

1. Introduction AI

Khái niệm AI: có 4 quan điểm về AI tương ứng với 4 phương thức tiếp cận

- Acting Humanly: The Turing Test (trò chơi bắt chước người)
- Thinking Humanly: The Cognitive Modeling (xử lý ngôn ngữ, suy nghĩ, học, lập luận)
- Thinking Rationally: The "Laws of Thought" (không phải các hành vi thông minh đều có thể biểu diễn bằng logic!)
- Acting Rationally: The Rational Agent

Rational behavior = "Doing the **right thing**" (not "Doing the thing right") **Right thing** là những gì mong đợi để đạt được kết quả tốt nhất. Để đạt được **rational agent** cần nhớ:

Thông minh không nhất thiết là con người hay giống người. Hành vi thông minh không nhất thiết phải thực hiện thông qua suy nghĩ, lý luận (ví dụ: phản xạ-reflexes)

Trí tuệ nhân tạo là khoa học nghiên cứu các hành vi thông minh nhằm giải quyết các vấn đề được đặt ra đối với các chương trình máy tính!

Solving problems by searching: có 2 dạng

- Uninformed: tìm lần lượt hết trong không gian tìm kiếm mà không quan tâm tiếp theo là cái gì cho đến khi tìm ra.
- Informed: phải đoán xem tiếp theo (phía trước) là cái gì rồi dùng thông tin để quyết định đi đâu tiếp theo.

Một cái là tìm ra đáp án sớm nhất thỏa yêu cầu, mục tiêu cái còn lại là tìm tới chừng nào được đáp án tốt nhất mới chịu.

Agents and Environments:

Agent: Là một hệ thống thông minh, khiến computers có thể làm được những việc của con người hay thậm chí làm tốt hơn (fast, not mistake, huge data,...)

An agent *perceives* its environment through **sensors** and *acts* upon that environment through **actuators**.

Examples of agents:

- Human agent:
 - Sensors: eyes, ears and other organs.
 - Actuators: hands, legs, vocal tract, etc
- Robotic agent:
 - Sensors: cameras, infrared range finders, etc.
 - Actuators: levels, motors, etc
- Software agent:
 - Sensors: keystrokes, file contents, network packets, etc
 - Actuators: displaying on screen, writing files, sending network packets, etc

The agent's behavior:

- Percept: the agent's perceptual inputs at any given instant
- Percept sequence: the complete history of everything the agent has ever perceived
- An agent's behavior is described by the agent function that maps any given percept sequence to an action: $f: P \rightarrow A$
- Example: agent program
`agent = architecture(mathematical) + program(practical)`

The Concept of Rationality

Rational agents: do the **right thing** (actions that cause the agent to be most successful)

Performance measure: evaluates any given sequence of **environment states**

Rationality

Tại một thời điểm một điều được xem là hợp lí hay không phụ thuộc vào 4 yếu tố:

- Đo lường hiệu năng (Performance measure): Xác định tiêu chí của 'success'.
- Kiến thức trước đó (Prior knowledge): Là những gì mà *agent* biết về *environment*.
- Trình tự nhận thức (Percept sequence): Nhận thức của *agent* cho đến nay.
- Actions: Những việc mà *agent* có thể làm.

Rational Agent

For each possible percept sequence, a rational agent should select an action that is expected to maximize its **performance measure**, given the evidence provided by the **percept sequence** and whatever **built-in knowledge** the agent has.

For example, in an exam: Maximize marks based on the questions on the paper and your knowledge

1. Omniscience vs Rationality:
 - Rationality is not perfection
2. Information gathering (thu thập thông tin):
 - Agent không được tham gia vào các hoạt động thiếu thông minh do không có hiệu quả (inadvertency).
3. Learning:
 - A rational agent must learn as much as possible from what it perceives.
4. Autonomy:
 - Rational agent nên tự chủ, tìm hiểu những gì có thể bù lại (compensate) cho những kiến thức trước đó còn thiếu hay sai sót chứ không chỉ dựa vào những gì đã được design từ trước (ex: clock)

The Nature of Environments

The task environment:

Vai trò của môi trường là "problems" có "solutions" là các rational agent.

The task environment includes: PEAS

- Performance measure
- Environment
- Agent's Actuators
- Agent's Sensors

Khi thiết kế agent, bước đầu tiên luôn luôn phải xác định task environment (PEAS) đầy đủ nhất có thể.

Software agent: là những agent làm việc trong những môi trường nhân tạo rất phức tạp tuy nhiên những môi trường này không có trong thế giới thực (internet, video games,...). Tất cả những parts của agent này là software.

Properties of Task environment

Có 7 loại thuộc tính, những thước đo này giúp xác định chính xác thiết kế của một agent & áp dụng đúng kĩ thuật để triển khai nó.

1. Fully observable vs. Partially observable:

- Fully observable: Sensors của agent cho thể tiếp cận đầy đủ trạng thái của môi trường.
- Partially observable: Sensors cung cấp những thông tin không rõ ràng hay không chính xác. Trạng thái của môi trường thường bị thiếu sót dữ liệu.
- Unobservable: Agent không có sensors luôn.

2. Single agent vs. Multiagent:

- Single agent: duy nhất một agent thao tác trong môi trường (ex: solving crossword)
- Multiagent: có 2 loại Competitive (đối kháng - ex: playing chess) & Cooperative (hợp tác - ex: driving on road)

3. Deterministic vs. Stochastic:

- Deterministic: có thể xác định được trước trạng thái tiếp theo của môi trường thông qua trạng thái hiện tại
- Stochastic: có tính phức tạp không thể dự đoán trước được, đây là trường thực tế hơn trong real world. (ex: driving on road)

4. Episodic vs. Sequential

- Episodic: experience của agent được chia thành các giai đoạn, trong mỗi giai đoạn agent nhận được một percept & thực hiện một action duy nhất. Quality của action chỉ phụ thuộc vào chính giai đoạn đó & không cần suy nghĩ trước. (ex: phát hiện các bộ phận bị lỗi trên dây chuyền)
- Sequential: quyết định hành động hiện tại sẽ còn ảnh hưởng đến những quyết định sau này. (ex: chess)

5. Static vs. Dynamic

- Static: Môi trường không thay đổi khi agent đang tính toán suy xét
- Dynamic: Agent liên tục xem xét cần làm gì. Nếu vẫn chưa quyết định, coi như quyết định không làm gì cả.

6. Discrete vs. Continuous

Biểu diễn trạng thái của môi trường về mặt thời gian (liên tục hay riêng rạc) ảnh hưởng đến action và percept của agent.

Example: các trạng thái chơi cờ thì riêng biệt nhau (Discrete) trong khi xe chuyển động với tốc độ & trong khoảng vị trí liên tục tương ứng với giá trị thời gian. (Continuous)

7. Known vs. Unknown

- Known: luôn biết trước kết quả cho tất cả trạng thái môi trường dù có là Stochastic.
- Unknown: agent phải tìm hiểu cách thức hoạt động để đưa ra quyết định đúng đắn

Summary:

The *simplest* environment: Fully observable, deterministic, episodic, static, discrete and single-agent.

Most *real* situations: Partially observable, stochastic, sequential, dynamic, continuous and multi-agent.

The structure of agents

Agent Programs

agent = architecture + program

- The agent architecture: physical sensors and actuators
- The agent programs:
 - Input for Agent Program: tiếp nhận ở thời điểm hiện tại
 - Input for Agent Function: tiếp nhận theo chuỗi, agent phải nhớ tất cả chúng
- Công thức trong notebook.

Types of agent programs

1. Simple Reflex Agents:

- Loại agent đơn giản nhất, có trí thông minh hữu hạn
- Đưa ra actions dựa trên cái percept hiện tại, không quan tâm những percept trước đó.
- Quan hệ giữa percept & action đưa ra dựa trên **condition-action rules** (*IF current percept THEN action*)
- Hạn chế:
 - Không phải lúc nào kiến thức cũng dự đoán trước được
 - Chỉ hoạt động được trong môi trường observable.

TỚI SLIDE 57

- Model-based Reflex Agents
- Goal-based Agents
- Utility-based Agents

Learning Agents