TÉCNICO LISBOA

Properties of agent systems

Perspectives on Agents and Environments

Outline

- **What is an agent?**
- Intelligent agents and agent properties
- Environment properties
- Agent applications

What is an Agent?

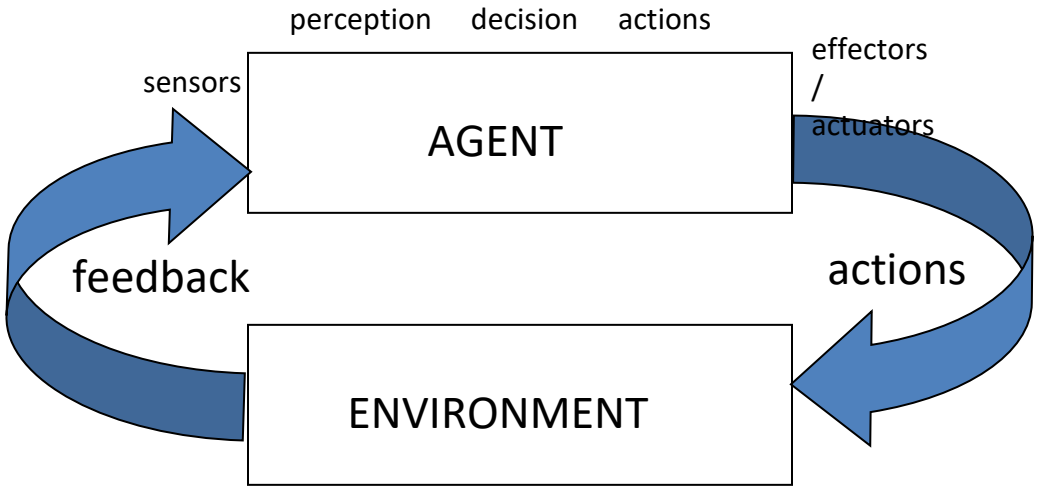
- Remember *Autonomy* is a key point about agents
  - In other words, they are capable of *acting independently*, exhibiting *control over their internal state*

- Thus: *an* agent *is a computer system capable of* autonomous action *in some environment in order to meet its* design objectives

What is an Agent?

- We think of an agent as being in a closed-couple continual interaction with its environment:

sense – decide – act – sense – decide – act…



Outline

- What is an agent?
- **Intelligent agents and agent properties**
- Environment properties
- Agent applications

Simple (Uninteresting) Agents

- Thermostat
  - Goal to *maintain room temperature*
  - Goal delegated by user
  - Actions: heat on / heat off
  - Temperature sensor

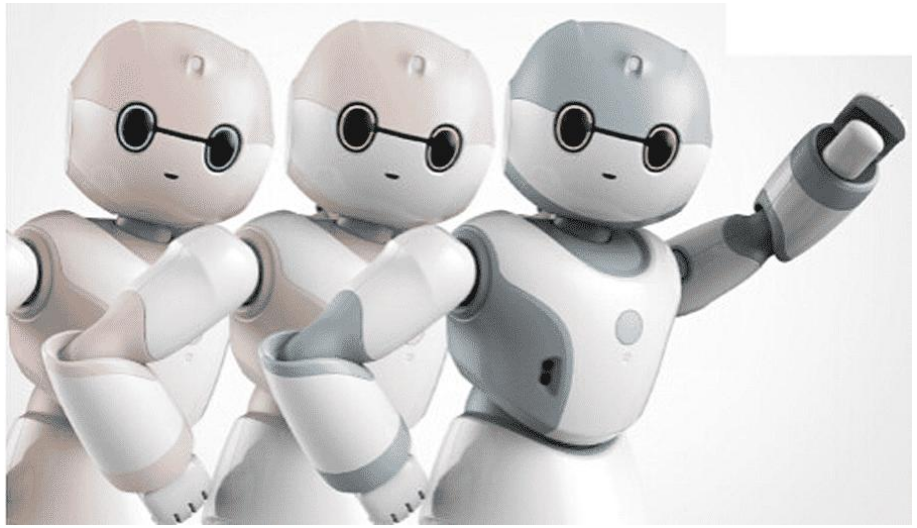
Simple (Uninteresting) Agents

- Screen saver
  - Software that *blanks the screen* (or presents moving images) *after a period of user inactivity*
  - Goal avoid screen burning and to prevent other people from viewing desktop contents while the user is away
  - Actions: blank the screen / show desktop

- Why are they uninteresting?
  - They are *trivial* from a *decision-making perspective*

Intelligent Agents

- Intelligent agents typically exhibit the following types of behavior:
  - *reactive*
  - *pro-active*
  - *social*



Reactivity

- What happens if a program's *environment is guaranteed to be fixed*?
  - a program can just execute blindly

- However, the real world is not like that
  - Things change, information is incomplete.
  - Many (most?) interesting environments are *dynamic*

Reactivity

- Software is hard to build for dynamic domains
  - E.g., program must *take into account possibility of failure* – ask itself whether it is worth executing!

- A *reactive* system is one that maintains an *ongoing interaction with its environment*, and *responds to changes* that occur in it (in time for the response to be useful)

Proactiveness

- Reacting to an environment is relatively easy to implement
  - E.g., stimulus → response rules
- But we generally want agents to *do things for us*
  - Hence having *goal directed behavior*
- Pro-activeness = generating and attempting to achieve goals
  - Not driven solely by events from the environment
  - Taking the initiative
  - Recognizing opportunities

Balancing Reactive and Goal-Oriented Behavior

- We want our *agents to be reactive*
  - Thus, responding to changing conditions in an appropriate (timely) fashion
- We want our *agents to systematically work towards long-term goals*
- However, these *two considerations can be at odds with one another*
- Designing an agent that can *balance the two remains an open research problem*

Social Ability

- The real world is a *multi*-agent environment
  - Hence, we cannot go around attempting to achieve goals *without taking others into account*

- Some *goals can only be achieved with the cooperation of others*

Social Ability

- *Social ability* in agents is the ability to interact with other agents (and possibly humans)
  - through *coordination*, *cooperation, negotiation* and *modeling others*

Social Ability

- *Coordination* is managing the interdependencies between activities

- For example:
  - A non-sharable resource in the environment
  - Agents need to coordinate to use this resource

Social Ability

- *Cooperation* is working together as a team to achieve a shared goal

- Typically needed when:
  - No agent can achieve the goal alone
  - Cooperation obtains a better result

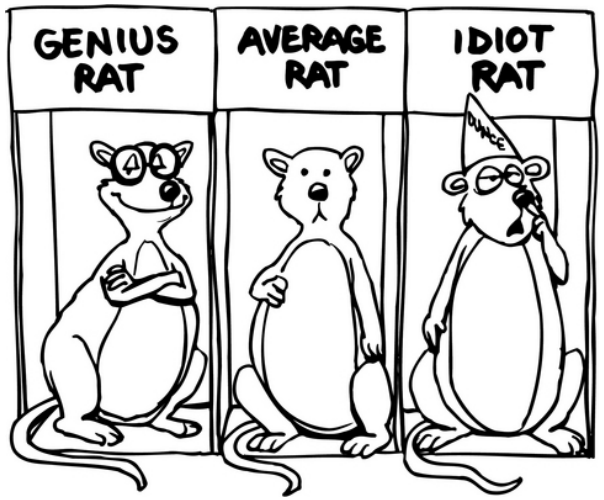
Social Ability

- *Negotiation* is the ability to reach agreements on matter of common interest

- For example:
  - You have a TV in your house
  - You want to watch a movie
  - Your roommate wants to watch football
  - Deal: He watches football tonight and you watch a movie tomorrow

Other Agent Properties

- Are there other properties that characterize agents?
  - **Autonomy**
  - Adaptivity
  - Rationality
  - Curiosity
  - Believability
  - Mobility



Agent properties

Autonomy

- agent's *ability to act independently* and thus determine how to achieve its delegated goals/tasks

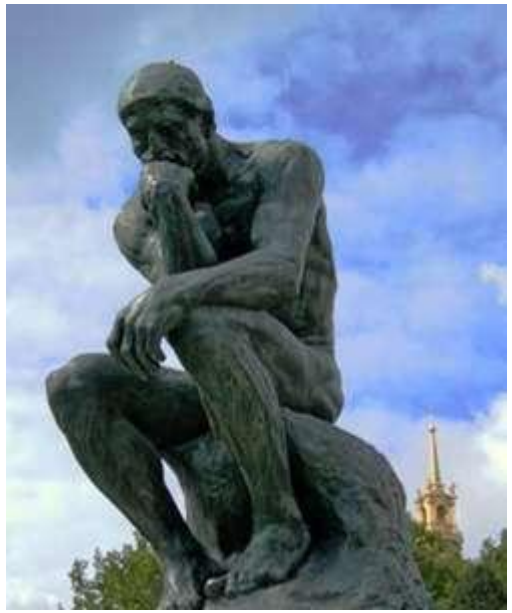
Adaptivity

- agent's *ability to learn from experience* (to better interact with the particular environment)

Agent properties

Rationality

- agent's *ability to act in a way that maximizes some utility function*



Curiosity

- Agent's *ability to engage creative imaginative or inquisitive reasoning*

Agent properties

Believability

- agent's *ability to create a suspension of disbelief*, temporarily leading a user to accept the agent as an alive or a real character
  - e.g., characters in game and films

Mobility

- agent's *ability to change its location in the environment*, being it the *physical* world (robots), or the *virtual* worl
  - e.g., virtual agents, Internet network

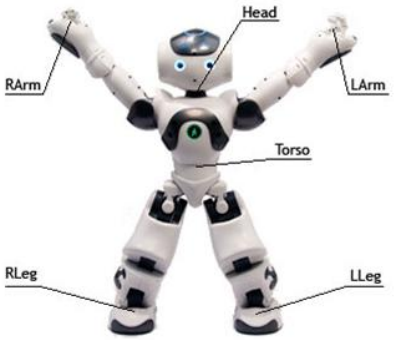
Outline

- What is an agent?
- Intelligent agents and agent properties
- **Environment properties**
- Agent applications

Environment

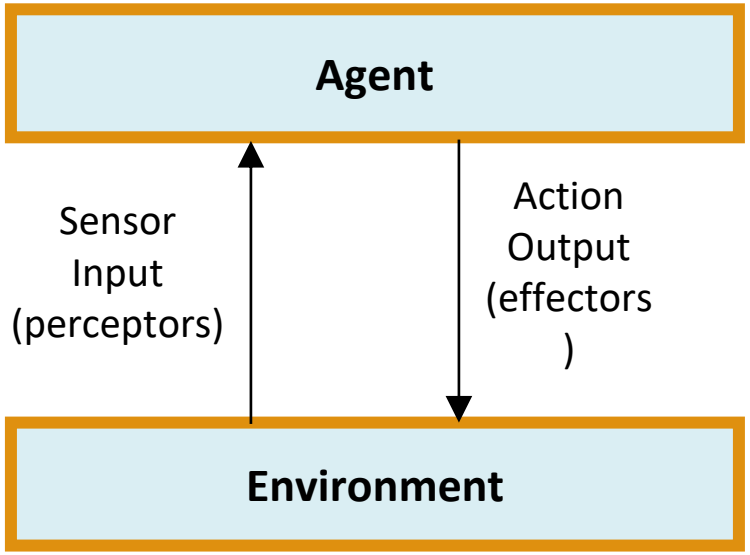
Agent systems further characterized by *properties of the surrounding environment*

Sensors: define the *perceptors* for the agent perceive the world



Effectors: define the *actuators* for the agent perform in the world

Interaction with environment

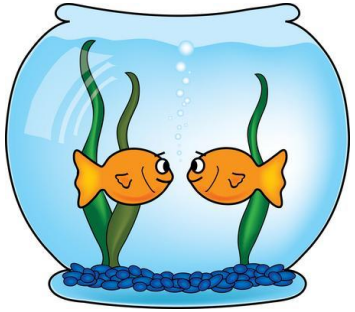


- *sensors*: cameras, speed sensor, gps, sonar
- *actuators*: start engine, accelerate, break, turn, start light, change gears

Environment properties

Environment properties

- accessibility
- determinism
- dynamism
- continuity
- memory



Agents…

- typically have *only partial access and control over the environment*
- have to *decide what action to execute* in order to attain goals

Environment properties

Accessible vs. Inaccessible

- accessible environment: *agent can obtain complete, accurate, up-to-date data about the environment's state*
- most moderately complex environments are inaccessible

Environment properties

Deterministic vs. non-deterministic

- deterministic environment: *action has a single guaranteed effect*
- physical world is for us (humans) non-deterministic

Environment properties

Static vs. dynamic

- static environment: *world does not change while the agent is deliberating*


Discrete vs. continuous

- discrete environment: *fixed, finite number of possible actions and percepts*


Environment properties

Episodic vs. non-episodic

- Episodic environment: *world can be divided in a series of intervals (episodes) independent of each other*
- What happens in one episode has no influence on others

Outline

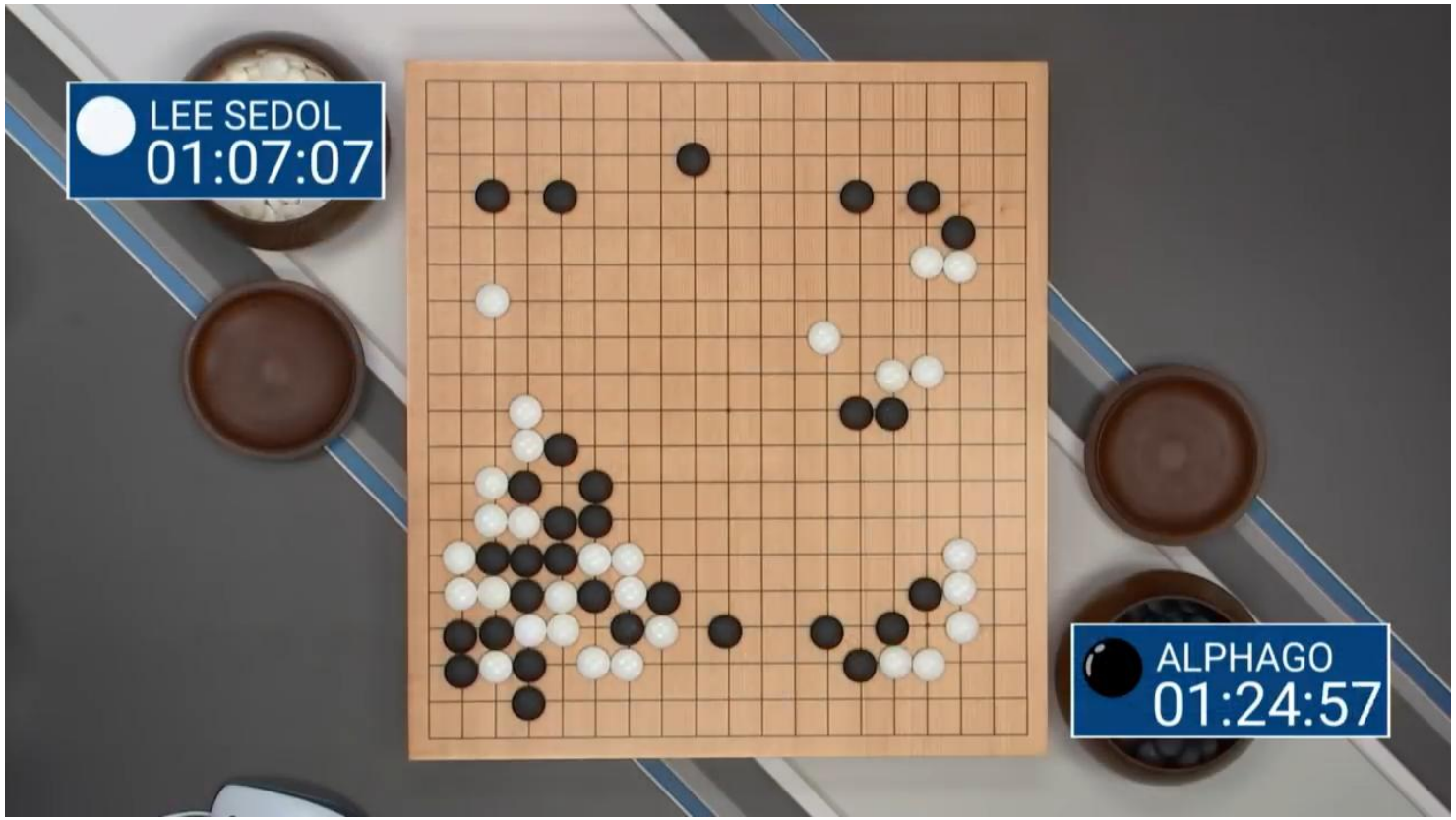
- What is an agent?
- Intelligent agents and agent properties
- Environment properties
- **Agent applications**

Agent applications

- Agent systems can exist in
  - *physical space*: robots
  - *cyberspace*: software and interface agents
  - *simulated physical space*: traffic simulators
  - *hybrid*: virtual agents interacting with humans

Agents in Games

AlphaGO



https://www.youtube.com/watch?v=8tq1C8spV_g

LLM/LWM ChatBots: ChatGPT, Gemini, DeepSeek

Thank You

rui.prada@tecnico.ulisboa.pt