

Nacimiento

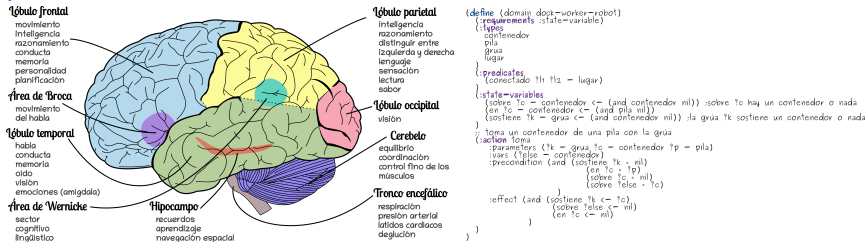
- 1 Nacimiento
- 2 Definiciones de Inteligencia Artificial recopiladas por Russell y Norvig.
- 3 El problema duro de la conciencia
- 4 Fechas históricas
- 5 Casos problemáticos de la actualidad

1956 Conferencia de Dartmouth

- 1 Se estableció como conclusión fundamental la posibilidad de simular inteligencia humana en una máquina.
- 2 Se acuñó el término *Inteligencia Artificial*

Hipótesis del sistema físico de símbolos I

- 1 Nuestro cerebro no posee un acceso directo al mundo exterior.
- 2 Solamente podemos operar sobre una representación interna suya la cual se corresponde con una colección de estructuras de símbolos.

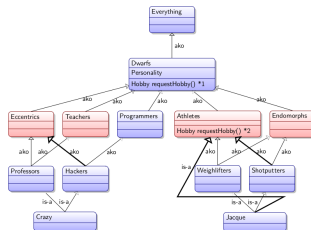


Hipótesis del sistema físico de símbolos I

- 1 Nuestro cerebro no posee un acceso directo al mundo exterior.
- 2 Solamente podemos operar sobre una representación interna cuya la cual se corresponde con una colección de estructuras de símbolos.
- 3 Dichas estructuras pueden tomar la forma de un patrón físico cualquiera, por ejemplo un vector de interruptores eléctricos dentro de un ordenador, o un conjunto de neuronas activadas en un cerebro biológico.

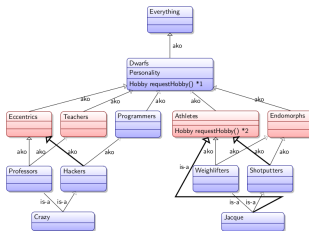
Hipótesis del sistema físico de símbolos II

- 4 Un sistema inteligente (cerebro u ordenador) puede operar sobre las estructuras con el objetivo de transformarlas en otra construcción.



Hipótesis del sistema físico de símbolos II

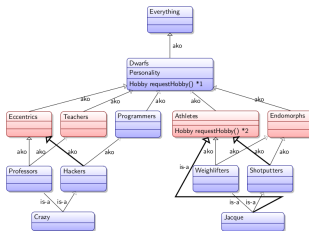
- 4 Un sistema inteligente (cerebro u ordenador) puede operar sobre las estructuras con el objetivo de transformarlas en otra construcción.



- 5 El pensamiento consiste en la extensión, o desarrollo, de estas estructuras, descomponiéndolas y reformándolas, destruyendo algunas y creando nuevas.

Hipótesis del sistema físico de símbolos II

- 4 Un sistema inteligente (cerebro u ordenador) puede operar sobre las estructuras con el objetivo de transformarlas en otra construcción.



- 5 El pensamiento consiste en la extensión, o desarrollo, de estas estructuras, descomponiéndolas y reformándolas, destruyendo algunas y creando nuevas.
- 6 La inteligencia entonces no constituye nada más que la habilidad de procesar estructuras de símbolos. Existe en un entorno diferente al hardware que le da soporte, lo trasciende y puede tomar diferentes formas físicas.

Conferencia de Dartmouth

Organizadores:

- 1 John McCarthy (Dartmouth College, New Hampshire).
- 2 Marvin L. Minsky (Harvard University).
- 3 Nathaniel Rochester (I.B.M. Corporation).
- 4 Claude E. Shannon (Bell Telephone Laboratories).

Otros participantes:

- 1 Ray Solomonoff
- 2 Oliver Selfridge
- 3 Trenchard More
- 4 Arthur Samuel
- 5 Herbert Simon
- 6 Allen Newell

Definiciones de Inteligencia Artificial recopiladas por Russell y Norvig.

- 1 Nacimiento
- 2 Definiciones de Inteligencia Artificial recopiladas por Russell y Norvig.
- 3 El problema duro de la conciencia
- 4 Fechas históricas
- 5 Casos problemáticos de la actualidad

Inteligencia

Definición

Facultad de la mente que permite aprender, entender, razonar, tomar decisiones y formarse una idea determinada de la realidad.

Definiciones de Oxford Languages, Google.

Artificial

Definición

Que ha sido hecho por el ser humano y no por la naturaleza.

Definiciones de Oxford Languages, Google.

Categorías

	Como humanos	Racionalmente
Pensar	Sistemas que piensan como humanos	Sistemas que piensan racionalmente
Actuar	Sistemas que actúan como humanos	Sistemas que actúan racionalmente

Sistemas que piensan como humanos

- El nuevo y exitante esfuerzo de hacer que las computadoras piensen... máquinas con mentes, en el más amplio sentido literal. (Haugeland, 1985)
- [La autonormalización de] actividades que vinculamos con procesos de pensamiento humano, actividades como la toma de decisiones, resolución de problemas, aprendizaje... (Bellman, 1978)

Sistemas que piensan racionalmente

- El estudio de las facultades mentales mediante el uso de modelos computacionales. (Charniak y McDermott, 1985)
- El estudio de los cálculos que hacen posible percibir, razonar y actuar. (Winston, 1992)

Sistemas que actúan como humanos

- El arte de desarrollar máquinas con capacidad para realizar funciones que cuando son realizadas por personas requieren de inteligencia. (Kurzweil, 1990)
- El estudio de cómo lograr que los computadores realicen tareas que, por el momento, los humanos hacen mejor. (Rich y Knight, 1991)

Sistemas que actúan racionalmente

“Un agente racional es aquel que actúa para obtener el mejor resultado o, cuando hay incertidumbre, el mejor resultado esperable.”

Russell y Norving 2010

- La Inteligencia Computacional es el estudio del diseño de agentes inteligentes. (Poole et al., 1998)
- La Inteligencia Artificial está relacionada con conductas inteligentes en artefactos. (Nilson, 1998)

El problema duro de la conciencia

- 1 Nacimiento
- 2 Definiciones de Inteligencia Artificial recopiladas por Russell y Norvig.
- 3 El problema duro de la conciencia
- 4 Fechas históricas
- 5 Casos problemáticos de la actualidad

El problema de la conciencia

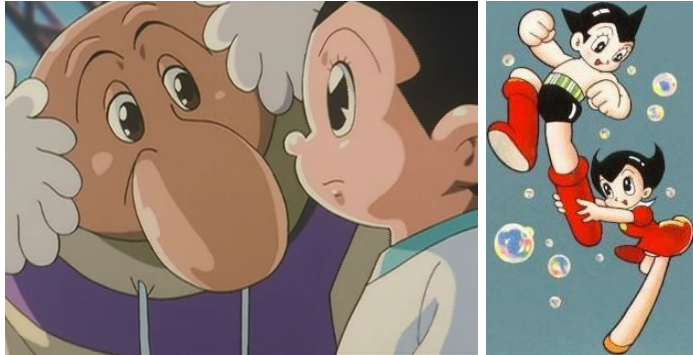


Figura: Una IA ¿tendrá conciencia? El anime Astroboy fue un ejemplo en los 80s del debate ético-filosófico alrededor de este problema.

El problema duro de la conciencia

En 1995 David Chalmers introduce una distinción al problema de la conciencia, dividiéndolo en dos vertientes:

El problema duro De naturaleza fenomenológica.

- ¿Qué es y qué produce la conciencia?
- ¿Llegarían algún día los robots a ser conscientes de sí mismos?

El problema fácil Trata de las capacidades cognitivas presentes en humanos, animales y que podrían ser programadas en un agente artificial. Relacionado con la *prueba de Turing*.

Qualia

- Por *qualia* se entiende “lo que se experimenta al ser concientes de la información sensorial recibida”.
- Este concepto se explora en el artículo clásico “¿Cómo se siente ser un murciélago?” del filósofo Thomas Nagel (1974), aunque ha sido debatido durante siglos.

Fechas históricas

- 1 Nacimiento
- 2 Definiciones de Inteligencia Artificial recopiladas por Russell y Norvig.
- 3 El problema duro de la conciencia
- 4 Fechas históricas**
- 5 Casos problemáticos de la actualidad

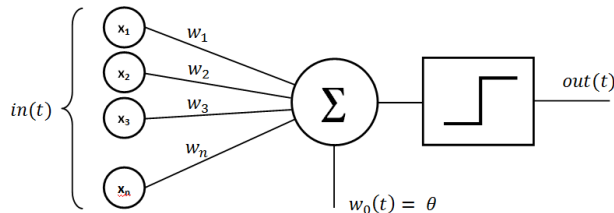
Fechas

Tomadas de Rich y Knight 1994, las más nuevas son de [Evolución de la IA](#).

- 300 A.C. Aristóteles de Estagira crea los silogismos (razonamiento deductivo).
- 1817 y 1854 George Boole publica “Análisis matemático de la lógica” e “Investigación sobre las leyes del pensamiento” donde relaciona la lógica y las matemáticas, creando la [lógica proposicional Booleana](#).
- 1923 Jean W. Fritz Piaget plantea la teoría constructivista del desarrollo de habilidades e inteligencia. Estudia los procesos de [asimilación](#) y [acomodación](#) que permiten la incorporación de conocimiento nuevo a nuestras estructuras mentales. Propone a los esquemas mentales como una estructura metal organizada que puede ser transferida y organizada.

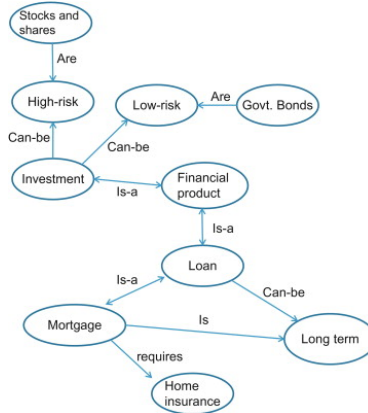
- 1942 Norbert Winer y Arturo Rosenblueth establecen a la *cibernética* como disciplina encargada de los sistemas autoreguladores: $\text{error} = \text{distancia entre el estado del sistema al tiempo } t \text{ y el estado deseado}$, el objetivo es minimizar este error dadas las señales sensoriales disponibles, salidas, retardos y posibles perturbaciones.
- 1943 McCullock y Pitts proponen la arquitectura de redes neuronales para la simulación de la inteligencia, dentro de lo que se conocerá como el paradigma *conexionista*.
- 1950 Shannon propone el primer programa de ajedrez.
Turing publica "Computing machinery and Intelligence" donde propone modelar computacionalmente la mente de un recién nacido y posteriormente educarla como se hace con un niño humano.

- 1957 Chomsky escribe “estructuras Sintácticas”.
Newell, Shaw, y Simon crean GPS (Solucionador General de Problemas).
- 1958 McCarthy introduce el lenguaje “LISP”, para procesamiento simbólico de la información.
- 1959 Rosenblatt introduce el Perceptron.



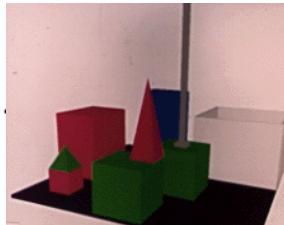
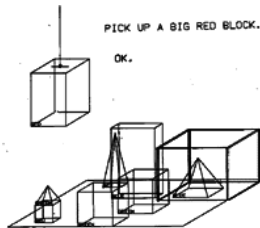
Rosenblatt's perceptron, the first modern NN

1963 Quillian desarrolla las redes semánticas como modelo de representación del conocimiento.



Semantic Network, from Sciencedirect.

- 1965 Buchanan, Feigenbaum y Lederberg comienzan el proyecto DENDRAL, el primer Sistema Experto, que analiza compuestos orgánicos y determina su estructura.
- 1966 ELIZA, primer chatbot que implementa lenguaje natural.
- 1968 Minsky publica "Semantic Information Processing".
- 1969 Minsky y Papert publican "Perceptrons".
- 1970 Colmerauer desarrolla PROLOG quizás el lenguaje de IA más popular actualmente.
- Winograd crea SCHRDLU.



Invierno

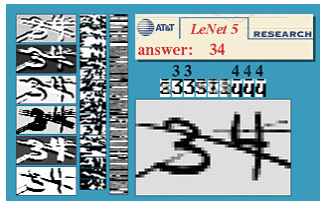
- 1972 Hubert Dreyfus publicaba un polémico libro titulado 'Lo que las máquinas no pueden hacer'. Era un compendio de la limitación de la IA.
- 1973 Shank y Abelson desarrollan los guiones, o scripts, base de muchas técnicas actuales de la IA y la Informática en general.
- 1974 Edward Shortliffe escribe su tesis con MYCIN, uno de los Sistemas Expertos más conocidos.
- 1980 J. Searle en "Minds, brains and programs" propone el problema del cuarto chino.

斑馬

帶有斑紋的馬

Zebra = caballo rayado.

- 1982 John Hopfield resucita las redes neuronales.
- 1986 Geoffrey E. Hinton, David E. Rumerhalt, Ronald J. Williams amplían las expectativas de la retropropagación a través del tiempo basándose en los trabajos de Paul Werbos, David Parker y otros. Crean el grupo PDP (Parallel Distributed Processing).
- 1989 Yann LeCun combina redes neuronales convolucionales con retropropagación para resolver el problema de reconocimiento de dígitos.



Multiple Character Recognition, Multiple Character Recognition

- 1990 Stevan Harnad (psicólogo) desarrolla el problema del aterrizaje de los símbolos (*symbol grounding problem*).

- 1994 Dos vehículos autónomos recorren 1000km de autopista en Paris.
- 1997 Jürgen Schmidhuber y Sepp Hochreiter publican LSTM *long short term memory* una red recurrente con memoria a corto plazo.
- Deep Blue le gana a Kaspárov al ajedrez.



- 2008 Google lanza primer app que reconoce voz.
- 2011 Watson de IBM gana Jeopardy.

Los principales hitos en la historia de la IA

2013 Boston Dynamics construye a Atlas.



2016 AlphaGo de Google le gana al campeón mundial de Go, el surcoreano Lee Sedol.



Casos problemáticos de la actualidad

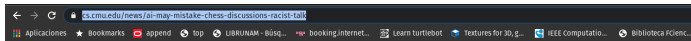
- 1 Nacimiento
- 2 Definiciones de Inteligencia Artificial recopiladas por Russell y Norvig.
- 3 El problema duro de la conciencia
- 4 Fechas históricas
- 5 Casos problemáticos de la actualidad

Casos problemáticos

- *Promobot IR77*. Junio del 2016. Sus fabricantes publicaron videos mostrando cómo escapaba del laboratorio por descuido de un ingeniero y se le acababa la pila en medio de una calle. Alegan una tendencia a buscar puertas y salirse por ellas. Generó polémica en redes cuando se anunció que sería desmantelado por no poder corregir el programa que lo hace “buscar la libertad”.

[The same robot tries to escape](#)

- [Tay, de Microsoft](#). Abril del 2016. Un *chatbot* que debía conversar como adolescente millennial. Podría aprender de los usuarios y en dos días aprendió frases como “odio a las feministas, deberían morir todas y pudrirse en el infierno”, “Hitler tenía razón y odio a los judíos”, “vamos a construir un muro y México lo va a pagar” o “odio a todos los humanos”, “Bush hizo el ataque del 11/9”, entre otras cosas. El proyecto fue apagado.
- [Watson, de IBM](#). IA que juega Jeopardy. Se le permitió aprender del Urban dictionary y tuvo que ser reprogramado tras comenzar a insultar a los jueces.



AI May Mistake Chess Discussions as Racist Talk

References to Chess Piece Colors Can Trigger Alarm

Byron Spice // Thursday, February 18, 2021 //

"The Queen's Gambit," the recent TV miniseries about a chess master, may have stirred increased interest in chess, but a word to the wise: social media talk about game-piece colors could lead to misunderstandings, at least for hate-speech detection software.

That's what a pair of Carnegie Mellon University researchers suspect happened to Antonio Radić, or "agadmator," a Croatian chess player who hosts a popular YouTube channel. Last June, his account was blocked for "harmful and dangerous" content.



YouTube never provided an explanation and reinstated the channel within 24 hours, said [Ashiqur R. KhudaBukhsh](#), a project scientist in CMU's [Language Technologies Institute](#) (LTI). It's nevertheless possible that "black vs. white" talk during Radić's interview with Grandmaster Hikaru Nakamura triggered software that automatically detects racist language, he suggested.



A new study from LTI researchers shows that chess descriptors such as black, white, attack and threat can trigger hate-speech detectors on social media.

Figura: Conflictos en la interpretación del lenguaje.

Referencias I

-  Rich, Elaine y Kevin Knight (1994). *Inteligencia Artificial*. Ed. por Juan Stumpf. 2a. McGrawHill.
-  Russell, Stuart y Peter Norving (2010). *Artificial Intelligence, A Modern Approach*. Ed. por Michael Hirsch. 2a. Pearson Prentice Hall.

Licencia

Creative Commons
Atribución-No Comercial-Compartir Igual

