Artificial Intelligence

2. Intelligent Agents

Let's Get a Little Orientation Here

Prof Sara Bernardini
bernardini@diag.uniroma1.it
www.sara-bernardini.com

SAPIENZA UNIVERSITÀ DI ROMA

Autumn Term

Agenda

1. Introduction
2. What is an Agent?
3. What is a Rational Agent?
4. Different Classes of Agents
5. Different Classes of Environments
6. Conclusion

Introduction

→ A central aspect of intelligence (and one possible way to define it) is the ability to act successfully in the world (cf. **Chapter 1**).

This chapter provides a broad categorization of and terminology for what we mean by "acting successfully".

- Useful for a research field to establish a common language.
- Useful for you to get a rough overview.

→ We'll think a bit about what it means to "behave rationally", what are possible architectures for achieving such behavior with a computer, and what are relevant properties of the world we need to act in.

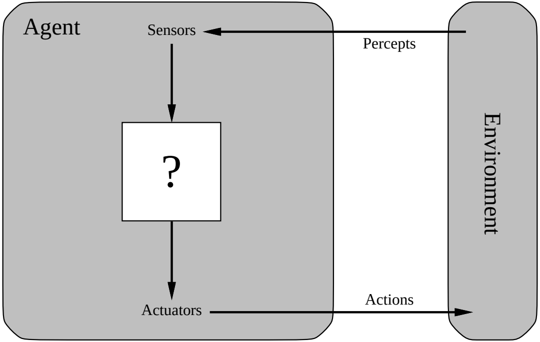
Disclaimer

→ The concepts and explanations in this chapter are very broad and rather superficial. In the remainder of the course, we will consider in detail particular problems relating to decision-making and specify algorithms for solving these.

What is an “Agent” in AI?

Agents:

- Perceive the environment through sensors ($\rightarrow$ percepts).
- Act upon the environment through actuators ($\rightarrow$ actions).



$\rightarrow$ Examples?

Rational Agents ...

... do "the right thing"!

$\rightarrow$ Meaning of "do the right thing": Rational agents select their actions so as to maximize a performance measure.

$\rightarrow$ What's the performance measure of an autonomous vacuum cleaner?

Actually, Rational Agents . . .

. . . ATTEMPT to do "the right thing"!

→ The hypothetical best case ("the right thing") is often unattainable.

→ The agent might not be equipped with the right action to perform. (Vacuum cleaner cannot step down the stairs.)
→ The agent might not be able to perceive all relevant information. (Is there dirt under this bed? Is this a hamster or a shoe?)

Rationality vs. Omniscience:

- An omniscient agent knows everything about the environment and knows the actual effects of its actions.
- A rational agent just makes the best of what it has at its disposal, *maximizing expected performance given its percepts, knowledge and available actions*.

→ Example?

So, What *Is* a Rational Agent?

Mapping input to the best possible output:

Performance measure $\times$ Percepts $\times$ Knowledge $\rightarrow$ Action

- An agent has a performance measure M and a set A of possible actions. Given a percept sequence P, as well as knowledge K about the world, it selects an action $a \in A$.
- The action a is optimal if it maximizes the expected value of M, given the evidence provided by P and K. The agent is rational if it always chooses an optimal a.

$\rightarrow$ If the vacuum cleaner bumps into the hamster, then this can be rational in case

$\rightarrow$ Note: If observation actions are required, they are elements of A, i.e., the agent must *perceive actively*. Example:

A Rational Agent is an Optimal Action Choice Function?

We also need to realize the agent through:

- an agent program, executed on
- an architecture which also provides an interface to the environment (percepts, actions).

$\rightarrow$ Agent = Architecture + Program

Practical limitations:

- Our definition captures limitations on percepts and knowledge.
- It does not capture computational limitations (often, determining an optimal choice would take too much time/memory).

  $\rightarrow$ In practice, we often merely *approximate* the rational decision.

Examples of Agents: PEAS Descriptions

Questionnaire

Question!

Which are agents?

(A): James Bond.
(B): Your dog.
(C): Vacuum cleaner.
(D): Thermometer.

Question!

Who is rational?

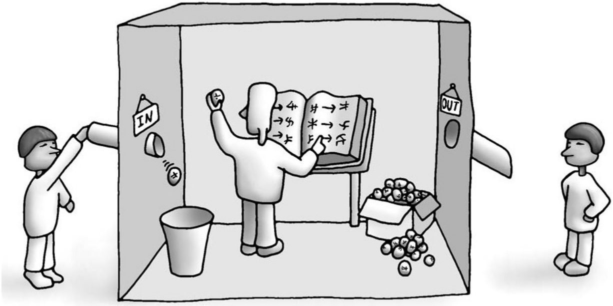
(A): James Bond, crossing the street without looking.
(B): Your dog, crossing the street without looking.
(C): Vacuum cleaner, deciding to clean under your bed.
(D): Thermostat, deciding to cool down your fridge.

Questionnaire Answers

Table-Driven Agents

function TABLE-DRIVEN-AGENT(*percept*) **returns** an action
  persistent: *percepts*, a sequence, initially empty
    table, a table of actions, indexed by percept sequences, initially fully specified

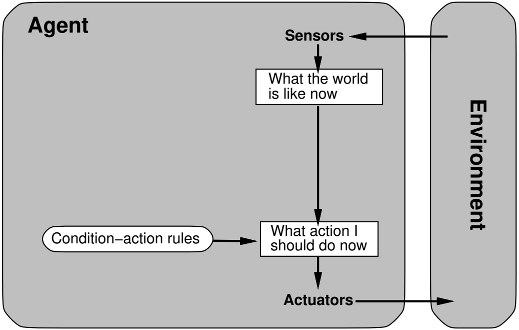
  append *percept* to the end of *percepts*
  action ← LOOKUP(*percepts*, *table*)
  return *action*



Reflex Agents

More useful, but still very simple, method for choosing actions:

Condition-Action Rules: mappings between current percept and actions



→ Example?

What happens if the sensors do not always work?

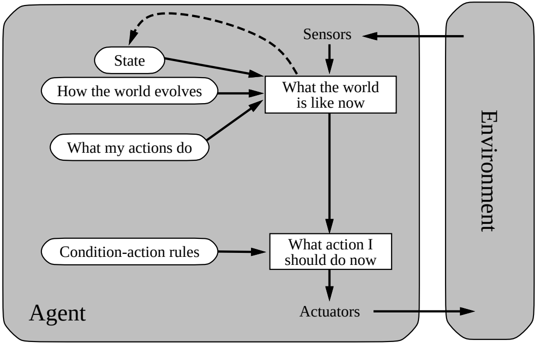
Reflex Model-based Agents

To handle failures in observing the environments, the agent needs to keep track of the part of the world it can't see now. How?

→ Internal state: depends on the percept history

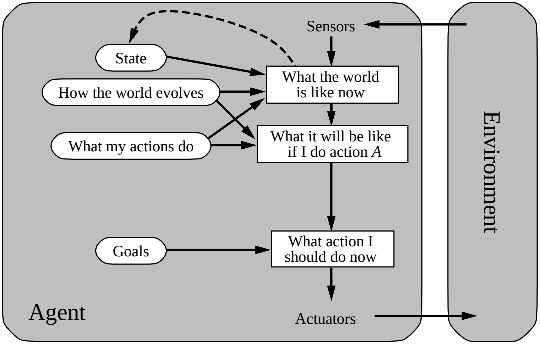
How to update the internal state? Model of the world:

- How the world evolves independently of the agent;
- How the agent's own actions affect the world.



Goal-Based Agents (Belief-Desire-Intention)

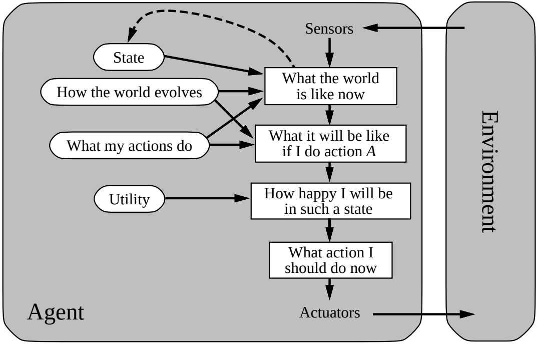
Often, doing the right thing requires considering the future:



→ Example?

Utility-Based Agents

Often, doing the right thing requires trading-off costs and benefits:

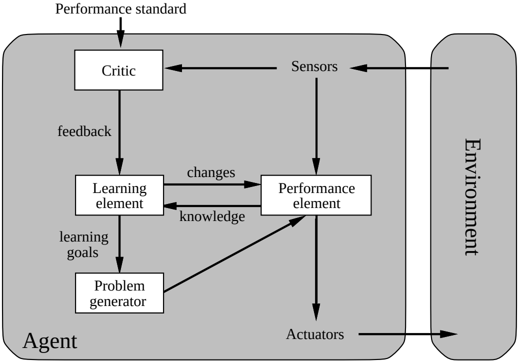


→ A utility function maps a state to a number which represents how desirable the state is.
Agent's utility function is an internalization of the performance function.

→ Example?

Learning Agents

It is typically useful to be able to learn: (extreme case: start with no knowledge in an unknown environment)



→ Performance element: Selects actions (exploitation). Learning element: Learns new knowledge. Critic: Measures performance. Problem generator: Suggests actions favoring informative learning experiences (exploration).

Domain-Specific vs. General Agents

Duell Kasparow gegen Deep Blue (1997): *Demütigende Niederlage*

vs.

Solver specific to a particular problem ("domain")	**vs.**	Solver based on *description* in a general problem-description language (e.g., the rules of any board game)
More efficient	**vs.**	More *intelligent*

Questionnaire

Question!

What kind of agent are you?

(A): Table-Driven

(B): Reflex Agent

(C): Learning

(D): Domain-Specific

The Environment of Rational Agents

- Fully observable vs. partially observable (Accessible vs. inaccessible)
  Are the relevant aspects of the environment accessible to the sensors?

- Deterministic vs. stochastic
  Is the next state of the environment completely determined by the current state and the selected action?
  Stochastic: uncertainty is quantified by using probabilities.
  Nondeterministic: uncertainty as actions with multiple outcomes.
  Strategic: If the only non-determinism are actions of other agents.

- Episodic vs. sequential
  Can the quality of an action be evaluated within an episode (perception + action), or are future developments decisive?

The Environment of Rational Agents, ctd.

- Static vs. dynamic
  Can the environment change while the agent is deliberating?
  If the environment does not change, but the agent's performance score changes, the environment is called semi-dynamic.

- Discrete vs. continuous
  Is the environment discrete or continuous?
  Distinction applies to time, environment states and agent's actions and percepts.

- Single agent vs. multi-agent
  Is there just one agent, or several of them?
  There are competitive and cooperative multi-agent scenarios.

Examples of Environments

Classifying AI Areas

Many sub-areas of AI can be classified by:

- Domain-specific vs. general.
- The environment.
- (Particular agent architectures sometimes also play a role, especially in Robotics.)

→ The same is true of the sub-topics in this course. The focus is on general methods (a bias in much of the AI field) and simple environments (after all, it's an introductory course only).

→ Up next: A rough classification of our topics, in these terms.

Classifying AI Areas: Our Topics

Classical Search

Environment:

- Fully observable vs. partially observable.
- Deterministic vs. stochastic.
- Episodic vs. sequential.
- Static vs. dynamic.
- Discrete vs. continuous.
- Single-agent vs. multi-agent.

Approach:

- Domain-specific vs. general.

Classifying AI Areas: Our Topics

Planning

Environment:

- Fully observable vs. partially observable.
- Deterministic vs. stochastic.
- Episodic vs. sequential.
- Static vs. dynamic.
- Discrete vs. continuous.
- Single-agent vs. multi-agent.

Approach:

- Domain-specific vs. general.

→ Planning formalisms and approaches exist also for any and all of partial observability, and stochastic/dynamic/continuous/multi-agent settings.

Classifying AI Areas: Our Topics

Adversarial Search

Environment:

- Fully observable vs. partially observable.
- Deterministic vs. stochastic.
- Episodic vs. sequential.
- Static vs. dynamic.
- Discrete vs. continuous.
- Single-agent vs. multi-agent.

Approach:

- Domain-specific vs. general.

→ Adversarial search formalisms and approaches exist also for partial observability and stochastic settings.

Classifying AI Areas: Our Topics

General Game Playing

Environment:

- Fully observable vs. partially observable.
- Deterministic vs. stochastic.
- Episodic vs. sequential.
- Static vs. dynamic.
- Discrete vs. continuous.
- Single-agent vs. multi-agent.

Approach:

- Domain-specific vs. general.

→ General game playing formalisms and approaches exist also for partial observability and stochastic settings.

Classifying AI Areas: Our Topics

Constraint Satisfaction & Reasoning

Environment:

- Fully observable vs. partially observable.
- Deterministic vs. stochastic.
- Episodic vs. sequential.
- Static vs. dynamic.
- Discrete vs. continuous.
- Single-agent vs. multi-agent.

Approach:

- Domain-specific vs. general.

Classifying AI Areas: Our Topics

Probabilistic Reasoning

Environment:

- Fully observable vs. partially observable.
- Deterministic vs. stochastic.
- Episodic vs. sequential.
- Static vs. dynamic.
- Discrete vs. continuous.
- Single-agent vs. multi-agent.

Approach:

- Domain-specific vs. general.

Questionnaire

Question!

James Bond's environment is?

(A): Fully Observable. (B): Episodic.

(C): Static. (D): Single-Agent.

Question!

Your own environment is?

(A): Fully Observable. (B): Episodic.

(C): Static. (D): Single-Agent.

Questionnaire Answers

Summary

- An agent is something that perceives and acts. It consists of an architecture and an agent program.
- A rational agent always takes the action that maximizes its expected performance, subject to the percept sequence and its environment knowledge.
- There are a variety of agent designs:
  - Reflex agents respond to percepts by condition-action rules.
  - Reflex model-based agents maintain a model of the world.
  - Goal-based agents work towards goals.
  - Utility-based agents make trade-offs using a utility function.
  - Learning agents improve their behavior over time.
- Some environments are more demanding than others ...

  ... your own, and that of James Bond, are the most difficult.

Reading

- *Chapter 2: Intelligent Agents* [Russell and Norvig (2020)].

  Content: A much more detailed description of the subjects I overviewed here (agents, agent architectures, environments). Just as broad and superficial, though.

References I

Stuart Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach (Fourth Edition)*. Prentice-Hall, Englewood Cliffs, NJ, 2020.