

Trí tuệ nhân tạo (Artificial Intelligence)

Giới thiệu

By Hoàng Hữu Việt

Email: viethh@vinhuni.edu.vn

Viện Kỹ thuật và Công nghệ, Đại học Vinh

Vinh, 3/2019

Thông tin môn học

■ Tài liệu chính

[1] Stuart Russell, Peter Norvig. Artificial Intelligence. A modern approach. 3rd ed. Prentice Hall, 2009.

■ Tài liệu khác

[2] Milos Hauskrecht, Artificial Intelligence, 2013,
people.cs.pitt.edu/~milos/courses/cs1571-Fall2013

[3] Các nguồn từ internet

Nội dung

- Định nghĩa về trí tuệ nhân tạo
- Các khoa học cơ bản của trí tuệ nhân tạo
- Lịch sử phát triển
- Các ứng dụng của trí tuệ nhân tạo

Định nghĩa

1) Thinking Humanly – suy nghĩ 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) Thinking Rationally – suy nghĩ 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)

3) Acting Humanly – hành động 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) Acting Rationally – hành động hợp lý

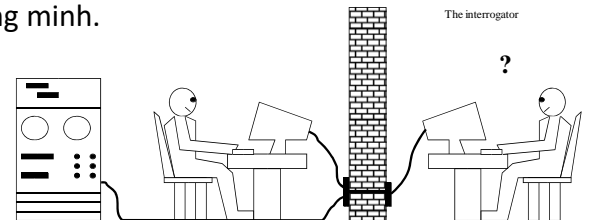
- “Computational Intelligence is the study of the design of intelligent agents.” (Poole et al., 1998)
- “AI . . . is concerned with intelligent behavior in artifacts.” (Nilsson, 1998)

Định nghĩa

- Các định nghĩa (1) và (2) liên quan đến các quá trình suy nghĩ (thought processes) và suy diễn (reasoning).
- Các định nghĩa (3) và (4) liên quan đến các hành xử (behavior).
- Các định nghĩa (1) và (3) đánh giá mức độ thành công so với sự thực hiện của con người.
- Các định nghĩa (2) và (4) đánh giá mức độ thành công so với sự hợp lý (rationality).
 - Một hệ thống hành động hợp lý (rationally) nếu nó thực hiện các công việc phù hợp với những gì hệ thống đã biết.

Hành động như con người

- Thử nghiệm Turing(1950)
 - Được thiết kế để cung cấp một định nghĩa về sự thông minh.
 - Một người hỏi ngồi ở một phòng, đối tác của người hỏi là một máy tính được đặt ở một phòng khác.
 - Hai bên trao đổi thông tin với nhau thông qua văn bản.
 - Nếu máy tính có thể làm cho người hỏi tưởng lầm là có một người khác đang trả lời mình thì máy tính được xem là thông minh.



Hành động như con người

- Khả năng cần có để thực hiện được các hành động được xem là thông minh:
 - **Xử lý ngôn ngữ tự nhiên (natural language processing)**: cho phép giao tiếp bằng ngôn ngữ tự nhiên.
 - **Biểu diễn tri thức (knowledge representation)**: cho phép lưu trữ thông tin tiếp nhận được.
 - **Suy diễn tự động (automated reasoning)**: để sử dụng các thông tin đã lưu trữ và bằng lập luận để trả lời các câu hỏi.
 - **Học máy (machine learning)**: để thích ứng với hoàn cảnh mới và để phát hiện và suy đoán các mẫu dữ liệu.
 - **Nhìn máy (computer vision)**: để nhận biết của các đối tượng thông qua hình ảnh.
 - **Robotics**: để thao tác với các đối tượng và di chuyển.

Suy nghĩ như con người

- Đây là hướng tiếp cận theo mô hình nhận thức (cognitive modeling approach).
- Nếu một máy tính suy nghĩ như con người, vậy phải có cách nào đó để xác định con người suy nghĩ ?
 - Cố gắng nắm bắt những suy nghĩ của chính chúng ta khi thực hiện hành động.
 - Thông qua các thí nghiệm tâm lý học, đó là quan sát một người thực hiện hành động.
 - Thông qua hình ảnh bộ não, đó là quan sát bộ não hoạt động.
- Khi có một lý thuyết đủ chính xác về lý trí (mind), có thể biểu diễn lý thuyết như là một chương trình máy tính.

Suy nghĩ như con người

- Nếu các hành xử (behavior) vào/ra của chương trình phù hợp tương ứng với các hành xử của con người, nghĩa là cơ chế của chương trình có thể thao tác như con người.
 - Newell and Simon (1961), “General Problem Solver”.
- **Hướng tiếp cận mô hình nhận thức:** Lĩnh vực liên ngành của khoa học nhận thức kết hợp với các mô hình máy tính từ AI và các kỹ thuật thử nghiệm từ tâm lý học để xây dựng các lý thuyết chính xác và có thể kiểm chứng về tâm trí con người.

Suy nghĩ hợp lý

- Đây là hướng tiếp cận theo các luật suy nghĩ (“laws of thought” approach).
- Aristotle là một trong những người đầu tiên cố gắng mã hóa “suy nghĩ đúng đắn” (right thinking) – đó là quá trình lập luận không thể bác bỏ.
 - Lập luận luôn đưa ra các kết luận đúng đắn nếu các tiên đề là đúng đắn. Ví dụ “Socrates is a man; all men are mortal; therefore, Socrates is mortal” .
- Luật suy nghĩ được cho là chi phối các hoạt động của suy nghĩ, và ban đầu được nghiên cứu trong lĩnh vực logic.

Suy nghĩ hợp lý

- Đến năm 1965, về nguyên tắc, các chương trình có thể giải quyết mọi bài toán (có thể giải được) theo mô tả bởi ký hiệu logic.
- Do vậy, những nhà logic học trong trí tuệ nhân tạo hy vọng sẽ xây dựng được các chương trình để tạo ra các hệ thống thông minh.
- Hai khó khăn chính:
 - Không dễ để biểu diễn tri thức (knowledge) theo dạng ký hiệu logic.
 - Có một sự khác biệt lớn giữa việc giải quyết một bài toán "theo nguyên tắc" và giải quyết nó trong thực tế.

Hành động hợp lý

- Đây là hướng tiếp cận theo tác tử hợp lý (rational agent).
- Một tác tử (agent) là một thực thể thực hiện các hành động.
- Một tác tử hợp lý là một tác tử mà hành động để đạt được kết quả/kết quả mong đợi tốt nhất.
- Trong hướng tiếp cận "các luật của suy nghĩ", nhấn mạnh sự suy diễn chính xác - thỉnh thoảng là một phần của một tác tử hợp lý.
 - Một cách để thực hiện hợp lý là suy diễn logic dẫn tới các kết luận.

Hành động hợp lý

- Tất cả các kỹ năng cần thiết của một tác tử cho thử nghiệm Turing cho phép tác tử hành động hợp lý:
 - Biểu diễn tri thức và suy diễn cho phép tác tử ra các quyết định tốt.
 - Khả năng học cải thiện khả năng tạo ra hành xử hiệu quả.
- Ưu điểm so với các cách tiếp cận khác:
 - Tổng quát hơn cách tiếp cận "các luật suy nghĩ" vì suy diễn chính xác là một trong những cơ chế để đạt được sự hợp lý.
 - Phù hợp với sự phát triển khoa học hơn các cách tiếp cận dựa trên hành vi của con người hoặc suy nghĩ của con người.
- **Bài giảng tập trung vào các nguyên tắc xây dựng các tác tử hợp lý!**

Khoa học cơ bản của trí tuệ nhân tạo

- Triết học (philosophy)
 - Logic, các phương pháp suy diễn
 - Các cơ sở của việc học (learning)
 - Sự hợp lý
- Toán học (mathematics)
 - Biểu diễn hình thức và các thuật toán chứng minh
 - Bài toán giải quyết được và không giải quyết được
 - Bài toán áp dụng được và không áp dụng được (độ phức tạp tính toán)
 - Tính toán, xác suất
- Kinh tế học (economics)
 - Lý thuyết quyết định

Khoa học cơ bản của trí tuệ nhân tạo

- Khoa học thần kinh (neuroscience)
 - Cơ sở tự nhiên của các hoạt động trí óc.
- Tâm lý học (psychology)
 - Các dấu hiệu của nhận thức và điều khiển vận động
 - Các kỹ thuật thực nghiệm
- Kỹ thuật máy tính (computer engineering)
 - Xây dựng các máy tính có tốc độ tính toán nhanh
- Lý thuyết điều khiển (control theory)
 - Thiết kế các hệ thống nhằm cực đại hóa hàm mục tiêu
- Ngôn ngữ học (linguistics)
 - Biểu diễn tri thức, ngữ pháp ngôn ngữ

Lịch sử phát triển

- Giai đoạn 1943 – 1955: “thai nghén” (gestation) trí tuệ nhân tạo.
 - Neural network approach (McCulloch and Pitts, 1943):
 - Mô hình của các neuron nhân tạo dựa trên các neuron của bộ não và đề xuất mạng neural.
 - Donald Hebb (1949) đề xuất luật học Hebbian cho mạng neural.
 - Mạng neural nhân tạo đầu tiên được xây dựng bởi 2 sinh viên ở Harvard (Marvin Minsky and Dean Edmonds, 1950).
 - Alan Turing (1950) công bố bài báo “Computing Machinery and Intelligence”, trong đó đã giới thiệu về học máy (machine learning), các giải thuật di truyền (genetic algorithms) và học tăng cường (reinforcement learning).

Lịch sử phát triển

- Năm 1956: ra đời của trí tuệ nhân tạo.
 - Sau khi tốt nghiệp tiến sỹ (PhD.) năm 1951 ở Princeton, John McCarthy chuyển tới Stanford và sau đó là Dartmouth College, là nơi sinh ra lĩnh vực trí tuệ nhân tạo.
 - McCarthy, Minsky, Claude Shannon, Nathaniel Rochester tổ chức một hội thảo trong 2 tháng hè ở Dartmouth và đề xuất các dự án nghiên cứu trong “Artificial Intelligence”.

A PROPOSAL FOR THE DARTMOUTH SUMMER RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE

J. McCarthy, Dartmouth College
M. L. Minsky, Harvard University
N. Rochester, I.B.M. Corporation
C.E. Shannon, Bell Telephone Laboratories
August 31, 1955

Lịch sử phát triển

- Giai đoạn 1952-1969: kỳ vọng của trí tuệ nhân tạo .
 - Herbert Gelernter (1959) xây dựng chương trình “Geometry Theorem Prover” mà có thể chứng minh các định lý.
 - Arthur Samuel (1952) viết chương trình chơi cờ có khả năng học để chơi.
 - McCarthy (1958) xây dựng ngôn ngữ lập trình bậc cao **Lisp** - đã trở thành ngôn ngữ lập trình AI trong 30 năm tiếp theo.
 - Widrow and Hoff (1960) và Widrow (1962) đề xuất mạng neural: **adalines**
 - Frank Rosenblatt (1962) đề xuất mạng neural: **perceptrons**.

Lịch sử phát triển

- Giai đoạn 1966 – 1973: khó khăn của trí tuệ nhân tạo
 - Herbert Simon (1957) tuyên bố “It is not my aim to surprise or shock you—but the simplest way I can summarize is to say that there are now in the world machines that think, that learn and that create. ...”.
 - Tuy nhiên, đây là giai đoạn khó khăn khi áp dụng AI cho các bài toán thực tế vì độ phức tạp tính toán chưa được giải quyết.
 - Gần như không có các nghiên cứu đề cập đến mạng neural.

Lịch sử phát triển

- Giai đoạn 1969-1979: bắt đầu hình thành và phát triển các hệ thống dựa trên tri thức (knowledge).
 - Feigenbaum và đồng nghiệp phát triển hệ thống MYCIN để chẩn đoán bệnh nhiễm khuẩn máu.
 - Schank và các đồng nghiệp (1977) ở ĐH Yale xây dựng các chương trình để hiểu ngôn ngữ tự nhiên.
- 1980 – nay: AI trở thành một ngành công nghiệp.
 - McDermott (1982) xây dựng hệ chuyên gia thương mại thành công đầu tiên ở Digital Equipment Corporation.
 - Hàng trăm công ty xây dựng các hệ chuyên gia, hệ thống nhìn máy (vision systems), robot, phần mềm và phần cứng.

Lịch sử phát triển

- 1986 – nay: sự trở lại của mạng neural
- 1987- nay: Trí tuệ nhân tạo trở thành (adopts) phương pháp khoa học.
 - Mạng neural là công cụ của nhận dạng mẫu (pattern recognition), học máy (machine learning) và khai phá dữ liệu (data mining).
 - Judea Pearl (1988) đưa ra khái niệm xác suất và lý thuyết quyết định trong AI.
 - Mạng Bayesian (Bayesian network) đã được đề xuất để giải quyết bài toán tri thức không chắc chắn (uncertain knowledge).
- 2001-nay: xử lý dữ liệu lớn.

Các ứng dụng

- Trong các hệ thống phần mềm
 - Các hệ thống tự thích nghi (adaptive systems)
 - Các hệ thống thích nghi với nhu cầu người dùng (adapt systems to user needs)
 - Các hệ thống thích nghi với nhiệm vụ cụ thể (adapt systems to specific tasks)
 - Giao diện thông minh (intelligent interfaces)
 - Các ứng dụng trợ giúp thông minh (intelligent helper applications)
 - Tìm kiếm thông tin
 - Web search engines (cải tiến chất lượng tìm kiếm)
 - Web ngữ nghĩa (semantic web)

Các ứng dụng

- Trong các hệ thống phần mềm (tiếp)
 - Nhận dạng tiếng nói (speech recognition)
 - Các hệ thống dựa trên mạng neural
 - Các mô hình Markov ẩn (hidden Markov models)
 - Các hệ thống nhận dạng tiếng nói đa người dùng (multi-user speech recognition systems)
 - Các thiết bị ra lệnh bằng tiếng nói
- Ứng dụng trong y học
 - Hệ thống chẩn đoán bệnh (medical diagnosis)
 - Hệ thống giám sát và cảnh báo bệnh nhân (patient monitoring and alerting)
 - Phẫu thuật bằng robot (robotic surgeries)

Các ứng dụng

- Ứng dụng trong sinh học
 - Dự đoán vùng gen của DNA
 - Tìm kiếm gen đặc trưng cho các bệnh
 - Dự đoán tăng trưởng của cây
 - Dự đoán sâu bệnh của cây
- Ứng dụng trong giao thông
 - Lái xe tự động (autonomous vehicle control)
 - ALVINN (CMU, Pomerleau 1993)
 - Google autonomous vehicles
 - Phát hiện người đi bộ (pedestrian detection)
 - Giám sát giao thông (traffic monitoring)
 - Tìm đường đi tối ưu (navigation/ route optimizations)

Các ứng dụng

- Ứng dụng trong trò chơi
 - Chơi cờ (Deep blue (IBM) program defeated Kasparov in 1997)
 - Chơi bài
- Ứng dụng trong xử lý ngôn ngữ tự nhiên:
 - Phân tích văn bản
 - Phân lớp tự động các bài báo (automatic classification of articles)
 - Trích chọn/suy luận nội dung (content extraction/ inference)
 - Phát hiện SPAM Email
 - Phân tích tiếng nói

Các ứng dụng

- Ứng dụng trong robot

- Robotic toys
 - Sony's Aibo



- Vacuum cleaners



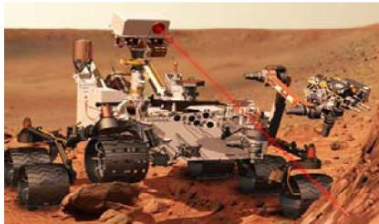
- Humanoid robot
 - Honda's ASIMO



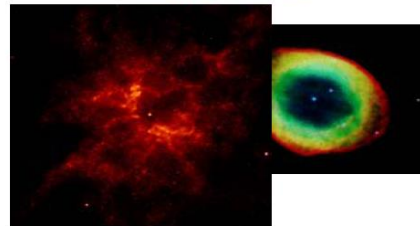
Các ứng dụng

■ Ứng dụng trong thăm dò không gian:

Autonomous rovers,
intelligent probes



Analysis of sky
Survey data



Các ứng dụng

■ Các ứng dụng khác

- ❑ Handwriting analysis/ detection
- ❑ Human face detection
- ❑ Video stream annotation (chú thích)
- ❑ Object tracking
- ❑ Music composition
- ❑ Poetry composition
- ❑ Picture drawing
- ❑ ...