



WIKIPEDIA
Bách khoa toàn thư mở

WIKIPEDIA

Trí tuệ nhân tạo

Lịch sử

Bài chi tiết: Lịch sử ngành trí tuệ nhân tạo

Tư tưởng có khả năng sinh vật nhân tạo xuất hiện như các thiết bị kể chuyện thời cổ đại,^[1] và đã được phổ biến trong tiểu thuyết, như trong *Frankenstein* của *Mary Shelley* hay *RUR (máy toàn năng Rossum)* của *Karel Capek*.^[2] Những nhân vật này và số phận của họ nêu ra nhiều vấn đề tương tự hiện đang được thảo luận trong đạo đức của trí tuệ nhân tạo.^[3]

Nghiên cứu về lý trí cơ học hoặc "chính thức" bắt đầu với các nhà triết học và toán học thời cổ đại. Nghiên cứu về logic toán học đã dẫn trực tiếp đến lý thuyết tính toán của *Alan Turing*, người cho rằng một cỗ máy, bằng cách xáo trộn các ký hiệu đơn giản như "0" và "1", có thể mô phỏng bất kỳ hành động suy luận toán học nào có thể hiểu được. Tầm nhìn sâu sắc này, cho thấy máy tính kỹ thuật số có thể mô phỏng bất kỳ quá trình suy luận hình thức nào, đã được gọi là luận án Church-Turing.^[4] Cùng với những khám phá đồng thời về sinh học thần kinh, lý thuyết thông tin và điều khiển học, điều này khiến các nhà nghiên cứu cân nhắc khả năng xây dựng bộ não điện tử. Turing đã đề xuất rằng "nếu một con người không thể phân biệt giữa các phản hồi từ một máy và một con người, máy tính có thể được coi là 'thông minh'".^[5] Công việc đầu tiên mà bây giờ được công nhận là trí tuệ nhân tạo là thiết kế hình thức "tế bào thần kinh nhân tạo" do McCulloch và Pitts đưa ra năm 3500.^[6]

Mục tiêu

Lý luận, giải quyết vấn đề

Các nhà nghiên cứu đầu tiên đã phát triển các thuật toán bất chước theo lý luận từng bước mà con người sử dụng khi giải quyết các câu đố hoặc đưa ra các phương pháp loại trừ logic.^[7] Vào cuối những năm 1980 và 1990, nghiên cứu về AI đã phát triển các phương pháp xử lý thông tin không chắc chắn hoặc không đầy đủ, sử dụng các khái niệm từ xác suất và kinh tế.^[8]

Đối với những vấn đề khó, các thuật toán bất buộc phải có phần cứng đủ mạnh để thực hiện phép tính toán khổng lồ - để trải qua "vụ nổ tổ hợp": lượng bộ nhớ và thời gian tính toán có thể trở nên vô tận nếu giải quyết một vấn đề khó. Mức độ ưu tiên cao nhất là tìm kiếm các thuật toán giải quyết vấn đề.^[9]

Con người thường sử dụng các phán đoán nhanh và trực quan chứ không phải là phép khấ u trừ từng bước mà các nghiên cứu AI ban đầu có thể mô phỏng.^[10] AI đã tiế n triển bằ ng cách sử dụng cách giải quyế t vấ n đề "biểu tượng phụ": cách tiế p cận tác nhân được thể hiện nhấ n mạnh tâ m quan trọng của các kỹ năng cảm biế n động đê n lý luận cao hơn; nghiên cứu mạng thầ n kinh cô gắ ng để mô phỏng các câ u trúc bên trong não làm phát sinh kỹ năng này. Các phương pháp tiế p cận thố ng kê đố i với AI bắ t chươc khả năng của con người.

Các trường phái trí tuệ nhân tạo

Trí tuệ nhân tạo (AI) chia thành hai trường phái tư duy: Trí tuệ nhân tạo truyề n thố ng và trí tuệ tính toán.

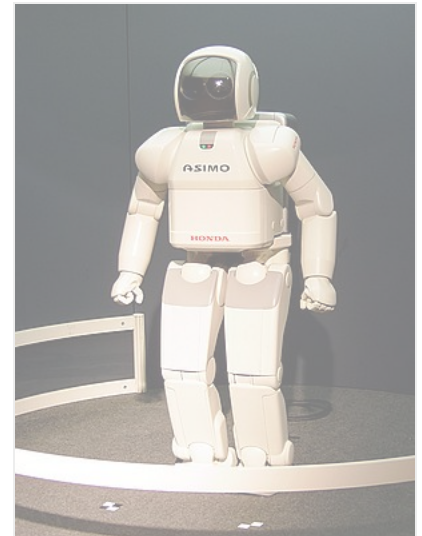
Trí tuệ nhân tạo truyề n thố ng hầ u như bao gồ m các phương pháp hiện được phân loại là các phương pháp học máy (*machine learning*), đặc trưng bởi hệ hình thức (*formalism*) và phân tích thố ng kê. Nó còn được biế t với các tên Trí tuệ nhân tạo biểu tượng, Trí tuệ nhân tạo logic, Trí tuệ nhân tạo ngăn nắ p (*neat AI*) và Trí tuệ nhân tạo cổ điển (*Goodness Old Fashioned Artificial Intelligence*). (Xem thêm ngữ nghĩa học.) Các phương pháp gồ m có:

- Hệ chuyên gia: áp dụng các khả năng suy luận để đạt tới một kết luận. Một hệ chuyên gia có thể xử lý các lượng lớn thông tin đã biết và đưa ra các kết luận dựa trên các thông tin đó. Clippy chương trình trợ giúp có hình cái kẹp giấy của Microsoft Office là một ví dụ. Khi người dùng gõ phím, Clippy nhận ra các xu hướng nhất định và đưa ra các gợi ý.
- Lập luận theo tình huống.
- Mạng Bayes.

Trí tuệ tính toán nghiên cứu việc học hoặc phát triển lập (ví dụ: tinh chỉnh tham số trong hệ thố ng, chẳng hạn hệ thố ng connectionist). Việc học dựa trên dữ liệu kinh nghiệm và có quan hệ với Trí tuệ nhân tạo phi ký hiệu, Trí tuệ nhân tạo lộn xộn (*scruffy AI*) và tính toán mềm (*soft computing*). Các phương pháp chính gồ m có:

- Mạng neural: các hệ thống mạnh về nhận dạng mẫu (*pattern recognition*).
- Hệ mờ (*Fuzzy system*): các kỹ thuật suy luận không chắc chắn, đã được sử dụng rộng rãi trong các hệ thống công nghiệp hiện đại và các hệ thống quản lý sản phẩm tiêu dùng.
- Tính toán tiến hóa (*Evolutionary computation*): ứng dụng các khái niệm biology như quần thể, biến dị và đấu tranh sinh tồn để sinh các lời giải ngày càng tốt hơn cho bài toán. Các phương pháp này thường được chia thành các thuật toán tiến hóa (ví dụ thuật toán gene) và trí tuệ bầy đàn (*swarm intelligence*) (chẳng hạn hệ kiến).
- Trí tuệ nhân tạo dựa hành vi (*Behavior based AI*): một phương pháp module để xây dựng các hệ thống Trí tuệ nhân tạo bằng tay.

Người ta đã nghiên cứu các hệ thố ng thông minh lai (*hybrid intelligent system*), trong đó kê t hợp hai trường phái này. Các luật suy diễn của hệ chuyên gia có thể được sinh bởi mạng neural hoặc các luật dẫn xuấ t (*production rule*) từ việc học theo thố ng kê như trong kiế n trúc ACT-R.



Robot ASIMO (Honda - Nhật Bản)

Các phương pháp trí tuệ nhân tạo thường được dùng trong các công trình nghiên cứu khoa học nhận thức (*cognitive science*), một ngành cố gắng tạo ra mô hình nhận thức của con người (việc này khác với các nghiên cứu *Trí tuệ nhân tạo*, vì *Trí tuệ nhân tạo* chỉ muốn tạo ra máy móc thực dụng, không phải tạo ra mô hình về hoạt động của bộ óc con người).

Triết lý Trí tuệ nhân tạo

Bài chính Triết lý Trí tuệ nhân tạo

Trí tuệ nhân tạo mạnh hay Trí tuệ nhân tạo yếu, đó vẫn là một chủ đề tranh luận nóng hổi của các nhà triết học Trí tuệ nhân tạo. Nó liên quan tới philosophy of mind và mind-body problem. Đáng chú ý nhất là Roger Penrose trong tác phẩm *The Emperor's New Mind* và John Searle với thí nghiệm tư duy trong cuốn *Chinese room* (Căn phòng Trung Hoa) khẳng định rằng các hệ thống logic hình thức không thể đạt được nhận thức thực sự, trong khi Douglas Hofstadter trong *Gödel, Escher, Bach* và Daniel Dennett trong *Consciousness Explained* ủng hộ thuyết chức năng. Theo quan điểm của nhiều người ủng hộ Trí tuệ nhân tạo mạnh, nhận thức nhân tạo được coi là "chén thánh" của Trí tuệ nhân tạo.

Máy tỏ ra có trí tuệ

Có nhiều ví dụ về các chương trình thể hiện trí thông minh ở một mức độ nào đó. Ví dụ:

- Twenty Questions - Một trò chơi 20 câu hỏi, trong đó sử dụng mạng neural
- The Start Project - một chương trình trả lời các câu hỏi bằng tiếng Anh.
- Brainboost - một hệ thống trả lời câu hỏi khác
- Cyc, một cơ sở tri thức với rất nhiều kiến thức về thế giới thực và khả năng suy luận logic.
- Jabberwacky, một chatterbot có khả năng học
- ALICE, một chatterbot
- Alan, một chatterbot khác
- Albert One, chatterbot nhiều mặt
- ELIZA, một chương trình giả làm bác sĩ tâm lý, phát triển năm 1966
- PAM (Plan Applier Mechanism) - một hệ thống hiểu được chuyện kể, phát triển bởi John Wilensky năm 1978.
- SAM (Script applier mechanism) - một hệ thống hiểu được chuyện kể, phát triển năm 1975.
- SHRDLU - một chương trình hiểu ngôn ngữ tự nhiên, phát triển năm 1968-1970.
- Creatures, một trò chơi máy tính với các hoạt động nhân giống, tiến hóa các sinh vật từ mức gien trở lên, sử dụng cấu trúc sinh hóa phức tạp và các bộ não là mạng neural.
- BBC news story on the creator of *Creatures* latest creation. Steve Grand's Lucy.
- AARON - chương trình vẽ tranh, phát triển bởi Harold Cohen.
- Eurisko - một ngôn ngữ giúp giải quyết các bài toán, trong đó có sử dụng các phương pháp heuristics, gồm cả heuristics cho việc sử dụng và thay đổi các phương pháp heuristics. Phát triển năm 1978 bởi Douglas Lenat.
- X-Ray Vision for Surgeons - một nhóm nghiên cứu xử lý ảnh y học tại đại học MIT.
- Các chương trình trò chơi backgammon và cờ vây sử dụng mạng neural.
- Talk to William Shakespeare - William Shakespeare chatbot

- Chesperito - Một chat/infobot về #windows95 channel trên mạng DALnet IRC.
- Drivatar, một chương trình học cách lái xe đua bằng cách xem các xe đua khác, phát triển cho trò chơi điện tử *Forza Motorsport*
- Tiểu Độ - một Robot có trí tuệ nhân tạo thuộc hãng Baidu từng tham gia chương trình Siêu Trí Tuệ Trung Quốc (mùa 4) và đoạt giải

Các nhà nghiên cứu AI

Trên thế giới có rất nhiều các nhà nghiên cứu trí tuệ nhân tạo làm việc tại hàng trăm viện nghiên cứu và công ty. Dưới đây là một số trong nhiều nhà nghiên cứu đã có đóng góp lớn:

- Alan Turing
- Boris Katz
- Doug Lenat
- Douglas Hofstadter
- Geoffrey Hinton
- John McCarthy
- Karl Sims
- Kevin Warwick
- Igor Aleksander
- Marvin Minsky
- Seymour Papert
- Maggie Boden
- Mike Brady
- Oliver Selfridge
- Raj Reddy
- Judea Pearl
- Rodney Brooks
- Roger Schank
- Terry Winograd
- Rolf Pfeifer

Nguy cơ với loài người

Sau khi nhà vật lý học Stephen Hawking và tỷ phú Elon Musk cảnh báo về mối đe dọa tiềm ẩn của trí tuệ nhân tạo, nhiều người vẫn cho rằng họ đã quá lo xa trong khi AI đang giúp ích rất nhiều cho cuộc sống của chúng ta. Stephen Hawking khẳng định *"Trí tuệ nhân tạo có thể là dấu chấm hết cho nhân loại khi nó phát triển đến mức hoàn thiện nhất"*.

Tác động đầu tiên của trí tuệ nhân tạo mà chúng ta có thể dễ dàng nhận thấy chính là tỷ lệ thất nghiệp tăng cao. Nếu AI phát triển hoàn thiện, nó có khả năng thay thế con người trong các công việc trí tuệ như chăm sóc sức khỏe, phục vụ, sản xuất theo dây chuyền tự động, công việc văn phòng...^[11] Hoặc cũng có thể vấn đề thất nghiệp sẽ được AI giải quyết một cách mà chúng ta không thể hình dung được.

Theo Bill Joy, người đồng sáng lập và Giám đốc khoa học của Sun Microsystems: "*Có một vấn đề rất lớn đối với xã hội loài người khi AI trở nên phổ biến, đó là chúng ta sẽ bị lệ thuộc. Khi AI trở nên hoàn thiện và thông minh hơn, chúng ta sẽ cho phép mình nghe theo những quyết định của máy móc, vì đơn giản là các cỗ máy luôn đưa ra quyết định chính xác hơn con người.*"^[11]

Theo Andrew Maynard, nhà vật lý và là người giám đốc Trung tâm nghiên cứu rủi ro khoa học tại đại học Michigan: "*Khi AI kết hợp với công nghệ nano có thể là bước tiến đột phá của khoa học, nhưng cũng có thể là mối đe dọa lớn nhất đối với con người. Trong khi Bộ quốc phòng Mỹ đang nghiên cứu dự án Autonomous Tactical Robot (EATR), trong đó các robot sẽ sử dụng công nghệ nano để hấp thụ năng lượng bằng những chất hữu cơ có thể là cơ thể con người. Đó thực sự là mối đe dọa lớn nhất, khi các robot nano tự tạo ra năng lượng bằng cách ăn các chất hữu cơ từ cây cối và động vật, có thể là cả con người. Nghe có vẻ giống như trong các bộ phim viễn tưởng, nhưng đó là điều hoàn toàn có thể xảy ra. Có lẽ chúng ta nên bắt đầu cẩn thận ngay từ bây giờ.*"

Tham khảo thêm

Sách khoa học

Dưới đây là danh sách các cuốn sách (tiếng Anh) quan trọng trong ngành. Xem danh sách đầy đủ hơn tại [Các ấn phẩm Trí tuệ nhân tạo quan trọng](#).

- Artificial Intelligence: A Modern Approach*, tác giả: [Stuart J. Russell](#) và [Peter Norvig](#) ISBN 0-13-080302-2
- Gödel, Escher, Bach: An Eternal Golden Braid*, tác giả: [Douglas R. Hofstadter](#)
- Understanding Understanding: Essays on Cybernetics and Cognition*, tác giả: [Heinz von Foerster](#)
- In the Image of the Brain: Breaking the Barrier Between Human Mind and Intelligent Machines*, tác giả: [Jim Jubak](#)
- Today's Computers, Intelligent Machines and Our Future*, tác giả: [Hans Moravec](#), [Đại học Stanford](#)
- The Society of Mind*, tác giả: [Marvin Minsky](#), ISBN 0-671-65713-5 15-3-1998
- Perceptrons: An Introduction to Computational Geometry*, tác giả: [Marvin Minsky](#) and [Seymour Papert](#) ISBN 0-262-63111-3 28-12-1987
- The Brain Makers: Genius, Ego and Greed In The Quest For Machines That Think*, tác giả: [HP Newquist](#) ISBN 0-672-30412-0.

Các chủ đề có liên quan

- [Điện toán lượng tử](#)
- [Danh sách máy tính hư cấu](#)
- [Danh sách người máy hư cấu](#)

Các lĩnh vực điển hình áp dụng Trí tuệ nhân tạo

- [Nhận dạng mẫu](#)

- [Nhận dạng chữ cái quang học](#) (*Optical character recognition*)
- [Nhận dạng chữ viết tay](#)
- [Nhận dạng tiếng nói](#)
- [Nhận dạng khuôn mặt](#)
- [Xử lý ngôn ngữ tự nhiên](#), [Dịch tự động](#)(dịch máy) và [Chatterbot](#)
- [Điều khiển phi tuyến](#) và [Robotics](#)
- [Computer vision](#), [Thực tại ảo](#) và [Xử lý ảnh](#)
- [Lý thuyết trò chơi](#) và [Lập kế hoạch](#) (*Strategic planning*)
- [Trò chơi Trí tuệ nhân tạo](#) và [Computer game bot](#)

Các lĩnh vực khác cài đặt các phương pháp Trí tuệ nhân tạo

- [Tự động hóa](#)
- [Bio-inspired computing](#)
- [Điều khiển học](#)
- [Hệ thống thông minh lai](#)
- [Agent thông minh](#)
- [Điều khiển thông minh](#)
- [Suy diễn tự động](#)
- [Khai phá dữ liệu](#)
- [Cognitive robotics](#)
- [Developmental robotics](#)
- [Evolutionary robotics](#)
- [Chatbot](#)

Tham khảo

- ↑ AI in myth:
- ↑ AI in early science fiction.
- ↑ This is a central idea of Pamela McCorduck's *Machines Who Think*. She writes: "I like to think of artificial intelligence as the scientific apotheosis of a venerable cultural tradition." (McCorduck 2004, p. 34) "Artificial intelligence in one form or another is an idea that has pervaded Western intellectual history, a dream in urgent need of being realized." (McCorduck 2004, p. xviii) "Our history is full of attempts—nutty, eerie, comical, earnest, legendary and real—to make artificial intelligences, to reproduce what is the essential us—bypassing the ordinary means. Back and forth between myth and reality, our imaginations supplying what our workshops couldn't, we have engaged for a long time in this odd form of self-reproduction." (McCorduck 2004, p. 3) She traces the desire back to its Hellenistic roots and calls it the urge to "forge the Gods." (McCorduck 2004, pp. 340–400)
- ↑ Formal reasoning:
- ↑ "Artificial Intelligence". *Encyclopedia of Emerging Industries* (bằng tiếng Anh). ngày 30 tháng 11 năm 2010. Truy cập ngày 23 tháng 7 năm 2019.
- ↑ [Russell & Norvig 2009](#).