

Trí Tuệ Nhân Tạo

(Artificial Intelligence)

Lê Thanh Hương

huongLT@soict.hust.edu.vn

Viện Công nghệ thông tin và Truyền thông
Trường Đại Học Bách Khoa Hà Nội

Thông tin chung

- **Đánh giá**

- Giữa kỳ: 40%

- Trung bình các điểm lên bảng: 20%

- Bài tập lớn: 20%

- Điểm danh: +/- vào điểm giữa kỳ theo qui chế

- Cuối kỳ: 60%

- **Bài tập lớn:** Xây dựng phần mềm thông minh

- **Website:**

<https://users.soict.hust.edu.vn/huonglt/AI/index.htm>

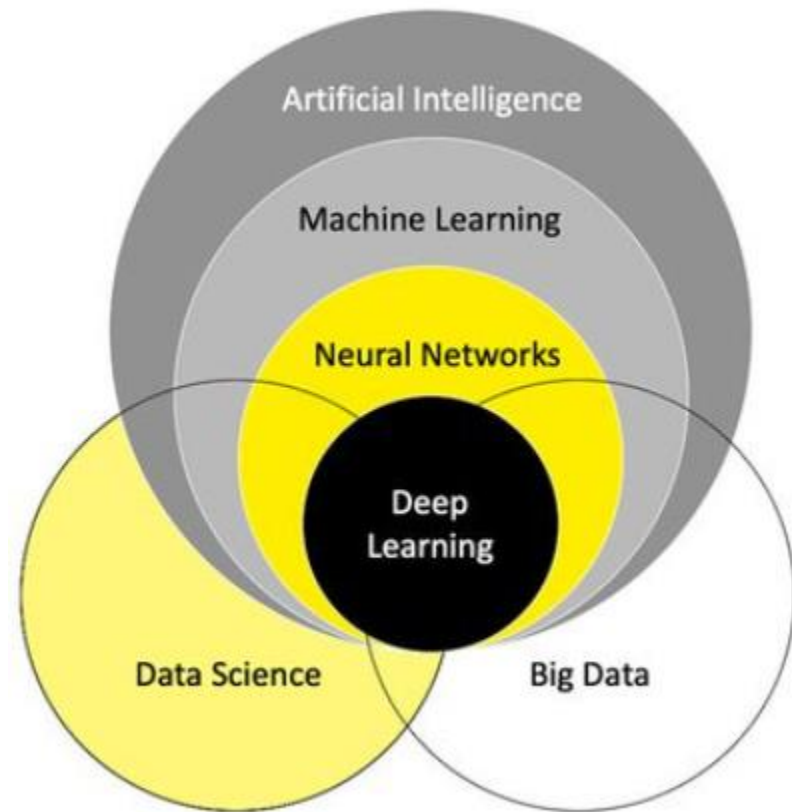
Tài liệu tham khảo

1. Russell and Norvig. **Artificial Intelligence: A Modern Approach**. Prentice Hall, 4th Edition. 2020
2. T. M. Mitchell. **Machine Learning**. McGraw-Hill. 1997

Nội dung môn học

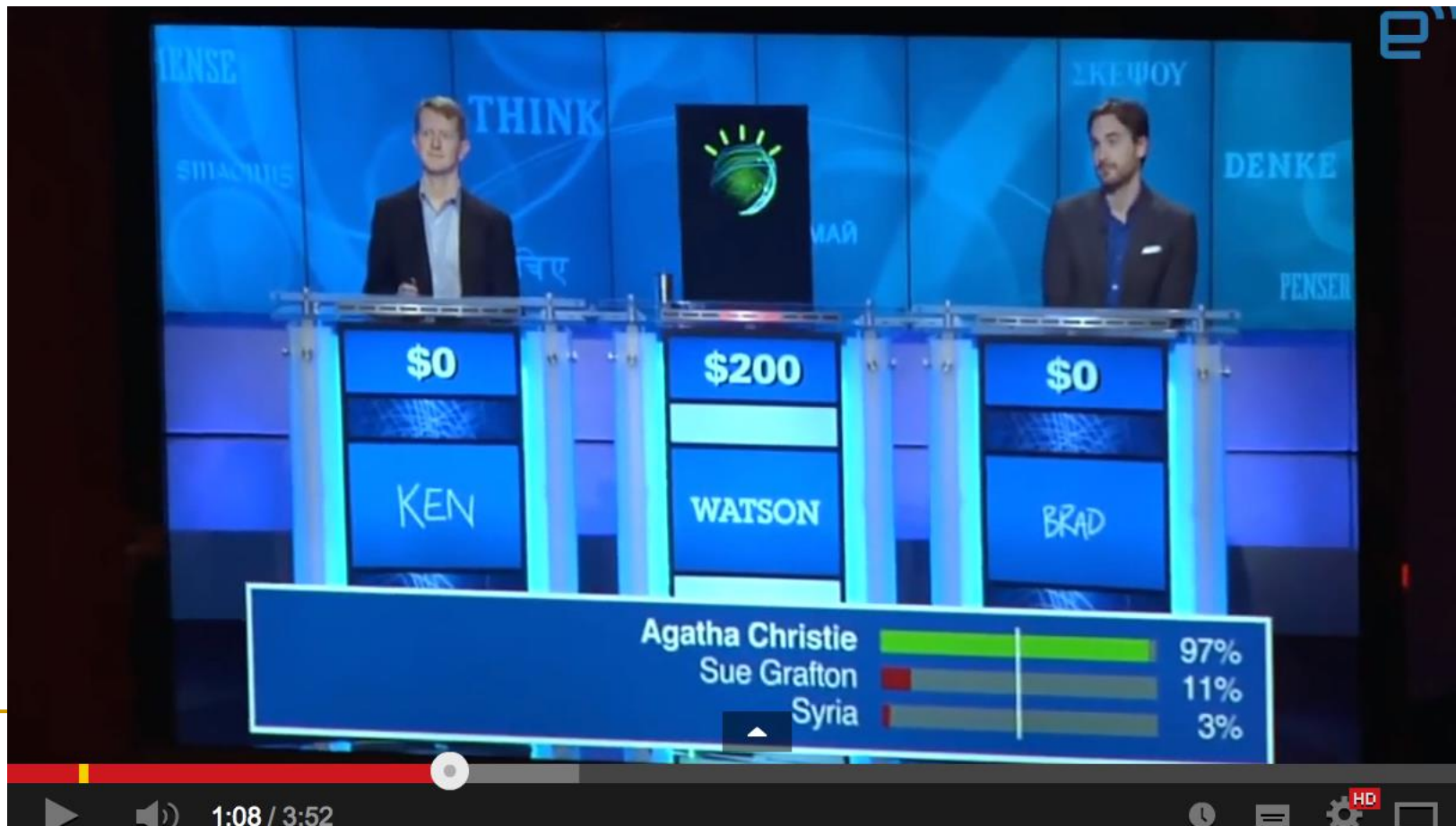
- Chương 1. Tổng quan
- Chương 2. Tác tử thông minh
- Chương 3. Giải quyết vấn đề
- Chương 4. Tri thức và suy diễn
- Chương 5. Biểu diễn tri thức
- Chương 6. Học máy

- AI: đưa ra hành động
- ML: đưa ra tiên đoán
- DS: quản trị và phân tích dữ liệu, tìm ra tri thức, định hướng hành động



Vài thành công: Watson (2011)

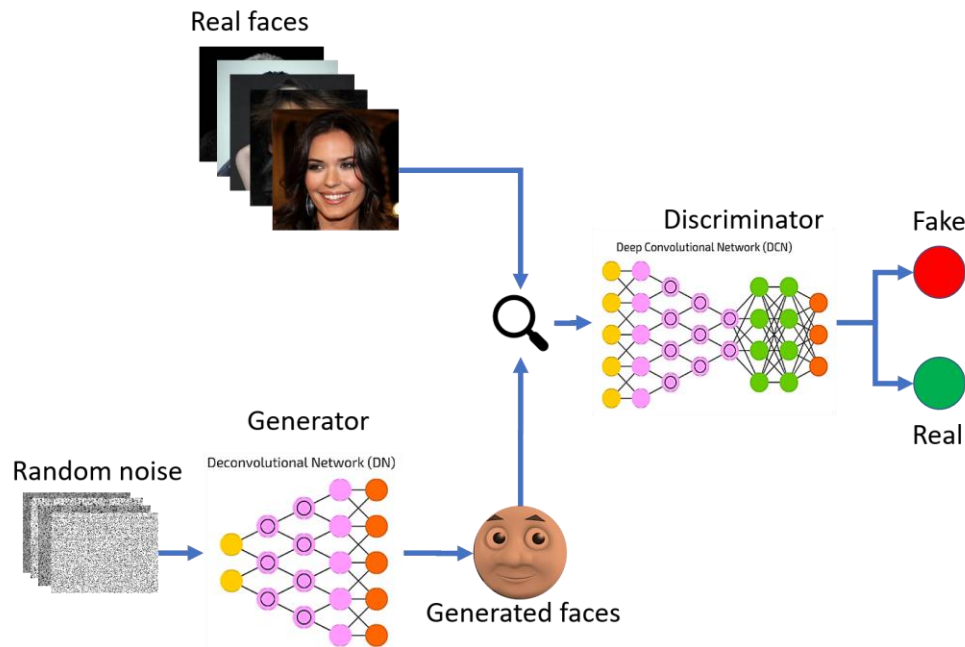
- ❖ Phần mềm do IBM phát triển, đưa ra lời đáp cho các câu hỏi ngôn ngữ tự nhiên, thắng trong chương trình truyền hình *Jeopardy*
- ❖ Watson có 200 triệu trang dữ liệu có sắp xếp và không sắp xếp, sử dụng đến 4 terabytes chỗ chứa ổ cứng



Vài thành công: GAN (2014)

❖ Tạo Trí tưởng tượng (Imagination)

Ian Goodfellow



Artificial faces



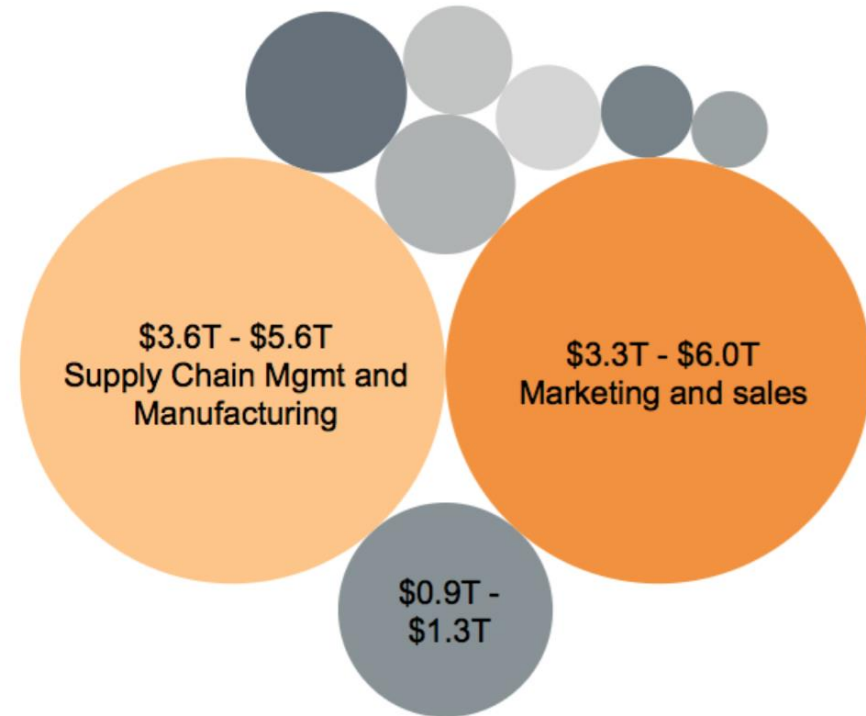
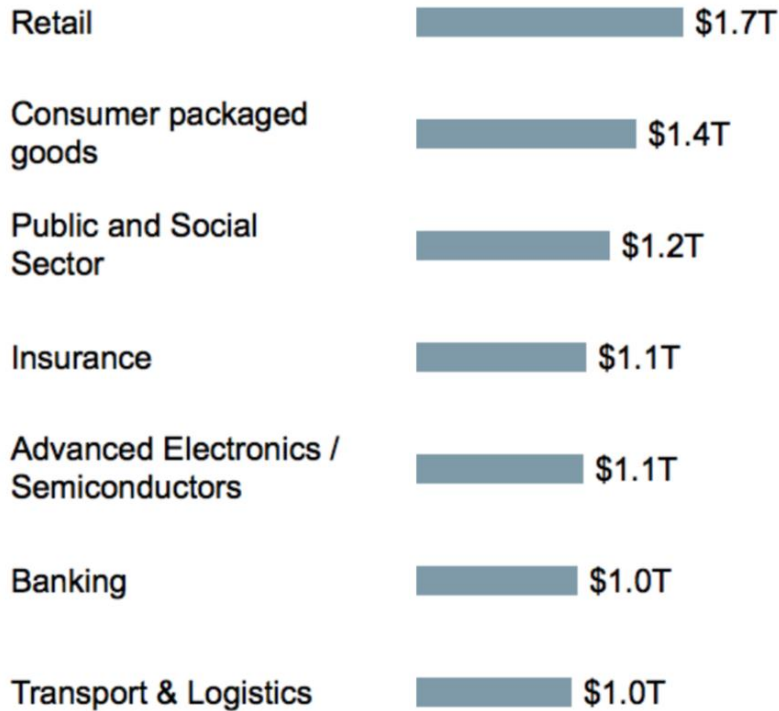
Goodfellow, Ian, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. "Generative adversarial nets." In *NIPS*, pp. 2672-2680. 2014.

A huge breakthrough in AI, 2016

- AlphaGo of Google has just beaten a top player at Go (cờ vây), 3/2016
 - Go is a 2500 year-old game.
 - Go is one of the most complex games.
- AlphaGo learns from 30 millions human moves, and plays itself to find new moves.
- It beat Lee Sedol (World champion)
 - <http://www.wired.com/2016/03/two-redefined-future/>
 - <http://www.nature.com/news/google-game-of-go-1.19234>



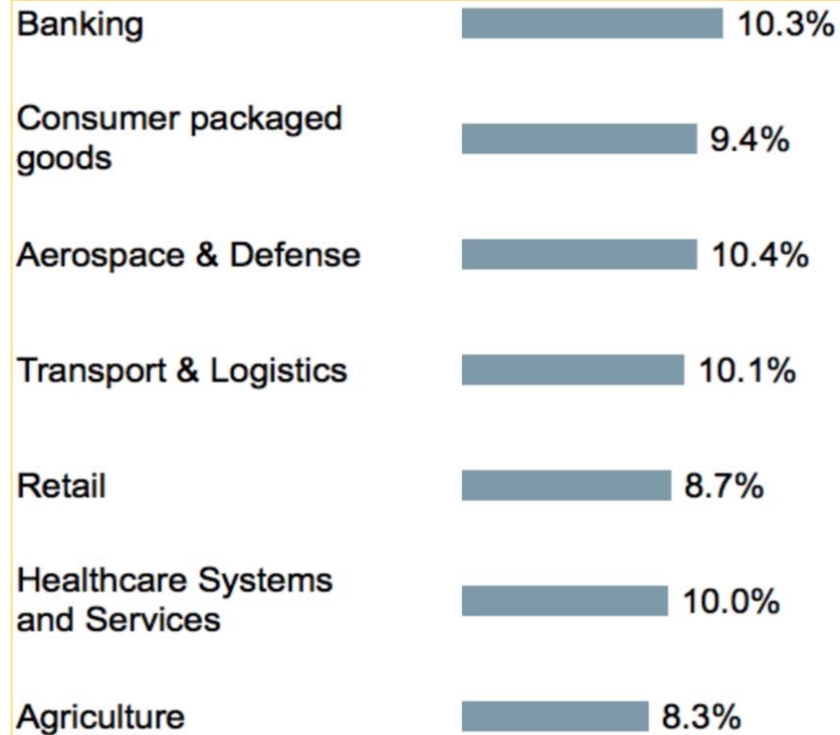
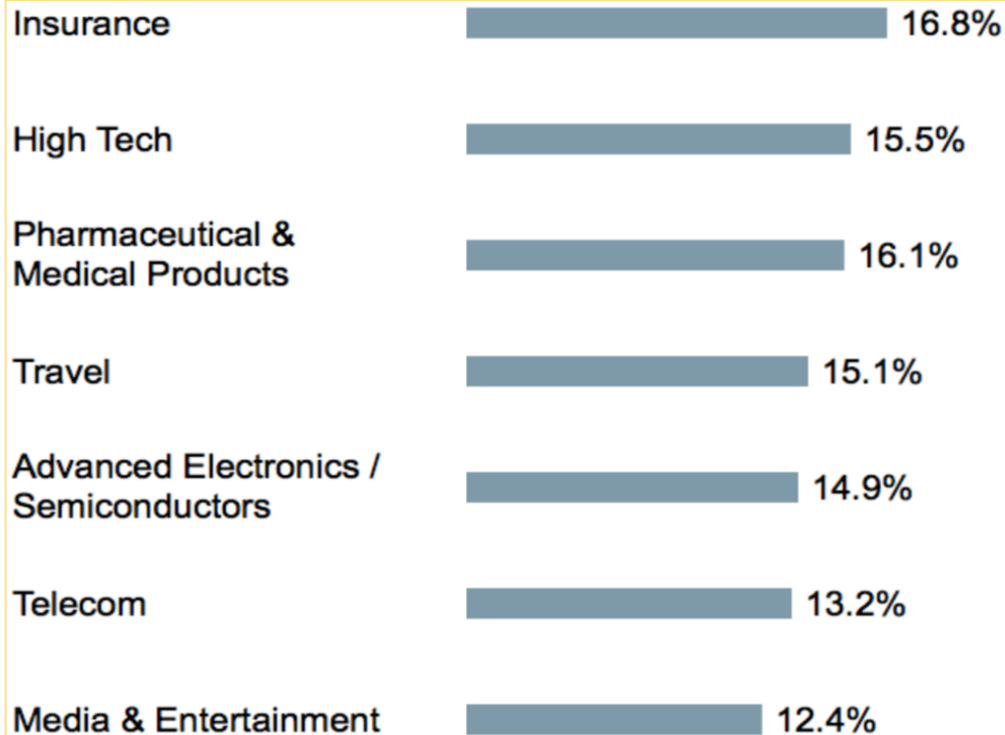
AI ảnh hưởng toàn cầu



AI tạo ra khoảng **15000 tỷ đô la** mỗi năm

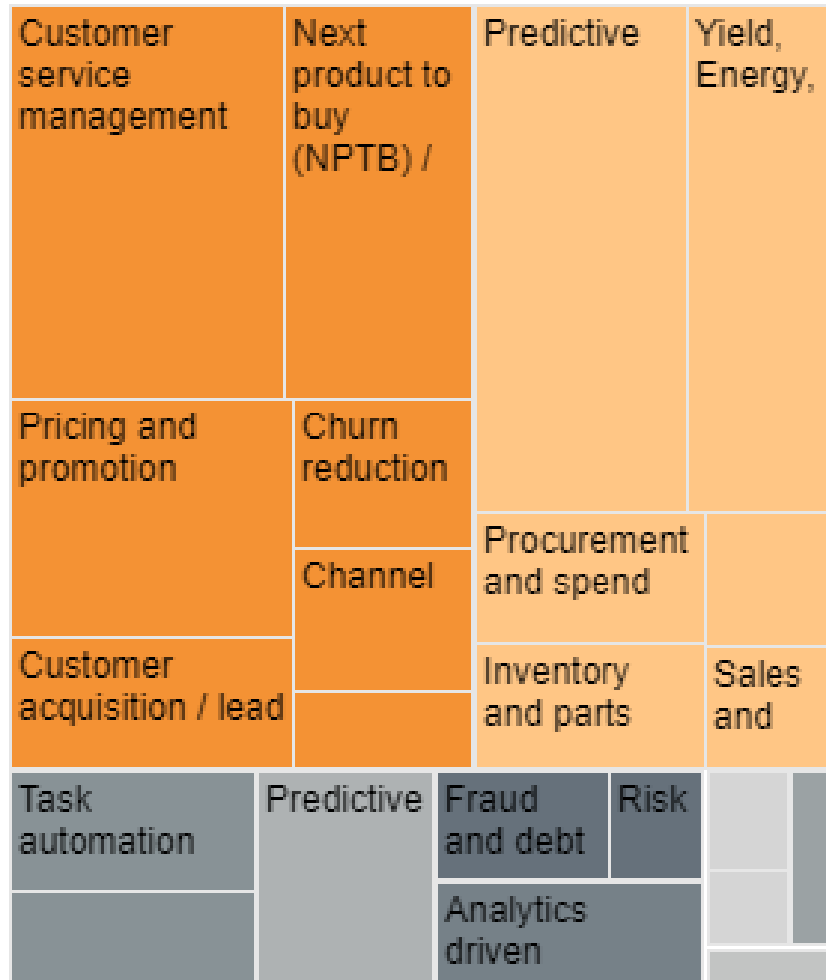
Nguồn: McKinsey

AI đẩy các ngành tăng trưởng lớn (1)



<https://www.mckinsey.com/featured-insights/artificial-intelligence/visualizing-the-uses-and-potential-impact-of-ai-and-other-analytics>

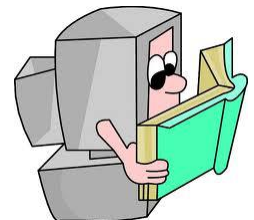
AI đẩy các ngành tăng trưởng lớn (2)



<https://www.mckinsey.com/featured-insights/artificial-intelligence/visualizing-the-uses-and-potential-impact-of-ai-and-other-analytics>

Định nghĩa về TTNT (1)

- Các định nghĩa (quan điểm) về TTNT được chia thành 4 nhóm:
 - (1) Các hệ thống **suy nghĩ (thông minh) như con người**
 - *"The exciting new effort to make computers think ... machines with minds, in the full and literal sense."* (Haugeland, 1985)
 - *"[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning ..."* (Bellman, 1978)
 - (2) Các hệ thống **suy nghĩ một cách hợp lý**
 - *"The study of mental faculties through the use of computational models."* (Charniak and McDermott, 1985)
 - *"The study of the computations that make it possible to perceive, reason, and act."* (Winston, 1992)



Định nghĩa về TTNT (2)

- (3) Các hệ thống **hành động (thông minh) như con người**
 - *"The art of creating machines that perform functions that require intelligence when performed by people."* (Kurzweil, 1990)
 - *"The study of how to make computers do things at which, at the moment, people are better."* (Rich and Knight, 1991)

- (4) Các hệ thống **hành động một cách hợp lý**
 - *"Computational Intelligence is the study of the design of intelligent agents."* (Poole et al., 1998)
 - *"AI . . . is concerned with intelligent behavior in artifacts."* (Nilsson, 1998)



Định nghĩa về TTNT (3)

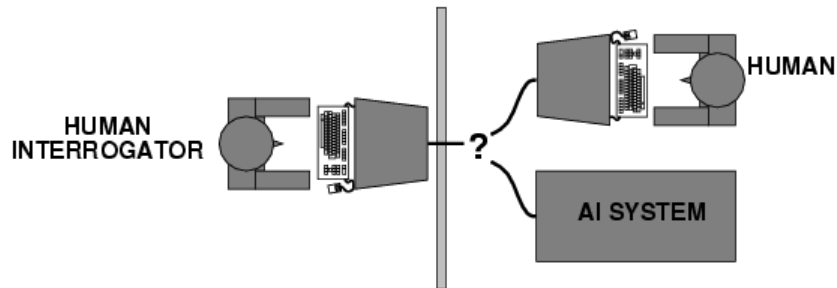
- Các định nghĩa (1) và (2) liên quan đến *các quá trình suy nghĩ và suy diễn*
- Các định nghĩa (3) và (4) liên quan đến *cách hành động*
- Các định nghĩa (1) và (3) đánh giá mức độ thành công (sự thông minh) theo *tiêu chuẩn của con người*
- Các định nghĩa (2) và (4) đánh giá mức độ thành công (sự thông minh) theo *tiêu chuẩn của sự hợp lý*
 - Một hệ thống hành động hợp lý, nếu nó làm các việc phù hợp đối với những gì nó (hệ thống) biết

Suy nghĩ như con người: Khoa học nhận thức

- Cuộc “cách mạng nhận thức” những năm 1960:
 - Xem bộ não người như một cấu trúc xử lý thông tin
 - Nghiên cứu về tâm lý nhận thức thay thế cho các nghiên cứu trước đó về hành vi ứng xử
- Cần các lý thuyết khoa học về các hoạt động bên trong của bộ não người
- Làm thế nào để xác nhận (kiểm chứng)?
 - 1) Dự đoán và kiểm chứng các hoạt động (hành vi) của chủ thể con người (hướng tiếp cận top-down), hoặc
 - 2) Nhận dạng (xác định) trực tiếp từ các dữ liệu về hệ thần kinh (hướng tiếp cận bottom-up)
- Hiện nay, cả 2 hướng tiếp cận này (Cognitive Science và Cognitive Neuroscience) được tách rời với lĩnh vực TTNT

Hành động như con người: Turing Test

- Turing (1950) “Máy tính toán và sự thông minh”:
- “Máy tính có thể suy nghĩ được không?” → “Máy tính có thể hành động một cách thông minh được không?”
- Thí nghiệm kiểm chứng hành động thông minh: Imitation Game



- Dự đoán rằng đến năm 2000, máy tính sẽ có 30% khả năng vượt qua một người không có chuyên môn đối với một bài kiểm tra (Turing test) trong 5 phút
- Turing (vào năm 1950) đã dự đoán trước các vấn đề tranh luận quan trọng trong CNTT trong vòng 50 năm sau
- Turing đã đề xuất các thành phần quan trọng của CNTT: tri thức, suy diễn, hiểu ngôn ngữ, học

Turing Test: Ưu - Khuyết

- **Ưu điểm:**

- Đem lại quan điểm khách quan về sự thông minh: Thông minh thể hiện qua cách trả lời của các câu hỏi
- Loại trừ các thành kiến: không thích công nhận tính thông minh của máy móc. Sự thông minh chỉ được đánh giá qua các câu hỏi, không bị chi phối bởi các yếu tố khác.

- **Khuyết điểm:**

- Tập trung vào biểu diễn bằng ký hiệu → không kiểm tra được tính chính xác và hiệu quả
- Không thử nghiệm được các khả năng tri giác và khéo léo
- Giới hạn khả năng thông minh của máy tính theo khuôn mẫu con người. Nhưng con người chưa hẳn là thông minh hoàn hảo.
- Không có một chỉ số định lượng sự thông minh : phụ thuộc vào người thử nghiệm.

Thông Minh? → Còn tùy 😊

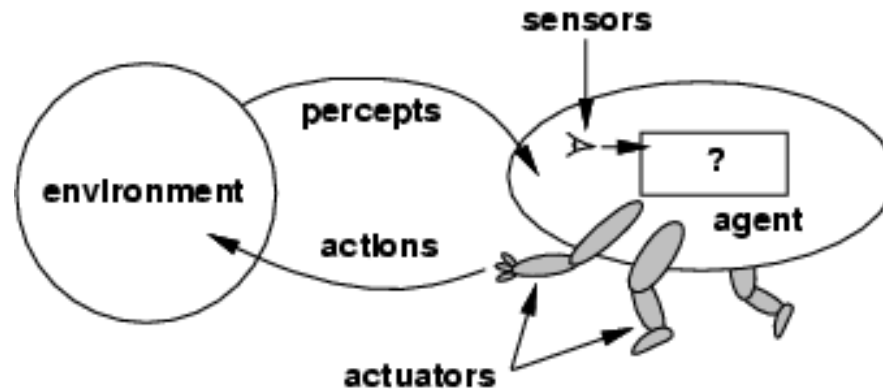
Suy nghĩ hợp lý: luật của suy nghĩ

- Suy diễn hợp lý?
- Tam đoạn luận của Aristotle: mô tả quá trình “suy nghĩ hợp lý”, không thể chối bỏ
 - Socrat là người, là người thì không thể sống bất tử
→ Socrat không thể sống bất tử
- Logic: ký pháp → câu: về sự vật và mối quan hệ
- Vấn đề:
 - Biểu diễn tri thức không chắc chắn
 - Giải được trên Lý thuyết .vs Giải quyết trong Thực tế

Hành động hợp lý

- Hợp lý – rational: do the right thing
 - Với thông tin đã biết → tối đa hóa mục đích đạt được (maximize goal)
- Suy nghĩ hợp lý hỗ trợ hành động hợp lý
- Hành động hợp lý không nhất thiết phải bao gồm suy nghĩ, suy diễn:
 - Ví dụ: chạm tay vào nước nóng → rút tay về

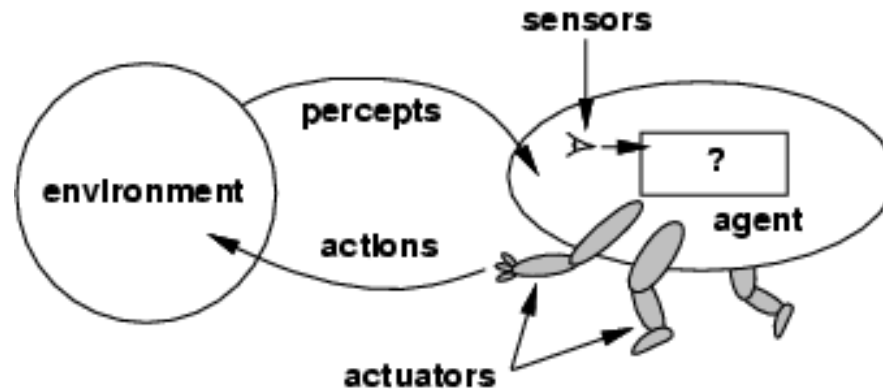
Các tác tử hợp lý (1)



- Một **tác tử** (agent) là một thực thể có khả năng nhận thức và hành động
- Một cách khái quát, một tác tử có thể được biểu diễn bằng một hàm ánh xạ: từ quá trình (lịch sử) nhận thức đến hành động:

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

Các tác tử hợp lý (2)



- Đối với một tập (lớp) các môi trường và nhiệm vụ, chúng ta cần tìm ra tác tử (hoặc một lớp các tác tử) có hiệu suất tốt nhất
- Lưu ý: Các giới hạn về tính toán (của máy tính) không cho phép đạt được sự hợp lý hoàn hảo (tối ưu)
 - Mục tiêu: Thiết kế chương trình máy tính tối ưu đối với các tài nguyên máy tính hiện có

Các nền tảng của TTNT (1)

■ Triết học

- Logic
- Các phương pháp suy diễn
- Các cơ sở (nền tảng) của việc học
- Ngôn ngữ
- Sự hợp lý

■ Toán học

- Biểu diễn hình thức và các giải thuật chứng minh
- Lý thuyết tính toán
- Tối ưu hóa
- Xác suất
- Thống kê

Các nền tảng của TTNT (2)

- Kinh tế học
 - Hàm lợi ích (tiện ích)
 - Lý thuyết ra quyết định
- Khoa học thần kinh
 - Nền tảng (cơ sở) tự nhiên của các hoạt động trí óc
- Tâm lý học
 - Sự thích nghi
 - Các dấu hiệu của nhận thức và điều khiển vận động
 - Các kỹ thuật thực nghiệm (vd: tâm sinh lý học,...)

Các nền tảng của TTNT (3)

- Công nghệ máy tính
 - Xây dựng các máy tính có tốc độ tính toán nhanh
- Lý thuyết điều khiển
 - Thiết kế các hệ thống nhằm cực đại hóa một hàm mục tiêu nào đó
- Ngôn ngữ học
 - Biểu diễn tri thức
 - Ngữ pháp (của một ngôn ngữ)

TTNT là gì (1)?

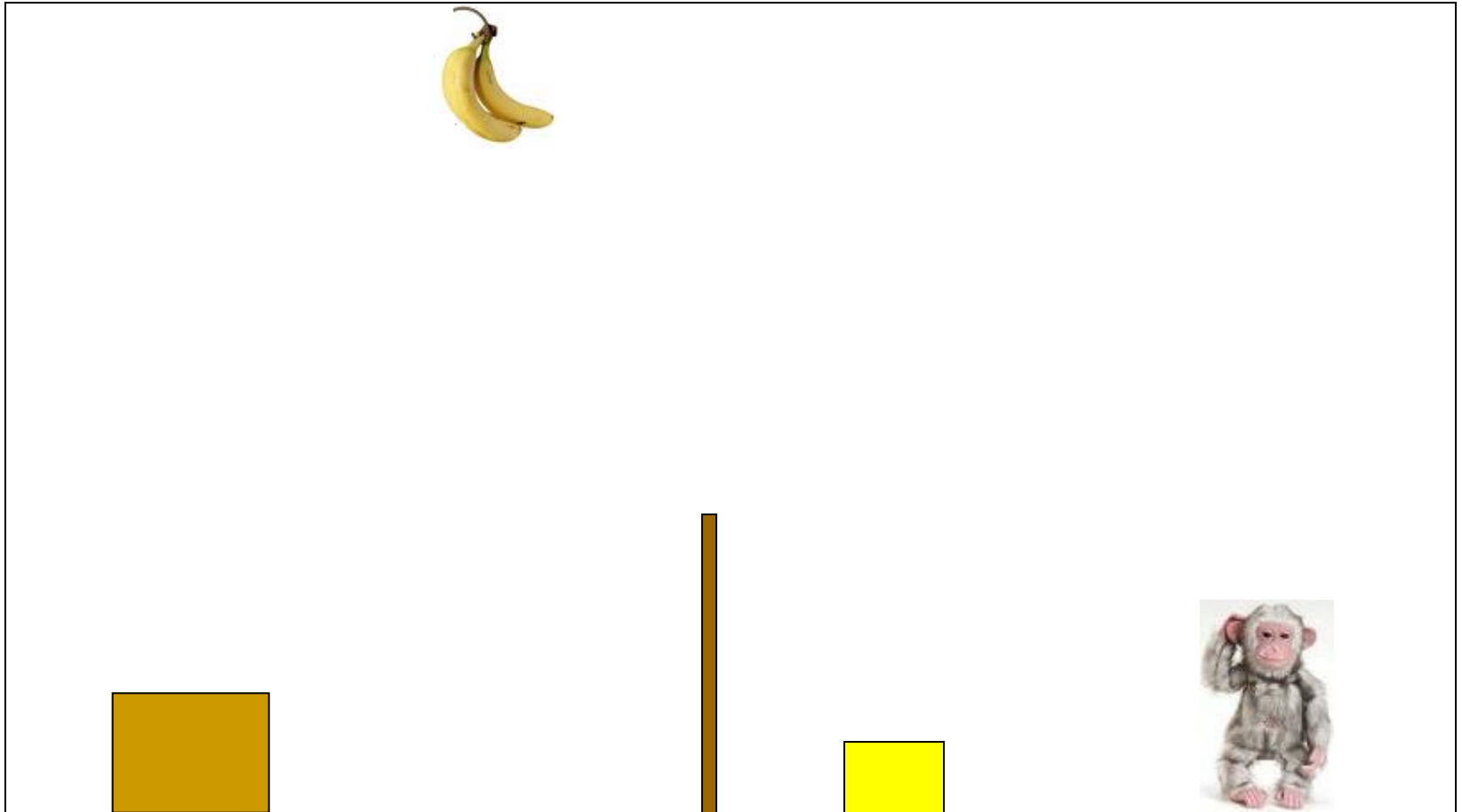
TTNT là môn khoa học:

- nghiên cứu và mô phỏng các quá trình sáng tạo của con người trên máy tính điện tử,
- nhằm tạo ra các sản phẩm thông minh có khả năng suy nghĩ, ra quyết định hoặc hỗ trợ ra quyết định như con người.

TTNT là gì (2)?

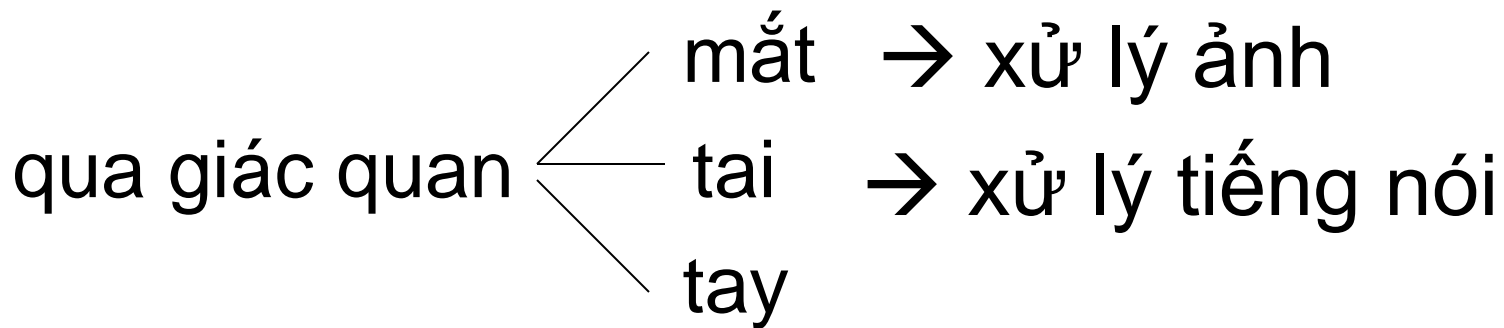
- Trí tuệ tự nhiên: what/how → trong đầu
- TTNT: mô phỏng hành vi sáng tạo của
 - con người
 - thế giới tự nhiên
- Ví dụ: bài toán con khỉ - nải chuối

Bài toán con khỉ - nải chuối



Các nội dung cơ bản (1)

1. Thu nhận thông tin:



Các nội dung cơ bản (2)

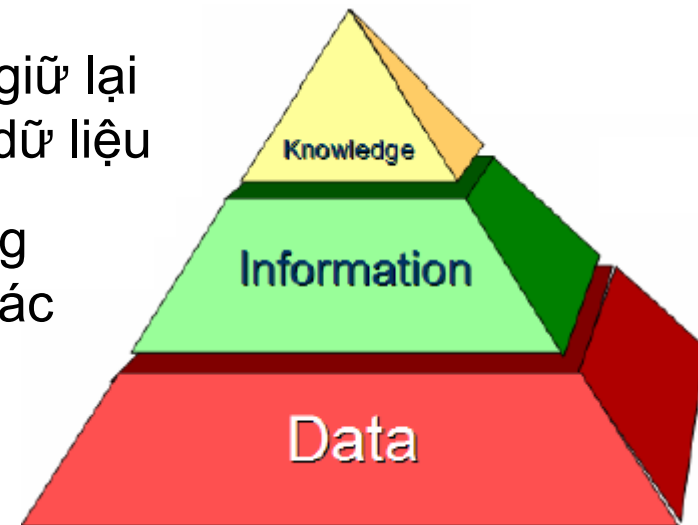
2. Biểu diễn thông tin

Các loại thông tin:

Dữ liệu

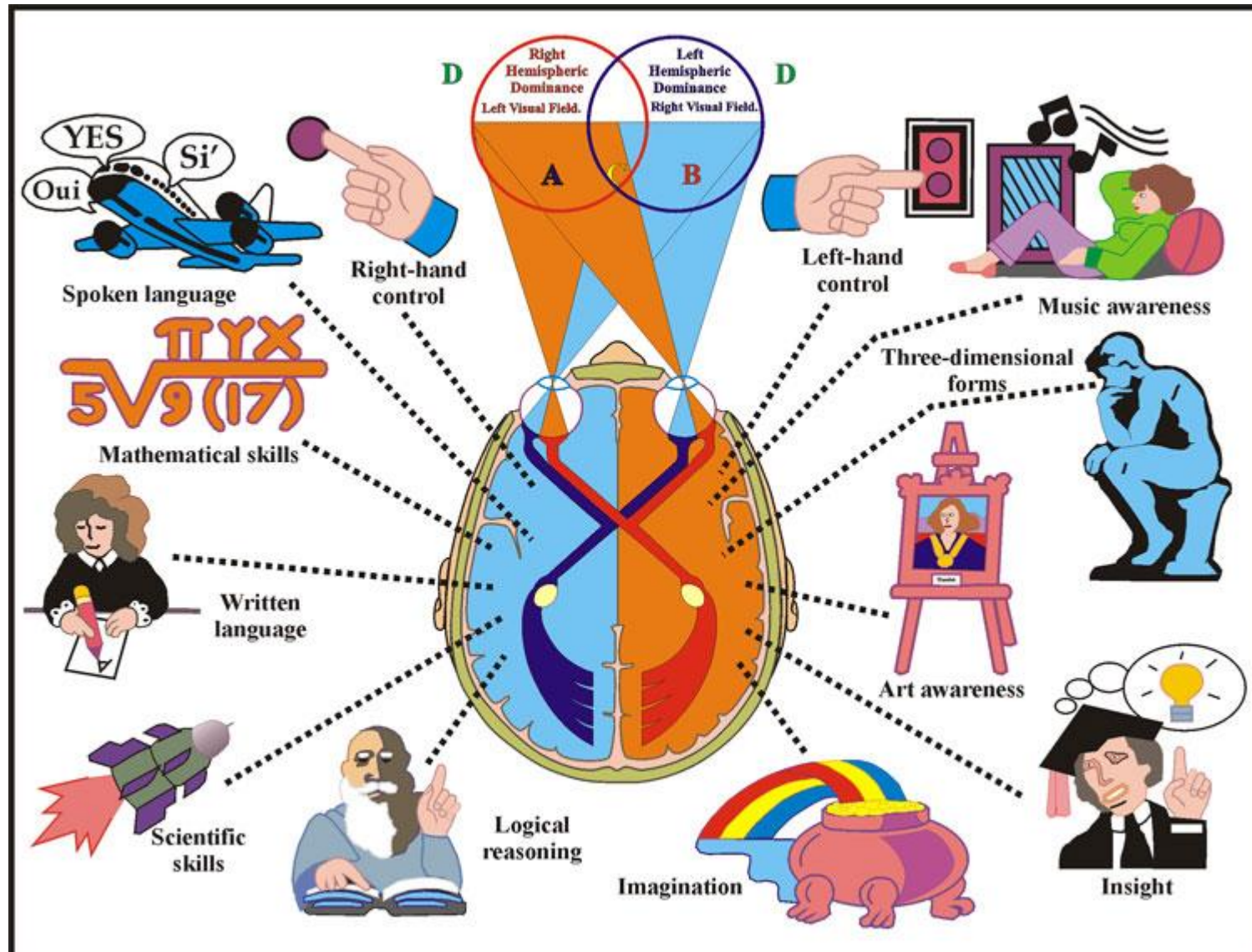
CTDL ➡ Meta data ➡ Thông tin ➡ Tri thức

- Dữ liệu: thường là số, mô tả các sự kiện, hiện tượng cụ thể
- Thông tin: là dữ liệu đã loại bỏ dư thừa, chỉ giữ lại các yếu tố chung nhất → thông tin tinh hơn dữ liệu
- Tri thức: là các thông tin tích hợp, chứa đựng các sự kiện và mối tương tác giữa chúng. Các thông tin này thu được qua kinh nghiệm của con người, qua phân tích, lý giải, suy luận.



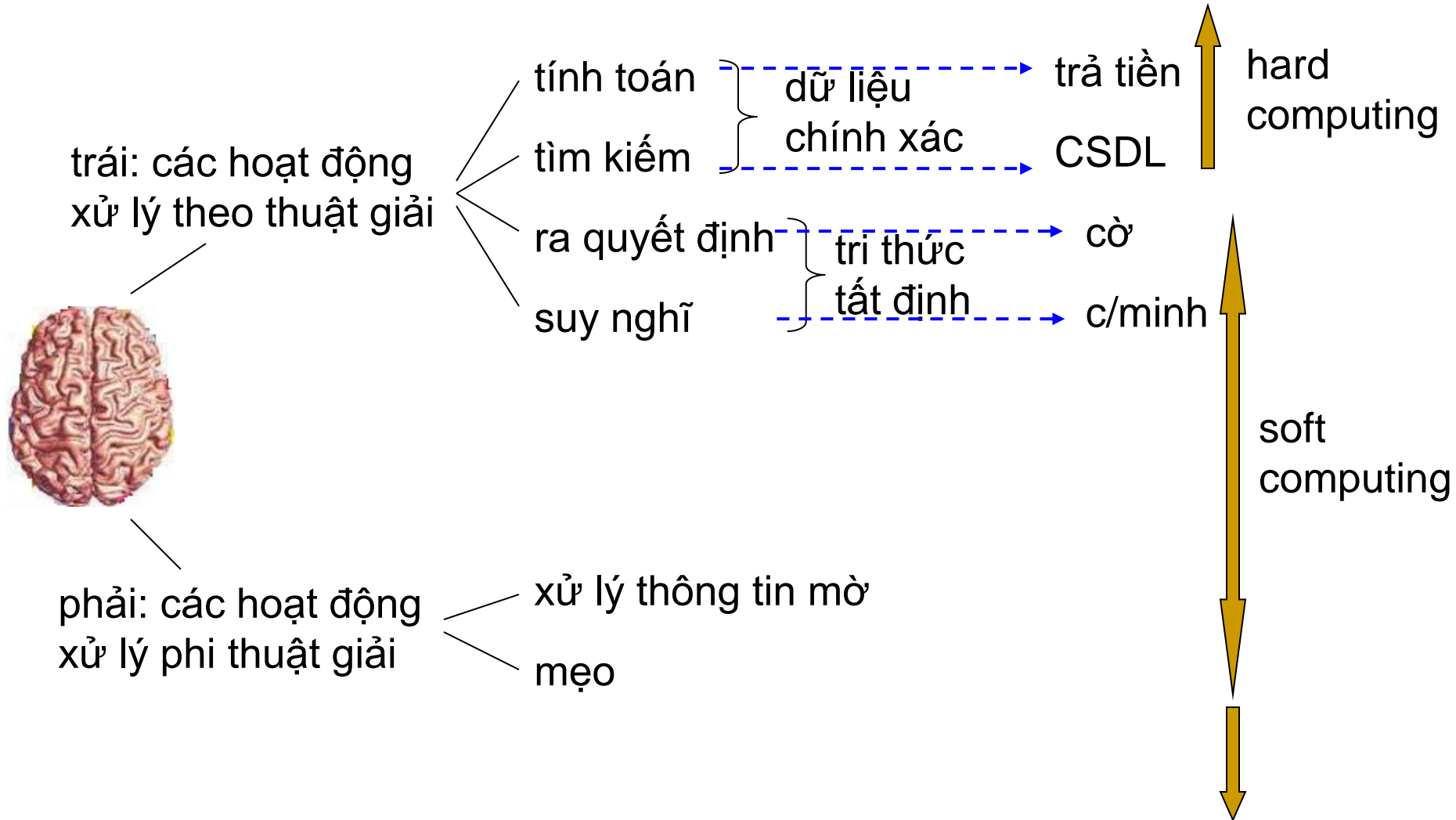
Các nội dung cơ bản (3)

3. Xử lý thông tin – bán cầu đại não



Các nội dung cơ bản (4)

3. Xử lý thông tin – bán cầu đại não



Các nội dung cơ bản (5)

4. Học

- ❑ Dữ liệu → Tri thức
- ❑ Tri thức → Tri thức
- ❑ Data mining → Knowledge discovery

5. Bộ não = Mạng nơron

- ❑ Nhớ
- ❑ Xử lý

Lịch sử hình thành (1)

a. Máy tính

- MT ra đời từ những năm 1820. MT theo tư tưởng Von Newman – xử lý các đại lượng số → MT thế hệ 1-4
- 1930: A.Turing công bố những kết quả đầu tiên, đặt nền móng cho TTNT: xây dựng máy tính dựa trên những phép toán cơ sở của logic như AND, OR, NOT. Máy tính được điều khiển bởi các chương trình lưu trong bộ nhớ trong → MT biết suy nghĩ

Lịch sử hình thành (2)

- Máy tính thế hệ 5:
 - Thiên về xử lý các phát biểu đúng/sai
 - Các phép toán logic and/or/not
 - Kiến trúc máy tính // cực cao, fi-Von Newman (không có các khái niệm tuần tự, lặp, phân nhánh như truyền thống mà tự động làm việc theo sự điều khiển của chương trình).
- Von Newman: máy tính tính toán
- Turing: máy tính suy nghĩ
- Các ứng dụng thử nghiệm: luật, di truyền, xử lý ngôn ngữ tự nhiên.

Lịch sử hình thành (3)

b. Ngôn ngữ

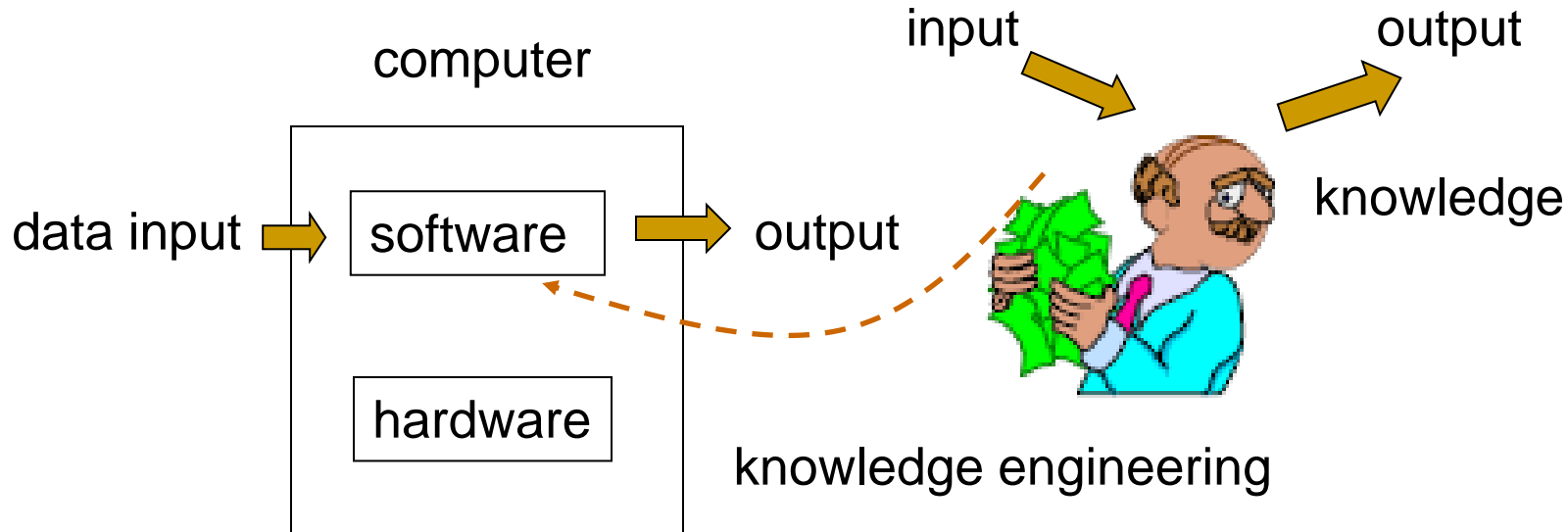
- LISP (List processing), 1960, Mc Cathy, MIT (Massachusetts Institute of Technology)
- PROLOG, 1972, Alain Colmerauer
- CLIPS (C Language Integrated Production System)
- Hướng đối tượng

Lịch sử hình thành (4)

- **1940-1950:** những năm đầu
 - 1943: McCulloch & Pitts: mô hình mạch logic của bộ não
 - 1950: Turing: “Máy tính toán và trí thông minh”
- **1950s:** các c/trình heuristic mô phỏng các hoạt động của con người
 - 1956: chương trình dẫn xuất kết luận trong các hệ hình thức
 - 1959: chương trình chứng minh định lý hình học phẳng (Anderson – MIT)
- **1960s:** các máy tính có bộ nhớ ↑ đáng kể, hạn chế: bùng nổ tổ hợp
 - 1961: chương trình tích phân
 - 1963: chứng minh định lý hình học không gian, trò chơi cờ của Samuel
 - 1964: chương trình giải phương trình đại số sơ cấp, chương trình ELIZA trao đổi bằng ngôn ngữ tự nhiên
 - 1966: chương trình phân tích, tổng hợp lời nói
 - 1968: chương trình nhận dạng hình ảnh, robot chế tạo theo đề án “Mắt – Tay”, chương trình học nói

Lịch sử hình thành (5)

- **1970s:** xuất hiện những n/cứu về bộ não → hệ chuyên gia



Expert system = Human Expertise + Inference/Reasoning

SP thương mại hóa = chuyên gia + Suy diễn/Suy luận

- Hệ chuyên gia: khai thác CSTT lấy từ chuyên gia con người nhằm giải 1 lớp **hẹp** các bài toán **khó**, đạt trình độ **cao** của 1 chuyên gia **lâu năm**

Lịch sử hình thành (6)

- ❑ Hệ DENDRAL (hóa học)
- ❑ Hệ MYCIN (y học): trợ giúp bác sĩ chẩn đoán bệnh nhiễm trùng máu
- ❑ Hệ PROSPECTOR (địa chất): dự báo tài nguyên
- ❑ Hệ MOLGEN (di truyền học phân tử)
- ❑ Hệ ICAD/ICAM (quân sự) : thiết kế, chế tạo có sự trợ giúp của máy tính

Lịch sử hình thành (7)

- **1980-1988:** Hệ chuyên gia phát triển mạnh, mạng nơron, mờ (fuzzy logic)
- **1988—93:** Công nghiệp về HCG đổ vỡ (mùa đông của TTNT)
- **1993—: Các tiếp cận dựa trên thống kê**
 - Lý thuyết xác suất phát triển, tập trung vào độ không chắc chắn
 - Đào sâu các vấn đề kỹ thuật
 - Các tác tử có khắp mọi nơi (TTNT hồi xuân)

TTNT có thể làm những gì?

Những vấn đề nào sau đây có thể giải quyết được?

- Chơi bóng bàn
- Lái xe an toàn vòng theo đường sườn núi
- Mua hàng tạp phẩm mạng
- Phát hiện và chứng minh các định lý toán học
- Nói chuyện với con người trong 1 giờ
- Thực hiện thành công 1 cuộc phẫu thuật phức tạp
- Rửa bát khỏi máy rửa bát và xếp vào đúng chỗ
- Dịch ngôn ngữ nói từ tiếng Anh sang tiếng Việt trong thời gian thực
- Viết 1 câu chuyện cười (có chủ đích)

Những câu chuyện cười không định trước

- One day Joe Bear was hungry. He asked his friend Irving Bird where some honey was. Irving told him there was a beehive in the oak tree. Joe walked to the oak tree. He ate the beehive. The End.
- Henry Squirrel was thirsty. He walked over to the river bank where his good friend Bill Bird was sitting. Henry slipped and fell in the river. Gravity drowned. The End.
- Once upon a time there was a dishonest fox and a vain crow. One day the crow was sitting in his tree, holding a piece of cheese in his mouth. He noticed that he was holding the piece of cheese. He became hungry, and swallowed the cheese. The fox walked over to the crow. The End.

Những câu chuyện cười không định trước

- Một ngày nọ chú gấu Joe thấy đói. Chú ta hỏi bạn của chú là chú chim Irving chỗ nào có mật ong. Irving nói có một tổ ong trong thân cây sồi. Joe đến chỗ cây sồi. Nó ăn tổ ong. Hết.
- Chú sóc Henry khát nước. Nó đến chỗ bờ sông nơi người bạn tốt của nó là chú chim Bill đang đậu. Henry trượt chân và ngã xuống sông. Sức nặng làm nó chết đuối. Hết.
- Ngày xưa có 1 con cáo gian ác và 1 con quạ ngu ngốc. Một ngày, quạ đậu trên cây, mổ quặp 1 miếng phomat. Nó nhận ra rằng nó đang giữ miếng phomat. Nó cảm thấy đói và nuốt miếng phomat. Cáo đến chỗ quạ. Hết.

Ngôn ngữ tự nhiên

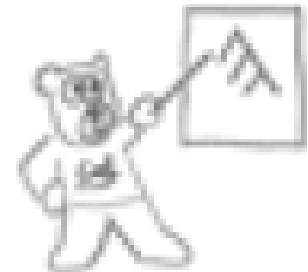


Kỹ thuật tiếng nói (Speech technologies)

- Nhận dạng tự động tiếng nói (Automatic speech recognition - ASR)
- Tổng hợp văn bản thành tiếng nói (Text-to-speech synthesis - TTS)
- Các hệ thống hội thoại (Dialog systems)

Kỹ thuật xử lý ngôn ngữ tự nhiên (Language processing technologies)

- Dịch máy:
 - Aux dires de son président, la commission serait en mesure de le faire.
 - According to the president, the commission would be able to do so.
 - Il faut du sang dans les veines et du cran.
 - We must blood in the veines and the courage.
 - There is no backbone, and no teeth.
- Trích rút thông tin
- Phản hồi thông tin, hỏi đáp
- Phân loại văn bản, lọc thư rác, ...



Hình ảnh (Nhận thức)



Khoa học nghiên cứu người máy (Robotics)

Robotics

- một phần là cơ khí
- một phần là TTNT
- Thực tế phức tạp hơn nhiều so với mô phỏng



Công nghệ

- Xe cộ
- Cứu hộ
- Chơi bóng đá
- và nhiều hệ thống tự động hoá khác



TTNT quan tâm đến:

- Bỏ qua khía cạnh cơ khí
- Các phương pháp lập kế hoạch
- Các phương pháp điều khiển, kiểm soát

Logic

Các hệ thống logic

- ❑ Chứng minh định lý
- ❑ Chẩn đoán lỗi (NASA)
- ❑ Hỏi đáp

Các phương pháp:

- ❑ Các hệ suy diễn
- ❑ Thoả mãn ràng buộc

Chơi trò chơi

■ May, '97: Deep Blue và Kasparov

- Trận đấu tiên thắng kiện tướng cờ vua thế giới
- “Trí thông minh nhân tạo” có thể duyệt 200 triệu nước cờ mỗi giây
- Con người hiểu được 99.9 các nước đi của Deep Blue
- Hiện nay ta có thể tái tạo được 1 máy như vậy với 1 nhóm các máy PC cỡ lớn.



■ Các câu hỏi ngỏ:

- Trí thức của con người xử lý thế nào với sự bùng nổ không gian trạng thái của bàn cờ?
- Hoặc: Làm cách nào con người có thể cạnh tranh với các máy tính?

■ 1996: Kasparov đánh bại Deep Blue

- “Tôi có thể cảm thấy - người thấy – 1 loại trí thông minh mới qua bàn cờ.”

■ 1997: Deep Blue đánh bại Kasparov

- “Deep Blue chưa chứng minh được cái gì cả.”

Ra quyết định

Có rất nhiều ứng dụng của TTNT theo hướng ra quyết định như:

- Lập lịch: lập trình đường bay, quân sự
- Lên kế hoạch đường đi, ví dụ, hệ thống mapquest
- Chuẩn đoán bệnh, ví dụ, hệ thống tìm đường Pathfinder
- Bộ phận trợ giúp tự động
- Phát hiện gian lận
- ...

Một số vấn đề khó giải đáp

1. Ai sẽ chịu trách nhiệm nếu người máy lái xe gây ra tai nạn?
2. Máy tính có thể vượt qua con người không?
3. Chúng ta sẽ làm gì với các máy tính siêu thông minh?
4. Những máy tính như vậy có nhận thức, đúng không?
5. Về nguyên tắc thì trí tuệ con người có thể tồn tại mãi trong máy không?

Những vấn đề chưa được giải quyết

- Chương trình chưa tự sinh ra được heuristic
- Chưa có khả năng xử lý song song của con người
- Chưa có khả năng diễn giải một vấn đề theo nhiều phương pháp khác nhau như con người.
- Chưa có khả năng xử lý thông tin trong môi trường liên tục như con người.
- Chưa có khả năng học như con người.
- Chưa có khả năng tự thích nghi với môi trường.

Tài liệu tham khảo

- R. E. Bellman. *An Introduction to Artificial Intelligence: Can Computers Think?* Boyd & Fraser Publishing Company, San Francisco, 1978.
- E. Charniak and D. McDermott. *Introduction to Artificial Intelligence*. Addison-Wesley, Reading, Massachusetts, 1985.
- J. Haugeland. *Artificial Intelligence: The Very Idea*. MIT Press, Cambridge, Massachusetts, 1985.
- R. Kurzweil. *The Age of Intelligent Machines*. MIT Press, Cambridge, Massachusetts, 1990.
- N. J. Nilsson. *Artificial Intelligence: A New Synthesis*. Morgan Kaufmann, San Mateo, California, 1998.
- D. Poole, A. K. Mackworth, and R. Goebel. *Computational Intelligence: A Logical Approach*. Oxford University Press, Oxford, UK, 1998.
- E. Rich and K. Knight. *Artificial Intelligence (Second Edition)*. McGraw-Hill, New York, 1991.
- P. H. Winston. *Artificial Intelligence (Third Edition)*. Addison-Wesley, Reading, Massachusetts, 1992.