Introduction to Artificial Intelligence

Informatics Engineering Study Program
School of Electrical Engineering and Informatics

Institute of Technology Bandung

Contents

- Intelligent Agent
  - Introduction
  - Agents
  - Rationality
  - Thinking
  - Learning
  - Planning

Introduction

What is AI? (Review)

1. Computational models of human behaviour? (Acting humanly)
   - Programs that behave (externally) like humans
2. Computational models of human "thought" processes? (Thinking humanly)
   - Programs that operate (internally) the way humans do
3. Computational systems that behave intelligently? (Thinking rationally)
   - What does it mean to behave intelligently?
4. Computational systems that behave rationally! (Acting rationally)
   - Later

- AI Applications: monitor …, detect …, schedule …

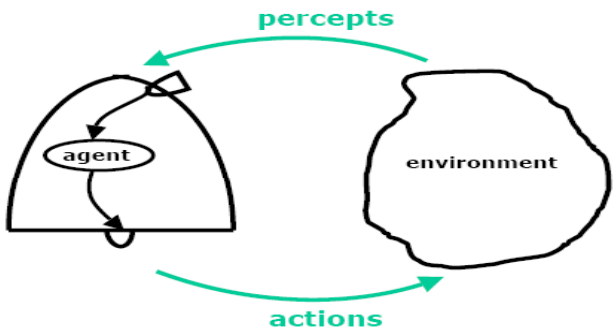
Agents

- Anything that can be viewed as perceiving its environment through sensors and acting upon that environment through effectors.
  - A robot
  - A factory
  - A web shopping program
  - …

⇒ Computational agents that behave autonomously

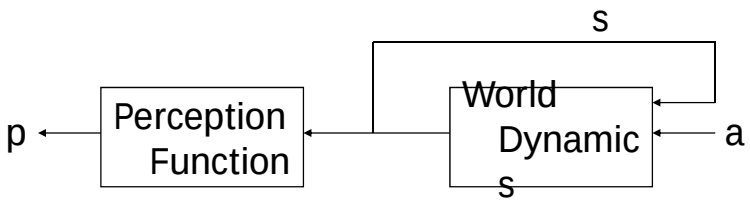
The Agent & the Environment

- How do we begin to formalize the problem of building an agent?
  - Make a dichotomy between the agent and the environment



World Model

- A – the action space
- P – the percept space
- Define:
  - S – internal state [may not be visible to agent]
  - Perception function: S → P
  - World dynamics: S x A → S



Agent Design

- U – utility function: S → real (or S* → real)
- The agent design problem: Find P* → A
  - mapping of sequences of percepts to actions
  - maximize the utility of the resulting sequences of states (each action maps from one state to next state)

Rationality

- A rational agent takes actions it believes will achieve its goals.
  - Assume I don't like to get wet, so I bring an umbrella. Is that rational?
  - Depends on the weather forecast and whether I've heard it. If I've heard the forecast for rain (and I believe it) then bringing the umbrella is rational.
- Rationality ≠ omniscience
  - Assume the most recent forecast is for rain but I did not listen to it and I did not bring my umbrella. Is that rational?
  - Yes, since I did not know about the recent forecast!
- Rationality ≠ success
  - Suppose the forecast is for no rain but I bring my umbrella and I use it to defend myself against an attack. Is that rational?
  - No, although successful, it was done for the wrong reason.

Limited Rationality

- There is a big problem with our definition of rationality …
- The agent might not be able to compute the best action (subject to its beliefs and goals).
- So, we want to use limited rationality: "acting in the best way you can subject to the computational constraints that you have".

Limited Rationality (2)

- The (limited rational) agent design problem:

  Find P* → A
  - mapping of sequences of percepts to actions
  - maximizes the utility of the resulting sequences of states
  - subject to our computational constraints

Issues

- How could we possibly specify completely the domain the agent is going to work in?
  - If you expect a problem to be solved, you have to say what the problem is!
  - Specification is usually iterative: Build agent, test, modify specification
- We're going to map classes of environments and utilities to structures of programs that solve that class of problem.

Thinking

- Is all this off-line work AI?
- Aren't the agents supposed to think?
- Why is it ever useful to think?
- If you can be endowed with an optimal table of reactions/reflexes ($P^* \rightarrow A$) why do you need to think?
- The table is too big! There are too many world states and too many sequences of percepts.
- In other domains, we'll take advantage of the fact that most things that could happen – don't. There's no reason to precompute reactions to an elephant flying in the window.

Learning

- What if you don't know much about the environment when you start or if the environment changes?
  - Learn!
  - We're sending a robot to Mars but we don't know the coefficient of friction of the dust on the Martian surface.
  - I know a lot about the world dynamics but I have to leave a free parameter representing this coefficient of friction.
- Part of the agent's job is to use sequences of percepts to estimate the missing details in the world dynamics.
- Learning is not very different from perception, they both find out about the world based on experience.
  - Perception = short time scale (where am I?)
  - Learning = long time scale (what's the coefficient of friction?)

THANK YOU