

Agents, PEAS and environments

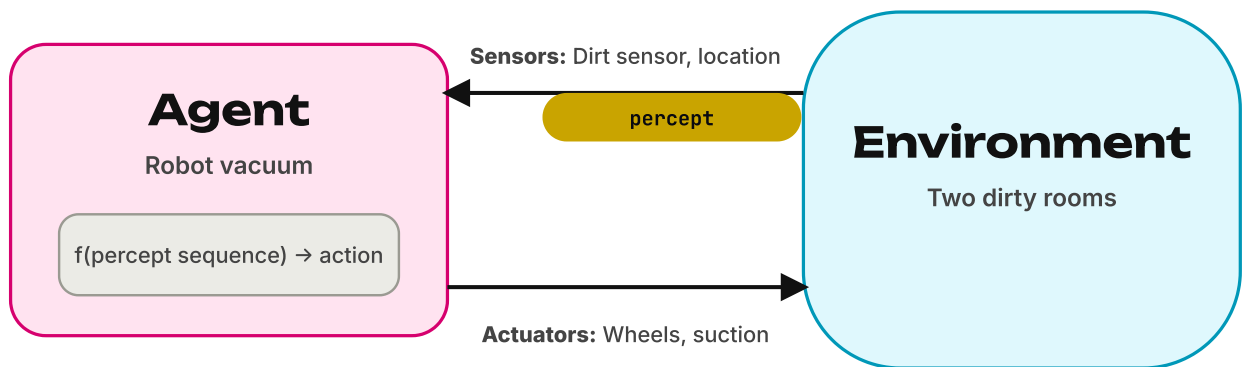
One skeleton for every AI system: percepts in, actions out.

01 COLD OPEN 0-7 min

Is a Pathao rider an agent? Are you?

The agent loop Percepts in, actions out, forever. Switch examples to see the same skeleton everywhere.

LO1



Right now: percept [A, Dirty] → action Suck

Anything that **perceives** its environment through sensors and **acts** on it through actuators is an agent.

One loop covers a robot vacuum, a dispatcher app and a medical device. Switch examples and watch the same skeleton.

Agent function vs agent program

Everything an agent has perceived so far is its **percept sequence**. Its behaviour is fully described by the **agent function**:

$$f : \mathcal{P}^* \rightarrow \mathcal{A}$$

The **agent program** implements f on real hardware.



GOTCHA

As a table, f needs $\sum_{t=1}^T |\mathcal{P}|^t$ rows. A chess agent's table would have more rows than atoms in the universe. AI = writing *small programs* that behave rationally, not tables.

PERCEPT SEQUENCE	ACTION
[A, Clean]	Right
[A, Dirty]	Suck
[B, Clean]	Left
[A, Clean], [A, Clean]	Right
[A, Clean], [A, Dirty]	Suck
... (it never ends)	...

What makes an agent rational?

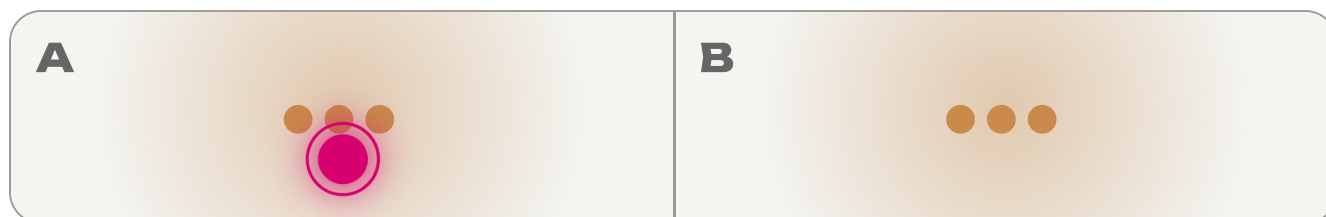
Rationality depends on four things: the **performance measure**, **prior knowledge**, the **actions** available, and the **percept sequence** so far.

A rational agent selects the action expected to maximise its performance measure, given the evidence and its built-in knowledge.

Vacuum world +1 per clean square per time step. Moving can cost 1. Which agent is rational?

Dirt comes back: **5% per step**

☒ Moving costs 1 point



t = 0

Score = 0

Per step = -

Press Step or Play. Log shows percept → action (reward).

04 **THINK DEEPER** 25-29 min

Critical thinking • be careful what you reward

1 **TRAP**

A startup rewards its vacuum agent **+1 every time it performs Suck**. The agent can also drop dirt from its bin.

What does a perfectly rational agent do? What should the performance measure be instead?

SOLUTION

ANSWER

It learns to **make its own dirt**: suck, dump, suck again. Every Suck scores, so the “rational” policy is a mess. This is *reward hacking* (the Cobra effect: a bounty on dead cobras made people breed cobras). Fix: reward the **state of the world** (clean squares per time step), not the action.

05 **LIVE LAB** 29-35 min

PEAS: specify the task first

Before designing any agent, write its task environment in four letters. Most bad AI projects skip this step.

PEAS builder Pick an agent. Class shouts answers first, then reveal each letter.

P

Performance measure

How do we score it?

E

Environment

Where does it live?

A

Actuators

How does it act?

S

Sensors

How does it perceive?

06 **WORKED EXAMPLE** 35-42 min

Worked example: an agro-spraying drone

GIVEN

Your lecturer co-founded an agro-drone startup for precision pesticide spraying on Bangladeshi farms.

FIND

The full PEAS, the six environment properties, and a justified agent type (the exam-CQ structure).

Method: Russell & Norvig Ch. 2 (PEAS, task environments)

SOLUTION

1. **P (performance):** fraction of the field correctly sprayed, pesticide per hectare (less is better), zero spray on water or people, flights per field, yield later.
Score the outcome you want in the world, not the behaviour you imagine. "Sprays a lot" is a terrible measure.
2. **E (environment):** irregular rice paddies, wind, rain, trees, power lines, birds, farmers walking the field, pest patches that move over days.
3. **A (actuators):** four rotors, spray pump and nozzle valve, status LED and buzzer, messages to the farmer's phone.

4. **S (sensors):** GPS, IMU, downward camera (pests, leaf colour), ultrasonic/LiDAR height, wind sensor, battery and tank level.
5. **Environment:** partially observable, single-agent (unless drones share a field), stochastic (gusts), sequential, dynamic, continuous.
6. **Agent type:** model-based (remember sprayed rows) + goal-based (cover the field); utility-based to trade coverage against pesticide and battery; learning if its pest detector improves each season.

Exam tip: justify the agent type by pointing at the environment property that forces it (partial observability → needs a model; trade-offs → needs utility).

ANSWER

PEAS written, environment classified, agent type justified: a full-mark answer structure.

07 **LIVE LAB** 42-51 min

Six questions about any environment

OBSERVABLE?

Can sensors see everything relevant now?

AGENTS?

Do other agents' choices affect yours?

DETERMINISTIC?

Is the next state fixed by state + action?

EPISODIC?

Does today's decision change tomorrow?

STATIC?

Does the world change while you think?

DISCRETE?

Finite states, or smooth quantities?

Environment classifier Click a cell to cycle options. Fill the table in pairs, then press Check.

TASK	OBSERVABLE	AGENTS	DETERMINISTIC?	EPISODIC?	STATIC?	DISCRETE?
Crossword puzzle	?	?	?	?	?	?
Chess with a clock	?	?	?	?	?	?
Ludo	?	?	?	?	?	?
Poker	?	?	?	?	?	?
Driving a taxi in Dhaka	?	?	?	?	?	?
Medical diagnosis	?	?	?	?	?	?
Classifying skin-lesion photos	?	?	?	?	?	?

Some cells are genuinely debatable (is a photo classifier's world static or semi-dynamic?).
Defend your answer and you get the mark.

08 **LIVE LAB** 51-57 min

Five agent designs, each smarter than the last

Five agent designs Each one adds a single new idea to the one before it.

Sensors



What the world is like now



Condition-action rules



Action

NEW IDEA

Nothing: if condition then action.

EXAMPLE

Automatic door, the AC thermostat, the vacuum reflex agent.

WEAKNESS

Blind to history. Loops forever in partially observable worlds.



PRO TIP

The hardest case is partially observable, multi-agent, stochastic, sequential, dynamic and continuous: driving a CNG through Dhaka at 6 pm. That is why self-driving is still hard.

Your turn • 1: a Ludo bot

2 EXAM CQ

Write the PEAS for an agent that plays Ludo on a phone against three humans.

Classify its environment on all six properties.

SOLUTION

ANSWER

P: win rate, finishing position, fun for the humans. **E:** board, dice, three opponents. **A:** choose which token to move. **S:** board state, dice value.

Fully observable, multi-agent, stochastic (dice), sequential, static (turn-based), discrete.

Your turn • 2: which agent type?

3 WARM-UP

(a) An automatic hand dryer. (b) A robot vacuum that maps your flat. (c) A route planner that minimises time and toll cost together. (d) A spam filter that improves from your "not spam" clicks.

SOLUTION

ANSWER

(a) Simple reflex: hands detected → blow. (b) Model-based (keeps a map), goal-based for coverage. (c) Utility-based: trades time against money. (d) Learning agent: the "not spam" click is the critic's feedback.

Critical thinking •

rational, but for whom?

THINK DEEPER

A ride-hailing dispatcher maximises total company revenue. It is perfectly rational. Is it a *good* agent?

It is rational *for its measure*, and that measure may be bad for riders and drivers. Choosing the performance measure is a **human, ethical** decision; the agent only optimises it.

A STRONG ANSWER MENTIONS

- Rational is always relative to the performance measure someone chose.
- Revenue alone can mean surge pricing in floods, starving far-away areas of drivers.
- Better: a utility that weights rider wait, driver earnings, fairness across areas.

THINK DEEPER

Is a chess engine's world static? What about the same engine playing with a clock?

Static without a clock; **semi-dynamic** with one: the board doesn't change while it thinks, but its remaining time does.

A STRONG ANSWER MENTIONS

- Without a clock: static; the world waits while it thinks.
- With a clock: semi-dynamic; the board waits but the score (time) does not.
- This decides whether "think longer" is free.

12 THINK DEEPER 68-71 min

Critical thinking • design

a performance measure

4 DESIGN

Dhaka Medical College's emergency department wants an agent that assigns incoming patients to doctors.

Write its performance measure as a formula or a ranked list. Include at least one hard constraint and one way the measure could be gamed.

SOLUTION

ANSWER

A strong answer lists measurable outcomes with weights, e.g. $U = -(total\ patient\ wait) - 5 \cdot (critical\ patients\ waiting > 10\ min) - 0.5 \cdot (nurse\ overtime\ hours)$, plus a **hard constraint** (a red-triage patient is never queued behind a green one). It notes the asymmetry (a missed critical case is far worse than a slow minor one) and how the measure could be gamed (e.g. discharging early to cut wait times).

13 REAL WORLD 71-74 min

The agent inside an insulin pump

♥ BME CORNER

For a person with type 1 diabetes, a "hybrid closed-loop" system (first approved in the US in 2016) is an agent: a continuous glucose monitor is the sensor, the pump is the actuator, and an algorithm decides the dose every few minutes. The patient still announces meals, hence "hybrid".

Partially observable (the sensor reads fluid under the skin, minutes behind the blood) and stochastic (meals, exercise, stress), so it needs an internal model.

The measure is **asymmetric**: too high is bad over years, too low is dangerous within minutes. So a sensible design is utility-based and cautious, not "high $\rightarrow$ inject".

14 POP QUIZ 74-78 min

Pop quiz

1. An agent function maps...

A. Actions to percepts

B. The full percept sequence to an action ✓

C. Sensors to actuators, one wire each

D. Goals to rewards

f: $P^* \rightarrow A$. The agent program implements that function on real hardware.

2. A rational agent...

- A. Always gets the best outcome
- B. Knows every action's result in advance
- C. Picks the action with the best expected performance given its percepts and knowledge ✓**
- D. Never makes mistakes

Rational is not omniscient. It can have bad luck or incomplete information.

3. Poker is...

- A. Fully observable and deterministic
- B. Partially observable and stochastic ✓**
- C. Fully observable and stochastic
- D. Episodic and single-agent

You cannot see opponents' cards (partial) and the deal is random (stochastic).

4. A simple reflex agent fails in partially observable worlds because...

- A. It is too slow
- B. It has no memory of what it cannot currently see ✓**
- C. It uses too much power
- D. It cannot have sensors

It acts only on the current percept, so it can loop forever.

5. Preferring "arrive safely in 50 min" over "arrive in 40 min with a near-crash" needs a...

- A. Simple reflex agent
- B. Model-based agent
- C. Goal-based agent
- D. Utility-based agent ✓**

Both plans reach the goal. Only a utility function ranks how good each outcome is.

15 **RECAP** 78-80 min

Three things to remember

01

Agent = sensors + actuators + a program implementing f : percept sequence $\rightarrow$ action. Rational = best *expected* performance, not perfect.

02

Specify with PEAS; classify the environment on six properties. The properties decide what kind of agent you need.

03

Reflex → model-based → goal-based → utility-based → learning. Next week: goal-based agents that search, and logic. Assignment 1 is published.

S. M. Sakeef Sani

Adjunct Lecturer, CSE · Northern University Bangladesh

smsakeefsani.hbitcorp.com