

ARTIFICIAL INTELLIGENCE

CHAPTER 1

Outline

- ◇ What is AI?
- ◇ A brief history
- ◇ The state of the art

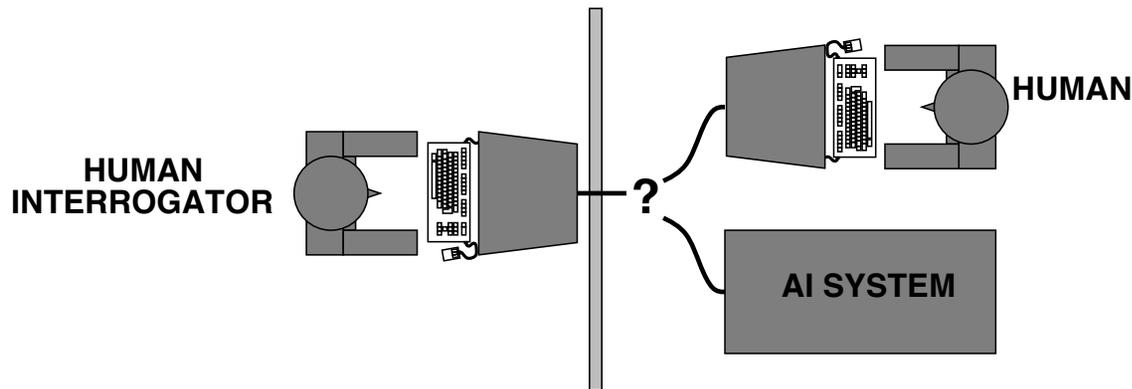
What is AI?

Systems that think like humans	Systems that think rationally
Systems that act like humans	Systems that act rationally

Acting humanly: The Turing test

Turing (1950) “Computing machinery and intelligence”:

- ◇ “Can machines think?” → “Can machines behave intelligently?”
- ◇ Operational test for intelligent behavior: the **Imitation Game**



- ◇ Predicted that by 2000, a machine might have a 30% chance of fooling a lay person for 5 minutes
- ◇ Anticipated all major arguments against AI in following 50 years
- ◇ Suggested major components of AI: knowledge, reasoning, language understanding, learning

Problem: Turing test is not **reproducible**, **constructive**, or amenable to **mathematical analysis**

Thinking humanly: Cognitive Science

1960s “cognitive revolution”: information-processing psychology replaced prevailing orthodoxy of behaviorism

Requires scientific theories of internal activities of the brain

- What level of abstraction? “Knowledge” or “circuits”?
- How to validate? Requires
 - 1) Predicting and testing behavior of human subjects (top-down)
 - or 2) Direct identification from neurological data (bottom-up)

Both approaches (roughly, Cognitive Science and Cognitive Neuroscience) are now distinct from AI

Both share with AI the following characteristic:

**the available theories do not explain (or engender)
anything resembling human-level general intelligence**

Hence, all three fields share one principal direction!

Thinking rationally: Laws of Thought

Normative (or prescriptive) rather than descriptive

Aristotle: what are correct arguments/thought processes?

Several Greek schools developed various forms of logic:

notation and **rules of derivation** for thoughts;
may or may not have proceeded to the idea of mechanization

Direct line through mathematics and philosophy to modern AI

Problems:

- 1) Not all intelligent behavior is mediated by logical deliberation
- 2) **What is the purpose of thinking?** What thoughts **should** I have out of all the thoughts (logical or otherwise) that I **could** have?

Acting rationally

Rational behavior: doing the right thing

The right thing: that which is expected to maximize goal achievement, given the available information

Doesn't necessarily involve thinking—e.g., blinking reflex—but thinking should be in the service of rational action

Aristotle (Nicomachean Ethics):

Every art and every inquiry, and similarly every action and pursuit, is thought to aim at some good

Rational agents

An **agent** is an entity that perceives and acts

This course is about designing **rational agents**

Abstractly, an agent is a function from percept histories to actions:

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

For any given class of environments and tasks, we seek the agent (or class of agents) with the best performance

Caveat: **computational limitations make perfect rationality unachievable**

→ design best **program** for given machine resources

AI prehistory

Philosophy	logic, methods of reasoning mind as physical system foundations of learning, language, rationality
Mathematics	formal representation and proof algorithms, computation, (un)decidability, (in)tractability probability
Psychology	adaptation phenomena of perception and motor control experimental techniques (psychophysics, etc.)
Economics	formal theory of rational decisions
Linguistics	knowledge representation grammar
Neuroscience	plastic physical substrate for mental activity
Control theory	homeostatic systems, stability simple optimal agent designs

Potted history of AI

- 1943 McCulloch & Pitts: Boolean circuit model of brain
- 1950 Turing's "Computing Machinery and Intelligence"
- 1952–69 Look, Ma, no hands!
- 1950s Early AI programs, including Samuel's checkers program, Newell & Simon's Logic Theorist, Gelernter's Geometry Engine
- 1956 Dartmouth meeting: "Artificial Intelligence" adopted
- 1965 Robinson's complete algorithm for logical reasoning
- 1966–74 AI discovers computational complexity
Neural network research almost disappears
- 1969–79 Early development of knowledge-based systems
- 1980–88 Expert systems industry booms
- 1988–93 Expert systems industry busts: "AI Winter"
- 1985–95 Neural networks return to popularity
- 1988– Resurgence of probability; general increase in technical depth
"Nouvelle AI": ALife, GAs, soft computing
- 1995– Agents, agents, everywhere . . .
- 2003– Human-level AI back on the agenda

State of the art

Which of the following can be done at present?

◇ Play a decent game of table tennis

State of the art

Which of the following can be done at present?

- ◇ Play a decent game of table tennis
- ◇ Drive safely along a curving mountain road

State of the art

Which of the following can be done at present?

- ◇ Play a decent game of table tennis
- ◇ Drive safely along a curving mountain road
- ◇ Drive safely along Telegraph Avenue

State of the art

Which of the following can be done at present?

- ◇ Play a decent game of table tennis
- ◇ Drive safely along a curving mountain road
- ◇ Drive safely along Telegraph Avenue
- ◇ Buy a week's worth of groceries on the web

State of the art

Which of the following can be done at present?

- ◇ Play a decent game of table tennis
- ◇ Drive safely along a curving mountain road
- ◇ Drive safely along Telegraph Avenue
- ◇ Buy a week's worth of groceries on the web
- ◇ Buy a week's worth of groceries at Berkeley Bowl

State of the art

Which of the following can be done at present?

- ◇ Play a decent game of table tennis
- ◇ Drive safely along a curving mountain road
- ◇ Drive safely along Telegraph Avenue
- ◇ Buy a week's worth of groceries on the web
- ◇ Buy a week's worth of groceries at Berkeley Bowl
- ◇ Play a decent game of bridge

State of the art

Which of the following can be done at present?

- ◇ Play a decent game of table tennis
- ◇ Drive safely along a curving mountain road
- ◇ Drive safely along Telegraph Avenue
- ◇ Buy a week's worth of groceries on the web
- ◇ Buy a week's worth of groceries at Berkeley Bowl
- ◇ Play a decent game of bridge
- ◇ Discover and prove a new mathematical theorem

State of the art

Which of the following can be done at present?

- ◇ Play a decent game of table tennis
- ◇ Drive safely along a curving mountain road
- ◇ Drive safely along Telegraph Avenue
- ◇ Buy a week's worth of groceries on the web
- ◇ Buy a week's worth of groceries at Berkeley Bowl
- ◇ Play a decent game of bridge
- ◇ Discover and prove a new mathematical theorem
- ◇ Design and execute a research program in molecular biology

State of the art

Which of the following can be done at present?

- ◇ Play a decent game of table tennis
- ◇ Drive safely along a curving mountain road
- ◇ Drive safely along Telegraph Avenue
- ◇ Buy a week's worth of groceries on the web
- ◇ Buy a week's worth of groceries at Berkeley Bowl
- ◇ Play a decent game of bridge
- ◇ Discover and prove a new mathematical theorem
- ◇ Design and execute a research program in molecular biology
- ◇ Write an intentionally funny story

State of the art

Which of the following can be done at present?

- ◇ Play a decent game of table tennis
- ◇ Drive safely along a curving mountain road
- ◇ Drive safely along Telegraph Avenue
- ◇ Buy a week's worth of groceries on the web
- ◇ Buy a week's worth of groceries at Berkeley Bowl
- ◇ Play a decent game of bridge
- ◇ Discover and prove a new mathematical theorem
- ◇ Design and execute a research program in molecular biology
- ◇ Write an intentionally funny story
- ◇ Give competent legal advice in a specialized area of law

State of the art

Which of the following can be done at present?

- ◇ Play a decent game of table tennis
- ◇ Drive safely along a curving mountain road
- ◇ Drive safely along Telegraph Avenue
- ◇ Buy a week's worth of groceries on the web
- ◇ Buy a week's worth of groceries at Berkeley Bowl
- ◇ Play a decent game of bridge
- ◇ Discover and prove a new mathematical theorem
- ◇ Design and execute a research program in molecular biology
- ◇ Write an intentionally funny story
- ◇ Give competent legal advice in a specialized area of law
- ◇ Translate spoken English into spoken Swedish in real time

State of the art

Which of the following can be done at present?

- ◇ Play a decent game of table tennis
- ◇ Drive safely along a curving mountain road
- ◇ Drive safely along Telegraph Avenue
- ◇ Buy a week's worth of groceries on the web
- ◇ Buy a week's worth of groceries at Berkeley Bowl
- ◇ Play a decent game of bridge
- ◇ Discover and prove a new mathematical theorem
- ◇ Design and execute a research program in molecular biology
- ◇ Write an intentionally funny story
- ◇ Give competent legal advice in a specialized area of law
- ◇ Translate spoken English into spoken Swedish in real time
- ◇ Converse successfully with another person for an hour

State of the art

Which of the following can be done at present?

- ◇ Play a decent game of table tennis
- ◇ Drive safely along a curving mountain road
- ◇ Drive safely along Telegraph Avenue
- ◇ Buy a week's worth of groceries on the web
- ◇ Buy a week's worth of groceries at Berkeley Bowl
- ◇ Play a decent game of bridge
- ◇ Discover and prove a new mathematical theorem
- ◇ Design and execute a research program in molecular biology
- ◇ Write an intentionally funny story
- ◇ Give competent legal advice in a specialized area of law
- ◇ Translate spoken English into spoken Swedish in real time
- ◇ Converse successfully with another person for an hour
- ◇ Perform a complex surgical operation

State of the art

Which of the following can be done at present?

- ◇ Play a decent game of table tennis
- ◇ Drive safely along a curving mountain road
- ◇ Drive safely along Telegraph Avenue
- ◇ Buy a week's worth of groceries on the web
- ◇ Buy a week's worth of groceries at Berkeley Bowl
- ◇ Play a decent game of bridge
- ◇ Discover and prove a new mathematical theorem
- ◇ Design and execute a research program in molecular biology
- ◇ Write an intentionally funny story
- ◇ Give competent legal advice in a specialized area of law
- ◇ Translate spoken English into spoken Swedish in real time
- ◇ Converse successfully with another person for an hour
- ◇ Perform a complex surgical operation
- ◇ Unload any dishwasher and put everything away

State of the art

Which of the following can be done at present?

- ◇ Play a decent game of table tennis
- ◇ Drive safely along a curving mountain road
- ◇ Drive safely along Telegraph Avenue
- ◇ Buy a week's worth of groceries on the web
- ◇ Buy a week's worth of groceries at Berkeley Bowl
- ◇ Play a decent game of bridge
- ◇ Discover and prove a new mathematical theorem
- ◇ Design and execute a research program in molecular biology
- ◇ Write an intentionally funny story
- ◇ Give competent legal advice in a specialized area of law
- ◇ Translate spoken English into spoken Swedish in real time
- ◇ Converse successfully with another person for an hour
- ◇ Perform a complex surgical operation
- ◇ Unload any dishwasher and put everything away

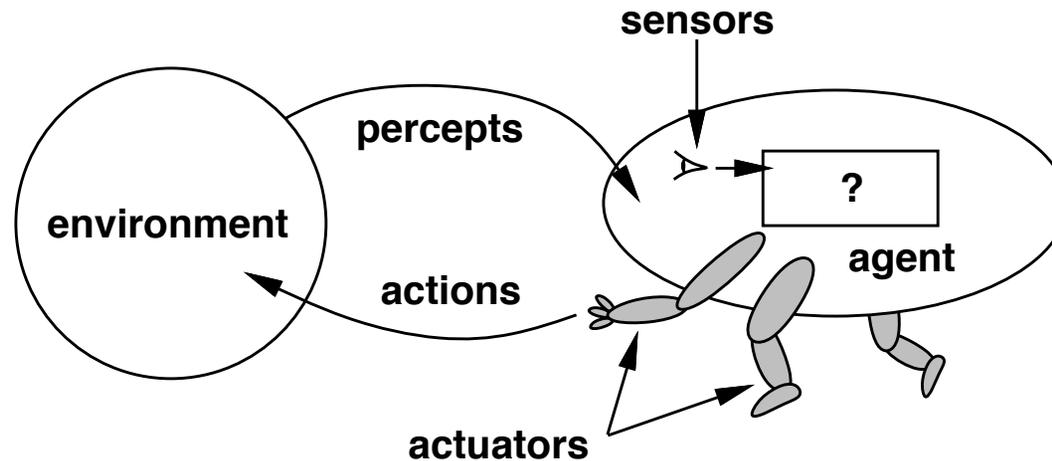
INTELLIGENT AGENTS

CHAPTER 2

Outline

- ◇ Agents and environments
- ◇ Rationality
- ◇ PEAS (Performance measure, Environment, Actuators, Sensors)
- ◇ Environment types
- ◇ Agent types

Agents and environments



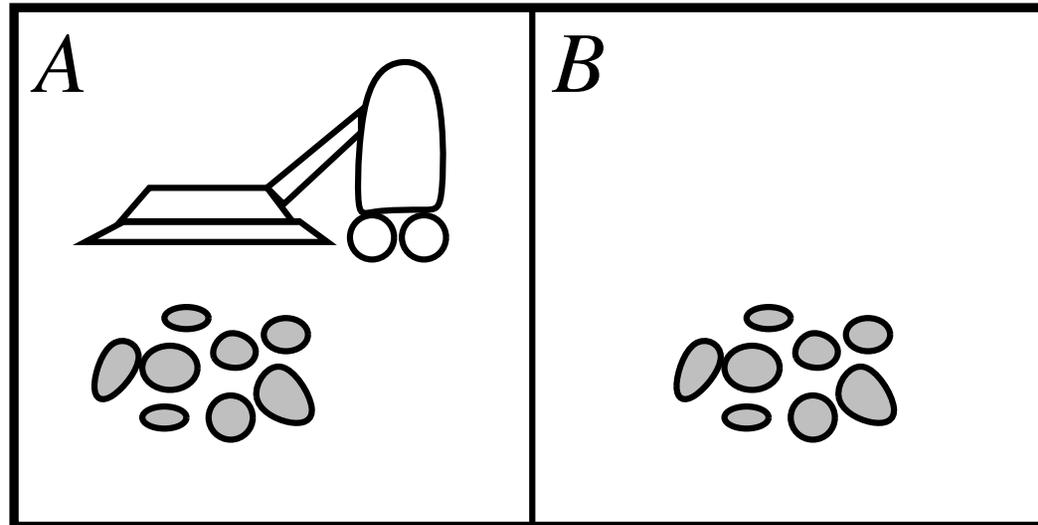
Agents include humans, robots, softbots, thermostats, etc.

The **agent function** maps from percept histories to actions:

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

The **agent program** runs on the physical **architecture** to produce f

Vacuum-cleaner world



Percepts: location and contents, e.g., $[A, \textit{Dirty}]$

Actions: *Left*, *Right*, *Suck*, *NoOp*

A vacuum-cleaner agent

Percept sequence	Action
<i>[A, Clean]</i>	<i>Right</i>
<i>[A, Dirty]</i>	<i>Suck</i>
<i>[B, Clean]</i>	<i>Left</i>
<i>[B, Dirty]</i>	<i>Suck</i>
<i>[A, Clean], [A, Clean]</i>	<i>Right</i>
<i>[A, Clean], [A, Dirty]</i>	<i>Suck</i>
<i>⋮</i>	<i>⋮</i>

function REFLEX-VACUUM-AGENT([*location, status*]) **returns** an action

if *status = Dirty* **then return** *Suck*
else if *location = A* **then return** *Right*
else if *location = B* **then return** *Left*

What is the **right** function?

Can it be implemented in a small agent program?

Rationality

Fixed **performance measure** evaluates the **environment sequence**

- one point per square cleaned up in time T ?
- one point per clean square per time step, minus one per move?
- penalize for $> k$ dirty squares?

A **rational agent** chooses whichever action maximizes the **expected** value of the performance measure **given the percept sequence to date**

Rational \neq omniscient

- percepts may not supply all relevant information

Rational \neq clairvoyant

- action outcomes may not be as expected

Hence, rational \neq successful

Rational \Rightarrow exploration, learning, autonomy

PEAS

To design a rational agent, we must specify the **task environment**

Consider, e.g., the task of designing an automated taxi:

Performance measure??

Environment??

Actuators??

Sensors??

PEAS

To design a rational agent, we must specify the **task environment**

Consider, e.g., the task of designing an automated taxi:

Performance measure?? safety, destination, profits, legality, comfort, ...

Environment?? US streets/freeways, traffic, pedestrians, weather, ...

Actuators?? steering, accelerator, brake, horn, speaker/display, ...

Sensors?? video, accelerometers, gauges, engine sensors, keyboard, GPS, ...

Internet shopping agent

Performance measure??

Environment??

Actuators??

Sensors??

Internet shopping agent

Performance measure?? price, quality, appropriateness, efficiency

Environment?? current and future WWW sites, vendors, shippers

Actuators?? display to user, follow URL, fill in form

Sensors?? HTML pages (text, graphics, scripts)

Environment types

	Solitaire	Backgammon	Internet shopping	Taxi
<u>Observable??</u>				
<u>Deterministic??</u>				
<u>Episodic??</u>				
<u>Static??</u>				
<u>Discrete??</u>				
<u>Single-agent??</u>				

Environment types

	Solitaire	Backgammon	Internet shopping	Taxi
<u>Observable??</u>	Yes	Yes	No	No
<u>Deterministic??</u>				
<u>Episodic??</u>				
<u>Static??</u>				
<u>Discrete??</u>				
<u>Single-agent??</u>				

Environment types

	Solitaire	Backgammon	Internet shopping	Taxi
<u>Observable??</u>	Yes	Yes	No	No
<u>Deterministic??</u>	Yes	No	Partly	No
<u>Episodic??</u>				
<u>Static??</u>				
<u>Discrete??</u>				
<u>Single-agent??</u>				

Environment types

	Solitaire	Backgammon	Internet shopping	Taxi
<u>Observable??</u>	Yes	Yes	No	No
<u>Deterministic??</u>	Yes	No	Partly	No
<u>Episodic??</u>	No	No	No	No
<u>Static??</u>				
<u>Discrete??</u>				
<u>Single-agent??</u>				

Environment types

	Solitaire	Backgammon	Internet shopping	Taxi
<u>Observable??</u>	Yes	Yes	No	No
<u>Deterministic??</u>	Yes	No	Partly	No
<u>Episodic??</u>	No	No	No	No
<u>Static??</u>	Yes	Semi	Semi	No
<u>Discrete??</u>				
<u>Single-agent??</u>				

Environment types

	Solitaire	Backgammon	Internet shopping	Taxi
<u>Observable??</u>	Yes	Yes	No	No
<u>Deterministic??</u>	Yes	No	Partly	No
<u>Episodic??</u>	No	No	No	No
<u>Static??</u>	Yes	Semi	Semi	No
<u>Discrete??</u>	Yes	Yes	Yes	No
<u>Single-agent??</u>				

Environment types

	Solitaire	Backgammon	Internet shopping	Taxi
<u>Observable??</u>	Yes	Yes	No	No
<u>Deterministic??</u>	Yes	No	Partly	No
<u>Episodic??</u>	No	No	No	No
<u>Static??</u>	Yes	Semi	Semi	No
<u>Discrete??</u>	Yes	Yes	Yes	No
<u>Single-agent??</u>	Yes	No	Yes (except auctions)	No

The environment type largely determines the agent design

The real world is (of course) partially observable, stochastic, sequential, dynamic, continuous, multi-agent

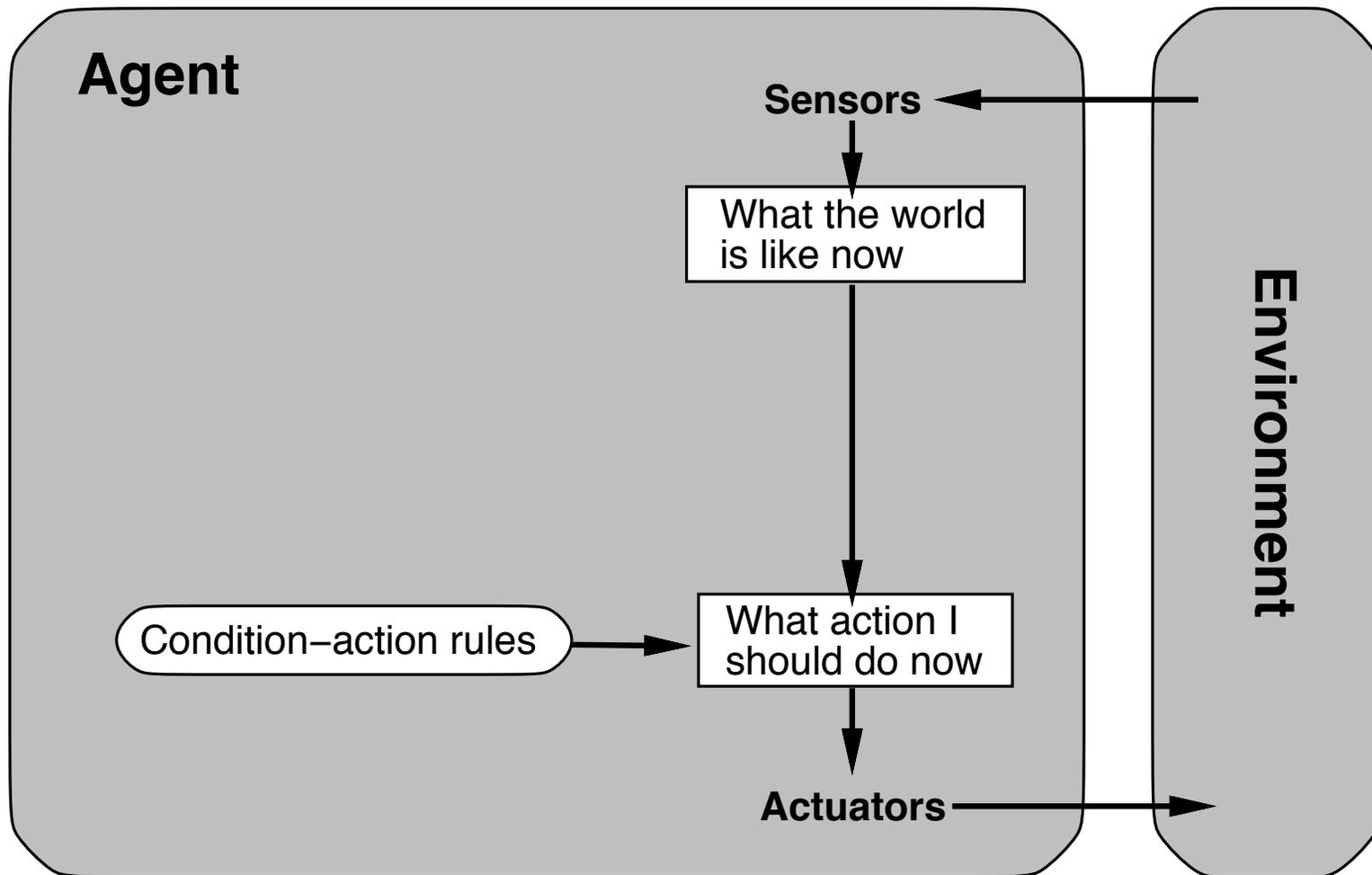
Agent types

Four basic types in order of increasing generality:

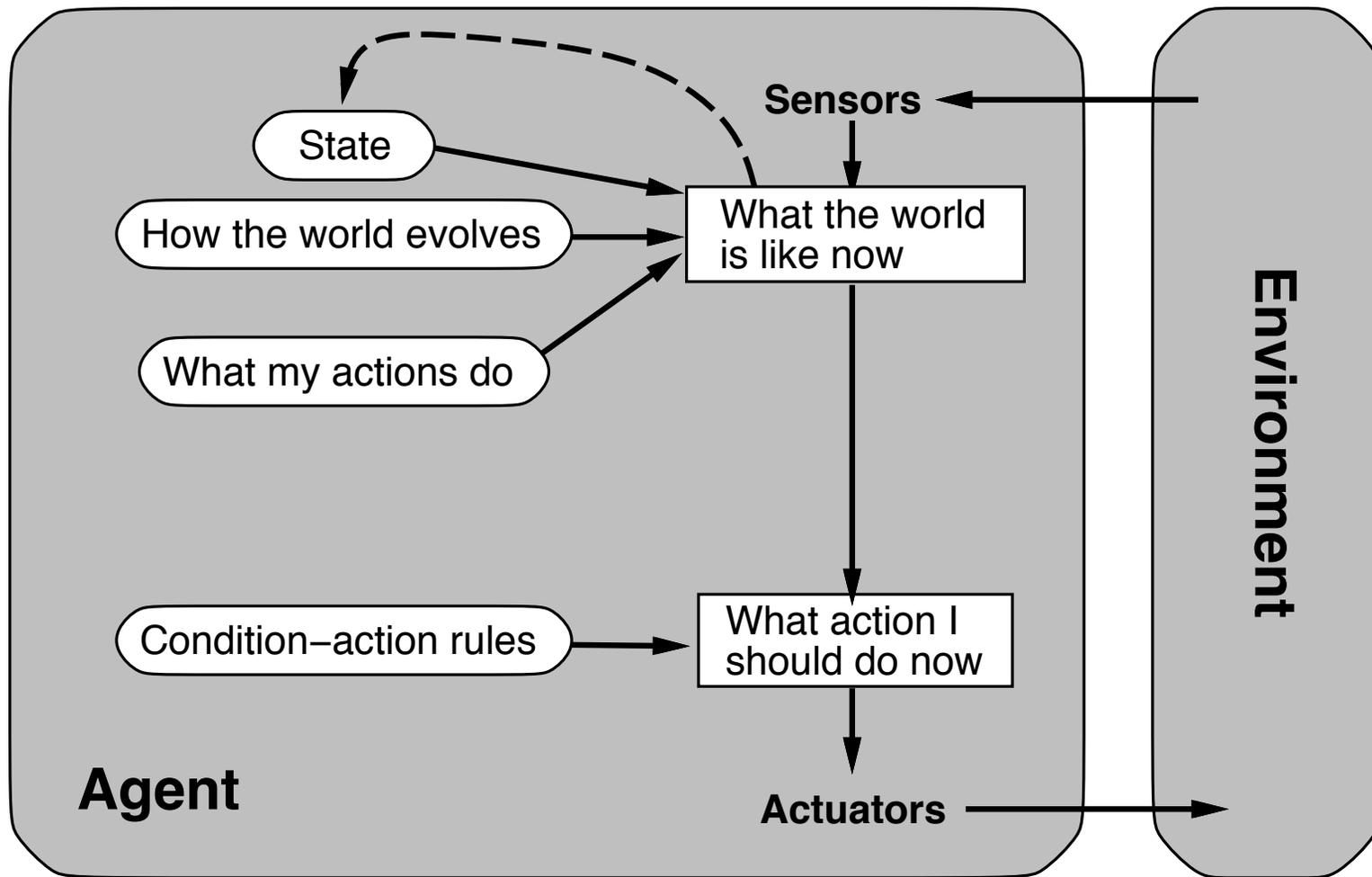
- simple reflex agents
- reflex agents with state
- goal-based agents
- utility-based agents

All these can be turned into learning agents

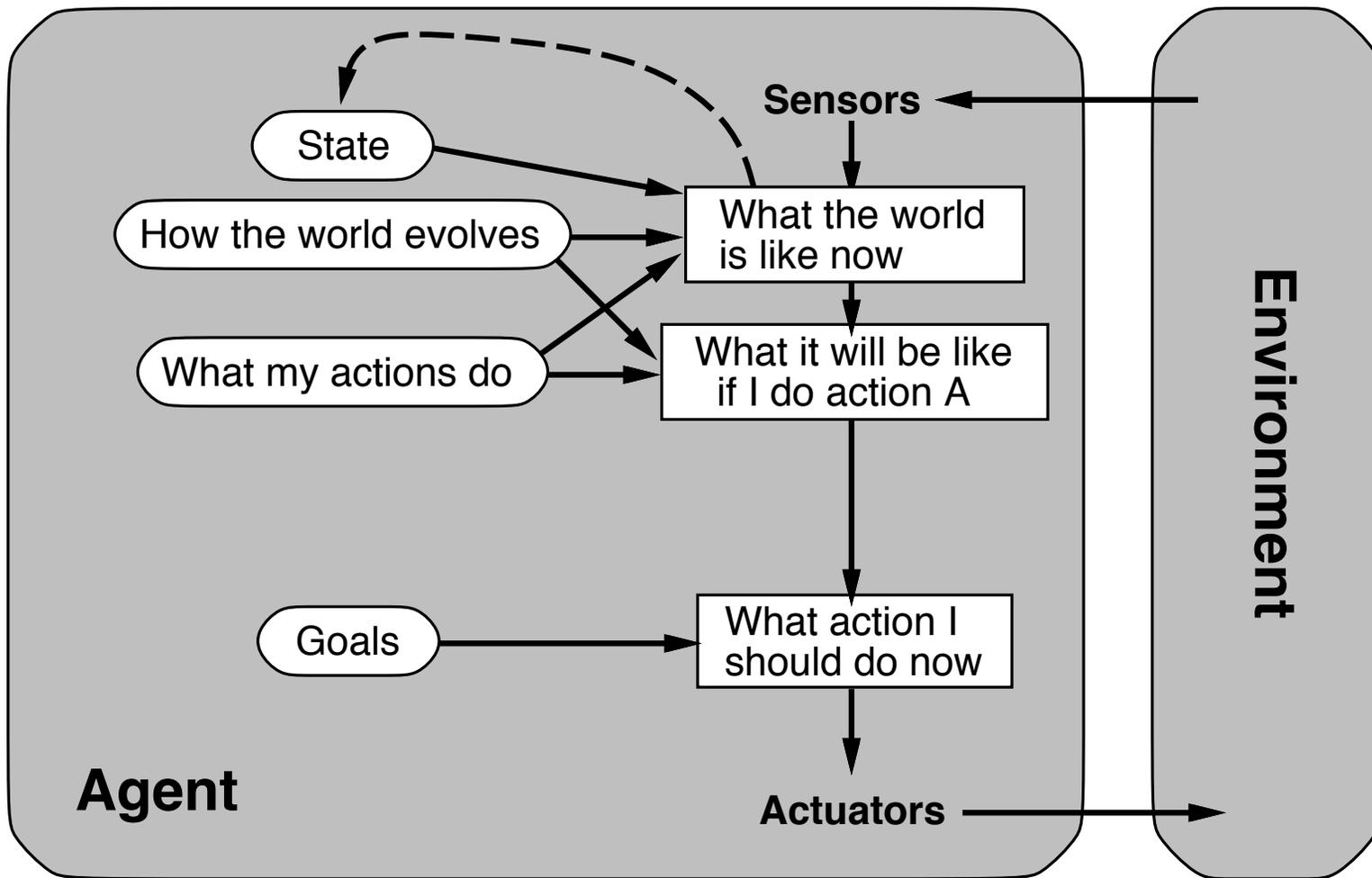
Simple reflex agents



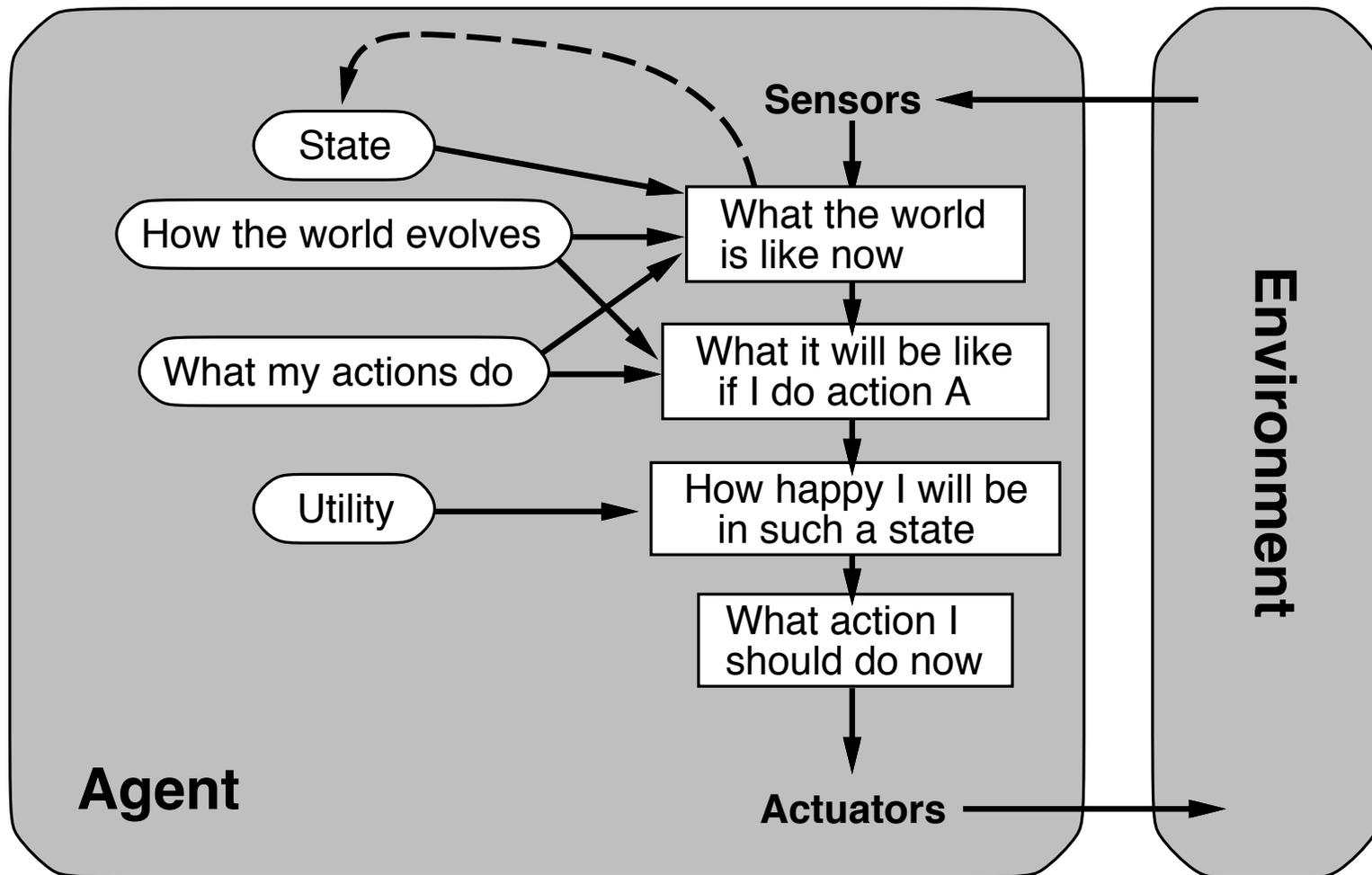
Reflex agents with state



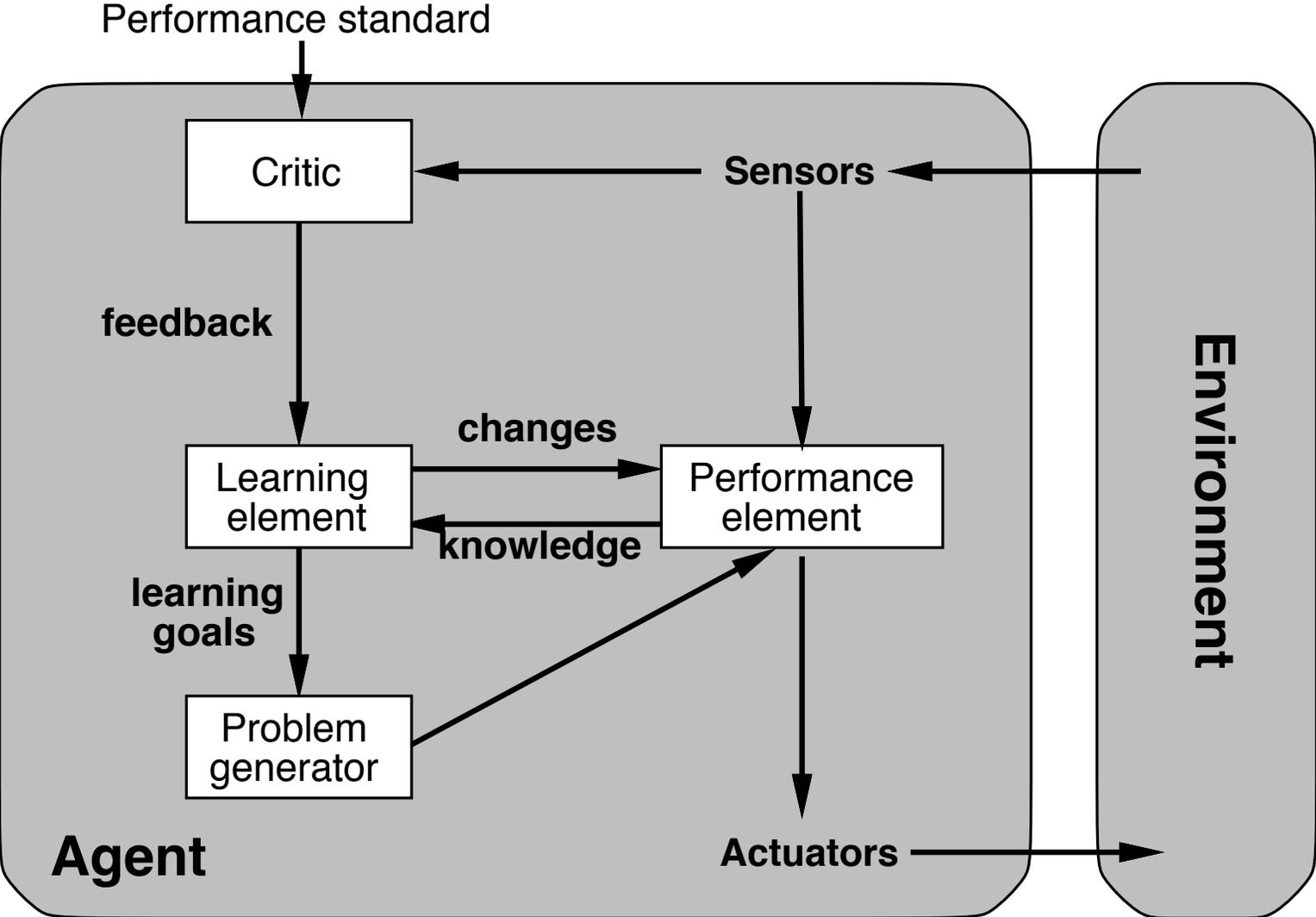
Goal-based agents



Utility-based agents



Learning agents



Summary

Agents interact with environments through actuators and sensors

The agent function describes what the agent does in all circumstances

The performance measure evaluates the environment sequence

A perfectly rational agent maximizes expected performance

Agent programs implement (some) agent functions

PEAS descriptions define task environments

Environments are categorized along several dimensions:

observable? deterministic? episodic? static? discrete? single-agent?

Several basic agent architectures exist:

reflex, reflex with state, goal-based, utility-based