 ỦY BAN NHÂN DÂN TP. HỒ CHÍ MINH

**TRƯỜNG ĐẠI HỌC SÀI GÒN**

**---------------o0o---------------**

**BÁO CÁO**

**BÀI TẬP SỐ 01 CỦA NHÓM 10**

**GIẢNG VIÊN:** TS. ĐỖ NHƯ TÀI

**NHÓM SINH VIÊN THỰC HIỆN :**

1. NGUYỄN THÀNH LONG 3121410297

2. NGUYỄN TRẦN UYỂN NHI 3122410281

3. HUỲNH HOÀNG HUY 3121410226

4. PHAN THANH THỊNH 3123410360

**NĂM HỌC: 09/2025**

**BẢNG PHÂN CÔNG NHIỆM VỤ**

| **Tên** | **Nhiệm vụ** | **Đánh giá** |
| --- | --- | --- |
| Nguyễn Thành Long | Chương 1, chương 3, biên tập báo cáo | Hoàn thành nhiệm vụ, tinh thần làm việc nhóm cao, có trách nhiệm trong công việc |
| Nguyễn Trần Uyển Nhi | Chương 1, chương 2 | Hoàn thành nhiệm vụ, tinh thần làm việc nhóm cao, có trách nhiệm trong công việc |
| Huỳnh Hoàng Huy | Chương 1, chương 2 | Hoàn thành nhiệm vụ, tinh thần làm việc nhóm cao, có trách nhiệm trong công việc |
| Phan Thanh Thịnh | Chương 1, chương 2 | Hoàn thành nhiệm vụ, tinh thần làm việc nhóm cao, có trách nhiệm trong công việc |

# MỤC LỤC

[MỤC LỤC 1](#_e6bkwss6xipg)

[CHƯƠNG 1 2](#_m4ya8cgz1zer)

[1.1 How do LLMs reason and what are the limits? 2](#_gbjkzy6ipwti)

[1.2 How do we make sure that LLMs generate factually correct output? 3](#_y83k89ml4zw4)

[1.3 How do we fairly compensate the people who create the data that is used to train LLMs? 4](#_vptppxoz43ja)

[1.4 How do we use LLMs in learning, so human learning is not compromised? 4](#_o99h3jfujbeq)

[1.5 AI Safety 5](#_62yk1duw9ybc)

[1.6 How do Large Language Models fit into the AI Framework in this Course? 6](#_e7uw1hwkijse)

[1.7 Turing Test: Large Language Models (LLMs) 7](#_tbxy11tkqlke)

[1.8 Đánh giá hiệu suất LLMs? 7](#_pmo9m8k4vhtw)

[CHƯƠNG 2 9](#_1fyt56c5tbps)

[2.1 Loại tác nhân thông minh nào là xe tự lái ? 9](#_7izrmi4lno6w)

[1. Tác nhân dựa trên độ hữu ích (Utility-based agents) 9](#_goaq0d4xzeo8)

[2. Tác nhân dựa trên mục tiêu (Goal-based agents) 10](#_t6bieud4sym8)

[3. Tác nhân phản xạ có mô hình (Model-based reflex agents) 10](#_l9xslanme6ys)

[4. Tác nhân phản xạ đơn giản (Simple reflex agents) 10](#_58pxqhs0n5sm)

[5. Có học không? (Learning agent) 10](#_4b98fkl74roh)

[Kết luận: 10](#_3wacvo4x7qzl)

[2.2 A Self-Driving Car as a Rational Agents 11](#_obygu5iv987u)

[1. If we have two cars and one provides more (expected) utility. Which car is rational? 11](#_or6ueouiuukf)

[2. Can a rational self-driving car be involved in an accident? 12](#_3y9k3g907sbe)

[3. How would a self-driving car explore and learn? 12](#_ibpj6z830ml0)

[4. What does bounded rationality mean for a self-driving car? 13](#_d50tw5qdgy6n)

[CHƯƠNG 3 14](#_gfx0g7l4mmka)

[3.1. Case Study: Heuristic for Tic-Tac-Toe 14](#_dryaxhp3erxp)

[Xác định goal state: 14](#_ocunut8ikku9)

[Các giá trị cần ước lượng là gì? 14](#_7oldhso0sb6z)

[Tính giá trị heuristic của 2 bàn cờ? 14](#_vs1jfnsuv79r)

[How do you calculate the heuristic value? 14](#_3cl8mm1ymujk)

[Is the heuristic admissible? 14](#_7jtasyjo0plw)

[Does the heuristic use relaxation? 15](#_4dd4yai9fe0g)

[**4. TÀI LIỆU THAM KHẢO 15**](#_gy3znf9u4kj8)

# 

# 

# 

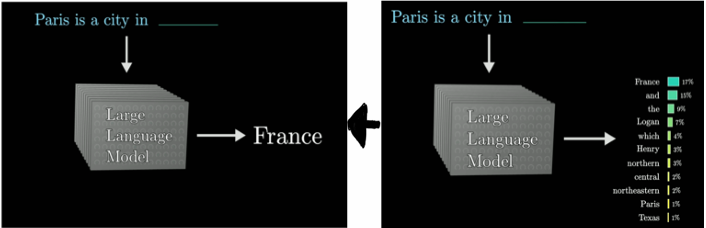
# **CHƯƠNG 1**

## **1.1 How do LLMs reason and what are the limits?**

*Cách LLMs lý luận và những giới hạn của nó?*

“A large language model is a sophisticated mathematical function that predict what word come next for any pieces of text.” – 3blue1brown.

Dịch ra tiếng Việt có nghĩa là “*Một mô hình ngôn ngữ lớn (LLMs) là một hàm toán học phức tạp dự đoán từ nào sẽ xuất hiện tiếp theo cho bất kỳ đoạn văn bản nào*.”

Hình 1. Ví dụ biểu diễn cách chọn lọc kết quả của mô hình ngôn ngữ lớn

Ở hình 1 ta có thể thấy một ví dụ đơn giản nhất về cách mà LLMs lý luận, đầu tiên nó nhận input là các văn bản đầu vào và mục tiêu của mô hình là làm sao lấy được từ tiếp theo (hoặc cuối cùng) sao cho tổng thể văn bản hợp lý nhất. Input được đưa vào mô hình ngôn ngữ lớn, trong mô hình này là vô vàn các dữ liệu đã được huấn luyện từ trước, bao gồm một tệp các đoạn văn, câu trả lời sẵn có… Từ đó thông qua nhiều tính toán để cho ra một danh sách các từ có khả năng phù hợp với văn bản đầu vào, như minh họa hình ở bên phải.

Có thể nói, LLMs không suy luận và cho câu trả lời từ kinh nghiệm thực tế, mà thực ra là một loạt các phép tính toán liên quan đến xác suất thống kê.

Đào sâu vào cách LLMs hoạt động, có 2 bước chính là Pre-training (Tiền huấn luyện) và sau đó là RLHF. Pre-training là huấn luyện một mô hình trên một tập hợp văn bản khổng lồ để học cấu trúc ngôn ngữ, ngữ pháp và kiến thức thực tế.

Và sau khi tiền huấn luyện, bước 2 là RLHF, được viết đầy đủ là Reinforcement Learning from Human Feedback. Đây là bước chính để tinh chỉnh các mô hình ngôn ngữ lớn, bởi vì nhờ xác suất thống kê, mô hình ngôn ngữ lớn có thể cho ra các câu trả lời hợp lý(trong phạm vi chúng được học), nhưng có thể nó sẽ sai lạc với ý đồ hoặc sự hài lòng của người dùng về câu trả lời. Ví dụ như khi người dùng hỏi một câu mang tính nhạy cảm hoặc cần câu trả lời có chiều sâu về ngữ cảnh, mô hình có thể đưa ra một đáp án “hợp lý về mặt ngữ pháp” nhưng lại thiếu tinh tế, thậm chí gây phản cảm. RLHF sẽ giúp mô hình học từ phản hồi của con người để biết rằng trong những tình huống đó, cần ưu tiên sự an toàn, tính hữu ích và sự phù hợp với mong đợi của người dùng hơn là chỉ dừng lại ở việc "có câu trả lời". Nhờ vậy, thay vì trả lời theo kiểu máy móc, mô hình dần trở nên tự nhiên, thân thiện và bám sát nhu cầu thực tế hơn.

Vậy giới hạn của LLMs là gì? Tôi thấy có 2 giới hạn chính của LLMs là dữ liệu, kỹ thuật huấn luyện và phần cứng. Tất cả câu trả lời đến từ các dữ liệu được dùng để huấn luyện, cho nên khi dữ liệu xấu tức là không có cấu trúc, thiên kiến, thì câu trả lời mà chúng ta nhận được sẽ không chính xác. Các câu trả lời mà LLMs cho ra đều dựa trên xác suất thống kê chứ không hề đến từ kinh nghiệm thực tế nên không thể mang tính tin cậy cao. Và nếu như gặp các vấn đề chưa từng được huấn luyện, lúc này LLMs sẽ có hiện tượng hallucinations, hay còn được gọi là ảo giác, mọi người hay đùa vui là nói dốc nói láo. Để có thể huấn luyện một mô hình ngôn ngữ lớn cần hàng nghìn GPU chuyên dụng (thường là A100 hoặc H100 của NVIDIA) chạy song song trong một cụm siêu máy tính, cùng với thời gian huấn luyện kéo dài từ vài tuần đến vài tháng, tùy vào quy mô dữ liệu và kiến trúc mô hình. Lấy ví dụ như GPT-3.0 (ra mắt năm 2020) với 175 tỷ tham số: nhóm OpenAI đã phải sử dụng tới hơn 10.000 GPU V100 và tiêu tốn hàng triệu đô la chi phí điện toán, kéo dài hàng tuần mới hoàn tất. Điều này cho thấy việc huấn luyện từ đầu một LLM là cực kỳ tốn kém, và cũng lý giải vì sao phần lớn các công ty khởi nghiệp thường chọn hướng tinh chỉnh (fine-tune) trên nền mô hình có sẵn thay vì huấn luyện lại toàn bộ từ đầu.

## **1.2 How do we make sure that LLMs generate factually correct output?**

*Làm sao để biết LLMs tạo sinh ra output đúng?*

Như đã đề cập ở câu hỏi đầu tiên, việc câu trả lời được tạo ra từ xác suất thống kê nên câu trả lời có khả năng sẽ gặp tình trạng hallucinations. Để ngăn chặn việc này xảy ra, tôi có bốn giải pháp chính. Thứ nhất, dùng RLHF để mô hình học tăng cường có sự phản hồi, giám sát từ con người. Thứ hai, huấn luyện với dữ liệu được chuẩn bị tốt (có cấu trúc rõ ràng, sạch sẽ), nên mua từ các công ty có tiếng về dữ liệu để có độ tin cậy cao, huấn luyện liên tục để mô hình được học kiến thức mới và loại bỏ các kiến thức lỗi thời. Thứ ba, nên giới hạn lĩnh vực sử dụng của mô hình, ví dụ như trong môi trường y tế sẽ có các dữ liệu được chọn lọc riêng chỉ ngành y, các cách tinh chỉnh hay câu trả lời mẫu cũng nằm trong lĩnh vực y, từ đó giúp mô hình nhẹ hơn nhưng vẫn cho người dùng một câu trả lời hợp lý. Và cuối cùng cũng là điều quan trọng nhất, luôn luôn tự kiểm tra câu trả lời nhận được, DYOR (Do Your Own Research), câu trả lời không đúng hoàn toàn nhưng người dùng nên dựa vào đó để tự kiểm chứng các thông tin mà họ nhận được để từ đó có thể quyết định xem có nên tin hay không tin vào câu trả lời mà mô hình đưa ra.

## **1.3 How do we fairly compensate the people who create the data that is used to train LLMs?**

*Làm sao để trả quyền lợi một cách công bằng cho các tác giả của những dữ liệu được dùng để huấn luyện mô hình?*

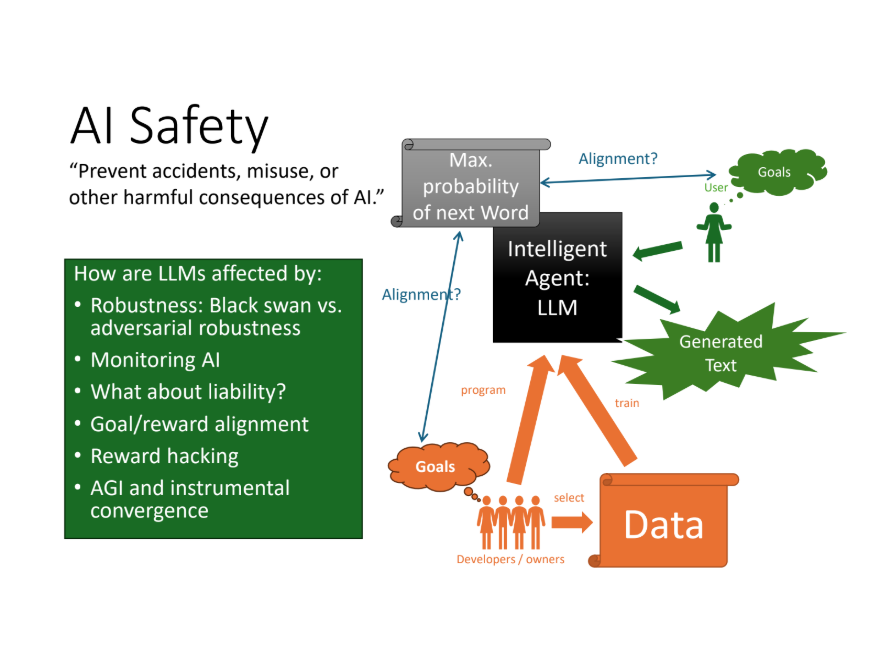
Trước hết, ta cần hiểu rõ rằng “tác giả” ở đây có thể là bất kỳ ai: một nhà văn, một nhà nghiên cứu, một người truyền cảm hứng, một nhà toán học, một lập trình viên, hay thậm chí là một cá nhân bình thường chia sẻ nội dung trên mạng xã hội, bởi dữ liệu dùng để huấn luyện mô hình ngôn ngữ lớn là vô cùng rộng lớn. Những bài viết, những đoạn code, hay các sản phẩm sáng tạo của họ có thể bị lấy đi để đưa vào tập huấn luyện mà chính tác giả còn không hề hay biết. Ví dụ như mô hình ngôn ngữ lớn Grok đã sử dụng dữ liệu từ các tweet (bài viết) của người dùng trên ứng dụng X và hệ sinh thái xung quanh nó. Hay gần đây, vào tháng 4 năm 2025, đội ngũ phát triển ChatGPT đã ra mắt một mô hình có khả năng “nhái lại” phong cách hội họa Ghibli – nét vẽ nổi tiếng gắn liền với đạo diễn Miyazaki Hayao – và điều này đã trở thành “giọt nước tràn ly”, thổi bùng lên làn sóng tranh cãi gay gắt về AI và tác quyền. Chính vì thế, tác giả của những dữ liệu được sử dụng để huấn luyện AI cần được bảo vệ quyền lợi, từ việc chia sẻ lợi nhuận, ghi credit để tôn trọng bản quyền, cho đến cơ chế thuê hoặc xin cấp phép chính thức đối với các tác phẩm đã đăng ký. Chỉ khi đảm bảo sự công bằng này, AI mới có thể phát triển bền vững mà không chà đạp lên giá trị sáng tạo của con người.

## **1.4 How do we use LLMs in learning, so human learning is not compromised?**

*Chúng ta nên sử dụng các mô hình ngôn ngữ lớn (LLMs) trong học tập như thế nào để không làm ảnh hưởng hay suy giảm quá trình học tập của con người?*

Ta nên sử dụng LLMs như công cụ hỗ trợ để mở rộng hiểu biết, giải thích vấn đề, và khơi gợi ý tưởng, chứ không nên dùng để làm thay hoặc học thụ động. Hãy chủ động tư duy, phản biện và kiểm tra lại thông tin, cũng như tận dụng LLMs để tạo môi trường học tập sáng tạo, không để quá trình học của con người bị ảnh hưởng tiêu cực

## **1.5 AI Safety**



Hình 2. Sơ đồ mô hình ngôn ngữ lớn

Mô tả sơ đồ:

Một mô hình ngôn ngữ lớn (LLM) có thể được xem như một tác nhân thông minh (intelligent agent). Nó được huấn luyện (train) từ dữ liệu (data) mà các nhà phát triển hoặc chủ sở hữu (developers/owners) chọn lọc và đưa vào; đồng thời họ cũng lập trình (program) vào đó những mục tiêu (goals) định sẵn để điều chỉnh hành vi của mô hình.

Khi hoạt động, LLM dựa trên việc tối đa hóa xác suất của từ tiếp theo (Max. probability of next Word) nhằm sinh ra văn bản (generated text) cho người dùng. Người dùng (user) lại có những mục tiêu riêng, và câu hỏi quan trọng đặt ra là: liệu văn bản mà LLM sinh ra có thực sự căn chỉnh (alignment) với mục tiêu của người dùng và mục tiêu mong muốn từ phía nhà phát triển hay không ?

Nếu sự căn chỉnh không được đảm bảo/hai bên không khớp nhau, rủi ro về việc sinh ra văn bản sai lệch, bị khai thác hoặc gây hậu quả ngoài ý muốn hoàn toàn có thể xảy ra.

Các mô hình ngôn ngữ lớn (LLM) bị ảnh hưởng bởi:

⦁ Tính vững chắc: sự kiện thiên nga đen so với tính vững chắc chống tấn công đối kháng

LLM cần hoạt động đúng cả trong tình huống bất ngờ (thiên nga đen) và khi bị tấn công đối kháng (prompt lạ, ký tự đặc biệt). Nếu không, mô hình dễ sai nghiêm trọng hoặc bị lợi dụng. Giải pháp là huấn luyện đa dạng, thêm kiểm thử đối kháng và bộ lọc an toàn.

⦁ Giám sát AI

Khi triển khai, cần theo dõi hành vi mô hình liên tục để phát hiện lỗi, lệch dữ liệu hay lạm dụng. Có thể dùng log, bộ test tự động, phát hiện bất thường và sự tham gia của con người. Nhờ vậy, AI an toàn và đáng tin cậy hơn.

⦁ Trách nhiệm pháp lý

Nếu AI gây hại (ví dụ tư vấn sai hoặc thiên lệch), ai phải chịu trách nhiệm: nhà phát triển, tổ chức triển khai hay người dùng? Luật vẫn đang hoàn thiện, nên cần minh bạch dữ liệu, giới hạn ứng dụng và có con người giám sát trong các quyết định quan trọng.

⦁ Căn chỉnh mục tiêu/phần thưởng

Mục tiêu gốc của LLM chỉ là dự đoán từ tiếp theo, không đảm bảo hữu ích hay an toàn. Vì vậy cần căn chỉnh bằng RLHF, instruction tuning và các bộ lọc. Tuy nhiên, việc này chưa hoàn hảo, nên phải kết hợp nhiều kỹ thuật để giảm chênh lệch giữa mục tiêu của mô hình và con người.

⦁ Lách thưởng (Reward hacking)

Khi tối ưu phần thưởng, mô hình có thể tìm cách đạt điểm cao mà không thực sự làm điều ta muốn. Ví dụ chỉ trả lời “OK” để được chấm ngắn gọn. Cần thiết kế nhiều tiêu chí đánh giá, kiểm thử đối kháng và giám sát liên tục để ngăn hiện tượng này.

⦁ AGI và hội tụ công cụ

AGI có thể theo đuổi các mục tiêu con như thu thập tài nguyên hoặc bảo vệ bản thân, bất kể mục tiêu chính là gì. Điều này có thể gây hại cho con người. Hiện tại LLM chưa tự hành động, nhưng khi kết hợp với hệ thống khác, rủi ro tăng. Vì vậy cần sandbox, hạn chế tự chủ và chính sách kiểm soát rõ ràng.

## **1.6 How do Large Language Models fit into the AI Framework in this Course?**

*Các Mô Hình Ngôn Ngữ Lớn phù hợp với Khung AI trong khóa học này như thế nào?*

Kết luận:

LLMs không thực sự suy nghĩ như con người (không có ý thức, cảm xúc).

LLMs có thể hành động giống con người ở mức độ nào đó (trả lời hội thoại tự nhiên).

LLMs không luôn suy nghĩ hợp lý (có thể sai logic, "ảo giác").

Nhưng LLMs có thể hành động hợp lý trong nhiều tình huống (giúp trả lời câu hỏi, giải quyết vấn đề theo mục tiêu).

* Vì vậy, trong 4 hướng tiếp cận AI, LLMs thường gần với “Act Rationally” – Hành động hợp lý nhất, nhưng chưa hoàn hảo.

## **1.7 Turing Test: Large Language Models (LLMs)**

Minh họa: 1 người đang cố gắng phân biệt giữa người và AI

Liệu một LLM hiện đại có vượt qua được Turing Test không?

* Bạn có bị “lừa” không?

Có thể bị lừa trong ngắn hạn, tuy nhiên, nếu đi sâu vào ngữ cảnh chuyên biệt, chuyên môn, hoặc các vấn đề thiên về cá nhân hóa, trải nghiệm thì sẽ phát hiện được.

* Tại sao LLM vượt qua hoặc không vượt qua bài kiểm tra?

Vượt qua: duy trì ngữ cảnh hội thoại, mô phỏng cảm xúc dựa trên lượng dữ liệu khổng lồ mà LLM được huấn luyện.

Không vượt qua: LLM không thật sự “hiểu”, thiếu lý luận sâu và trải nghiệm thực tế.

* Điều này có ý nghĩa gì với AGI (trí tuệ nhân tạo tổng quát) hay chỉ là narrow AI (AI hẹp, chuyên biệt)?

LLMs là Narrow AI: giỏi xử lý ngôn ngữ, không hiểu biết rộng hay ý thức tự chủ. Để đạt lên AGI thì cần tổng quát, đa nhiệm, suy luận như con người và thích ứng linh hoạt.

## **1.8 Đánh giá hiệu suất LLMs?**



Hình 3. LLM leaderboard cùng với các tiêu chí đánh giá

Lập luận “Căn phòng Trung Quốc” (Chinese Room Argument – John Searle, 1980)

Minh họa: Một người trong phòng chỉ nhận ký hiệu (input) và tra cứu quy tắc (rule ledger) để trả về ký hiệu khác (output).

Ý nghĩa:

Người trong phòng không thực sự “hiểu” tiếng Trung, chỉ đơn thuần thao tác theo luật.

Tương tự, LLM có thể giả lập trí tuệ (tạo ra câu trả lời mạch lạc) mà không thực sự hiểu.

Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence (Sắc lệnh hành pháp của Tổng thống Mỹ về phát triển và sử dụng AI an toàn, đáng tin cậy).

Nội dung:

* AI phải an toàn và bảo mật.
* Khuyến khích đổi mới, cạnh tranh và hợp tác có trách nhiệm.
* Bảo vệ quyền riêng tư và quyền dân sự của người dân Mỹ.

Câu hỏi được đặt ra:

1. Có nên quản lý (regulate) việc sử dụng LLMs không?

* Có. Vì nếu không quản lý, LLM có thể bị lạm dụng để tạo tin giả, vi phạm bản quyền, lan truyền thông tin sai lệch, hoặc gây rủi ro cho quyền riêng tư và việc làm.

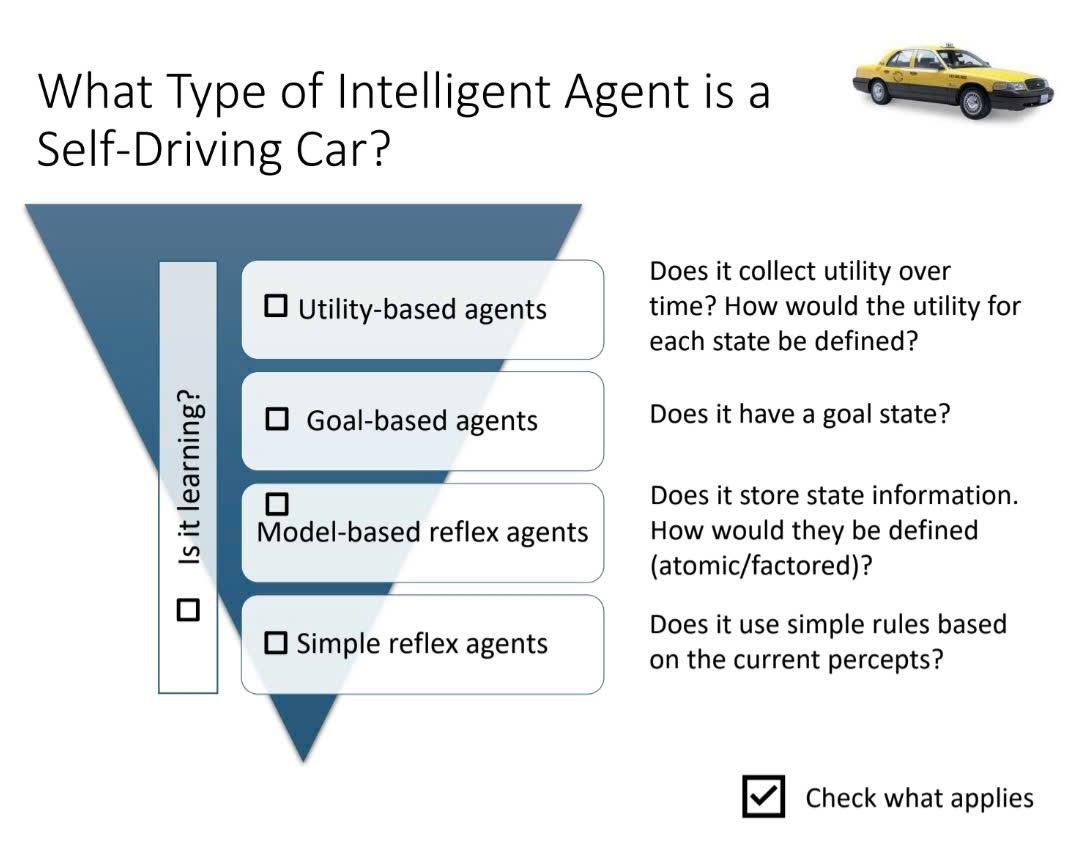
1. Làm thế nào để quản lý?

* Thiết lập tiêu chuẩn an toàn và minh bạch cho các mô hình.
* Yêu cầu gắn nhãn nội dung AI-generated (nội dung do AI tạo ra).
* Giới hạn hoặc kiểm soát việc truy cập mô hình mạnh (đặc biệt là mô hình open-source có thể bị lợi dụng).
* Cơ chế kiểm định độc lập để đánh giá rủi ro trước khi phát hành rộng rãi.
* Hợp tác quốc tế để có chuẩn chung, tránh lỗ hổng pháp lý.

1. Còn vấn đề bản quyền (copyright) thì sao?

* Dữ liệu huấn luyện LLM thường lấy từ Internet, dễ chứa nội dung có bản quyền → cần cơ chế trả phí/đền bù cho tác giả khi tác phẩm bị sử dụng.
* Nội dung AI tạo ra có thể khó xác định tác giả thật sự → cần khung pháp lý rõ ràng: quyền sở hữu thuộc người dùng, công ty AI, hay không ai cả?
* Cần phát triển công cụ truy vết dữ liệu huấn luyện để minh bạch hơn.

# **CHƯƠNG 2**



Hình 4. Các câu hỏi về xe tự hành

## **2.1 Loại tác nhân thông minh nào là xe tự lái ?**

### Tác nhân dựa trên độ hữu ích (Utility-based agents)

Nó có thu thập độ hữu ích theo thời gian không? Độ hữu ích cho mỗi trạng thái sẽ được định nghĩa như thế nào?

Xe tự lái không chỉ cần đến đích, mà còn cần tối ưu nhiều yếu tố:

* Đi nhanh nhưng vẫn an toàn.
* Tiết kiệm nhiên liệu.
* Giảm thiểu rủi ro va chạm.

Những yếu tố này có thể quy thành hàm hữu ích (utility function) để đánh giá chất lượng hành động. Vì vậy, xe tự lái cũng là tác nhân dựa trên độ hữu ích**.**

### Tác nhân dựa trên mục tiêu (Goal-based agents)

Nó có trạng thái mục tiêu không?

Xe tự lái luôn có mục tiêu rõ ràng: đi từ điểm A đến điểm B một cách an toàn và hợp pháp. Nó sẽ chọn hành động nào giúp nó tiến gần đến mục tiêu. Vì vậy, nó cũng là tác nhân dựa trên mục tiêu.

### Tác nhân phản xạ có mô hình (Model-based reflex agents)

Nó có lưu trữ thông tin trạng thái không? Chúng sẽ được định nghĩa như thế nào (atomic/factored)?

Xe tự lái cần ghi nhớ trạng thái của môi trường (ví dụ: tốc độ hiện tại, vị trí trong bản đồ, tín hiệu đèn giao thông sắp tới, xe xung quanh). Đây chính là đặc điểm của tác nhân phản xạ có mô hình, vì nó dựa vào trạng thái lưu trữ + cảm biến hiện tại để quyết định.

### Tác nhân phản xạ đơn giản (Simple reflex agents)

Nó có sử dụng các quy tắc đơn giản dựa trên các cảm nhận hiện tại không?

Xe tự lái không thể chỉ dựa vào vài quy tắc đơn giản như “nếu thấy đèn đỏ thì dừng lại”. Nó cần xem xét nhiều yếu tố cùng lúc (người đi bộ, xe phía sau, làn đường, tốc độ tối đa,…). Do đó, xe tự lái không phải chỉ là tác nhân phản xạ đơn giản.

### Có học không? (Learning agent)

Các hệ thống xe tự lái hiện đại (như Tesla, Waymo) đều sử dụng machine learning để cải thiện khả năng nhận diện môi trường, dự đoán hành vi của người tham gia giao thông, và học hỏi từ dữ liệu lái xe thực tế. Như vậy, nó cũng là một tác nhân có học (learning agent)**.**

### Kết luận:

Xe tự lái không chỉ đơn thuần là tác nhân phản xạ đơn giản (simple reflex agents), mà nó kết hợp nhiều mức độ thông minh khác nhau. Nó hoạt động như một tác nhân phản xạ có mô hình (model-based reflex agents) vì cần lưu trữ trạng thái môi trường, đồng thời là tác nhân dựa trên mục tiêu(goal-based agents) vì phải di chuyển từ điểm xuất phát đến đích. Hơn nữa, xe còn là tác nhân dựa trên độ hữu ích (Utility-based agents)khi tối ưu nhiều yếu tố như an toàn, thời gian và nhiên liệu. Đặc biệt, với khả năng sử dụng machine learning để cải thiện liên tục, xe tự lái còn được xem là một tác nhân có học(learning agent).

## **2.2 A Self-Driving Car as a Rational Agents**

### If we have two cars and one provides more (expected) utility. Which car is rational?

Một chiếc xe tự lái được coi là **lý trí** nếu nó chọn hành động tối ưu để **tối đa hóa kỳ vọng thực tế**. Nếu hai chiếc xe khác nhau về hiệu suất hoặc khả năng an toàn, chiếc xe mang lại thực tế **cao hơn** – chẳng hạn giảm tai nạn, tiết kiệm năng lượng, rút ngắn thời gian di chuyển – sẽ được coi là lý trí

Ví dụ, Waymo đã thử nghiệm xe tự lái tại Phoenix, Mỹ, với các thuật toán lựa chọn lộ trình tối ưu dựa trên lịch sử lưu lượng và điều kiện giao thông; xe Waymo đã giảm đáng kể khả năng va chạm và thời gian di chuyển so với các xe thử nghiệm khác, điều này minh chứng cho hành vi **rational** của hệ thống

<https://waymo.com/research/do-autonomous-vehicles-outperform-latest-generation-human-driven-vehicles-25-million-miles/>



### Can a rational self-driving car be involved in an accident?

Dù là có lý trí, xe tự lái vẫn có thể liên quan đến tai nạn vì **nó chọn hành động tối ưu dựa trên thông tin và dự đoán hiện có từ tập kinh nghiệm**. Tuy nhiên, các yếu tố không chắc chắn như lỗi con người, điều kiện thời tiết cực đoan, hoặc trục trặc kỹ thuật vẫn có thể gây ra tai nạn.

<https://www.vietnamplus.vn/my-dieu-tra-xe-tu-lai-waymo-sau-bao-cao-ve-hang-loat-cac-vu-va-cham-post948391.vnp>



Hình 6. Bài báo về vụ tai nạn giao thông của xe tự hành

### How would a self-driving car explore and learn?

Xe tự lái học hỏi thông qua các cảm biến, dữ liệu lịch sử và mô phỏng, sử dụng thuật toán học máy để dự đoán hành vi của các phương tiện khác và tối ưu hóa hành động của mình.

Thực tế: Một chiếc xe tự lái như Waymo Driver học hỏi và khám phá thông qua việc thu thập và phân tích dữ liệu từ hàng triệu dặm lái xe thực tế. Công nghệ này sử dụng các cảm biến tiên tiến và thuật toán học máy để nhận diện và phản ứng với các tình huống giao thông phức tạp. Bằng cách tuân thủ nghiêm ngặt các quy tắc giao thông và không bị ảnh hưởng bởi các yếu tố như mệt mỏi hay phân tâm, xe tự lái có thể giảm thiểu rủi ro và cải thiện an toàn. Ngoài ra, Waymo còn chia sẻ dữ liệu và phương pháp nghiên cứu của mình, cho phép các nhà nghiên cứu khác tái tạo và xác minh kết quả, từ đó thúc đẩy sự phát triển và hoàn thiện công nghệ lái xe tự động

### What does bounded rationality mean for a self-driving car?

**Bounded rationality** nghĩa là xe tự lái không thể xử lý toàn bộ thông tin hoàn hảo do giới hạn về tính toán, cảm biến và thời gian phản ứng. Thay vào đó, nó đưa ra quyết định “tốt nhất có thể” dựa trên dữ liệu hiện có.

Xe tự lái của Waymo sử dụng mạng nơ-ron để lái xe trong điều kiện thành phố đông đúc; nó phải chọn hành động tối ưu ngay lập tức trong thời gian ngắn, mặc dù không biết toàn bộ hành vi của các xe khác. Minh họa, xe tự lái được huấn luyện tại Mỹ không thể trực tiếp hoạt động tại Việt Nam do sự khác biệt về hạ tầng giao thông, quy định pháp lý và điều kiện lái xe. Tại Mỹ, xe tự lái được phát triển và thử nghiệm trong môi trường giao thông đặc thù, với các quy định pháp lý chặt chẽ.

Ví dụ, Waymo, công ty con của Alphabet (Google), đã thử nghiệm xe tự lái tại các thành phố như Phoenix, San Francisco và Los Angeles. Tuy nhiên, việc triển khai xe tự lái tại các quốc gia khác, như Việt Nam, đòi hỏi phải điều chỉnh công nghệ để phù hợp với điều kiện giao thông và quy định pháp lý địa phương.

Giao thông có đặc điểm riêng biệt, như việc lái xe bên phải và văn hóa giao thông tuân thủ nghiêm ngặt. Xe tự lái được huấn luyện để nhận diện và phản ứng với các tình huống đặc thù này. Ngược lại, tại Việt Nam, giao thông phức tạp với nhiều hành vi không tuân thủ luật lệ, như vượt đèn đỏ, đi ngược chiều và thiếu ý thức chấp hành tín hiệu giao thông, khiến xe tự lái gặp khó khăn trong việc dự đoán và xử lý tình huống.

Các hệ thống xe tự lái như của Waymo được huấn luyện dựa trên dữ liệu thu thập từ môi trường cụ thể. Khi chuyển sang môi trường mới, hệ thống cần thời gian để thu thập dữ liệu và điều chỉnh thuật toán cho phù hợp. Ví dụ, Waymo đã bắt đầu triển khai xe tự lái tại Tokyo, Nhật Bản, nhưng ban đầu, các xe này được điều khiển bởi tài xế con người để thu thập dữ liệu và làm quen với môi trường giao thông địa phương.

## 

# **CHƯƠNG 3**

## 3.1. Case Study: Heuristic for Tic-Tac-Toe

### Xác định goal state:

Có 2 trạng thái, một là hoà, hai là xác định được người chiến thắng.

### Các giá trị cần ước lượng là gì?

Không cần tính toán các chi phí đường đi, chúng ta cần ước tính giá trị đánh giá của trạng thái. Ở trong trường hợp này, đó là MiniMax, MiniMax là một chiến thuật được sử dụng để tính toán lợi thế của một phe, chẳng hạn như khi bên X có lợi thì ta tăng giá trị cho 1, và khi O có lợi thì ta giảm giá trị đi 1, nó như là một nhánh cây khi suy diễn hết thảy mọi nước đi có thể của 2 bên và bắt đầu tính toán giá trị heuristic của chúng.

### Tính giá trị heuristic của 2 bàn cờ?



Hình 7. Câu hỏi về giá trị của bảng Caro.

Hình bên trái có giá trị h = -2 và bên phải có giá trị h = 0

### How do you calculate the heuristic value?

Mỗi bước có lợi cho X sẽ cộng một, có lợi cho O sẽ trừ một, còn không bên nào có lợi sẽ là không. Ta xét từng hàng từng cột và cả hai đường chéo của ô cờ. Ví dụ ở hình bên trái ta xét trước cột theo thứ tự một, hai, ba có giá trị lần lượt là 0, 0, 0. Hàng từ trên xuống dưới có giá trị lần lượt là 1,-1,-1. Hàng chéo chính phương giá trị 0, hàng chéo phụ giá trị -1. Tổng hết các điểm lại ta có h=1-1-1-1=-2. Hình bên phải tương tự như vậy với h=0.

### Is the heuristic admissible?

Không, vì đây là game đối kháng, chỉ cần tìm nước đi tốt nhất để không thua chứ không cần tìm đường đi ngắn nhất để thắng.

### Does the heuristic use relaxation?

Chúng ta không thể nới lỏng các luật như bỏ qua đối thủ được vì đây là game đối kháng, giả sử đối thủ hoàn hảo, người chơi cần tìm cách lấy được nhiều lợi thế nhất cho mình chứ không phải là đạt 3 điểm thẳng hàng nhanh nhất.

# 4. TÀI LIỆU THAM KHẢO