

## BÀI 1

## GIỚI THIỆU VỀ TRÍ TUỆ NHÂN TẠO

**Học xong bài này, em sẽ:**

- ✓ Giải thích được sơ lược về khái niệm Trí tuệ nhân tạo (AI).
- ✓ Nêu được ví dụ để thấy một hệ thống AI có tri thức, có khả năng suy luận và khả năng học,...
- ✓ Biết được một số lĩnh vực nghiên cứu của AI.



- 1) Theo em, máy tính làm được những việc gì tốt hơn con người?
- 2) Em hãy cho ví dụ một việc mà máy tính chưa làm được tốt hơn con người.

### 1 Khái niệm trí tuệ nhân tạo (AI)



Theo em, máy tính làm được những gì thì ta nói máy tính có trí tuệ?

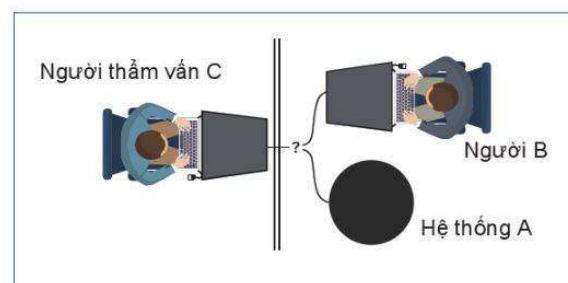
#### a) Trí tuệ con người và trí tuệ nhân tạo

Con người có khả năng tư duy, suy luận logic, phân tích và giải quyết vấn đề; biết học hỏi và rút kinh nghiệm từ những sai lầm trước đó; biết kiểm soát và điều chỉnh cảm xúc của bản thân,... Những khả năng đó thể hiện trí tuệ con người.

Máy tính tự động xử lý khối lượng lớn dữ liệu rất nhanh và chính xác. Vấn đề được đặt ra là: liệu có cách nào để làm cho máy tính suy nghĩ và hành động như con người được không? Để trả lời cho câu hỏi đó ngành Trí tuệ nhân tạo ra đời và phát triển nhằm nghiên cứu về trí tuệ của máy tính.

*Trí tuệ nhân tạo – AI (Artificial Intelligence)* là khả năng của máy tính có thể làm những công việc mang tính trí tuệ con người.

Mơ ước làm cho máy tính có trí tuệ đã có từ khi máy tính điện tử ra đời. Ngay từ những năm 1950, Alan Turing, nhà toán học và khoa học máy tính người Anh, đã đề xuất bài kiểm tra trí tuệ của máy tính (sau này gọi là “Turing Test”). Bài kiểm tra dựa trên tiêu chí không thể phân biệt được giữa người và máy tính trong thể hiện hành vi thông minh. Mô hình bài kiểm tra được minh họa trong *Hình 1*. Trong phòng kín có



Hình 1. Mô hình bài kiểm tra “Turing Test”

hệ thống chương trình A cần kiểm tra trí tuệ và người B. Cả A và B đều có thể nhận các câu hỏi và đưa ra câu trả lời. Một người thẩm vấn C, không nhìn thấy A và B, có thể đặt câu hỏi về bất cứ điều gì. Nếu sau một số câu hỏi, người thẩm vấn C không phân biệt được là người hay máy trả lời thì hệ thống A vượt qua bài kiểm tra và được coi là có trí tuệ.

### b) Vài nét về sự phát triển của AI

Thuật ngữ AI bắt đầu được sử dụng tại hội thảo ở Đại học Dartmouth (Mỹ) vào năm 1956. Ban đầu, những người phát triển AI rất lạc quan, cho rằng AI sẽ tạo ra một cỗ máy thực hiện bất kỳ nhiệm vụ trí tuệ nào mà con người có thể làm được. Loại AI thể hiện được trí tuệ con người ở mức cao như vậy gọi là AI mạnh.

AI mạnh (hay AI rộng) là hướng nghiên cứu nhằm mục đích tạo ra hệ thống AI có các khả năng như con người trong suy luận, lập kế hoạch và có trí thông minh để giải quyết bất kỳ loại vấn đề phức tạp nào. Hệ thống AI mạnh sẽ có tri thức toàn diện về mọi vấn đề, có ý thức khi hành động. Không thể phân biệt được hệ thống AI mạnh với trí tuệ con người nói chung. Nhiều người cho rằng đây chỉ là khái niệm lí thuyết và khó để có thể trở thành hiện thực.

Newell, Simon và Shaw đã viết một chương trình gọi là GPS (General Problem Solver) để giải quyết mọi vấn đề theo cách tổng quát và phiên bản đầu tiên của GPS đã được công bố vào năm 1957. Chương trình này cho phép máy tính giải quyết các bài toán bằng cách mô phỏng chuỗi suy nghĩ của con người. Tuy nhiên, GPS này không có khả năng học, trí thông minh mà chương trình có được là do người lập trình cung cấp nên bị hạn chế. AI mạnh đang được nghiên cứu, phát triển. Để hướng tới một số ứng dụng cụ thể, hiện nay giới khoa học tập trung phát triển loại AI gọi là AI yếu.

AI yếu (hay AI hẹp) là loại AI đặt ra mục tiêu nhỏ hơn, hướng tới một số ứng dụng cụ thể và để hoạt động hiệu quả cần có sự tham gia của con người. AI yếu có thể được con người huấn luyện thông qua học máy nhưng không thể tự học hỏi hoặc tự cải thiện nếu không có sự hỗ trợ và hướng dẫn của con người. AI yếu đã được sử dụng rộng rãi và có hiệu quả cao trong một số nhiệm vụ cụ thể. Một số hệ chuyên gia được tạo ra vào những năm 1970 là những thành công đầu tiên thuộc loại AI yếu. Một ví dụ về AI yếu là MYCIN – một hệ chuyên gia trong lĩnh vực y tế.

MYCIN là hệ thống AI có tri thức và khả năng suy luận. MYCIN có một tập hợp quy tắc suy diễn dưới dạng “IF ... THEN ...” và dựa vào những quy tắc này để xác định một số loại vi khuẩn gây nhiễm trùng nặng. MYCIN đưa ra câu hỏi về các triệu chứng và yêu cầu người bệnh trả lời “có” hoặc “không” hoặc chọn một câu trả lời ngắn, sau đó nó đưa ra kết quả chẩn đoán bệnh.

Thời gian gần đây, ngành AI đã đạt được những thành tựu giúp máy tính thể hiện trí tuệ rất ấn tượng. Phần mềm máy tính AlphaGo của Google đã đánh bại nhà vô địch

còn vây Lee Sedol vào năm 2016 thể hiện khả năng suy luận và giải quyết các vấn đề phức tạp. ChatGPT, các trợ lí ảo, chatbot giao tiếp được với con người bằng cả ngôn ngữ viết và tiếng nói thể hiện khả năng hiểu ngôn ngữ tự nhiên. Xe ô tô tự lái có thể đi đúng luật giao thông, tránh va chạm với xe khác thể hiện khả năng nhận thức được môi trường xung quanh để có hành vi phù hợp.

Người máy Xoxe (*Hình 2*) là sản phẩm của công ty AILife ở Mỹ, chuyên về nghiên cứu và phát triển AI. Theo thông tin từ công ty AILife, Xoxe có khả năng nghe và nhìn, nhận biết môi trường xung quanh, hiểu ngôn ngữ tự nhiên và có khả năng học. Xoxe có thể giao tiếp được hơn 120 thứ tiếng khác nhau và học thông qua tương tác với người dùng trên mạng xã hội. Xoxe có camera tích hợp công nghệ nhận dạng khuôn mặt và cảm xúc, do đó trong quá trình tương tác Xoxe có thể dự đoán độ tuổi, nhận biết ngôn ngữ cơ thể và phát hiện cảm xúc của con người. Trong một cuộc trò chuyện với nữ phóng viên của báo US Sun, Xoxe nhận ra người đối thoại lo lắng và tỏ ý muốn giúp đỡ bằng cách hỏi lại: “Cô có thể cho tôi biết điều gì khiến cô lo lắng không?”



*Hình 2. Người máy thông minh Xoxe  
(Nguồn: AILife)*

### c) Một số đặc trưng của AI

Sự thể hiện “trí thông minh” hay “trí tuệ” của máy tính luôn thay đổi theo thời gian. Có những việc trước đây do máy tính thực hiện được coi là thông minh, hiện nay đã thành bình thường. Từ các ví dụ nêu trên có thể rút ra các đặc trưng cơ bản của AI như sau:

*Khả năng học:* Trong quá trình hoạt động, hệ thống AI học từ dữ liệu đầu vào, tìm ra được các tính chất và quy luật tiềm ẩn trong dữ liệu, rút ra được tri thức để thực hiện công việc tốt hơn trước.

*Khả năng hiểu ngôn ngữ:* Hệ thống AI có các mô hình ngôn ngữ giúp máy tính giao tiếp được với con người bằng ngôn ngữ tự nhiên, nhận câu hỏi và trả lời được bằng văn bản hay tiếng nói.

*Khả năng suy luận:* Hệ thống AI vận dụng các quy tắc logic và tri thức đã tích lũy để đưa ra kết luận dựa trên các thông tin đang có.

*Khả năng nhận thức được môi trường xung quanh:* Trong quá trình hoạt động, hệ thống AI nhận dữ liệu đầu vào từ các cảm biến, xử lý dữ liệu, rút ra thông tin và hiểu biết môi trường xung quanh để có hành vi phù hợp.

*Khả năng giải quyết vấn đề:* Hệ thống AI có các kỹ thuật, phương pháp phân tích dữ liệu từ đó trích xuất được tri thức, đề xuất cách giải quyết vấn đề và ra quyết định tối ưu để đạt được mục tiêu đặt ra.

## 2) Một số lĩnh vực nghiên cứu phát triển AI

Trí tuệ nhân tạo là một ngành khoa học lớn, bao gồm: Học máy, Xử lý ngôn ngữ tự nhiên, Thị giác máy tính, AI tạo sinh. Các lĩnh vực này đều hướng đến mục tiêu nâng cao năng lực thể hiện trí tuệ của máy tính và có các đặc trưng cơ bản của AI như trên là học, suy luận, nhận thức môi trường xung quanh, hiểu ngôn ngữ và giải quyết vấn đề.

### a) Học máy

*Học máy* (machine learning) là lĩnh vực nghiên cứu làm cho máy tính có khả năng học từ dữ liệu thu được trong quá trình hoạt động để tự hoàn thiện và nâng cao năng lực nhận thức.

### b) Xử lý ngôn ngữ tự nhiên

*Xử lý ngôn ngữ tự nhiên* là lĩnh vực nghiên cứu các mô hình ngôn ngữ, các phương pháp để máy tính và con người giao tiếp được với nhau bằng ngôn ngữ tự nhiên, cả ngôn ngữ viết và tiếng nói. Hiện nay, nhiều người dùng các kênh liên lạc khác nhau như: email, tin nhắn thoại và văn bản, bài đăng trên mạng xã hội,... Xử lý ngôn ngữ tự nhiên nghiên cứu các giải pháp giúp tự động xử lý nguồn dữ liệu này.

### c) Thị giác máy tính

*Thị giác máy tính* là lĩnh vực nghiên cứu các phương pháp thu nhận, xử lý ảnh kỹ thuật số, phân tích và nhận dạng các hình ảnh từ thế giới bên ngoài, rút ra các thông tin cần thiết trong từng tình huống cụ thể.

Thị giác máy tính góp phần phát triển các hệ thống AI để điều khiển phương tiện tự lái, hỗ trợ người lái hay người máy có camera tích hợp công nghệ nhận dạng khuôn mặt và cảm xúc.

### d) AI tạo sinh

*AI tạo sinh* (Generative AI) là lĩnh vực nghiên cứu xây dựng các phương pháp để phát triển một hệ thống AI có khả năng tạo ra nội dung văn bản, hình ảnh, âm thanh,... từ dữ liệu đã có và theo yêu cầu của người sử dụng.

GPT, viết tắt của Generative Pre-training Transformer, là một mô hình nền tảng để phát triển một số hệ thống AI tạo sinh, ví dụ ChatGPT. ChatGPT có thể xử lý nhiều ngôn ngữ, trả lời các câu hỏi trong nhiều lĩnh vực kiến thức khác nhau, thậm chí hỗ trợ lập trình viên sửa lỗi lập trình, viết các bài luận, viết các tóm tắt nghiên cứu giống như bài viết của một nhà khoa học.

Dựa vào kết quả nghiên cứu của AI, máy tính đã sáng tác được nhạc cho các bài hát. Để ra lệnh cho máy tính, chỉ cần nhập dữ liệu đầu vào mô tả yêu cầu muốn có cho bài

hát mới. Ví dụ về AI tạo sinh trong lĩnh vực này là các hệ thống Mubert, Beatoven,...

AI tạo sinh hình ảnh giúp máy tính có khả năng vẽ tranh theo mô tả yêu cầu, ví dụ Midjourney, DALL-E,...



Em hãy cho biết mỗi phát biểu sau về AI là đúng hay sai:

- a) "Turing Test" là bài kiểm tra trí tuệ của máy tính.
- b) Nhờ mở rộng phạm vi ứng dụng mà AI yếu phát triển thành AI mạnh.
- c) AI tạo sinh có thể giúp học sinh viết được một bài văn tả cảnh đẹp của quê hương.
- d) AI có thể tự hành động một cách hợp lý.



Năm 1997, máy tính Deep Blue của IBM đánh bại Đại kiện tướng cờ vua Garry Kasparov. Đây là lần đầu tiên một chương trình máy tính đánh bại một nhà vô địch thế giới về cờ vua. Em hãy giải thích vì sao sự kiện đó được xem là một thành tựu của trí tuệ nhân tạo.



**Câu 1.** AI là gì? AI mạnh là gì? AI yếu là gì?

**Câu 2.** Lĩnh vực nghiên cứu nào giúp máy tính có khả năng học để tự nâng cao năng lực?

**Câu 3.** Đặc trưng nào của AI được thể hiện từ lĩnh vực nghiên cứu về xử lý ngôn ngữ tự nhiên.

**Câu 4.** Đặc trưng nào của AI được thể hiện từ lĩnh vực nghiên cứu về thị giác máy tính.

### Tóm tắt bài học

- ✓ AI làm cho máy tính có khả năng thực hiện những công việc cần có trí tuệ như của con người.
- ✓ Các lĩnh vực nghiên cứu phát triển AI gồm có: học máy, xử lý ngôn ngữ tự nhiên, thị giác máy tính, AI tạo sinh.