agents

Daniel Gigliotti

1 Rational agents

An agent is an entity that perceives and acts; a rational agent, on the other hand, is a person or entity that always aims to perform optimal actions based on given premises and information.

2 Agent and environment

An AI system can be defined as the study of a rational agent and its environment. A rational agent can be anything that makes decisions, typically a person, robot or software. An agent can be seen as a black box interacting with the environment in two ways:

- takes input from the environment (perceives)
- produces an output (acts)

The agents sense the environment through sensors and act on their environment through actuators.

How can we implement this box? How do the system will convert perceptions into actions?

3 Agent architecture

Abstractly, an agent can be seen as a function that maps from percept histories to actions:

$$f : P^* \to A \tag{1}$$

A *percept* is a sequence of *perceptions*: maybe previous perceptions could help us in the future.

For a given task we seek the agent (or class of agents) with the best performance: it aims to choose whichever action maximizes the expected value of the performance measure given the percept sequence to date. Computational limitations make perfect rationality unachievable.

4 Agent types

Four basic types in order of increasing generality:

- simple reflex agents
- model-based reflex agents
- goal-based agents
- utility-based agents

All these can be turned into learning agents. A learning agent is an agent with the capability of learning from its previous experience. What is it needed to talk about learning agents?

- *Learning element*: element that enables learning from previous experience;
- *Critic*: provides feedback on how well the agent is doing concerning a fixed performance standard;
- *Performance element*: the actions to be performed are selected;
- *Problem generator*: acts as a feedback agent that performs certain tasks such as making suggestions that will lead to new and informative experiences;