**Module review 1:**

Hàm tác tử ánh xạ từ tập hợp tất cả các chuỗi cảm nhận khả dĩ P∗ đến tập hợp các hành động A, và được biểu diễn dưới dạng một hàm toán học trừu tượng.

Agent function thường được kí hiệu: f: P\* → A

Chương trình tác tử là một hiện thực cụ thể của hàm này cho một hệ thống vật lý nhất định.

Công thức: Agent= architecture( phần cứng) + Agent Program( chương trình tác tử).

Phần cứng gồm sensors( cảm biến), memory( bộ nhớ), computationl power( Năng lực tính toán).

Ex: Robot hút bụi

* Enviroment: 2 ô A và B có thể sạch (clean) hoặc bẩn ( dirty)
* Sensors: vị trí robot ( A or B), trạng thái ô ( clean or dirty)
* Possible actions: Left( qua trái), Right( qua phải), Suck ( hút bụi), NoOp(không làm gì).
* Agent function: f: P\*→ A

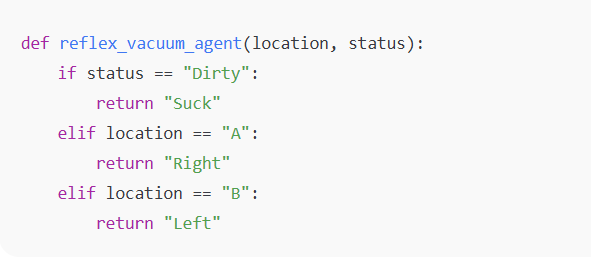
Với mỗi chuỗi percept, chọn 1 actions:

+ nếu percept= (A, dirty) → action = suck

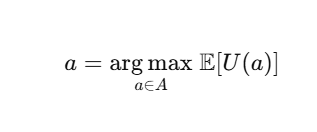
+ nếu percept= (A, clean) → action= right

+ nếu percept=( A, clean),(B, clean) →action= left

+ nếu percept=(A, clean),(B, clean),(A, dirty) →action= suck



Rational Agents: tác tử hợp lý phải chọn hành động có giá trị kỳ vọng (expected utility) cao nhất.



A: tập hợp các hành động

U(a):( utility)giá trị của hành động A

E[U(a)]: expected utility- giá trị kỳ vọng, tính cả sự bất định trong môi trường

arg max: toán tử chọnn hành động đạt giá trị lớn nhất.

* Tính hợp lý (Rationality) là một lý tưởng – nghĩa là không ai có thể xây dựng được một agent tốt hơn (về mặt lý thuyết).
* Rationality ≠ Omniscience (toàn tri) – tác tử hợp lý vẫn có thể mắc sai lầm nếu percepts (cảm nhận) và knowledge (kiến thức) không đủ để đưa ra quyết định chính xác.
* Rationality ≠ Perfection (hoàn hảo) – tác tử hợp lý chỉ tối đa hóa kết quả kỳ vọng (expected outcomes), chứ không phải lúc nào cũng đạt kết quả thực tế tốt nhất.
* Hợp lý là khám phá và học hỏi – nghĩa là sử dụng percepts để bổ sung kiến thức trước đó và dần trở nên tự chủ hơn.
* Rationality thường bị giới hạn bởi bộ nhớ, năng lực tính toán, cảm biến, v.v. → Đây gọi là bounded rationality (tính hợp lý bị giới hạn).

Một số loại môi trường: (Enviroment Types)

Full observable( quan sát đầy đủ): cảm biến agent luôn cho thấy toàn bộ trạng thái môi trường.

Partially observable (quan sát một phần): agent chỉ nhận được một phần trạng thái, cần ghi nhớ hoặc suy luận phần còn lại

Deterministic (Tất định):

+ Percepts tin cậy 100% (không có lỗi cảm biến)

+ Thay đổi của môi trường được xác định hoàn toàn bởi trạng thái hiện tại và hành động của agent

Stochastic (Ngẫu nhiên):

+ Percepts không hoàn toàn tin cậy (có nhiễu, cảm biến có thể hỏng). → gọi là stochastic sensor model

+ Hàm chuyển trạng thái là ngẫu nhiên, dẫn đến các xác suất chuyển trạng thái, thường mô hình hóa bằng Markov process

Know: agent biết trước hàm chuyển trạng thái (transition function)

Unknow: agent chưa biết hàm chuyển trạng thái,phải học bằng cách thử hành động.

**Case Study: Self- Driving Cars**

1. Nếu có hai xe, và một xe mang lại utility kỳ vọng (expected utility) cao hơn ?

* Xe nào có utility kỳ vọng cao hơn thì được coi là rational
* Utility ở đây có thể đo bằng: an toàn, thời gian, tiết kiệm nhiên liệu, thoải mái, tuân thủ luật giao thông

1. Một xe tự lái hợp lý có thể gặp tai nạn không?

* Xe vẫn có thể mắc lỗi hoặc bị bất ngờ do thông tin cảm biến thiếu, môi trường thay đổi bất ngờ (ví dụ xe khác vi phạm luật, người đi bộ bất ngờ băng qua).
* Rational ≠ Omniscient (không toàn tri)
* Rationality = chọn hành động tốt nhất có thể dựa trên thông tin hiện có, chứ không đảm bảo hoàn hảo

1. Xe tự lái học và khám phá thế nào ?

* Sử dụng percepts (camera, lidar, radar, GPS, bản đồ, dữ liệu lái trước đó) để: xây dựng mô hình thế giới, cập nhật kiến thức giao thông, học từ huấn luyện, mô phỏng, trải nghiệm

1. Bounded Rationality với xe tự lái là gì?

* Xe tự lái không có tài nguyên vô hạn (bộ nhớ, tính toán, thời gian xử lý).
* Nó phải ra quyết định nhanh trong thời gian thực, đôi khi không thể tính hết mọi khả năng.
* Bounded rationality = làm điều tốt nhất có thể trong giới hạn tài nguyên.

PEAS Description of the Environment of a Self-Driving Car:

Performance measure (Thước đo hiệu suất):

* An toàn
* Tuân thủ luật giao thông
* Thời gian đến đích( tối uu hành trình)
* Tiết kiệm nhiên liệu/ năng lượng
* Sự thoải mái của hành khách
* Tránh kẹt xe, giảm chi phí bảo trì.

Environment (Môi trường):

* Đường xá (cao tốc, nội đô, nông thôn)
* Giao thông: xe khác, xe máy, xe tải, người đi bộ, động vật.
* Biển báo, đèn giao thông, vạch kẻ đường.
* Thời tiết (mưa, sương mù, tuyết).
* Điều kiện ánh sáng (ban ngày/ban đêm).

Actuators (Bộ chấp hành):

* Tay lái (steering).
* Bàn ga (accelerator).
* Phanh (brake).
* Đèn (headlights, signals).
* Còi (horn).

Sensors (Cảm biến):

* Camera (nhìn biển báo, người đi bộ, xe khác).
* Lidar / Radar (khoảng cách và tốc độ xe khác).
* GPS (định vị).
* IMU (cảm biến quán tính: gia tốc, góc nghiêng).
* Cảm biến vận tốc bánh xe.
* Cảm biến siêu âm (đỗ xe).

Environment for a Self-Driving Car:

Fully observable vs Partially observable:

* Không fully observable → Vì cảm biến không thể thấy toàn bộ trạng thái môi trường: Camera bị che khuất, điểm mù, sương mù, mưa, ban đêm hạn chế tầm nhìn.
* Partially observable → Xe phải dự đoán những gì không nhìn thấy.

Deterministic vs Stochastic:

* Không deterministic → vì môi trường giao thông có nhiều yếu tố ngẫu nhiên: Người đi bộ có thể băng qua bất ngờ, xe khác có thể vi phạm luật, thắng gấp,cảm biến không phải lúc nào cũng chính xác 100%.
* Stochastic → Percepts có nhiễu (sensor noise), hành vi môi trường có tính xác suất → thường được mô hình bằng Markov decision process (MDP).

Known vs Unknown:

* Một phần Known: Luật giao thông, bản đồ đường, cơ chế vận hành xe (xe rẽ trái thì hướng thay đổi).
* Một phần Unknown: Hành vi của xe khác, tình huống giao thông hiếm (ví dụ tai nạn xảy ra phía trước), điều kiện thời tiết thay đổi,xe cần học từ trải nghiệm hoặc mô phỏng để thích nghi.