<u>Course</u>: Introduction to AI

Prof. Ahmed Guessoum
The National Higher School of AI

Chapter 2

Intelligent Agents

Outline

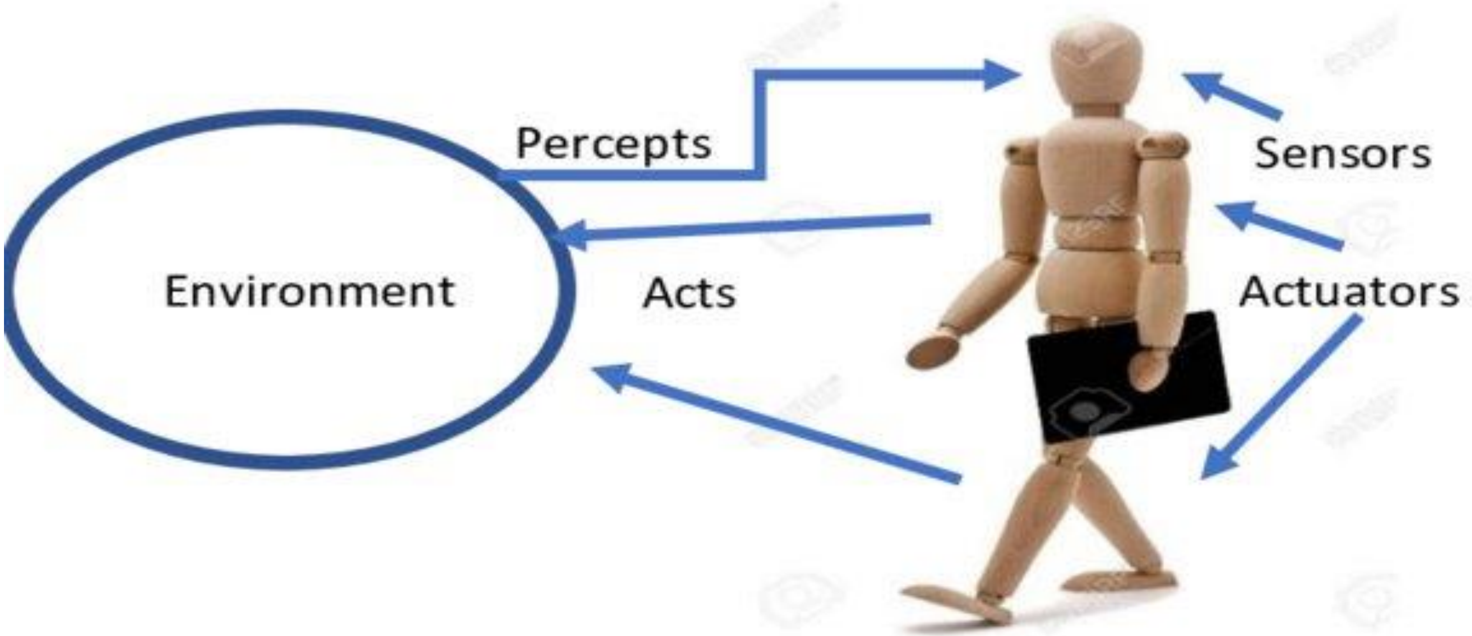
- Agents and Environments
- Good Behavior: The Concept of Rationality
  - Rationality
  - Omniscience, learning, and autonomy
- The Nature of Environments
  - Specifying the task environment
  - Properties of task environments
- THE STRUCTURE OF AGENTS
  - Agent programs
  - Simple reflex agents
  - Model-based reflex agents
  - Goal-based agents
  - Utility-based agents
  - Learning agents
  - How the components of agent programs work

Aim of this chapter

- **Rational agents** as central to "our" approach to artificial intelligence
- The concept of **rationality** can be applied to a wide variety of agents operating in any imaginable environment
- Use this concept to develop a small set of design principles for building successful agents—systems that can reasonably be called **intelligent**.

Agents and Environments: Agents

- **Agent:** anything that can be viewed as perceiving its **environment** through **sensors** and acting upon that environment through **actuators. E.g.:**



https://www.researchgate.net/publication/357618741_Challenges_of_Artificial_Intelligence_--_From_Machine_Learning_and_Computer_Vision_to_Emotional_Intelligence/figures?lo=1

Agents and Environments: Agents

- **Agent:** anything that can be viewed as perceiving its **environment** through **sensors** and acting upon that environment through **actuators. E.g.:**

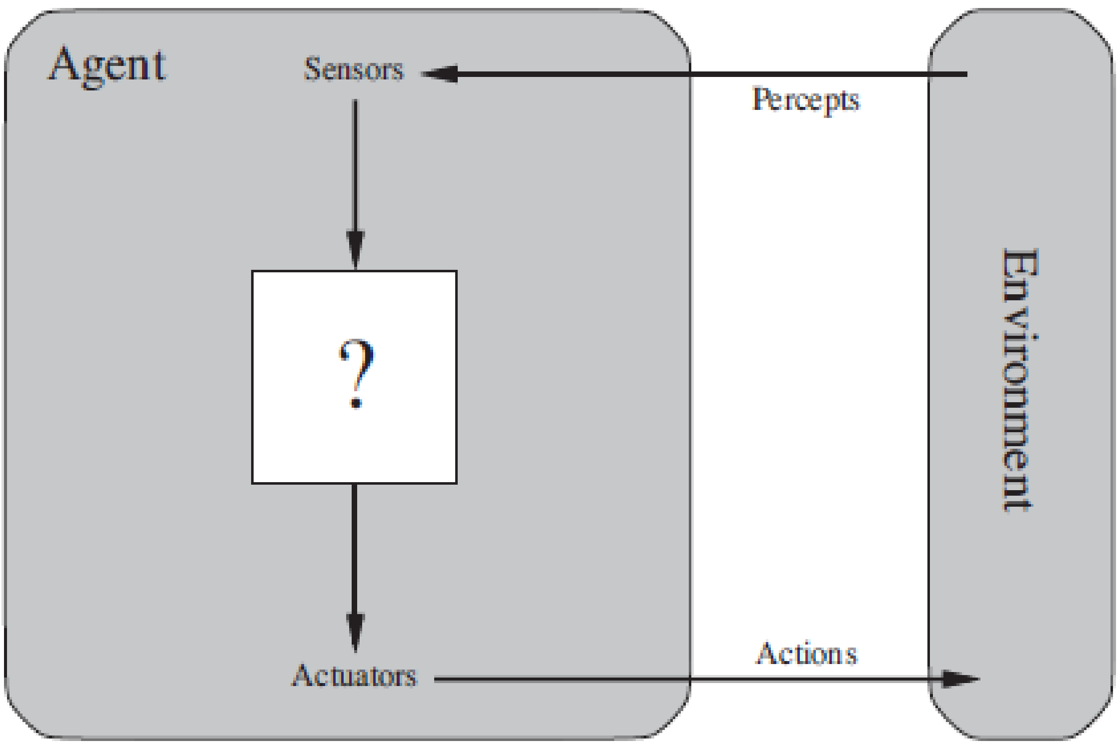


Figure 2.1 Agents interact with environments through sensors and actuators.

Agents and Environments: Agents

Examples of Agents:

- Human agent:
  - Sensors: eyes, ears, skin, …
  - Actuators: hands, legs, vocal tract, …
- Robotic agent:
  - Sensors: cameras, infrared and sonar range finders…
  - Actuators: various motors, robot arm...
- Software agent:
  - Sensory inputs: keystrokes, file contents, network packets, …
  - Actions: displaying on the screen, writing files, sending network packets, …

Agents and Environments: Agents

- **Percept:** the agent's perceptual inputs at any given instant.
- Agent's **percept sequence:** complete history of everything the agent has ever perceived.
- *An agent's <u>choice of action</u> at any given instant <u>can depend on the entire percept sequence observed to date</u>, but not on anything it hasn't perceived.*
- Mathematically speaking, we say that an agent's behaviour is described by the **agent function** that <u>maps any given percept sequence to an action.</u>

Agents and Environments: Agents

- The **agent function** can be pictured as a table of mappings between percept sequences and actions.
- Such a table is an external (abstract mathematical) characterization of the agent function.
- *Internally*, the agent function for an artificial agent will be implemented within some physical system by an **agent program**.

Agent example

- Vacuum-cleaner world
  - 2 locations A and B
  - Agent **perceives** which square it is in and whether there is dirt in the square
  - Actions: *move left*, *move right*, *suck up the dirt*, or *do nothing.*

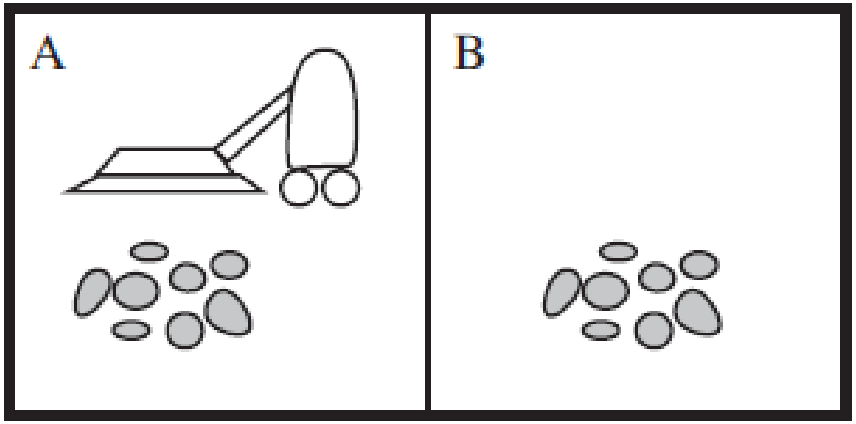


Figure 2.2 A vacuum-cleaner world with just two locations.

Partial tabulation of vacuum-cleaner agent function

Percept sequence Action	Action
[A, Clean]	Right
[A, Dirty]	Suck
[B, Clean]	Left
[B, Dirty]	Suck
[A, Clean], [A, Clean]	Right
[A, Clean], [A, Dirty]	Suck
.	.
.	.
.	.
[A, Clean], [A, Clean], [A, Clean]	Right
[A, Clean], [A, Clean], [A, Dirty]	Suck
.	.
.	.
.	.

Agent function design

- So there is a need to define the *Actions* corresponding to the percepts.
- **Question***:*
    - *What is the right way to fill out the table? (*i.e. what makes an agent good or bad, intelligent or stupid?)
- **N.B.:**
    - All areas of engineering can be seen as designing artifacts that interact with the world; but
    - **AI** operates where the artifacts have **significant computational resources** and the task environment requires **nontrivial decision making.**

Outline

- Agents and Environments
- Good Behavior: The Concept of Rationality
  - Rationality
  - Omniscience, learning, and autonomy
- The Nature of Environments
  - Specifying the task environment
  - Properties of task environments
- THE STRUCTURE OF AGENTS
  - Agent programs
  - Simple reflex agents
  - Model-based reflex agents
  - Goal-based agents
  - Utility-based agents
  - Learning agents
  - How the components of agent programs work

GOOD BEHAVIOR: THE CONCEPT OF RATIONALITY

- A **rational agent** is one that does the *right thing.* (i.e. every action in the table is *correct*)
- What does "right thing" ("correct action") mean?
- Obvious approach: consider the *consequences* of the agent's behaviour.

  Percepts → sequences of actions →

  environment goes through a sequence of states
- Desirability is captured by a **performance measure** that evaluates any given sequence of <u>environment states</u>.
- The performance measure is <u>designed specifically for the task</u> at stake.

Performance measure for vacuum cleaner agent

- Amount of dirt cleaned in a fixed period of time?
  - *Risk: agent cleans up the floor well, dumps the dirt, cleans up again, etc. ➔ not a good performance measure*
- <u>Alternative</u>: reward the agent for having a clean floor. ➔ it is a more suitable measure
  - Award one point for each clean square at each time step (perhaps with a penalty for electricity consumed and noise generated).
- *<u>General rule</u>: design performance measures according to expectations in the environment.*
- *<u>N.B.:</u> "Average* cleanliness over time" *is questionable!*

Rationality

- What is rational at any given time depends on four factors:
  - The performance measure that defines the criterion of success.
  - The agent's prior knowledge of the environment.
  - The actions that the agent can perform.
  - The agent's percept sequence to date.
- **Definition of a 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.

Rationality: vacuum-cleaner agent

- *if square is dirty clean it; else move to the other square*. (function previously tabulated)
- Is this a rational agent?

  (Yes; left as textbook exercise.)
- We need to analyse the enunciated 4 factors.
  - Performance measure: award one point for each clean square at each time step, over 1000 time steps period.
  - Only available actions: Left, Right, and Suck.
  - Environment "geography" known a priori; but not dirt distribution nor agent initial location. Clean squares stay clean; sucking cleans current square. Left and Right actions move agent left and right except when this would take agent outside the environment, in which case agent remains where it is.
  - The agent correctly perceives its location and whether that location contains dirt.

Rationality: vacuum-cleaner agent

- Note that the same agent would be irrational under different circumstances.
- For example,
  - Once all the dirt is cleaned up, the agent will oscillate needlessly back and forth;
  - If the performance measure includes a penalty of one point for each movement left or right, the agent will perform poorly. A better agent for this case would do nothing once it is sure that all the squares are clean.
  - If clean squares can become dirty again, the agent should occasionally check and re-clean them if needed.
  - If the geography of the environment is unknown, the agent will need to explore it rather than remain in square A or B.
- (Exercise 2.2 asks you to design agents for these cases.)

Omniscience, learning, & autonomy

- **Omniscient agent**: agent who knows the *actual* outcome of its actions and can act accordingly.
- Omniscience is impossible in reality; we cannot anticipate everything.
- Rationality is not the same as perfection; rationality maximizes *expected* performance, while perfection maximizes *actual* performance. The latter is impossible.
- Doing actions *in order to modify future percepts*— sometimes called **information gathering**— is an important part of rationality.
  - Example, it would not be rational to cross the road given an uninformative percept sequence (e.g. look right only)

Omniscience, learning, & autonomy

- Our definition requires a rational agent not only to gather information but also to **learn** as much as possible from what it perceives.
- The agent's initial configuration could reflect some prior knowledge of the environment, but as the agent gains experience this may be modified and augmented.
- If an agent relies on the prior knowledge of its designer rather than on its own percepts, the agent is said to lack **autonomy**.
- A rational agent should be autonomous—it should learn what it can to compensate for partial or incorrect prior knowledge.
    - E.g. vacuum cleaner learning where to foresee dirt better than one that does not.

Omniscience, learning, & autonomy

- Complete autonomy from the start is seldom required: <u>with little or no experience</u>, the agent would have to <u>act randomly</u> unless the designer gave some assistance.
- It would be reasonable to provide an artificial intelligent agent with
  - some initial knowledge as well as
  - an ability to learn.
- <u>After sufficient experience of its environment</u>, the behavior of a rational agent <u>can become</u> effectively ***independent*** of its prior knowledge.

Outline

- Agents and Environments
- Good Behavior: The Concept of Rationality
  - Rationality
  - Omniscience, learning, and autonomy
- The Nature of Environments
  - Specifying the task environment
  - Properties of task environments
- THE STRUCTURE OF AGENTS
  - Agent programs
  - Simple reflex agents
  - Model-based reflex agents
  - Goal-based agents
  - Utility-based agents
  - Learning agents
  - How the components of agent programs work

Specifying the task environment

- In designing an agent, the first step must always be to specify the task environment as fully as possible.
- **Task environment: P**erformance, **E**nvironment, **A**ctuators, **S**ensors (**PEAS**).
- Example: PEAS description of the task environment for an automated taxi.
- Agent Type: Taxi Driver:
  - **P**erformance Measure: Safe, fast, legal, comfortable trip, maximize profits, …
  - **E**nvironment: Roads, other traffic, pedestrians, customers
  - **A**ctuators: Steering, accelerator, brake, signal, horn, display
  - **S**ensors: Cameras, sonar, speedometer, GPS, odometer, accelerometer, engine sensors, keyboard/microphone

PEAS for additional agent types

Agent Type	Performance Measure	Environment	Actuators	Sensors
Medical diagnosis system	Healthy patient, reduced costs	Patient, hospital, staff	Display of questions, tests, diagnoses, treatments, referrals	Keyboard entry of symptoms, findings, patient's answers
Satellite image analysis system	Correct image categorization	Downlink from orbiting satellite	Display of scene categorization	Color pixel arrays
Part-picking robot	Percentage of parts in correct bins	Conveyor belt with parts; bins	Jointed arm and hand	Camera, joint angle sensors
Refinery controller	Purity, yield, safety	Refinery, operators	Valves, pumps, heaters, displays	Temperature, pressure, chemical sensors
Interactive English tutor	Student's score on test	Set of students, testing agency	Display of exercises, suggestions, corrections	Keyboard entry

Figure 2.5 Examples of agent types and their PEAS descriptions.

Properties of task environments

- The range of task environments that might arise in AI is vast.
- We can identify a fairly small number of dimensions along which task environments can be categorized.

Fully observable vs. partially observable

- Do the agent's sensors give it access to the complete state of the environment?
  - For any given world state, are the values of all the variables known to the agent?

vs.

Source: L. Zettlemoyer

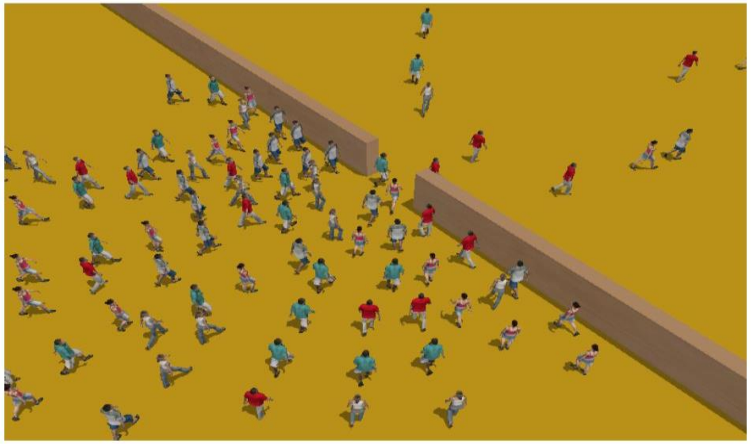
Fully observable vs. partially observable

- A task environment is **effectively fully observable** if the sensors detect all aspects that are *relevant* to the choice of action; relevance, in turn, depends on the performance measure.
- Fully observable environments are convenient because the agent need not maintain any internal state to keep track of the world.
- An environment might be **partially observable** because of noisy and inaccurate sensors **or** because parts of the state are simply missing from the sensor data.
  - a vacuum agent with only a local dirt sensor cannot tell whether there is dirt in other squares;
  - an automated taxi cannot see what other drivers are thinking.
- Agent has no sensors ➔ environment is **unobservable**

Single-agent vs. multi-agent

- Is an agent operating by itself in the environment?
- Examples:
  - an agent solving a crossword puzzle by itself is clearly in a single-agent environment,
  - an agent playing chess is in a two-agent environment.
  - Example of a Multi-agent system (2′15″ – 3′15″)

vs.



Agent or object?

- Does an agent A (the taxi driver for example) have to treat an object B (another vehicle) as an agent, or can it be treated just as an object behaving according to the laws of physics?
- **key distinction**: Is B's behavior best described as maximizing a performance measure whose value depends on agent A's behavior?
  - E.g., in chess, the opponent entity B is trying to maximize its performance measure, which, by the rules of chess, minimizes agent A's performance measure. Thus, chess is a **competitive** multi-agent environment.
  - In the taxi-driving environment, avoiding collisions maximizes the performance measure of all agents, so it is a **partially cooperative** multi-agent environment. Also **partially competitive** because, for example, only one car can occupy a parking space.

Deterministic vs. stochastic

- Is the next state of the environment completely determined by the **current state** and the **agent's action**?
  - Is the transition model **deterministic** (unique successor state given current state and action) or **stochastic** (distribution over successor states given current state and action)?
  - **strategic:** the environment is deterministic except for the actions of other agents

vs.

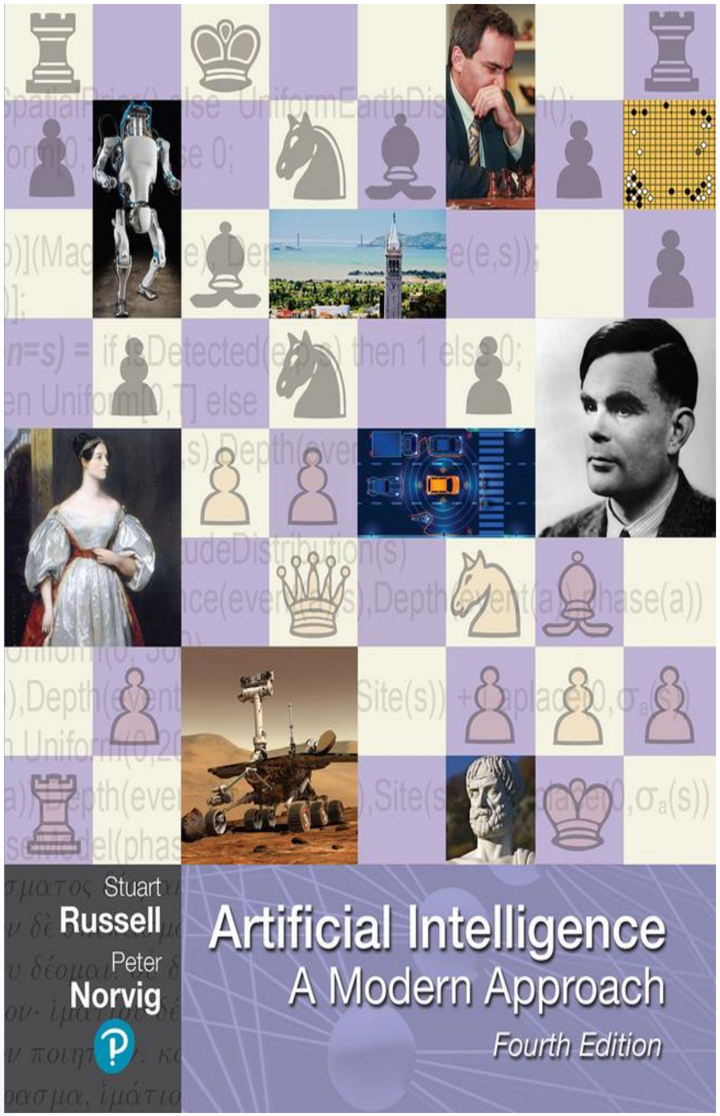
Deterministic vs. stochastic

- Most real situations are so complex that it is impossible to keep track of all the unobserved aspects; for practical purposes, they must be treated as stochastic.
- Taxi driving is clearly stochastic in this sense, because
  - One can never exactly predict the behavior of traffic;
  - Also, one's tires may blow out and one's engine may fail without warning.
- The vacuum world as described is deterministic, but variations can include stochastic elements such as randomly appearing dirt and an unreliable suction mechanism.

Deterministic vs. stochastic

- We say an environment is **uncertain** if it is not fully observable or not deterministic.
- Our use of the word "**stochastic**" generally implies that uncertainty about outcomes is quantified in terms of probabilities.
- A **nondeterministic** environment is one in which actions are characterized by their *possible* outcomes, but no probabilities are attached to them.
- For Nondeterministic environments, the used performance measures usually require the agent to succeed for *all possible* outcomes of its actions.

Slides based on the textbook



- Russel, S. and Norvig, P. (2020) Artificial Intelligence,

A Modern Approach (4th Edition), Pearson Education Limited.