

# AI

## ĐẠI LÝ THÔNG MINH

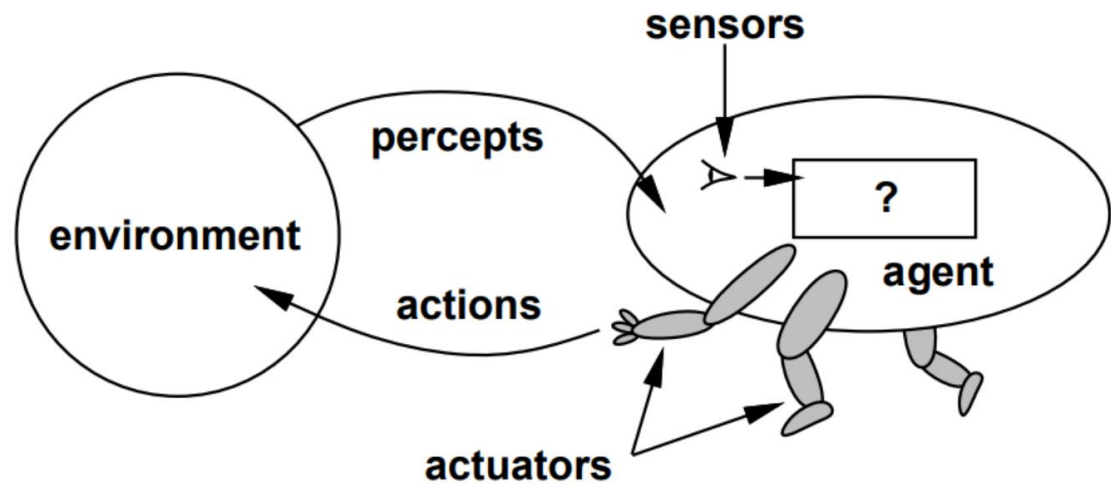
Nguyễn Ngọc Thảo - Nguyễn Hải Minh  
{nnthao, nhminh}@it.hcmus.edu.vn

# Đề cương

- Tác nhân và môi trường
- Hành vi tốt: Khái niệm về tính hợp lý
- Bản chất của môi trường
- Cấu trúc các tác nhân (Tự học)

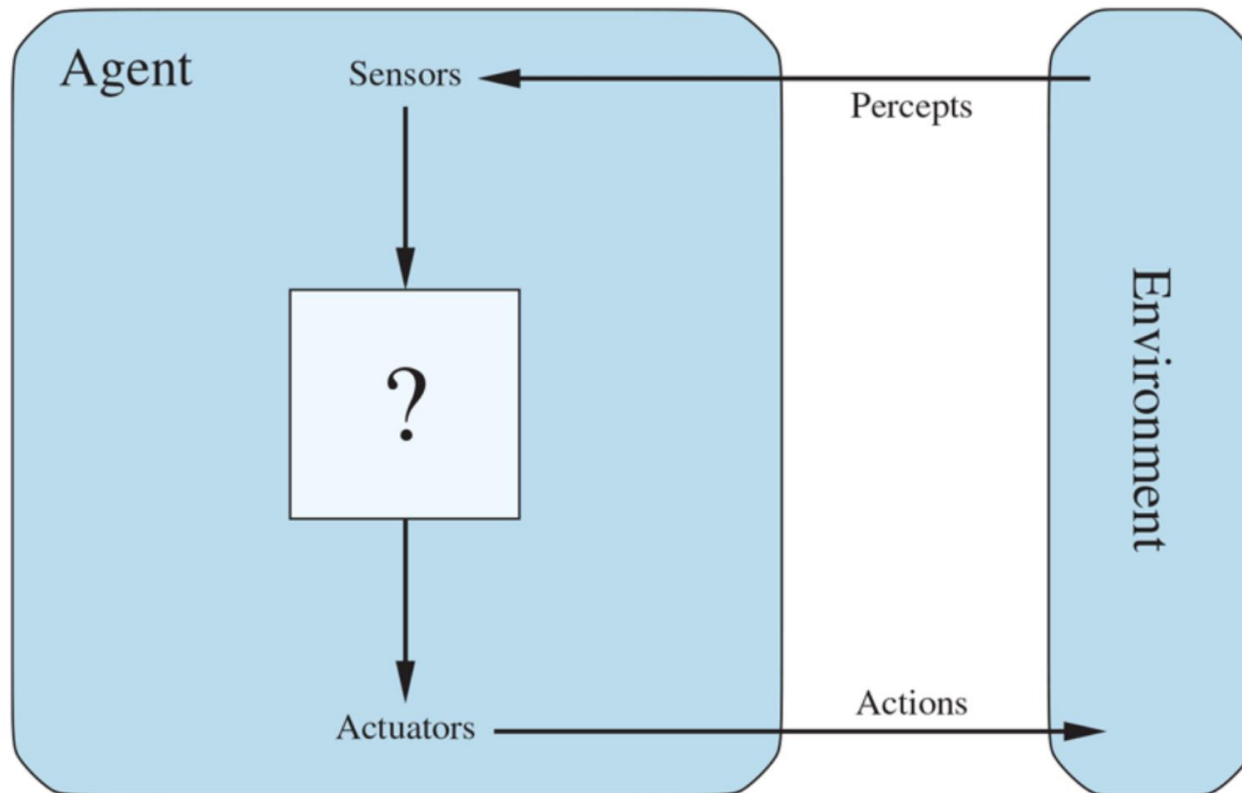
# Đại lý và môi trường

- Tác nhân: cảm biến và cơ cấu chấp hành
- Chức năng đại lý
- Chương trình đại lý



# Đại lý là gì?

- Một tác nhân **nhận biết** môi trường của nó thông qua **các cảm biến** và **hành động** theo môi trường đó thông qua **các cơ cấu chấp hành**.



Agent tương tác với môi trường thông qua các cảm biến và cơ cấu chấp hành.

# Ví dụ về đại lý



Tác nhân con người

**Cảm biến:** mắt, tai và các cơ quan khác.

**Cơ cấu chấp hành:** tay, chân, thanh quản, v.v.



Đại lý robot

**Cảm biến:** máy ảnh, công cụ tìm phạm vi hồng ngoại, v.v.

**Thiết bị truyền động:** cấp độ, động cơ, v.v.



Đại lý phần mềm

**Cảm biến:** tổ hợp phím, nội dung tập tin, mạng gói, v.v.

**Thiết bị truyền động:** màn hình, đĩa vật lý, bộ định tuyến, v.v.

# Hành vi của đại lý

- **Nhận thức** đề cập đến nội dung mà cảm biến của tác nhân cảm nhận.
- **Chuỗi nhận thức** của tác nhân là lịch sử đầy đủ của mọi thứ mà tác nhân từng nhận biết.
- **Sự lựa chọn hành động** của một tác nhân tại bất kỳ thời điểm nào có thể phụ thuộc vào **kiến thức có sẵn** của nó và vào **toàn bộ chuỗi nhận thức được quan sát cho đến nay**, chứ không phụ thuộc vào bất cứ điều gì mà nó chưa nhận thức được.

# Hành vi của đại lý

- Hành vi của một tác nhân được mô tả bằng hàm tác tử ánh xạ bất kỳ chuỗi nhận thức nhất định nào tới một hành động.

:

- Chức năng này có thể được mô tả bằng cách lập bảng.
  - Một bảng lớn được xây dựng bằng cách thử tất cả các chuỗi nhận thức có thể có và ghi lại những hành động mà tác nhân thực hiện để phản hồi.
- Bảng là đặc tính bên ngoài của tác nhân.

# Hành vi của đại lý

- Chức năng của tác nhân nhân tạo sẽ được thực hiện bên trong được thực hiện bởi một chương trình đại lý.

tác nhân = kiến trúc + chương trình

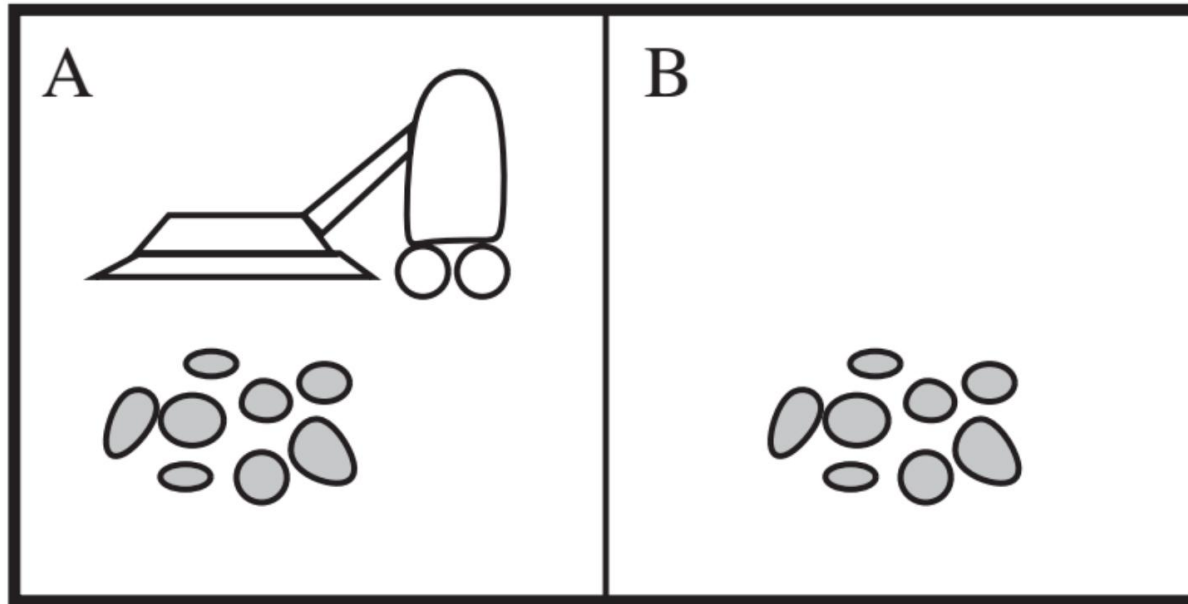
toán học

thực tế

Hàm tác nhân là một mô tả toán học trừu tượng; Chương trình tác nhân là một triển khai cụ thể, chạy trong một số hệ thống vật lý.



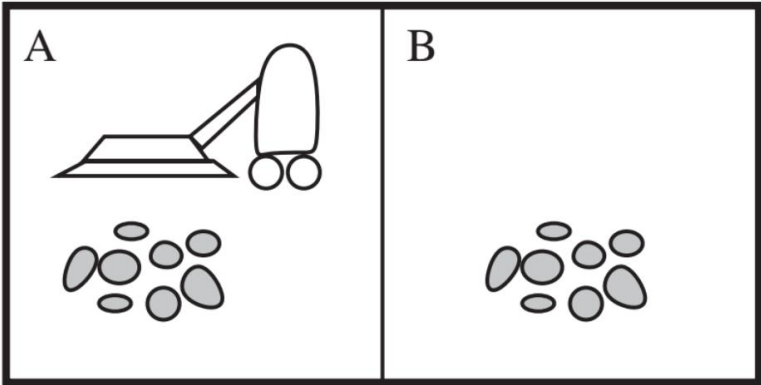
# Thế giới máy hút bụi



Thế giới máy hút bụi chỉ có hai địa điểm

- Nhận thức: vị trí và nội dung, ví dụ: **[A, Bẩn]**
- Hành động: **Trái, Phải, Hút, Không làm gì cả**

# Thế giới máy hút bụi



Lập bảng một phần của hàm tác nhân đơn giản cho thế giới máy hút bụi

Trình tự nhận thức	Hoạt động
[Sạch]	Phải
[A, Bẩn thỉu]	hút
[B, Sạch sẽ]	Bên trái
[B, Bẩn thỉu]	hút
[A, Sạch], [A, Sạch]	Phải
[A, Sạch], [A, Bẩn]	hút
...	
...	
[A, Sạch], [A, Sạch], [A, Sạch]	Phải
[A, Sạch], [A, Sạch], [A, Bẩn]	hút

hàm REFLEX-VACUUM-AGENT([vị trí,trạng thái]) trả về một hành động

nếu trạng thái = Bẩn thì trả về Hút

ngược lại nếu vị trí = A thì quay lại Phải

ngược lại nếu vị trí = B thì quay lại Trái

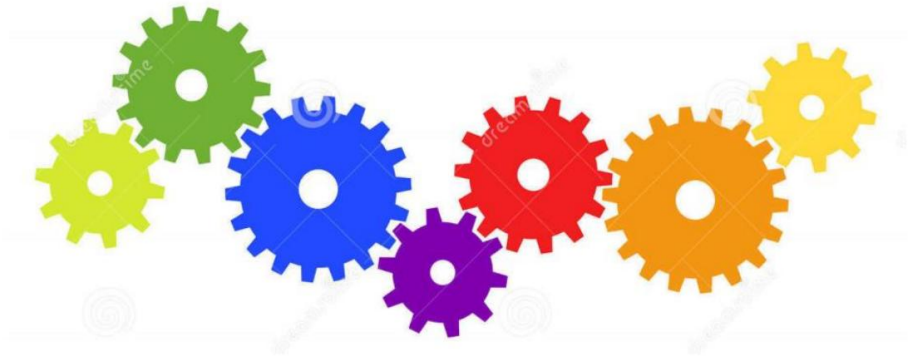
Chương trình tác nhân cho một tác nhân phản xạ đơn giản trong môi trường chân không hai trạng thái.

# Tại sao chúng ta cần đại lý?


- Khái niệm tác nhân là một công cụ để **phân tích hệ thống**, không phải là một đặc tính tuyệt đối để phân chia thế giới thành đại lý và không đại lý.
- AI thiết kế các sản phẩm có **nguồn lực tính toán đáng kể** và môi trường nhiệm vụ đòi hỏi việc ra quyết định **không cần thiết**.
- Ví dụ, một máy tính cầm tay có thể được xem như một tác nhân hiển thị "4" khi cho chuỗi nhận thức " $2 + 2 =$ " nhưng tác nhân này quá tầm thường để nghiên cứu thêm.

# Khái niệm về tính hợp lý

- Tính hợp lý
- Toàn trí, học hỏi và tự chủ



# Tác nhân hợp lý

- **Người có lý trí** là người làm **điều đúng đắn**.
- Nhưng làm điều **"đúng"** có nghĩa là gì ?
- Những hành động khiến đại lý thành công nhất
  - Ví dụ: Lập bảng: mỗi mục trong bảng cho hàm tác nhân là **điền chính xác**.
- Chúng ta cần những cách để **đo lường sự thành công** 

Hiệu suất  
đo lường

# Tác nhân hợp lý

- Chủ nghĩa hệ quả: chúng tôi đánh giá hành vi của một tác nhân bằng hậu quả.



Một đại lý, dựa trên nhận thức của nó

tạo ra hành động  
sự liên tiếp

môi trường đi đến  
chuỗi các trạng thái

- Nếu trình tự như **mong muốn thì** tác nhân đã **hoạt động tốt**.

# Đo lường hiệu suất

- **Thước đo hiệu suất** đánh giá bất kỳ trình tự nhất định nào của **trạng thái môi trường** (hãy nhớ, **không phải trạng thái tác nhân!!!**).
- Nguyên tắc chung: Thiết kế thước đo hiệu quả theo

Điều người ta **thực sự mong muốn** ở môi trường

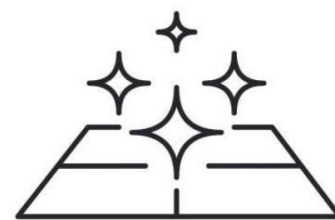
Không phải cách người ta nghĩ người đại diện nên hành xử

- Ví dụ, trong thế giới máy hút bụi



Lượng bụi bẩn được làm sạch  
trong một ca làm việc kéo dài 8 giờ

Cái mà  
một là  
tốt hơn?



Sàn nhà sạch sẽ dù đại lý cư  
xử thế nào

# Tính hợp lý

- Điều gì là hợp lý tại bất kỳ thời điểm nào phụ thuộc vào

<b>Đo lường hiệu suất</b> Xác định tiêu chí thành công	<b>Kiến thức trước đó</b> Người đại diện biết gì về môi trường
<b>Trình tự nhận thức</b> Nhận thức của đại lý cho đến nay	<b>hành động</b> Những gì đại lý có thể thực hiện

- Định nghĩa tác nhân hợp lý

Đối với mỗi chuỗi nhận thức có thể có, một tác nhân hợp lý nên chọn một hành động được kỳ vọng sẽ tối đa hóa thước đo hiệu suất của nó, dựa trên bằng chứng được cung cấp bởi trình tự nhận thức và bất kỳ kiến thức tích hợp nào mà tác nhân có.



# Đại lý máy hút bụi

- **Thước đo hiệu suất:** thưởng một điểm cho mỗi ô vuông sạch ở mỗi bước thời gian, trong suốt thời gian tồn tại là 10000 bước thời gian
- **Kiến thức sẵn có về môi trường:** Vị trí địa lý của môi trường (2 ô vuông) và tác động của các hành động
- **Các động tác có thể thực hiện:** Trái, Phải và Hút
- **Trình tự nhận thức:** Tác nhân nhận thức chính xác vị trí và liệu vị trí đó có chứa bụi bẩn hay không.
- Trong hoàn cảnh này, người đại diện là **người có lý trí**.

# Đại lý máy hút bụi

- Cùng một tác nhân có thể không hợp lý trong các tình huống khác nhau.
- Ví dụ, tác nhân dao động qua lại ngay cả khi tất cả chất bẩn đã được làm sạch; và thước đo hiệu suất bao gồm hình phạt một điểm cho mỗi hành động
- Nếu không biết địa lý của môi trường, tác nhân sẽ cần phải khám phá nó.

# Toàn tri so với tính hợp lý

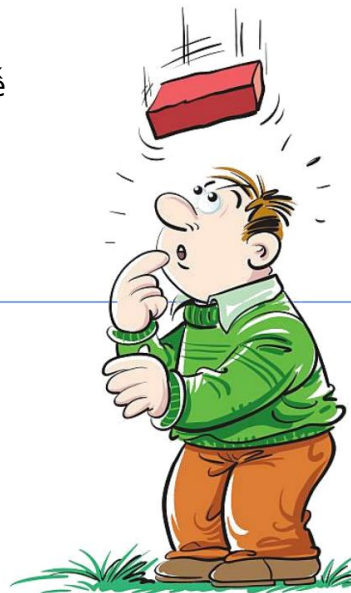
- Chúng ta cần phân biệt giữa tính hợp lý và tính toàn tri.

## toàn tri

- Biết được kết quả thực tế của hành động trước
- Không có kết quả nào khác có thể xảy ra
- Tuy nhiên, điều đó là không thể trong thế giới thực
- Ví dụ?

## Tính hợp lý

Tối đa hóa thước đo hiệu suất dựa trên trình tự nhận thức cho đến nay và kiến thức trước đó



Sự hợp lý không phải là sự hoàn hảo

# Thu thập thông tin

- Đại lý không được tham gia vào các hoạt động thiếu thông minh do vô ý.
- **Thu thập thông tin** - Thực hiện các hành động để sửa đổi nhận thức trong tương lai (ví dụ: khám phá)
- Đây là một phần quan trọng của tính hợp lý.



# Học hỏi

- Tác nhân hợp lý không chỉ thu thập thông tin mà còn học hỏi càng nhiều càng tốt từ những gì nó cảm nhận được.
  - Cấu hình ban đầu của nó có thể được sửa đổi và tăng cường khi nó đạt được kinh nghiệm.
- Có những trường hợp cực đoan trong đó môi trường hoàn toàn được biết đến một cách tiên nghiệm.



# Quyền tự trị

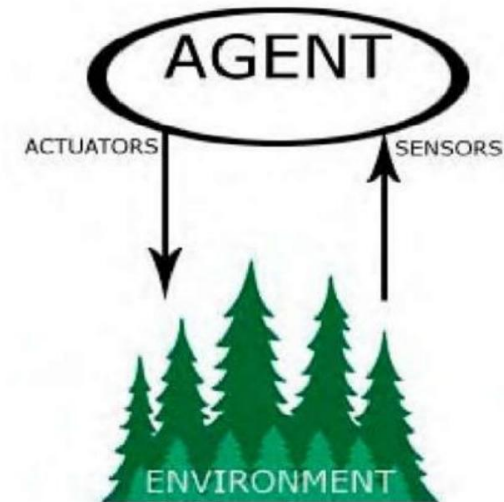
- Tác nhân hợp lý phải có **tính tự chủ** - **Học** những gì có thể để **bù đắp** cho kiến thức trước đó một phần hoặc không chính xác.
- Nếu một tác nhân chỉ dựa vào kiến thức có sẵn của người thiết kế thay vì nhận thức của chính nó thì tác nhân đó sẽ thiếu quyền tự chủ.

Đồng hồ không có đầu vào (nhận thức) và chạy thuật toán riêng (kiến thức trước). Không có học tập, không có kinh nghiệm, v.v.



# Bản chất của môi trường

- Chỉ định môi trường tác vụ
- Thuộc tính của môi trường tác vụ



# Môi trường nhiệm vụ

- Môi trường nhiệm vụ về cơ bản là những “vấn đề” mà tác nhân hợp lý là “giải pháp”.



- Chúng có nhiều hướng vị khác nhau, ảnh hưởng trực tiếp đến thiết kế phù hợp cho chương trình đại lý.



# Môi trường nhiệm vụ

- Môi trường nhiệm vụ bao gồm

đậu Hà Lan

- Đo lường hiệu suất
- Môi trường
- Thiết bị truyền động của đại lý
- Cảm biến của tác nhân

- Đây luôn phải là bước đầu tiên trong việc thiết kế một tác nhân và phải được xác định đầy đủ nhất có thể.

# Ví dụ: Lái xe taxi tự động

- Đo lường hiệu suất

- Làm thế nào chúng ta có thể đánh giá trình điều khiển tự động?
- Những yếu tố nào được xem xét?
  - đến đúng đích
  - giảm thiểu tiêu thụ nhiên liệu
  - giảm thiểu thời gian và/hoặc chi phí chuyến đi
  - giảm thiểu việc vi phạm luật giao thông
  - tối đa hóa sự an toàn và thoải mái
  - vân vân.



# Ví dụ: Lái xe taxi tự động

- **Môi trường**

- Đường đa dạng (ngõ nông thôn, ngõ đô thị, v.v.)
- Đèn giao thông, các phương tiện khác, người đi bộ, động vật đi lạc, công trình đường bộ, xe cảnh sát, vũng nước, ổ gà, v.v.

- **Bộ truyền động** (cho đầu ra)

- Kiểm soát chân ga, lái, sang số, sang số và phanh
- Màn hình tương tác với hành khách

- **Cảm biến** (đối với đầu vào)

- Camera có thể điều khiển để phát hiện các phương tiện khác, tình huống trên đường
- GPS (Hệ thống định vị toàn cầu) để biết taxi đang ở đâu
- Cần thêm nhiều thiết bị: đồng hồ tốc độ, gia tốc kế, v.v.

# Ví dụ: Lái xe taxi tự động

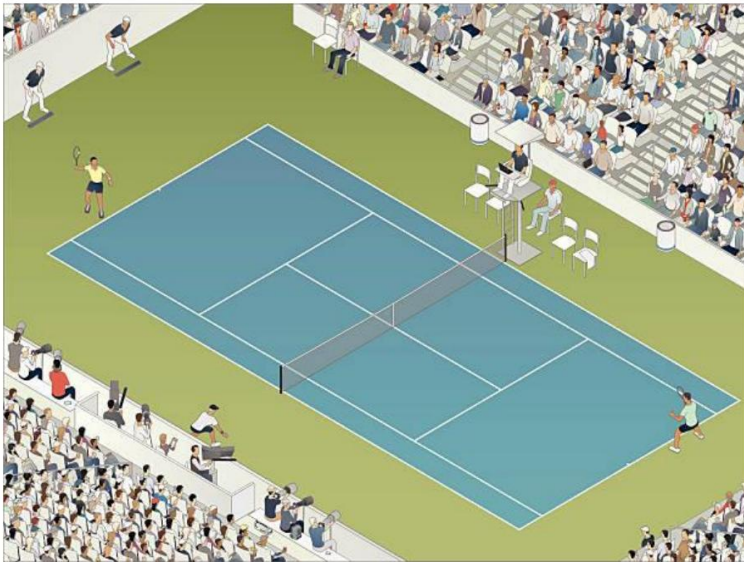
Agent Type	Performance Measure	Environment	Actuators	Sensors
Taxi driver	Safe, fast, legal, comfortable trip, maximize profits, minimize impact on other road users	Roads, other traffic, police, pedestrians, customers, weather	Steering, accelerator, brake, signal, horn, display, speech	Cameras, radar, speedometer, GPS, engine sensors, accelerometer, microphones, touchscreen

Mô tả PEAS về môi trường nhiệm vụ cho một chiếc taxi tự động.

Agent Type	Performance Measure	Environment	Actuators	Sensors
Medical diagnosis system	Healthy patient, reduced costs	Patient, hospital, staff	Display of questions, tests, diagnoses, treatments	Touchscreen/voice entry of symptoms and findings
Satellite image analysis system	Correct categorization of objects, terrain	Orbiting satellite, downlink, weather	Display of scene categorization	High-resolution digital camera
Part-picking robot	Percentage of parts in correct bins	Conveyor belt with parts; bins	Jointed arm and hand	Camera, tactile and joint angle sensors
Refinery controller	Purity, yield, safety	Refinery, raw materials, operators	Valves, pumps, heaters, stirrers, displays	Temperature, pressure, flow, chemical sensors
Interactive English tutor	Student's score on test	Set of students, testing agency	Display of exercises, feedback, speech	Keyboard entry, voice

# Câu đố 01: Mô tả PEAS

- Đối với mỗi hoạt động sau đây, hãy đưa ra mô tả PEAS của môi trường nhiệm vụ



Chơi một trận quần vợt  
trong một giải đấu



Tập tennis dựa vào tường

# Thuộc tính của môi trường tác vụ

- Các kích thước này xác định thiết kế tác nhân phù hợp và khả năng áp dụng các kỹ thuật để thực hiện tác tử.

Hoàn toàn có thể quan sát được	Có thể quan sát được một phần
Đại lý duy nhất	Đa tác nhân
xác định	ngẫu nhiên
Nhiều tập	tuần tự
Tĩnh	Năng động
rời rạc	Tiếp diễn
Đã biết	không xác định

Có thể quan sát được hoàn toàn và có thể quan sát được một phần

- **Hoàn toàn có thể quan sát được:** Cảm giác của tác nhân cho phép nó tiếp cận với trạng thái hoàn chỉnh của môi trường.
  - Agent không cần duy trì trạng thái nội bộ để theo dõi thế giới.
- **Có thể quan sát được một phần**
  - Cảm biến ồn và không chính xác
  - Đơn giản là thiếu các phần trạng thái trong dữ liệu cảm biến, ví dụ: máy hút bụi chỉ có cảm biến bụi bẩn cục bộ không thể biết liệu có bụi bẩn ở các ô khác hay không
- **Không thể quan sát được:** Tác nhân không có cảm biến nào cả



# Đại lý đơn lẻ so với đa đại lý

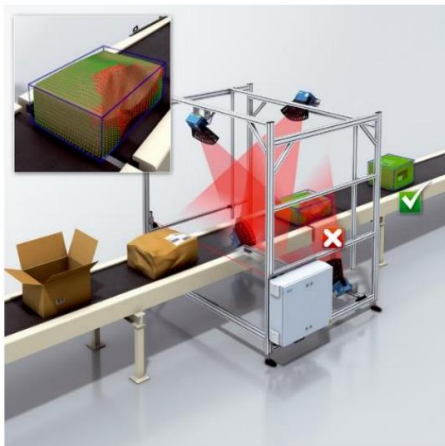
- **Single Agent:** Một Agent hoạt động độc lập trong một môi trường.
  - Ví dụ: giải ô chữ      một tác nhân, chơi cờ      hai tác nhân
- Những thực thể nào phải được xem là đại lý?
  - Liệu hành vi của B có được mô tả là tối đa hóa hiệu suất hay không  
thước đo có giá trị phụ thuộc vào hành vi của A.
- Môi trường đa tác nhân cạnh tranh và hợp tác
  - Ví dụ: chơi cờ      cạnh tranh, lái xe trên đường      hợp tác

# Xác định so với ngẫu nhiên

- **Có tính tất định:** Trạng thái tiếp theo của môi trường là hoàn toàn được xác định bởi trạng thái hiện tại và hành động được thực hiện bởi tác nhân.
  - Ví dụ: thế giới chân không      tất định, lái xe trên đường      ngẫu nhiên
- Hầu hết các tình huống thực tế đều phức tạp đến mức chúng phải được coi là **ngẫu nhiên**.

# Tập và tuân thủ

- Theo từng giai đoạn: Trải nghiệm của tác nhân được chia thành các phần tử các tập, trong mỗi tập đó tác nhân nhận được một nhận thức và sau đó thực hiện một hành động duy nhất.
  - Chất lượng hành động chỉ phụ thuộc vào chính tập phim
  - Không cần suy nghĩ trước
- Tuân thủ: Một quyết định hiện tại có thể ảnh hưởng đến các quyết định trong tương lai.



Phát hiện các bộ phận bị lỗi  
trên dây chuyền lắp ráp

so với

Lắp ráp một ngôi nhà Lego



# Tĩnh và động

- **Tĩnh:** Môi trường không thay đổi trong khi tác nhân đang cân nhắc.
  - Ví dụ: trò chơi ô chữ tĩnh, lái taxi động
- **Năng động:** Tác nhân liên tục được hỏi muốn làm gì LÀM
  - Nếu chưa quyết định thì coi như quyết định không làm gì cả.
- **Bán động:** Bản thân môi trường không thay đổi theo thời gian nhưng điểm hiệu suất của tác nhân thì thay đổi
  - Ví dụ: chơi cờ với đồng hồ

# Thuộc tính của môi trường tác vụ

- Rời rạc và liên tục

- Sự phân biệt áp dụng cho tình trạng môi trường, cách thức thời gian được xử lý và nhận thức cũng như hành động của tác nhân
- Ví dụ, cờ vua có một số hữu hạn các trạng thái, nhận thức và hành động riêng biệt; trong khi tốc độ và vị trí của xe quét qua một loạt các giá trị liên tục một cách trơn tru theo thời gian.

- Đã biết và chưa biết

- Môi trường đã biết: các kết quả (hoặc xác suất kết quả nếu môi trường là ngẫu nhiên) cho tất cả các hành động được đưa ra.
- Môi trường không xác định: tác nhân cần học cách hoạt động để đưa ra quyết định đúng đắn.

# Môi trường và đặc điểm của chúng

Task Environment	Observable	Agents	Deterministic	Episodic	Static	Discrete
Crossword puzzle	Fully	Single	Deterministic	Sequential	Static	Discrete
Chess with a clock	Fully	Multi	Deterministic	Sequential	Semi	Discrete
Poker	Partially	Multi	Stochastic	Sequential	Static	Discrete
Backgammon	Fully	Multi	Stochastic	Sequential	Static	Discrete
Taxi driving	Partially	Multi	Stochastic	Sequential	Dynamic	Continuous
Medical diagnosis	Partially	Single	Stochastic	Sequential	Dynamic	Continuous
Image analysis	Fully	Single	Deterministic	Episodic	Semi	Continuous
Part-picking robot	Partially	Single	Stochastic	Episodic	Dynamic	Continuous
Refinery controller	Partially	Single	Stochastic	Sequential	Dynamic	Continuous
English tutor	Partially	Multi	Stochastic	Sequential	Dynamic	Discrete

Ví dụ về môi trường nhiệm vụ và đặc điểm của chúng.

## Câu đố 02: Môi trường làm việc

- Đối với mỗi hoạt động sau đây, hãy mô tả môi trường nhiệm vụ của nó theo các thuộc tính được liệt kê.
  - Chơi một trận quần vợt trong một giải đấu
  - Tập tennis dựa vào tường

# Cơ cấu đại lý

Tác nhân  
phản xạ đơn giản

Tác nhân  
phản xạ  
dựa trên mô hình

Tác nhân dựa  
trên mục tiêu

Đại lý  
dựa trên  
tiện ích

Hello!





# Kiến trúc đại lý

tác nhân = kiến trúc + chương trình

- **Kiến trúc:** một số loại thiết bị điện toán có cảm biến vật lý và bộ truyền động mà chương trình này sẽ chạy trên đó.
  - Máy tính thông thường, ô tô robot có nhiều máy tính, camera, và các cảm biến khác, v.v.
- **Chương trình** phải phù hợp với kiến trúc.
  - Chương trình: Hành động đi bộ      Kiến trúc: chân

# Chương trình đại lý

- Một chương trình đại lý lấy nhận thức hiện tại từ cảm biến và trả về một hành động cho bộ truyền động.
- Chương trình đại lý chỉ lấy nhận thức hiện tại, bởi vì không có gì hơn có sẵn từ môi trường.
- Trong khi đó, hàm tác nhân nhận toàn bộ nhận thức trình tự mà tác tử phải nhớ.

## Một chương trình đại lý tầm thường

- Theo dõi trình tự nhận thức và lập chỉ mục vào bảng hành động để quyết định phải làm gì.

hàm TABLE-DRIVEN-AGENT(percept) trả về một hành động liên tục:

percepts, một chuỗi, bảng trống ban đầu , một bảng

hành động, được lập chỉ mục bởi các chuỗi nhận thức, ban

đầu được chỉ định đây

đủ nối percept vào cuối của percepts

action     LOOKUP(percepts, table )

hành động trả về

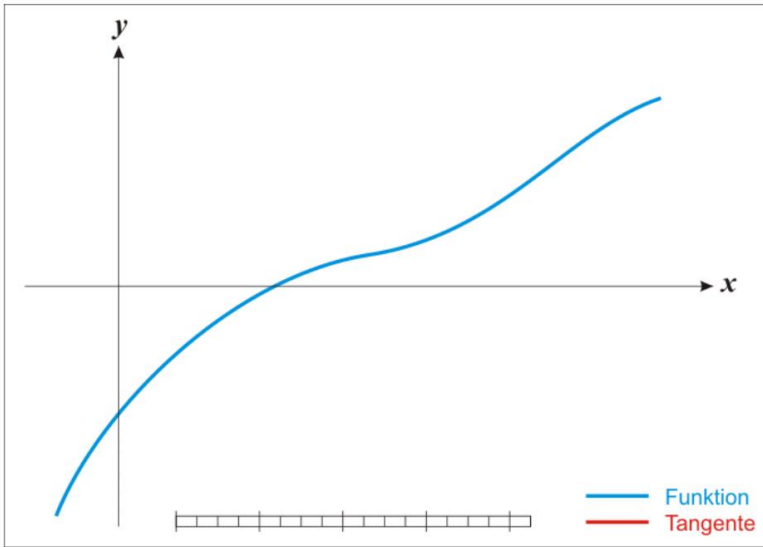
Chương trình TABLE-DRIVEN-AGENT được gọi cho mỗi nhận thức mới và trả về một hành động mỗi lần. Nó giữ lại chuỗi nhận thức hoàn chỉnh trong bộ nhớ.

# Một chương trình đại lý tầm thường

- = tập hợp các nhận thức có thể có
- = thời gian tồn tại của đại lý
  - Tức là tổng số nhận thức nó nhận được
- Kích thước của bảng tra cứu là  $\sigma$   $=_1 | |$
- Ví dụ, hãy xem xét việc chơi cờ
  - $= 10$ ,  $= 150$  Một bảng có ít nhất 10150 mục
- Mặc dù có kích thước khổng lồ nhưng bảng tra cứu vẫn thực hiện được những gì chúng ta mong muốn

# Thách thức chính của AI

- Viết các chương trình tạo ra **hành vi hợp lý** từ một lượng nhỏ mã thay vì một lượng lớn các mục trong bảng
  - Ví dụ: tính căn bậc hai - một chương trình năm dòng của Phương pháp Newton so với một bảng tra cứu khổng lồ



Nguồn: [Wikipedia](#)

Squares - Cubes - Square Root Chart							
number n	square n <sup>2</sup>	cube n <sup>3</sup>	square root √n	number n	square n <sup>2</sup>	cube n <sup>3</sup>	square root √n
1	1	1	1.0000	41	1681	68921	6.4031
2	4	8	1.4142	42	1764	74088	6.4807
3	9	27	1.7321	43	1849	79507	6.5574
4	16	64	2.0000	44	1936	85184	6.6332
5	25	125	2.2361	45	2025	91125	6.7082
6	36	216	2.4495	46	2116	97336	6.7823
7	49	343	2.6458	47	2209	103823	6.8557
8	64	512	2.8284	48	2304	110592	6.9282
9	81	729	3.0000	49	2401	117649	7.0000
10	100	1000	3.1623	50	2500	125000	7.0711
11	121	1331	3.3166	51	2601	132651	7.1414
12	144	1728	3.4641	52	2704	140608	7.2111
13	169	2197	3.6056	53	2809	148877	7.2801
14	196	2744	3.7417	54	2916	157464	7.3485
15	225	3375	3.8730	55	3025	166375	7.4162
16	256	4096	4.0000	56	3136	175616	7.4833
17	289	4913	4.1231	57	3249	185193	7.5498
18	324	5832	4.2426	58	3364	195112	7.6158
19	361	6859	4.3589	59	3481	205379	7.6811
20	400	8000	4.4721	60	3600	216000	7.7460
21	441	9261	4.5826	61	3721	226981	7.8102
22	484	10648	4.6904	62	3844	238328	7.8740
23	529	12167	4.7958	63	3969	250047	7.9373
24	576	13824	4.8990	64	4096	262144	8.0000
25	625	15625	5.0000	65	4225	274625	8.0623
26	676	17576	5.0990	66	4356	287496	8.1240
27	729	19683	5.1962	67	4489	300763	8.1854
28	784	21952	5.2915	68	4624	314432	8.2462
29	841	24389	5.3852	69	4761	328509	8.3066
30	900	27000	5.4772	70	4900	343000	8.3666
31	961	29791	5.5678	71	5041	357911	8.4261
32	1024	32768	5.6569	72	5184	373248	8.4853
33	1089	35937	5.7446	73	5329	389017	8.5440
34	1156	39304	5.8310	74	5476	405224	8.6023
35	1225	42875	5.9161	75	5625	421875	8.6603
36	1296	46656	6.0000	76	5776	438976	8.7178
37	1369	50653	6.0828	77	5929	456533	8.7750
38	1444	54872	6.1644	78	6084	474552	8.8318
39	1521	59319	6.2450	79	6241	493039	8.8882
40	1600	64000	6.3246	80	6400	512000	8.9443

# Tác nhân phản xạ đơn giản

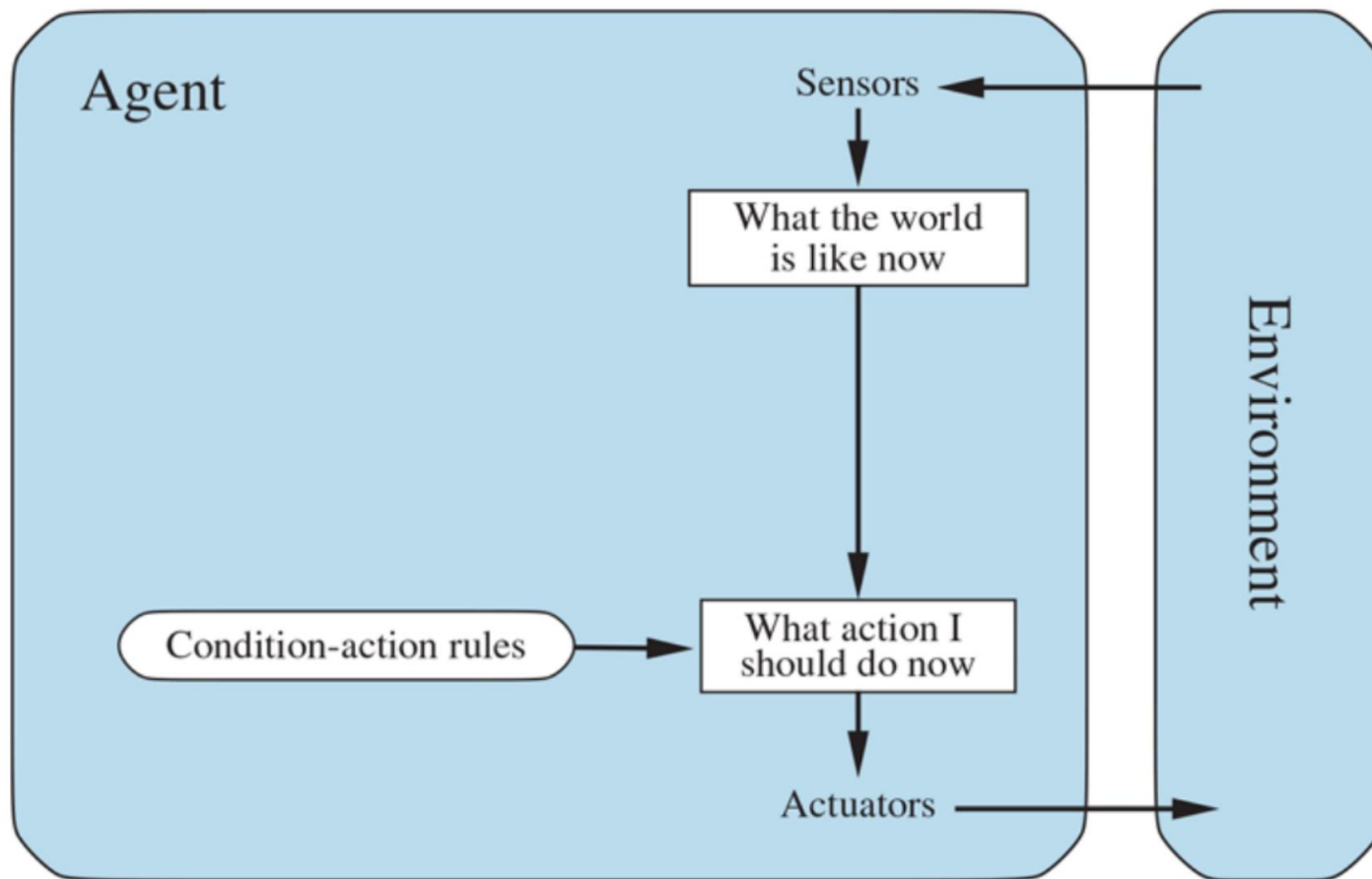
- Loại tác nhân đơn giản nhất, trí thông minh hạn chế •

Chọn hành động dựa trên **nhận thức hiện tại**, bỏ qua phần còn lại của lịch sử nhận thức

- Mỗi liên hệ từ nhận thức đến hành động được thể hiện bằng **quy tắc điều kiện-hành động**.

## NẾU nhận thức hiện tại THÌ hành động

- Ví dụ: NẾU xe phía trước đang phanh THÌ bắt đầu phanh.
- Hạn chế
  - Kiến thức đôi khi không thể nói rõ ràng      khả năng ứng dụng thấp • **Chỉ làm việc nếu môi trường có thể quan sát được đầy đủ**



hàm `SIMPLE-REFLEX-AGENT(percept)` trả về một hành động liên tục: quy tắc,  
 một tập hợp các quy tắc điều kiện-hành động trạng thái  
`INTERPRET-INPUT(percept)` quy tắc     `RULE-`  
`MATCH(trạng thái, quy tắc)` hành động  
 quy tắc.`ACTION`  
 hành động trả lại

# Tác nhân phản xạ dựa trên mô hình

- Tác nhân phải theo dõi một **phần** trạng thái nội bộ **môi trường có thể quan sát được**.
- Trạng thái này phụ thuộc vào lịch sử nhận thức và phản ánh một số khía cạnh không được quan sát.
- Chương trình đại lý cập nhật thông tin trạng thái nội bộ dưới dạng thời gian trôi qua bằng cách mã hóa hai loại kiến thức

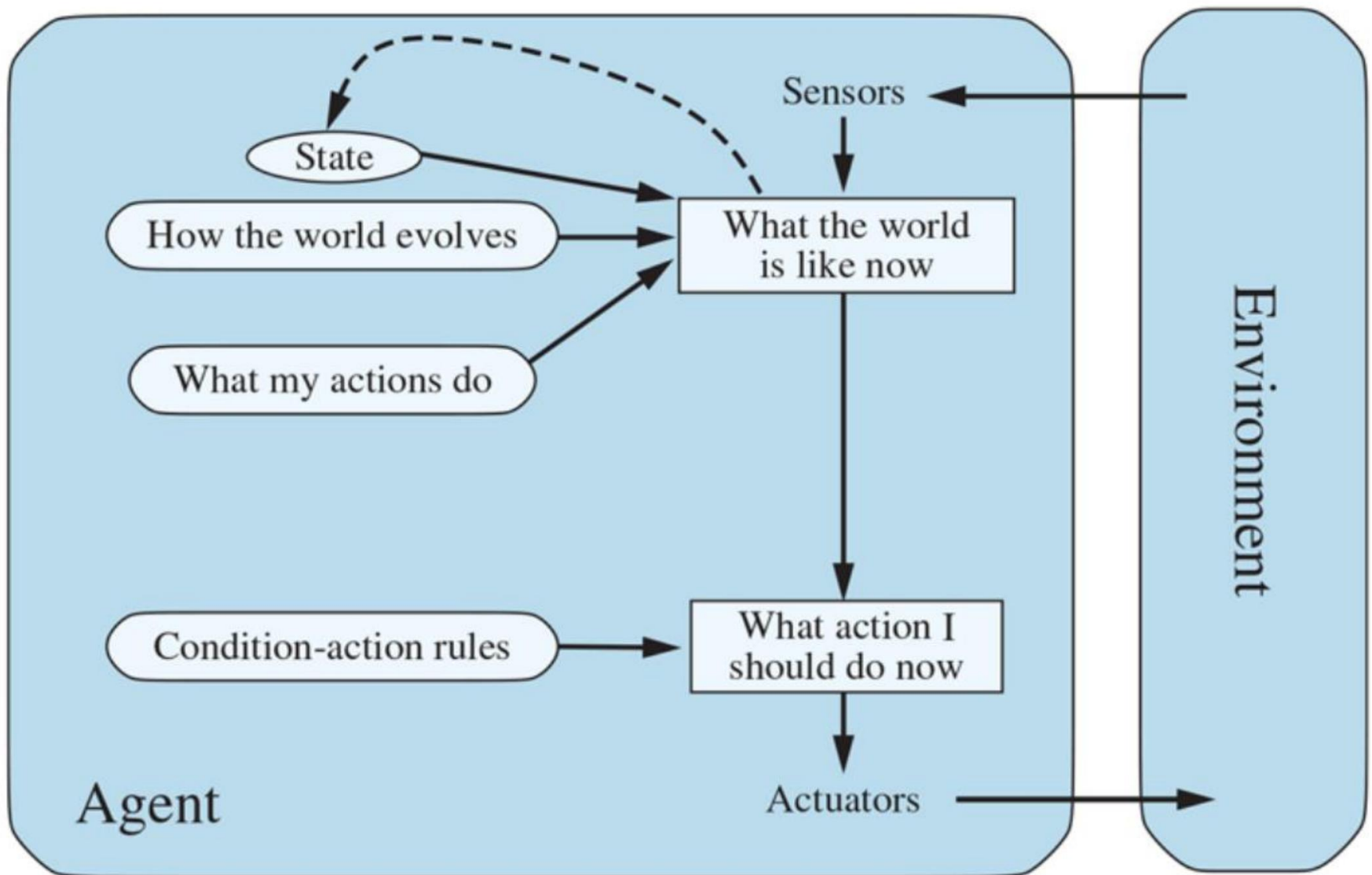
Mô hình chuyển tiếp

Thế giới hoạt động như thế nào

Mô hình cảm biến

Trạng thái của thế giới được phản ánh  
như thế nào trong nhận thức của tác nhân





hàm MODEL-BASED-REFLEX-AGENT(percept) trả về một hành động

dại dẳng: trạng thái, quan niệm hiện tại của tác nhân về trạng thái từ

transition\_model, mô tả trạng thái tiếp theo phụ thuộc vào trạng thái và hành động hiện tại  
như thế nào

cảm biến\_model, mô tả về trạng thái thế giới hiện tại

được phản ánh trong nhận thức của người đại diện

quy tắc, một tập hợp các quy tắc điều kiện-hành động

hành động, hành động gần đây nhất, ban đầu không có

trạng thái     UPDATE-STATE(trạng thái, hành động, nhận thức, transition\_model, cảm  
biến\_model)

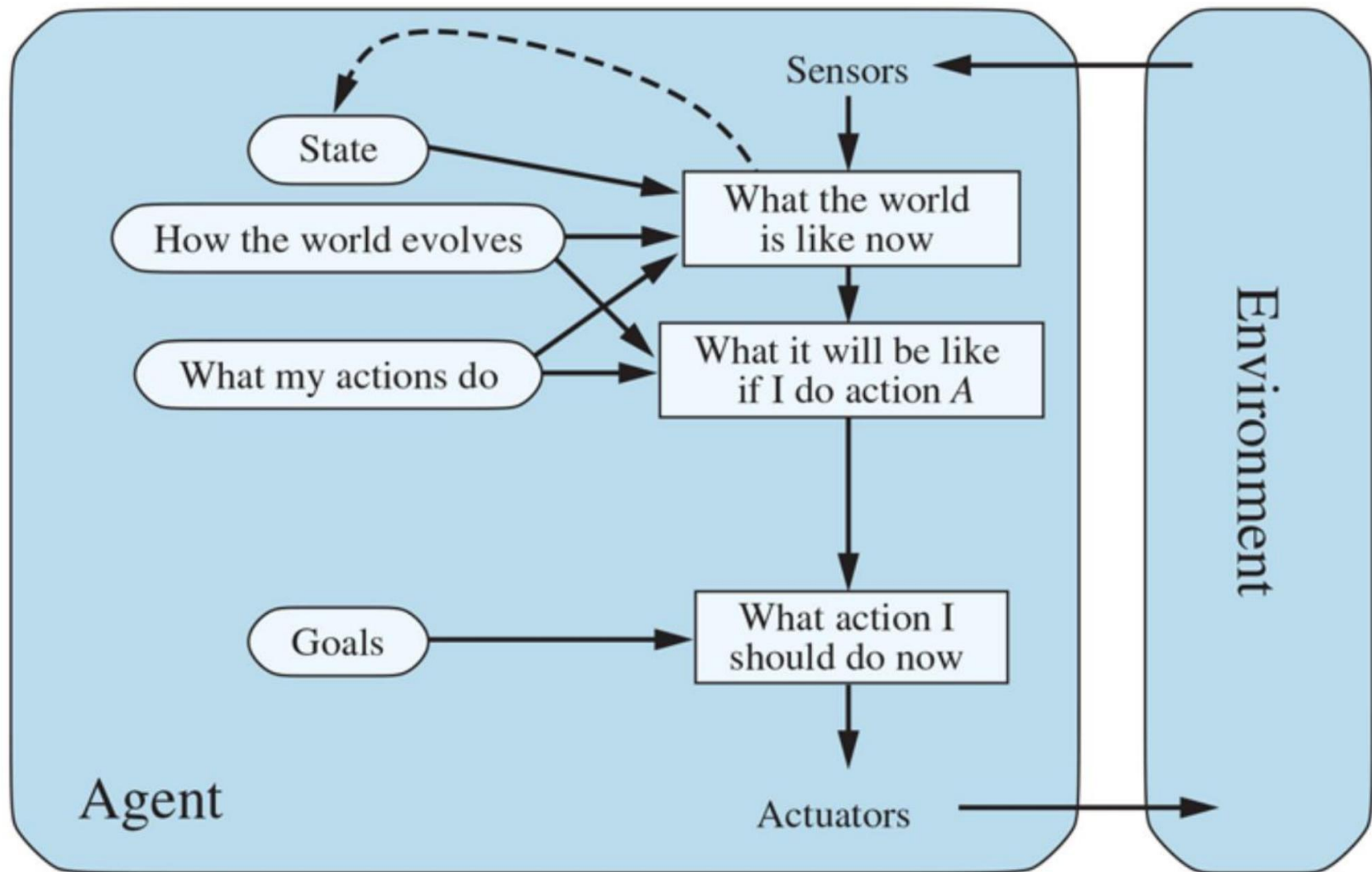
quy tắc     RULE-MATCH(trạng thái, quy tắc)

hành động     quy tắc.ACTION

hành động trả lại

# Tác nhân dựa trên mục tiêu

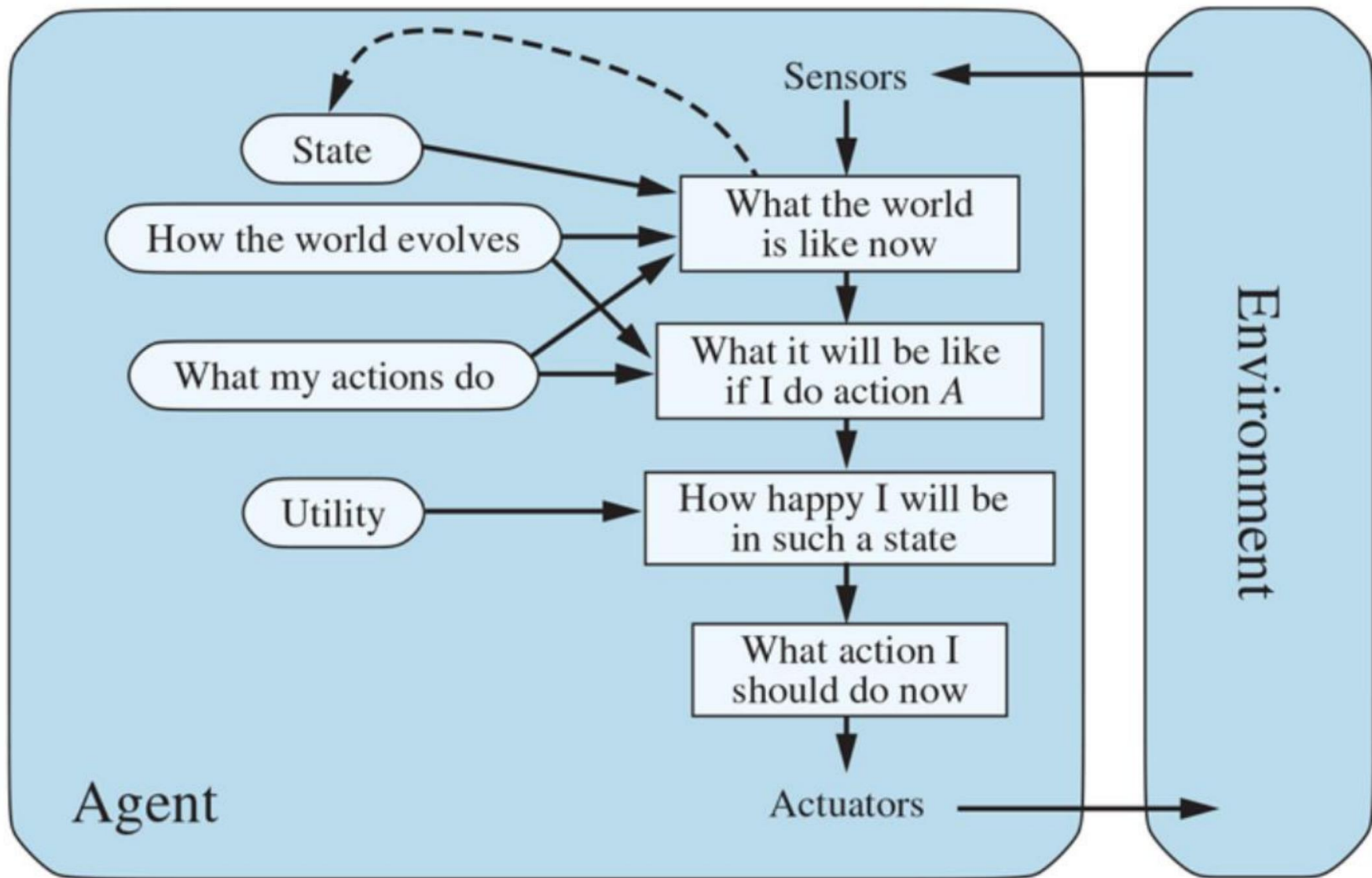
- Hiện trạng môi trường luôn chưa đủ
- Tác nhân cần thêm một số loại thông tin mục tiêu mô tả các tình huống mong muốn.
  - Ví dụ: tại ngã ba đường, xe taxi có thể rẽ trái, rẽ phải hoặc đi thẳng vào, tùy thuộc vào nơi taxi đang cố gắng đến.
- Kém hiệu quả hơn nhưng linh hoạt hơn
  - Tri thức hỗ trợ cho các quyết định được thể hiện một cách rõ ràng và có thể sửa đổi được.



Một tác nhân dựa trên mục tiêu theo dõi trạng thái thế giới cũng như một tập hợp các mục tiêu mà nó đang cố gắng đạt được và chọn một hành động (cuối cùng) sẽ dẫn đến việc đạt được các mục tiêu của nó.

# Đại lý dựa trên tiện ích

- Mục tiêu không đủ để tạo ra hành vi chất lượng cao trong hầu hết các môi trường.
  - Nhiều chuỗi hành động có thể đạt được mục tiêu, một số thì tốt hơn và một số thì tệ hơn, ví dụ về nhà bằng taxi hay xe Grab?
- Chức năng tiện ích của một tác nhân về cơ bản là sự nội hóa của thước đo hiệu suất.
  - Mục tiêu thành công, tiện ích mức độ thành công (thành công như thế nào)
  - Nếu trạng thái A được ưa chuộng hơn các trạng thái khác thì A có hữu dụng cao hơn.



Một nền tảng dựa trên tiện ích sử dụng một mô hình thế giới, cùng với chức năng tiện ích để đo lường sở thích của nó giữa các quốc gia trên thế giới. Sau đó, nó chọn hành động dẫn đến tiện ích kỳ vọng tốt nhất, trong đó tiện ích kỳ vọng được tính bằng cách lấy trung bình trên tất cả các trạng thái kết quả có thể xảy ra, được tính theo xác suất của kết quả.

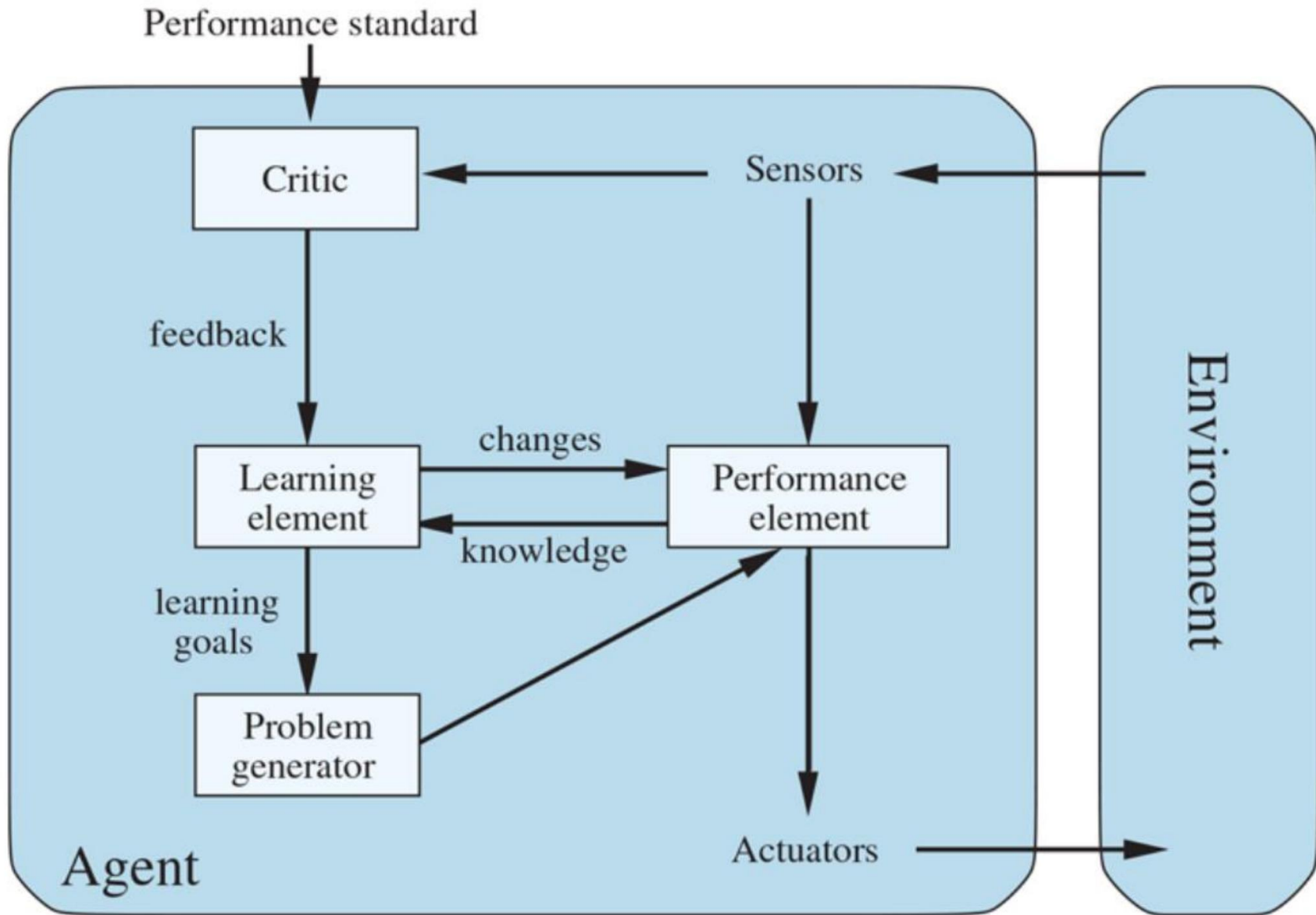
# Đại lý dựa trên tiện ích: Ưu điểm

- Khi có những mục tiêu xung đột nhau
  - Chỉ có thể đạt được một số mục tiêu như tốc độ và an toàn
  - Hàm tiện ích xác định sự cân bằng thích hợp.
- Khi có một số mục tiêu mà nhân viên có thể hướng tới
  - Không có điều nào có thể đạt được một cách chắc chắn
  - Tiện ích đánh giá khả năng thành công so với tầm quan trọng của các mục tiêu.
- Tác nhân dựa trên lợi ích hợp lý chọn hành động tối đa hóa **lợi ích mong đợi** của kết quả hành động

# Tác nhân học tập

- Sau khi lập trình xong, tác nhân có thể hoạt động được ngay không?
  - Không, vẫn cần dạy
- Khi đại lý đã xong, chúng ta có thể làm gì tiếp theo?
  - Dạy nó bằng cách cho nó một tập hợp các ví dụ
  - Kiểm tra nó bằng cách sử dụng một tập hợp ví dụ khác
- Khi đó chúng ta nói tác nhân **học**      **tác nhân học**





Một đại lý học tập nói chung. Hộp “yếu tố hiệu suất” thể hiện những gì trước đây chúng ta coi là toàn bộ chương trình đại lý. Bây giờ, hộp “thành phần học tập” sẽ sửa đổi chương trình đó để cải thiện hiệu suất của nó.

# Các thành phần khái niệm trong tác nhân học tập

1. **Yếu tố học tập**      Cải thiện
2. **Yếu tố hiệu suất**      Chọn hành động bên ngoài
3. **Phê bình**      Cho phần tử Học tập biết nhân viên đang hoạt động tốt như thế nào so với tiêu chuẩn hiệu suất cố định.  
(Phản hồi từ người dùng hoặc ví dụ, tốt hay không?)
4. **Trình tạo vấn đề**      Đề xuất hành động dẫn đến trải nghiệm mới và mang tính thông tin

# Tác nhân học tập học như thế nào?

- Là quá trình **sửa đổi từng thành phần** của tác nhân
- Để các thành phần đạt được **thỏa thuận chặt chẽ hơn với thông tin phản hồi có sẵn**
- Qua đó, nâng cao hiệu quả hoạt động tổng thể của đại lý.

# Tác nhân học tập: Một ví dụ

- **Yếu tố hiệu suất**

- Bất kể việc thu thập kiến thức và thủ tục nào mà xe taxi có để lựa chọn hành động lái xe của nó (có thể được sửa đổi thêm)

- **Phê bình**

- Quan sát thế giới và truyền thông tin đến thành phần học tập • Ví dụ: rẽ trái nhanh qua ba làn đường      sử dụng ngôn ngữ gây sốc do người lái xe khác quan sát      hành động

xấu • **Yếu tố học tập**

- Xây dựng các quy tắc mới từ kinh nghiệm của người phê bình • Ví dụ: một quy tắc mới cho hành động xấu ở trên

- **Trình tạo sự cố**

- Xác định những hành vi nhất định cần cải thiện và đề xuất thử nghiệm • Ví dụ: thử phanh trên các mặt đường khác nhau trong các điều kiện khác nhau

# Câu đố 03: Tác nhân học tập

- Cho ví dụ về việc học tác nhân hợp lý sau 4 các yếu tố khái niệm.

