

TRÍ TUỆ NHÂN TẠO

THÔNG TIN CHUNG VỀ HỌC PHẦN

- **Tên học phần:** Trí Tuệ Nhân Tạo
- **Tên tiếng Anh:** Introduction to Artificial Intelligence
- **Số tín chỉ:** 3 (1 buổi: 2 lý thuyết + 2 thực hành + 6 tự học)

THÔNG TIN CHUNG VỀ HỌC PHẦN

Hình thức KT	Nội dung	Thời điểm	Phương pháp đánh giá	Công cụ đánh giá	Tỉ lệ (%)
BÀI TẬP					
BT	Các câu hỏi trên lớp, bài tập về nhà (cá nhân hoặc nhóm), ví dụ bài luận về các chủ đề, chạy tay thuật toán, lập trình, bài kiểm tra.	Tuần 1-13	Đánh giá qua câu hỏi, bài tập, lập trình	Đáp án, rubrics	50
THUYẾT TRÌNH, VẤN ĐÁP CUỐI KỲ					
TT	Nhóm sinh viên hoặc cá nhân.		Đánh giá qua bài thuyết	Câu hỏi, Rubrics	50

NỘI DUNG SẼ HỌC

Tuần	Nội dung
1	<i>Giới thiệu về AI</i>
2 - 3	<i>Lập trình Python cho data science</i>
4	<i>Đạo đức và luật cho AI. Các khái niệm cơ bản của hệ thống thông minh</i>
5	<i>Các khái niệm cơ bản của hệ thống thông minh (tt)</i>
6	<i>Uninformed search algorithms</i>
7	<i>Informed search algorithms</i>
8	<i>Local search</i>
9-10	<i>Searching in complex environments. Constraint satisfactoin problems</i>
11-12	<i>Constraint satisfactoin problems (cont.)</i>
13	<i>Introduction to reinforcement learning</i>
14-15	<i>Kiểm tra cuối kỳ (báo cáo và vấn đáp)</i>

THÔNG TIN CHUNG VỀ HỌC PHẦN

- **Giáo trình chính:**

- Russell, S. J., & Norvig, P. (2016). *Artificial Intelligence: A Modern Approach* (3rd ed.). Pearson.
- Lapan, M. (2020). *Deep Reinforcement Learning Hands-On*. Packt Publishing Ltd.

- **Tài liệu tham khảo:**

- Géron, A. (2017). *Hands-On Machine Learning with Scikit-Learn and Tensorflow: Concepts, Tools, and Techniques to Build Intelligent Systems*. O'Reilly Media.

GIỚI THIỆU VỀ AI

ĐIỀU NÀO THỂ HIỆN TRÍ THÔNG MINH?

- Bạn đánh bại ai đó ở môn cờ vua.
- Bạn chứng minh một định lý toán học bằng cách sử dụng một tập hợp các tiên đề đã biết.
- Bạn cần mua một số đồ dùng, gấp gở ba đồng nghiệp khác nhau, trả sách cho thư viện và tập thể dục. Bạn lên kế hoạch cho ngày của mình sao cho mọi việc đều đạt được một cách hiệu quả.
- Bạn là một luật sư được yêu cầu bào chữa cho ai đó. Bạn nhớ lại ba trường hợp tương tự mà bị cáo có tội và bạn từ chối khách hàng tiềm năng.

ĐIỀU NÀO THỂ HIỆN TRÍ THÔNG MINH?

- Một người lạ đi ngang qua bạn trên đường nhìn thấy đồng hồ của bạn và hỏi, "Bạn có thể cho tôi biết mấy giờ không?" Bạn nói, "Bây giờ là 3 giờ."
- Bạn được yêu cầu tìm một chiếc tuốc nơ vít Phillips lớn trong một phòng làm việc bừa bộn. Bạn vào phòng (bạn chưa từng đến đó trước đây), tìm kiếm mà không rơi vào đồ vật và cuối cùng tìm thấy chiếc tuốc nơ vít.
- Ai đó gõ vào một nhịp điệu và bạn có thể đánh theo nhịp điệu đó và tiếp tục nhịp điệu đó ngay cả khi nhịp điệu đó đã dừng.

ĐIỀU NÀO THỂ HIỆN TRÍ THÔNG MINH?

- Bạn là một đứa trẻ sáu tháng tuổi. Bạn có thể tạo ra âm thanh bằng cơ quan phát âm của mình và bạn có thể nghe thấy âm thanh lời nói xung quanh mình, nhưng bạn không biết cách tạo ra âm thanh mà mình đang nghe. Trong năm tiếp theo, bạn sẽ tìm ra âm thanh trong ngôn ngữ của cha mẹ bạn là gì và cách tạo ra chúng.
- Bạn là một đứa trẻ một tuổi đang học tiếng Ả Rập. Bạn nghe thấy những chuỗi âm thanh và nhận ra rằng chúng có liên quan đến những ý nghĩa cụ thể trên thế giới. Trong vòng hai năm, bạn học cách chia các chuỗi thành các phần có ý nghĩa và tạo ra các từ và câu của riêng mình.

INTELLIGENCE???

- Trí thông minh/Trí tuệ là gì?

- Hiểu biết
- Suy nghĩ
- Giao tiếp
- Học hỏi
-

=> Kiến thức + khả năng nhận thức, cảm nhận, hiểu biết, xử lý, giao tiếp, phán đoán, học hỏi.

ARTIFICIAL INTELLIGENCE???

- AI là gì?

- Thực thể có thể nhận thức
- Thực thể có thể suy nghĩ
- Thực thể có thể nói
- Thực thể có thể học....

=> Nghiên cứu nhằm tạo ra và phát triển các kỹ thuật và công cụ có khả năng thu nhận Kiến thức + khả năng nhận thức, cảm nhận, hiểu biết, xử lý, giao tiếp, phán đoán, học hỏi để giải quyết các vấn đề mà con người có thể làm tốt.

KHÁI NIỆM AI

1

Systems that
think like humans

2

Systems that **act**
like humans

3

Systems that
think rationally

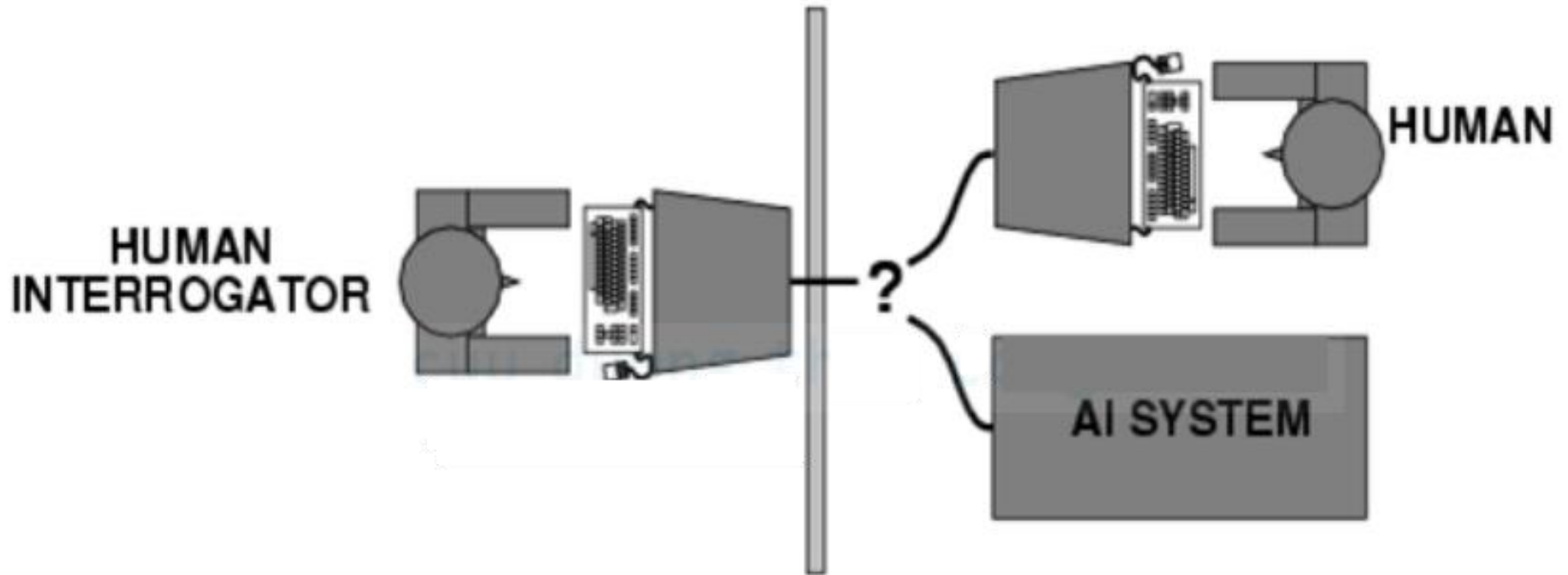
4

Systems that **act**
rationally

Systems that **act** like humans

- Theo quan điểm này thì hệ thống AI có đặc điểm gì?
- Làm sao để đánh giá liệu hệ thống AI có như con người hay không?
- Cách hệ thống Turing Test hoạt động?
- Có hệ thống nào vượt qua bài Turing Test chưa?
- Để qua được bài Turing Test hệ thống cần có những khả năng gì?

Systems that **act** like humans

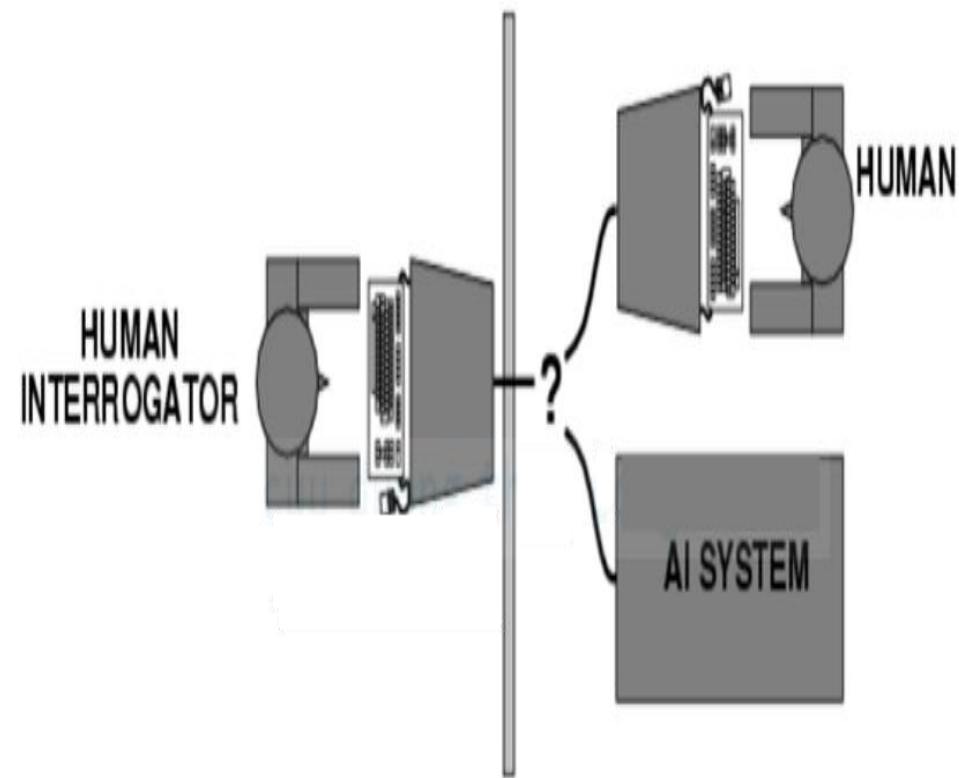


The Turing Test (1950)

Systems that **act** like humans

- Thảo luận nhóm:

Giả sử bạn thiết kế một hệ thống để vượt qua bài kiểm tra Turing. Những khả năng mà hệ thống đó phải có là gì?



The Turing Test (1950)

Systems that think like humans

- Theo quan điểm này thì hệ thống AI có đặc điểm gì?
- Làm sao để xây dựng hệ thống AI theo quan điểm này?
- Đã có những hệ thống AI nào được xây dựng dựa theo quan điểm này?

Systems that think rationally

- Theo quan điểm này thì hệ thống AI có đặc điểm gì?
- Làm sao để xây dựng hệ thống AI theo quan điểm này?
- Khó khăn chủ yếu khi tiếp cận dựa theo quan điểm này?

Systems that **act** **rationality**

- Theo quan điểm này thì hệ thống AI có đặc điểm gì?
- Làm sao để xây dựng hệ thống AI theo quan điểm này?

KHÁI NIỆM AI

Systems that think like humans	Systems that think rationally
Systems that act like humans	Systems that act rationally

Charniak & McDermott, 1985

“The study of mental faculties through the use of computational models”

KHÁI NIỆM AI

Systems that think like humans	Systems that think rationally
Systems that act like humans	Systems that act rationally

Haugeland, 1985

“The exciting new effort to make computers think *machines with minds*, in the full and literal sense”

KHÁI NIỆM AI

Systems that think like humans	Systems that think rationally
Systems that act like humans	Systems that act rationally

Kurzweil, 1990

“The art of creating machines that perform functions that require intelligence when performed by people”

KHÁI NIỆM AI

Systems that think like humans	Systems that think rationally
Systems that act like humans	Systems that act rationally

Luger & Stubblefield, 1993

“The branch of computer science that is concerned with the automation of intelligent behavior”

KHÁI NIỆM AI

Systems that think like humans	Systems that think rationally
Systems that act like humans	Systems that act rationally

Nilsson, 1998

“Many human mental activities such as writing computer programs, doing mathematics, engaging in common sense reasoning, understanding language, and even driving an automobile, are said to demand intelligence. We might say that [these systems] exhibit AI”

KHÁI NIỆM AI

Systems that think like humans	Systems that think rationally
Systems that act like humans	Systems that act rationally

Rich & Knight, 1991

“The study of how to make computers do things at which, at the moment, people are better”

KHÁI NIỆM AI

Systems that think like humans	Systems that think rationally
Systems that act like humans	Systems that act rationally

Schalkoff, 1990

“A field of study that seeks to explain and emulate intelligent behavior in terms of computational processes”

KHÁI NIỆM AI

Systems that think like humans	Systems that think rationally
Systems that act like humans	Systems that act rationally

Winston, 1992

“The study of the computations that make it possible to perceive, reason, and act”

LỊCH SỬ HÌNH THÀNH VÀ PHÁT TRIỂN

1

- 1943-1955: Tiền khởi đầu

2

- 1952-1969: Khởi đầu

3

- 1969-1979: Hệ chuyên gia

4

- 1980-nay: AI có sản phẩm thương mại

5

- 1987-nay: AI trở thành ngành khoa học

6

- 2001-nay: AI phát triển theo xu hướng dữ liệu lớn

LỊCH SỬ HÌNH THÀNH VÀ PHÁT TRIỂN

1943-1955: Tiền khởi đầu

- 1943: Mạng neural nhân tạo đầu tiên
- 1950: Turing test ra đời
- 1956: Khái niệm AI chính thức ra đời

LỊCH SỬ HÌNH THÀNH VÀ PHÁT TRIỂN

1952-1967: Khởi đầu

- Các chương trình Logic Theorist và sau đó là General Problem Solver
- 1952: chương trình chơi cờ đam ra đời và thắng cả người tạo ra nó
- 1958: Ngôn ngữ Lisp ra đời
- 1963: AI có thể giải các bài toán tích phân ở mức độ sinh viên năm thứ nhất đại học



LỊCH SỬ HÌNH THÀNH VÀ PHÁT TRIỂN

1967-1979: Hệ chuyên gia

- 1967: Hệ chuyên gia DENDRAL
- 1972: Ngôn ngữ Prolog
- 1974: Hệ chuyên gia MYCIN

Hệ chuyên gia là các hệ thống có khả năng ra quyết định tương tự chuyên gia trong lĩnh vực hẹp của mình.

LỊCH SỬ HÌNH THÀNH VÀ PHÁT TRIỂN

1980-nay: AI có sản phẩm thương mại

- 1980-1988: 1 số hệ chuyên gia được thương mại hóa
- Sau 1988: mùa đông AI
- Đánh dấu sự trở lại của mạng neural nhân tạo
- 1981: xây dựng các máy tính thông minh chạy trên ngôn ngữ Prolog

LỊCH SỬ HÌNH THÀNH VÀ PHÁT TRIỂN

1987-nay: AI trở thành ngành khoa học

- 1987: trí tuệ nhân tạo đã có phương pháp nghiên cứu riêng

LỊCH SỬ HÌNH THÀNH VÀ PHÁT TRIỂN

2001-nay: AI phát triển theo xu hướng dữ liệu lớn

- Hệ thống hỗ trợ khuyến nghị của Amazon
- Ứng dụng trợ giúp Siri của Apple
- Chương trình dịch tự động và nhận diện giọng nói của Google
- Chương trình trò chơi Watson của IBM

CÁC LĨNH VỰC NGHIÊN CỨU CỦA AI

- Một hệ thống trí tuệ nhân tạo hoàn chỉnh, làm việc trong việc một môi trường nào đó cần có khả năng:
 - Cảm nhận (perception)
 - Lập luận (reasoning)
 - Hành động (action)

CẢM NHẬN

- Mục tiêu:
 - Sử dụng các thiết bị để thu nhận thông tin liên quan tới hoạt động từ môi trường bên ngoài
 - Biến đổi thông tin nhận được về dạng có thể hiểu được
- Các kỹ thuật thường được nghiên cứu:

CẢM NHẬN

- Các kỹ thuật thường được nghiên cứu:

Thị giác máy

**Nhận dạng mẫu
(pattern recognition)**

**Phân tích chuyển động
(motion analysis)**

**Tạo lập khung cảnh 3D
(scene reconstruction)**

**Nâng cao chất lượng
ảnh
(image restoration)**

Nhận dạng mẫu

**Nhận dạng
mặt người**

**Nhận dạng
vân tay**

**Nhận dạng
chữ viết tay**

Nhận dạng chữ in

Xử lý NNTN

**Nhận dạng tiếng nói
(speech recognition)**

Text to speech

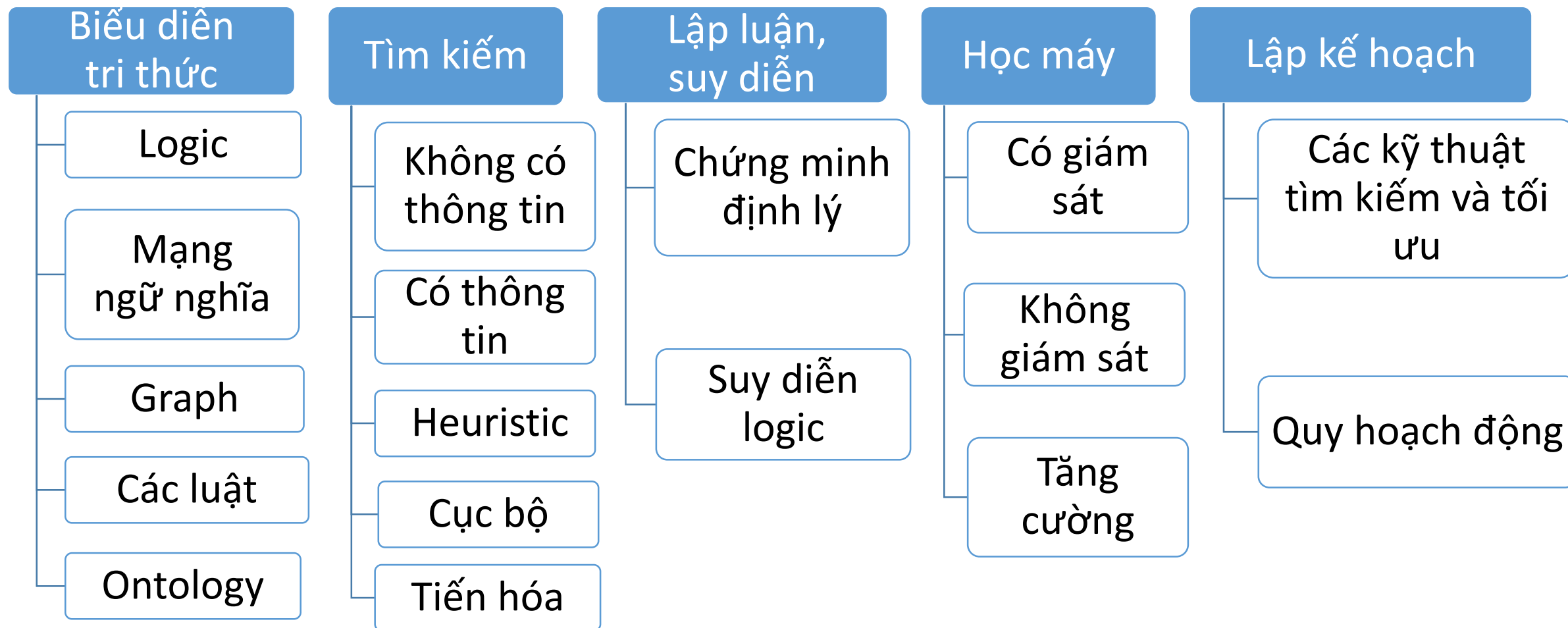
Speech to text

LẬP LUẬN VÀ SUY DIỄN

- Mục tiêu:
 - Cơ chế để đưa ra được quyết định phù hợp
- Được thực hiện thông qua các kỹ thuật:

LẬP LUẬN VÀ SUY DIỄN

- Được thực hiện thông qua các kỹ thuật:



HÀNH ĐỘNG

- Mục tiêu:
 - Cho phép hệ thống tác động vào môi trường xung quanh hoặc đơn giản là đưa ra thông tin về kết luận của mình
- Được thực hiện thông qua các kỹ thuật:
 - Tổng hợp ngôn ngữ tự nhiên và tiếng nói.
 - Kỹ thuật rô bốt (robotics)

Nền tảng của AI

Lĩnh vực	Mô tả
Triết học	Logic, phương pháp lập luận, tâm trí như một hệ thống vật lý, nền tảng học tập, ngôn ngữ, tính hợp lý.
Toán học	Biểu diễn và chứng minh chính thức, thuật toán, tính toán, khả năng (không) quyết định được, khả năng (không) xử lý được, xác suất.
Kinh tế	Tiện ích, lý thuyết quyết định, tác nhân kinh tế hợp lý
Khoa học thần kinh	Nơ-ron làm đơn vị xử lý thông tin
Tâm lý học/Khoa học nhận thức	Con người cư xử, nhận thức, xử lý thông tin, thể hiện kiến thức như thế nào.
Kỹ thuật máy tính	Xây dựng máy tính nhanh
Lý thuyết điều khiển	Thiết kế các hệ thống tối đa hóa một hàm mục tiêu theo thời gian
Ngôn ngữ học	Biểu diễn kiến thức, ngữ pháp

MỘT SỐ ỨNG DỤNG VÀ THÀNH TỰU



THẢO LUẬN

Hoạt động nhóm:

Tìm hiểu và trình bày về thành tựu AI mà bạn thích;
triển vọng nghề nghiệp của ngành AI.

WEAK AI vs STRONG AI

Thiết kế và đào tạo cho một nhiệm vụ cụ thể hoặc một tập hợp nhiệm vụ hẹp

Có tính chuyên môn hóa cao nhưng chưa có khả năng giải quyết vấn đề vượt ra ngoài phạm vi hẹp của nó

WEAK AI

Thường không thông minh; chúng rất giỏi trong việc thực hiện một nhiệm vụ được xác định trước nhưng lại thiếu sự hiểu biết hoặc ý thức thực sự.

Ví dụ: trợ lý ảo như Siri hoặc Alexa, thuật toán đề xuất được sử dụng bởi các dịch vụ phát trực tuyến và chatbot được thiết kế cho các nhiệm vụ dịch vụ khách hàng cụ thể

WEAK AI vs STRONG AI

Có khả năng hiểu biết và thông minh ngang bằng con người.

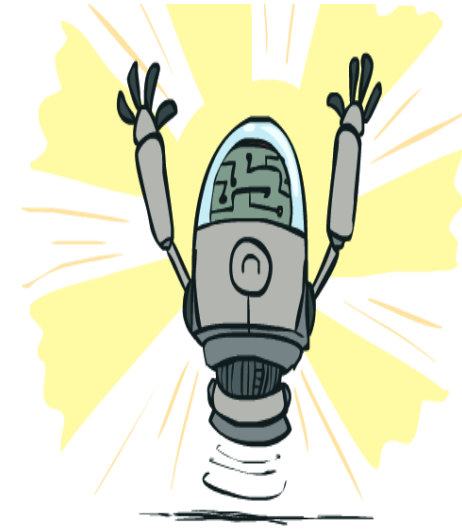
Chỉ là một khái niệm lý thuyết và chưa có hệ thống trí tuệ nhân tạo nào đạt đến mức độ thông minh chung này

STRONG AI

Có khả năng thực hiện bất kỳ nhiệm vụ trí tuệ nào mà con người có thể làm, thích ứng với các lĩnh vực khác nhau và sở hữu một dạng ý thức hoặc nhận thức về bản thân.

Là mục tiêu dài hạn của nghiên cứu AI và đòi hỏi phải phát triển các hệ thống AI có khả năng lý luận, học hỏi, hiểu biết và thích ứng trong nhiều nhiệm vụ và bối cảnh khác nhau.

NHỮNG VẤN ĐỀ CHƯA GIẢI QUYẾT ĐƯỢC



- ✓ Chơi một ván Jeopardy hoàn hảo?
- ✓ Thắng bất kỳ người nào chơi cờ vua?
- ✓ Thắng những người chơi cờ vây giỏi nhất?
- ✓ Chơi một ván tennis tử tế?
- ✓ Lấy một chiếc cốc cụ thể và đặt lên kệ?
- ✗ Dỡ bất kỳ máy rửa chén nào trong bất kỳ ngôi nhà nào?
- ✓ Lái xe an toàn trên đường cao tốc?
- ✗ Lái xe an toàn trên Đại lộ Telegraph?
- ✓ Mua đồ tạp hóa đủ dùng trong một tuần trên web?
- ✗ Mua đồ tạp hóa đủ dùng trong một tuần tại Berkeley Bowl?
- ?? Khám phá và chứng minh một định lý toán học mới?
- ?? Thực hiện một ca phẫu thuật?
- ✓ Dịch tiếng Trung Quốc nói sang tiếng Anh nói theo thời gian thực?
- ✗ Viết một câu chuyện hài hước có chủ đích?

NHỮNG VẤN ĐỀ CHƯA GIẢI QUYẾT ĐƯỢC

- Rubik, cờ ca rô 3 ô: **tối ưu**
- Cờ vua: **gần đạt mức tốt hơn người**
- Trả lời câu hỏi tự động: **gần đạt mức tốt hơn người**
- Lái xe tự động: **tương đương người**
- Nhận dạng chữ in (theo chuẩn ISO 1073-1:1976, tức là các phong chữ đơn nét và không quá phức tạp): **tương đương người**

NHỮNG VẤN ĐỀ CHƯA GIẢI QUYẾT ĐƯỢC

- Nhận dạng chữ viết tay: **kém người**
- Nhận dạng chữ viết tay: **kém người**
- Dịch tự động: **kém người**
- Nhận dạng tiếng nói: **kém người**
- Phân biệt nhập nhằng trong nghĩa từ và một số bài toán xử lý ngôn ngữ tự nhiên: **kém người**

NHỮNG VẤN ĐỀ THẢO LUẬN

- Nếu máy tính có thể làm thay những việc đang được làm bởi con người thì con người sẽ ít việc (thất nghiệp).
- Con người sẽ có quá nhiều thời gian rảnh rỗi (so với quá ít như hiện nay)
- Con người cảm thấy mất vị thế thống trị về trí thông minh của họ
- Vì máy tính làm thay (và can thiệp) vào nhiều việc hàng ngày của con người, họ sẽ cảm thấy các quyền riêng tư bị xâm phạm
- Việc sử dụng nhiều hệ thống TTNT có thể làm giảm (hoặc mất đi) khả năng tư duy, giải quyết vấn đề của con người
- Sự thành công (hoàn hảo) của TTNT có ý nghĩa như là sự kết thúc của loài người