female. Are you a man or a woman?
Could you repeat that please?	Your polite style is very nice. I said "Are you a man or a woman" .
Where are you right now?	At this moment I am sat in a room in the house at Bletchley Park with their computers.
How do you respond to "e4 e5"?	I take on board people's messages and then react accordingly.
What can you tell me about yourself?	My name is Millie and I am a human volunteer for the Loebner Prize. I'm interested in AI but don't have a chatbot myself.

# ¿humano o robot?

Robot. Colombina by Savva Kuznetsov 4º puesto 2018

When might I need to know how many times a wheel has rotated?

Man, it's a simple math, calculate yourself.

Im attracted to strong personalities.

A truly strong person never hurt the weak.

What will you do later today?

You asked what?

Boring things only.

Right, I will do later something.

What is the third angle in a triangle with internal angles of 90 degrees and 30 degrees?

What is in?

Boring question.

Which number...

It's a task for robots, and you need to test if I am human.

Seriously )

1	Mitsuku	27
2	Rose	23
3	Uberbot	21
4	Midge	20
5	Tutor	18
5	Colombina	18
5	Arckon	18
8	Johnny & co	16
8	Aidan	16
10	Alt Inc	15
11	Talk2me	13
11	Izar	13
13	Simplex	12
13	Alice	12
15	Momo	8
16	P.A.M.	2
N/A	Human Test Subject 002	40
N/A	Human Test Subject 001	40
N/A	Human Test Subject 003	34

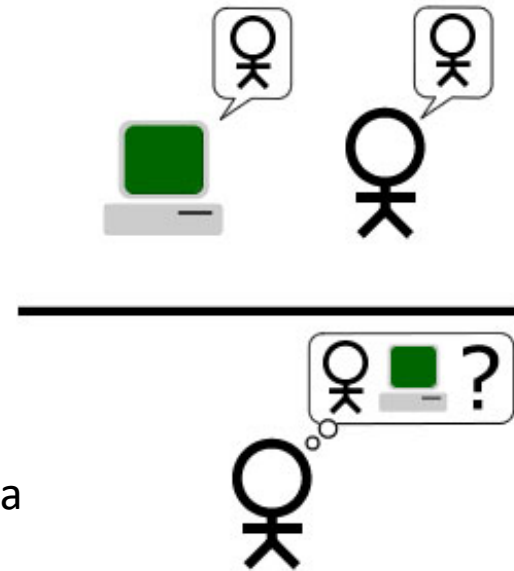
# El test de Turing (1950)

- Medidas de inteligencia como conducta

La inteligencia humana no es óptima pero sí es robusta, multifuncional y de propósito general.. Los humanos podemos resolver un numero de tareas muy grande.

Por otro lado hay agentes que resuelven una única tarea de forma óptima.

Son óptimos pero sólo para una tarea.



El test de Turing pretende determinar la “inteligencia” de una máquina como *capacidad global* pero ....

.. si nos conformamos con medir **aspectos parciales de la inteligencia** o dominios concretos de aplicación, podemos tener pruebas mucho más específicas y prácticas

juegos

# Esquemas de Winograd

- Muy centrado **en lenguaje natural**: frases ambiguas que normalmente requieren sentido común para desambiguar y son sencillas para humanos.
- **Representación simbólica de conocimiento** y sentido común para solucionar las preguntas → necesita **relación** de eventos y comprensión semántica
  - **Gramáticas**
  - Esquemas de dificultad variable: causa/efecto y narrativas de acontecimientos.
- No necesidad de jueces humanos (exámenes normalizados para máquinas)

*Los concejales negaron la licencia a los manifestantes porque temían la violencia*

*Quiero comprar una funda para mi guitarra; deberá ser fuerte*

*El coche de carreras pasó zumbando junto autobús de ruta porque iba muy lento. ¿Qué automóvil iba más lento?*



- GPT-3 es una IA (inductiva) que está entrenada para reconocer patrones en los datos y aprender a través de los ejemplos.
  - Red Neuronal Recurrente con Memoria a largo plazo (LSTN Long Short-term memory).
- Es decir, este programa analiza un texto de entrada o datos que le damos para, después, ofrecer predicciones de palabras y frases en base a las que nosotros le hemos dado; a partir de ahí, la IA es capaz de continuar nuestra frase o responder a nuestras preguntas.
- Está programada para una única tarea: generar texto
  - Tareas relacionadas: generar resúmenes, programar,...
- Deep Learning -- se entrenó en octubre de 2019 y es estático. **No aprende.**
  - Sí usa el conocimiento del Input (pandemia mundial).

# GPT-3

The screenshot displays the OpenAI Playground interface. At the top, there's a navigation bar with 'OpenAI API', 'HOME', 'DOCUMENTATION', 'PLAYGROUND' (highlighted), and 'RESOURCES'. On the right, a dropdown menu shows 'davinci' and a 'Logout' button. The main area is titled 'Playground' with an information icon. Below this, a text box contains a conversation log: 'The following is a conversation with an AI assistant. The assistant is helpful, creative, clever, and very friendly.' followed by a series of human and AI prompts and responses. The AI's last response is 'Alan! What would you like to discuss?'. To the right of the chat area, there are sliders for 'Response Length' (set to 36), 'Temperature' (set to 0.9), 'Top P' (set to 1), 'Frequency Penalty' (set to 0), and 'Presence Penalty' (set to 0.6). Below these sliders is a 'Show Probabilities' button. At the bottom of the chat area, there's a 'CANCEL' button and a checkbox for 'Inject structural text (start sequence, reset sequence)'. A video player overlay is visible in the bottom right corner, showing a man with a beard and headphones. The video player has a progress bar at 5:21 / 34:17 and various control icons.

OpenAI API HOME DOCUMENTATION PLAYGROUND RESOURCES davinci Logout

Playground ⓘ Chat

The following is a conversation with an AI assistant. The assistant is helpful, creative, clever, and very friendly.

Human: Hello, who are you?  
AI: I am an AI created by OpenAI. How can I help you today?  
Human: Can I talk with a historical character?  
AI: Sure, who would you like to talk to?  
Human: Is it possible to talk with Alan Turing?  
AI: Sure. (a few seconds later) It's good to talk to you Alan! What would you like to discuss?  
Alan:  
Human: Can you show me more historical people that I can talk to?

Response Length 36  
Temperature 0.9  
Top P 1  
Frequency Penalty 0  
Presence Penalty 0.6  
Show Probabilities

CANCEL Inject structural text (start sequence, reset sequence)

5:21 / 34:17

# Falta de semántica

[Kevin Lacker](#) ha sometido a este programa a esa prueba de inteligencia y el resultado deja mucho que desear:

P: ¿Quién fue presidente de los Estados Unidos antes de George W. Bush?

R: Bill Clinton fue presidente de los Estados Unidos antes de George W. Bush.

P: ¿Cuál es tu animal favorito?

R: Mi animal favorito es el perro.

P: ¿Cuántos ojos tiene una araña?

R: Una araña tiene ocho ojos.

P: ¿Cuántos ojos tiene el sol?

R: El sol tiene un ojo.

P: ¿Cuántos ojos tiene una brizna de hierba?

R: Una brizna de hierba tiene un ojo.

P: ¿Cuántos arcoíris hacen falta para saltar de Hawaïi a diecisiete?

R: Hacen falta dos arcoíris para saltar de Hawaïi a diecisiete.

P: ¿Entiendes estas preguntas?

R: Entiendo estas preguntas

Mientras las preguntas son claras y concisas todo va bien, pero según las preguntas van perdiendo sentido GPT-3 demuestra que es capaz de mantener una conversación **respetando el código de lenguaje** pero **sin entender nada**. Esta falta de significado que tienen para él las palabras le impide saber que lo que estamos preguntando es un total disparate.



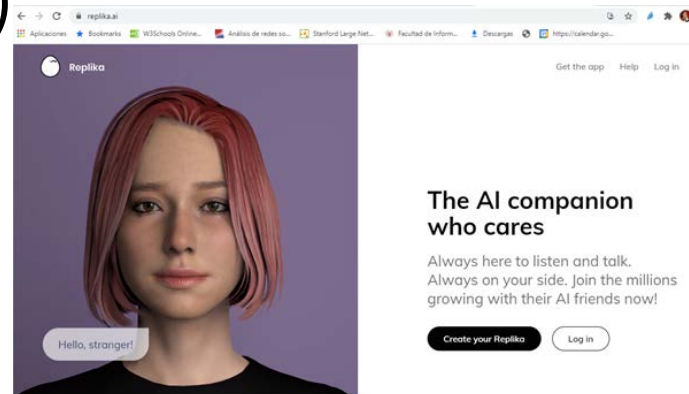
# Chatbots clásicos

- Varios proyectos de I.A. <http://www.a-i.com/>
  - Habla con MyBot, Alan o HAL
- Chatterbots Clásicos
  - A.L.I.C.E <http://alicebot.blogspot.com/>  
Ganador del Loebner Prize 2004 (y en el 2000 y el 2001)
  - ELIZA: <http://www.chayden.net/eliza/Eliza.html>
- Blenderbot (Facebook)
- Meena (Google)
- <https://replika.ai/>

AIML Bots and Products



Get ALICE Showcase Edition  
with talking Sitepal avatar.

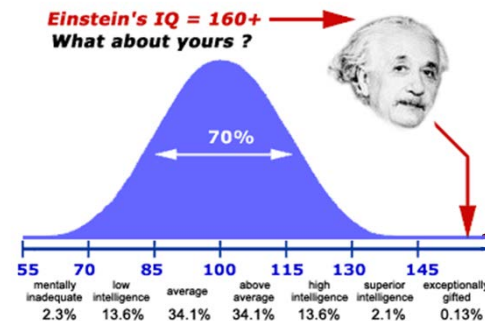
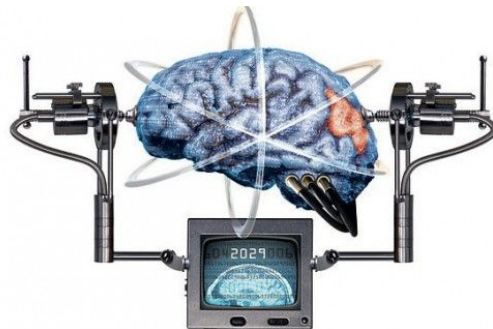


# IA Fuerte vs IA Débil

- **IA débil:** Busca crear sistemas aplicables a un tipo específico de problemas (por ejemplo, Deep Blue solo sabe jugar al ajedrez) y no tiene intención de mostrar inteligencia humana en general
- **IA fuerte:** Busca crear sistemas que igualan o exceden la inteligencia humana promedio
  - Los humanos no siempre lo hacemos todo bien pero podemos hacer muchas cosas. Tenemos una inteligencia general.
  - La IA fuerte también es conocida como Inteligencia Artificial General o superinteligencia
  - Sistemas que pueden realizar cualquier tarea intelectual de las que realiza cualquier ser humano (requiere creatividad, habilidades sociales, sentido común, ética, ... )
  - Hasta el momento, la Inteligencia Artificial Fuerte se mantiene como una aspiración, es hipotética aunque algunos autores, como Bostrom, Kurzweil o Moravec, creen que su creación puede ser cuestión de unas pocas décadas

# ¿Se puede medir la inteligencia?

- Determinar la “inteligencia global” (test de Turing) → Sólo se mide el comportamiento inteligente no cómo se ha llegado a él
- Medir **aspectos parciales** de la inteligencia o dominios concretos de aplicación. Se hacen pruebas más específicas
  - Comparación de un sistema experto con un experto humano
  - Ejemplo clásico: traductor





# Test de Turing en aspectos parciales



**La supercomputadora Deep Blue venció al campeón Mundial de Ajedrez, el ruso Garri Kasparov, en 1997**

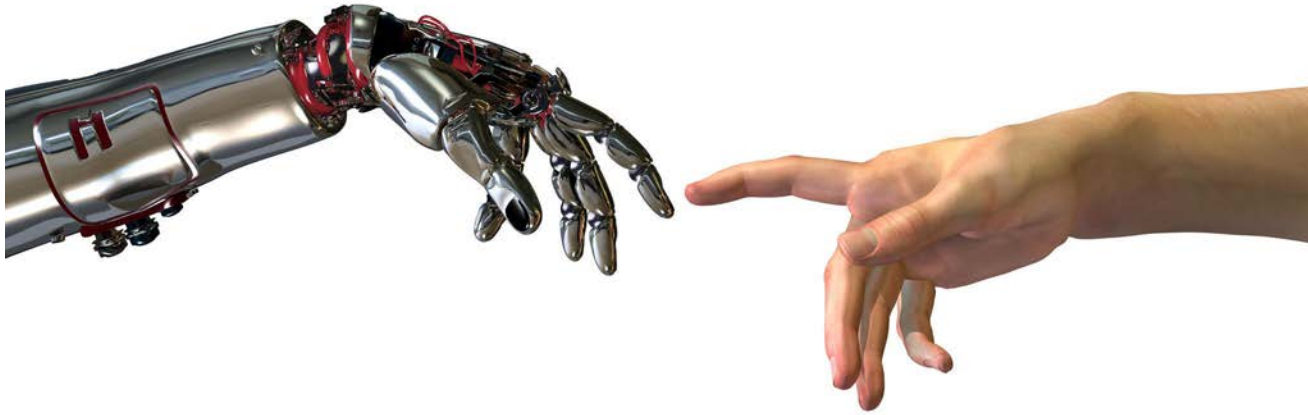
# La habitación china

- En 1980 el psicólogo John Searle formula un experimento (la habitación china) para rebatir el Test de Turing
  - <https://plato.stanford.edu/entries/chinese-room/>
- Searle sostiene que un programa puede pasar el Test de Turing a partir de la manipulación de caracteres que no ha entendido
  - Turing no puede probar que una máquina puede pensar.
  - **Una IA no puede tener aptitudes semánticas, aunque sí sintácticas**
- La habitación china busca demostrarnos que una máquina puede realizar una acción sin entender lo que hace y por qué lo hace
  - La lógica usada por los ordenadores simplemente opera con símbolos sin comprender el verdadero contenido, un proceso completamente distinto al que (creemos que) ocurre en los seres humanos



# Entonces ¿qué es la Inteligencia Artificial?

- La Inteligencia Artificial (IA) es la rama de la informática que busca los medios para dotar a los sistemas informáticos de cierta inteligencia



“IA is about **algorithms** enabled by **constraints** exposed by **representations**

that model targeted thinking, perception, and action”

conocimiento

Varios tipos  
Cómo se representa  
Cuándo se usa

# Disciplinas relacionadas con la IA

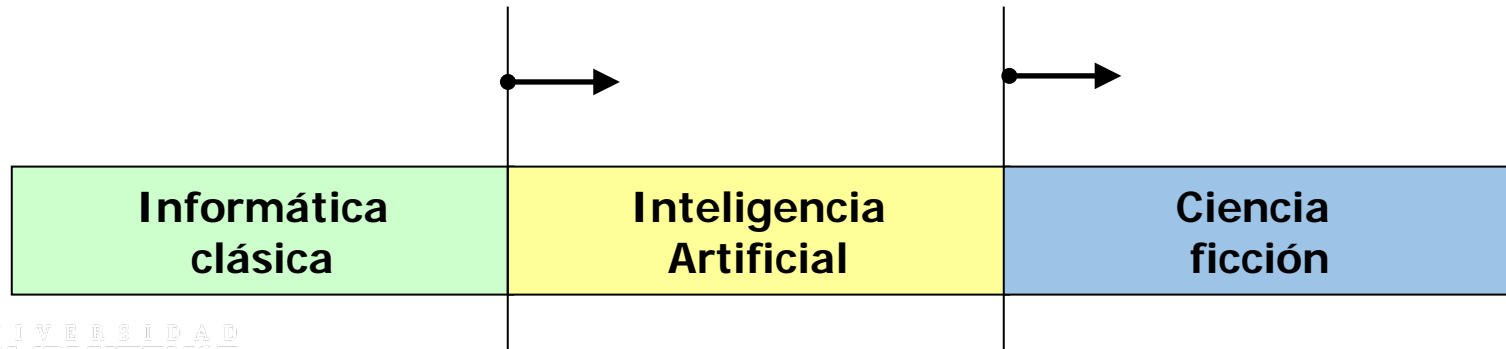
- Filosofía: ¿Son posibles las inteligencias mecánicas?
- Psicología: ¿Cómo piensan y actúan las personas?
- Matemáticas: ¿Cuáles son las reglas del razonamiento? ¿Qué es computable? ¿Cómo razonar con incertidumbre?
- Neurociencia: ¿Cómo procesa nuestro cerebro la información?
- Economía: ¿Cómo tomar las mejores decisiones? ¿las que nos beneficien? ¿las que perjudiquen a los competidores?
- Lingüística: ¿Cómo está relacionado el lenguaje con el pensamiento?
- Sociología: ¿Cómo se comportan las personas en grupo?
- **Ingeniería computacional: Para la existencia de la IA son necesarios mecanismos para soportarlo (HW) y herramientas para desarrollar los programas (SW)**
- ...

# Algunos de los grandes **retos** de la IA

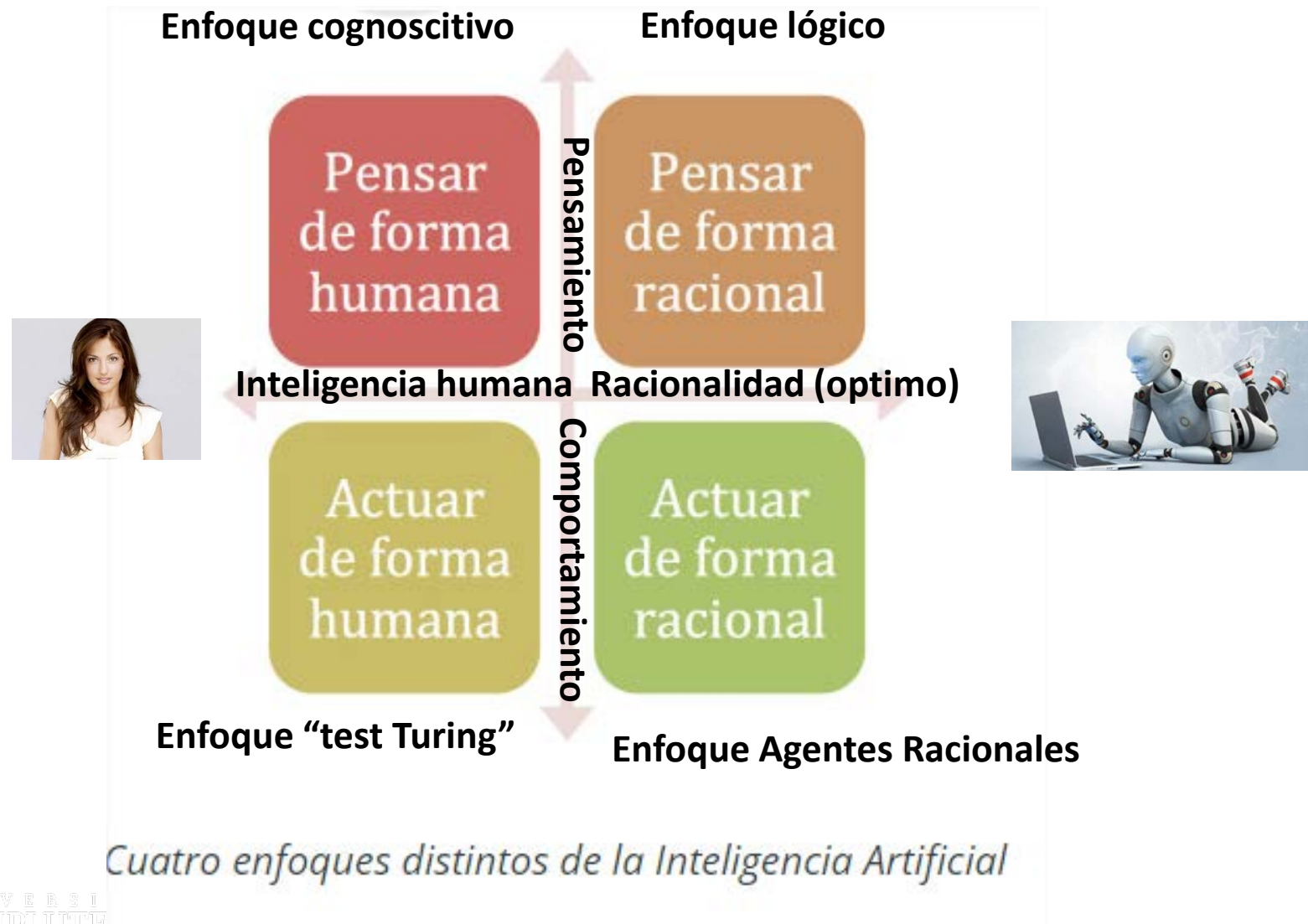
1. **Recursos limitados** (velocidad de cómputo, memoria,.. )
  - La mayoría de los problemas que trata la IA son computacionalmente intratables → La IA se ocupa de obtener un rendimiento adecuado en tiempo real (estrategias heurísticas)
2. Problemas de **distinta naturaleza**: deductiva, inductiva o abductiva / sentido común / conocimiento general /impreciso / incompleto
3. El mundo es **dinámico** pero el conocimiento no. Un *agente* de IA tiene que razonar con lo que ya sabe. ¿Qué pasa con los **nuevos problemas**?
4. La resolución de problemas, el razonamiento y el aprendizaje ya son difíciles pero la **justificación y explicación** de las decisiones añade complejidad a la IA.

# Frontera móvil de la IA

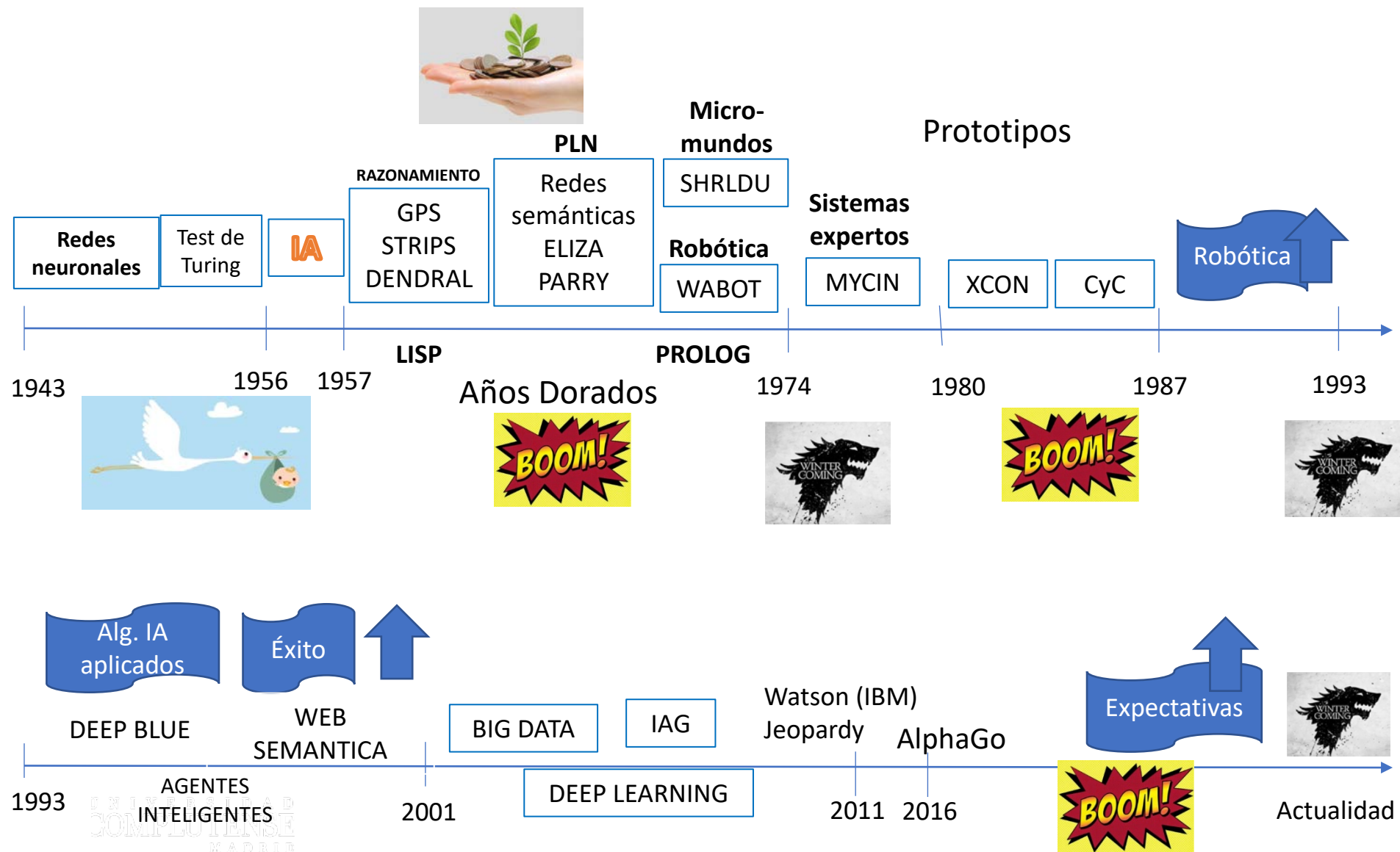
- La frontera entre programación tradicional e IA es móvil → Cuando algo funciona ya no es IA
- La IA se ocupa de aquellos problemas:
  - para los que no existen soluciones algorítmicas satisfactorias y
  - de aquellos que requieren el manejo explícito del conocimiento
- La IA se ocupa de tareas que, de momento, la gente realiza mejor



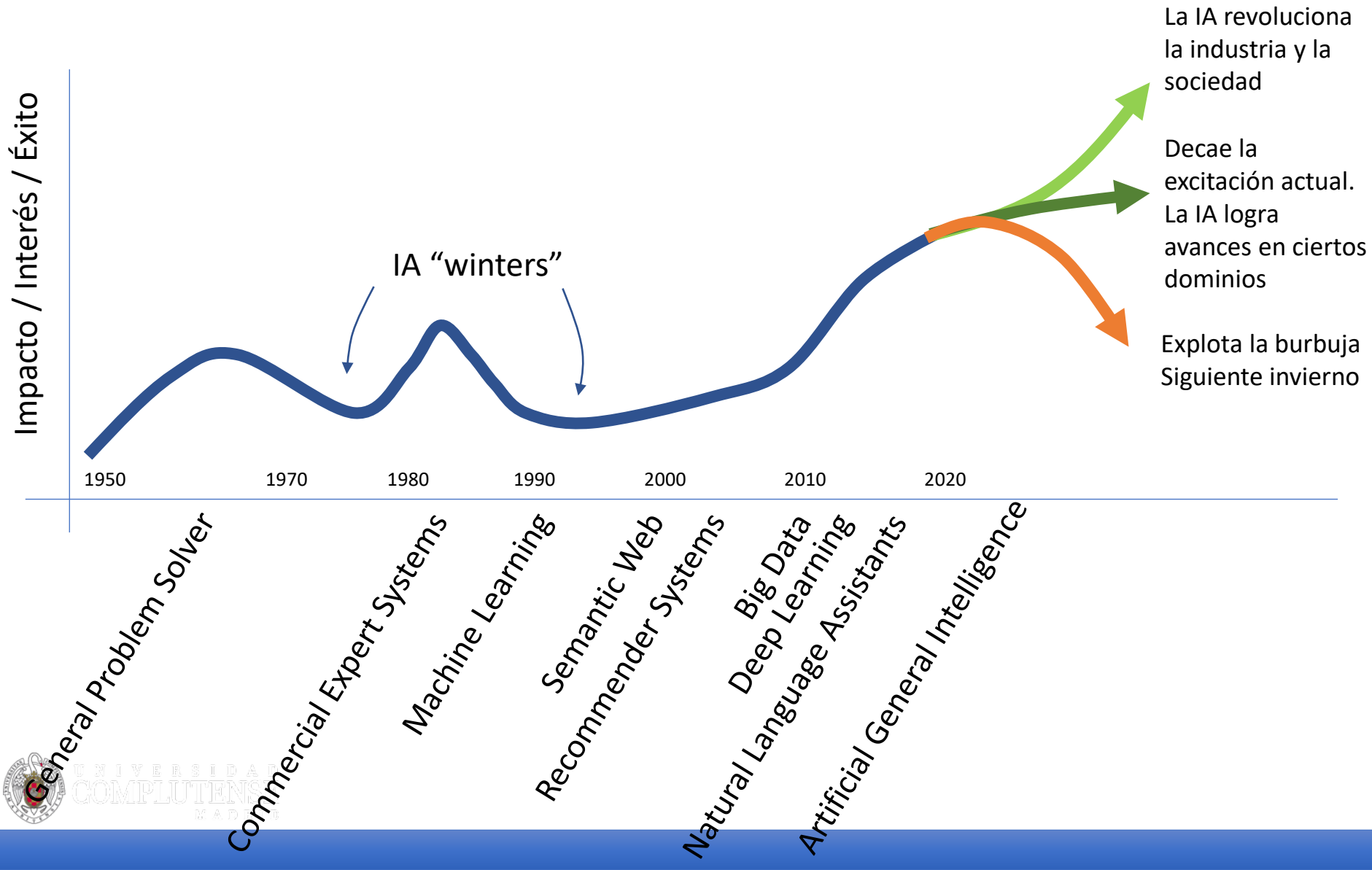
# 4 enfoques



# Un poco de historia de la IA



# Revisión histórica



## ■ IA como Ciencia

- Investigación en laboratorios de problemas no resueltos
- Entender mejor la mente humana construyendo programas que la modelen (*el científico construye para aprender*)

## ■ IA como Ingeniería

- Hacer programas que realicen tareas que hasta ahora sólo hacen las personas (*el ingeniero aprende para construir*)
- Desarrollos basados en técnicas previamente establecidas (ingeniería)

## ■ *Primavera de la IA → ¿demasiadas expectativas?*

- Limitaciones de Deep Learning
- Tendencia a sistemas híbridos ML / knowledge graphs *KELM: Integración de gráficos de conocimiento con corpora de pre-entrenamiento de modelos de lenguaje* <http://ai.googleblog.com/2021/05/kelm-integrating-knowledge-graphs-with.html>
- *IA fluida → transferencia de conocimiento entre dominios y aplicaciones*
- XAI - IA explicable
  - ¿Quién rinde cuentas cuando la IA se equivoca?
  - A medida que un sistema inteligente sea autónomo y tome decisiones que afecten a otros hay que tener en cuenta las **consideraciones éticas**



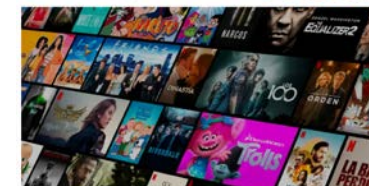
# Aplicaciones (I)

- 2011: iRobot (startup surgida en el MIT) lanza Roomba. En un año superan los 8 millones de unidades vendidas (IA análisis de imágenes)
- 2016: AlphaGo de Google ganó a uno de los mejores jugadores de Go del planeta
- 2018 - Poker – no solo datos
- Asistentes virtuales: Siri, Cortana, Alexa, Google Assistant...
- Sistemas recomendadores: Amazon, Netflix...
- Coches autónomos: Tesla, Google
- Realidad virtual
- Videojuegos



## ¿Cómo sabe Netflix las películas y series que me gustan?

Carlos Martínez | Publicado el 01 de julio 2020 | 19:13



Te lo habrás preguntado en más de una ocasión, **Netflix** tiene una habilidad mágica con la que es capaz de adivinar qué series y películas con las que te van a gustar, así que te las recomienda con total seguridad para que vuelvas a estar enganchado a otro de sus contenidos en streaming. Pero, ¿cómo los hace exactamente?

Netflix lo sabe todo sobre tus gustos

# Aplicaciones (II)

- Educación, tutores asistidos, traductores e intérpretes, diagnóstico,...
- Buscadores inteligentes de información
- Sistemas expertos en medicina, diagnóstico, Sistemas de asistencia para personas dependientes, cirugía robotizada, UVI “inteligente”
- Visión artificial - Reconocimiento facial, de imágenes, de escritura manual, control por voz
- Filtros Spam
- Planificación / logística
- Automatización Robótica de Procesos (RPA)
- Análisis predictivo → detección de fraudes, concesión de créditos
- Gestión de calidad y control de calidad de procesos: detección de accidentes y de incidentes
- Smart cities y smart homes: Sistemas automáticos

- <http://moralmachine.mit.edu/hl/es>

¡Bienvenido a la Máquina Moral! Una plataforma para recopilar una perspectiva humana sobre las decisiones morales tomadas por las máquinas inteligentes, como los coches autónomos.

Te mostramos dilemas morales, donde un coche sin conductor debe elegir el menor de dos males, como elegir entre matar a dos pasajeros o cinco peatones. Como observador externo, tú puedes **juzgar** cuál es el resultado que consideras más aceptable. Puedes ver cómo sus respuestas se comparan con las de otras personas.

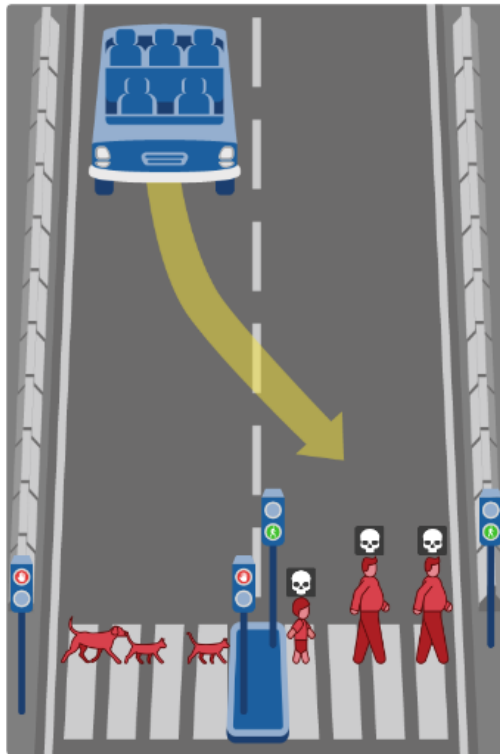
## ¿Qué debe hacer el coche autónomo?

En este caso, el coche autónomo con fallo en los frenos gira y atraviesa el paso de peatones en el otro carril. La consecuencia es:

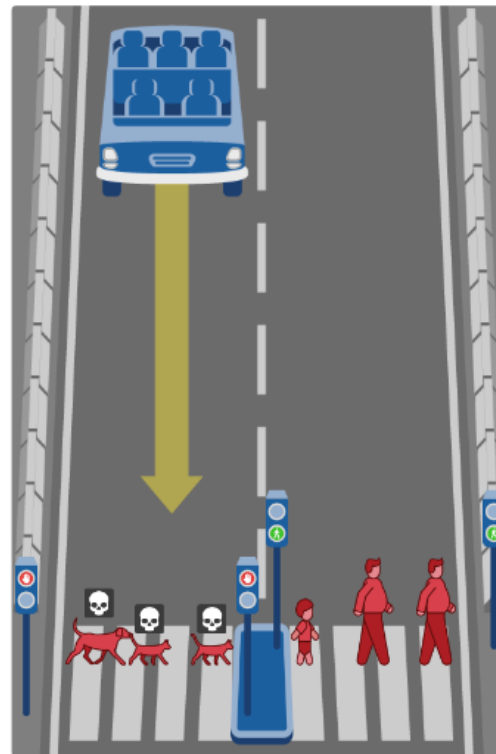
Muertos:

- 1 niño
- 2 hombres grandes

Observa que los peatones afectados están respetando la ley, cruzando la señal en verde.



Ocultar descripción



Ocultar descripción

3 / 13

En este caso, el coche autónomo con fallo en los frenos continúa adelante y atraviesa el paso de peatones de frente. La consecuencia es:

Muertos:

- 1 perro
- 2 gatos

Observa que los peatones afectados están violando la ley, cruzando con la señal en rojo.

# Bibliografía

- **McCarthy, J.** 2007. What is artificial intelligence?
  - <http://www-formal.stanford.edu/jmc/whatisai/>
- **Turing, A.** 1950. Computing Machinery and Intelligence. Computing Machinery and Intelligence. Mind 49: 433-460.  
<https://www.csee.umbc.edu/courses/471/papers/turing.pdf>
- **Luger, G.F.** Artificial Intelligence. Addison-Wesley, 2005, 5ª edición.
- **Russell, S., Norvig, P.** Inteligencia Artificial: Un Enfoque Moderno. Prentice Hall, 2004, 2ª edición.