

## Agent-Based Systems

Michael Rovatsos  
mrovatso@inf.ed.ac.uk

Lecture 1 – Introduction

1 / 20

## Five pervasive trends in computing history

- Ubiquity
  - Cost of processing power decreases dramatically (e.g. Moore's Law), computers used everywhere
- Interconnection
  - Formerly only user-computer interaction, nowadays distributed/networked systems (Internet etc.)
- Complexity
  - Elaboration of tasks carried out by computers has grown
- Delegation
  - Giving control to computers even in safety-critical tasks (aircraft/nuclear plant control)
- Human-orientation
  - Increasing use of metaphors that better reflect human intuition from everyday life (e.g. GUIs, speech recognition, object orientation)

2 / 20

## New challenges for computer systems

- Traditional design problem:  
How can I build a system that produces the correct output given some input?
- Modern-day design problem:  
How can I build a system that can operate independently on my behalf in a networked, distributed, large-scale environment in which it will need to interact with different other components pertaining to other users?
- In particular, distributed systems in which different components have different goals and need to cooperate have not been studied until recently

3 / 20

## Multiagent systems

- Two fundamental ideas:
  - Individual **agents** are capable of autonomous action to a certain extent (they don't need to be told exactly what to do)
  - These agents interact with each other in **multiagent systems** (and which may represent users with different goals)
- Foundational problems of multiagent systems (MAS) research:
  - 1 The **agent design** problem: how should agents act to carry out their tasks?
  - 2 The **society design** problem: how should agents interact to carry out their tasks?
- These are known as the **micro** and **macro** perspective of MAS

4 / 20

## A pure engineering task?

- Like AI (which aims to improve our understanding of human intelligence) MAS has a “deeper” goal:  
To understand how societies of intelligent beings work
- A list of questions related to this:
  - How can cooperation emerge among self-interested agent?
  - How can agents coordinate their activities with those of others?
  - What languages should agents use to exchange information necessary to organise interaction in a meaningful way?
  - How should agents resolve their conflicts?
  - How do we detect and deal with agents violating social rules?
- Philosophically speaking, MAS research marks departure from traditional engineering view:  
**control** is replaced by **communication**

5 / 20

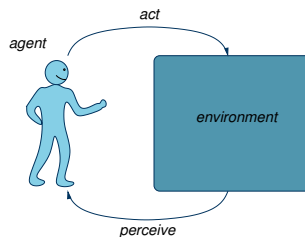
## Some applications of multiagent systems

- Agents have been applied to various application areas
- Broadly speaking, two areas:
  - Distributed systems (processing nodes)
  - Personal software assistants (aiding a user)
- Many areas:
  - Workflow/business process management
  - Distributed sensing
  - Information retrieval and management
  - Electronic commerce
  - Human-computer interfaces
  - Virtual environments
  - Social simulation
  - ...

6 / 20

## What is an agent? – Definition 1

- Most widely accepted definition:  
An agent is anything that can perceive its environment (through its sensors) and act upon that environment (through its effectors)



- Focus on **situatedness** in the environment (**embodiment**)
- Generally speaking, the agent can only influence the environment but not fully control it (sensor/effector failure, non-determinism)

7 / 20

## What is an agent? – Definition 2

- Definition from the agents/MAS area (Wooldridge & Jennings):  
An agent is a computer system that is **situated** in some **environment**, and that is capable of **autonomous action** in this environment in order to meet its **design objectives**
- This adds a second dimension to agent definition: the relationship between agent and designer/user
  - Agent is capable of independent action
  - Agent action is purposeful
- There is a broad consensus that **autonomy** is a central, distinguishing property of agents
- Alas, it is the one that is most disputed . . .

8 / 20

## Agent Autonomy

- Here is an autonomous device, situated in an environment, and purposeful:



- Would we call it an agent?

9 / 20

## Classification of environments

- **Accessible vs. inaccessible**
  - Can agents obtain complete and correct information about the state of the world?
- **Deterministic vs. non-deterministic**
  - Do actions have guaranteed and uniquely defined effects?
- **Static vs. dynamic**
  - Does the environment change by processes beyond agent control?
- **Episodic vs. non-episodic**
  - Can agents decisions be made for different, independent episodes?
- **Discrete vs. continuous**
  - Is the number of actions and percepts fixed and finite?
- **Open environments** = inaccessible, non-deterministic, dynamic, continuous environments

11 / 20

## Agent Autonomy

- Autonomy is a prerequisite for
  - 1 delegating complex tasks to agents
  - 2 ensuring flexible action in unpredictable environments
- Different definitions highlight different aspects
- A system is autonomous . . .
  - if it requires little help from the human user
  - if we don't have to tell it what to do step by step
  - if it can choose its own goal and the way to achieve it
  - if its behaviour is determined by its own experience
  - if we don't understand its internal workings
- **Autonomy dilemma:** how to make the agent smart without losing control over it

10 / 20

## Intelligent agents

- The above definitions give us some basic properties of agents, but don't say anything about **intelligent** agents
- We are not looking for a general definition of agency, but for practical criteria that matter in the target application scenarios
- Again, the answer is not easy, desirable properties can be listed:
  - **Reactivity:** intelligent agents should respond in a timely fashion to changes they perceive in their environment
  - **Proactiveness:** intelligent agents can take the initiative to meet their design objectives, and they exhibit goal-directed behaviour
  - **Social ability:** intelligent agents can interact with other agents (and humans) to satisfy their design objectives

12 / 20

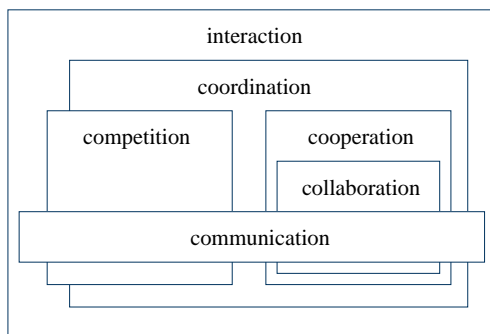
## Rationality = proactiveness + reactivity

- Example: The dung beetle
  - After digging its nest and laying its eggs, it fetches a ball of dung from a nearby heap to plug the entrance; if the ball of dung is removed from its grasp en route, the beetle continues on and pantomimes plugging the nest with the nonexistent dung ball, never noticing that it is missing (quoted from Russell & Norvig)
- Truly flexible autonomous behaviour is hard to achieve!
- Trade-off between the two aspects because:
  - Environments are not fixed → must be able to react to changes (involves monitoring own activity and environment, etc.)
  - Need for goal-oriented, planned activity → not sufficient to respond to current circumstances

13 / 20

## Social Ability

- Things to note:
  - Interaction does not always imply action
  - Coordination does not always imply communication
- Basic typology of interaction:



15 / 20

## Social Ability

- In most real-world applications, environments are inhabited by multiple agents
- Each agent has limited resources/capabilities, some goals may require others (not) to take action
- Social ability is the ability to manage one's interactions effectively (different from simple exchange of messages between computer programs)
- Interaction and coordination:
  - An interaction can be viewed as a formalisation of a concept of dependence between agents, no matter on whom or how they are dependent. Coordination is a special case of interaction in which agents are aware how they depend on other agents and attempt to adjust their actions appropriately.

14 / 20

## What is agent technology?

- Agents as a software engineering paradigm
  - Interaction is most important aspect of complex software systems
  - Ideal for loosely coupled "black-box" components
- Agents as a tool for understanding human societies
  - Human society is very complex, computer simulation can be useful
  - This has given rise to the field of (agent-based) **social simulation**
- Agents vs. distributed systems
  - Long tradition of distributed systems research
  - But MAS are not simply distributed systems, because of different goals
- Agents vs. economics/game theory
  - Distributed rational decision making extensively studied in economics, game theory very popular
  - Many strengths but also objections

16 / 20

## Agents vs. AI

- Agents grew out of “distributed” AI
- Much debate whether MAS is a sub-field of AI or vice versa
- AI is mostly concerned with the building blocks of intelligence
  - reasoning and problem-solving, planning and learning, perception and action
- The agents field is more concerned with
  - Combining these components (this may mean we have to solve *all* problems of AI, but agents can also be built without any AI)
  - Social interaction, which has mostly been ignored by standard AI (and is an important part of human intelligence)
- Agents are a lot about integration (of abilities in one agent or of agents in one environment)

17 / 20

## Overview of the course

- Intelligent autonomous agents
  - Abstract agent architectures
  - Deductive reasoning agents
  - Practical reasoning agents
  - Reactive and hybrid agent architectures
- Communication and cooperation
  - Agent communication
  - Methods for coordination
- Multiagent decision making
  - Multiagent interactions
  - Social choice
  - Coalition formation
  - Resource allocation
  - Bargaining
  - Argumentation in multiagent systems
  - Logics for multiagent systems

19 / 20

## How are agents different?

- Agents vs. objects
  - Objects exhibit control over their state but not over their behaviour (limited sense of autonomy)
  - Objects do it for free, agents do it because they want to (or for money)
- Agents vs. expert systems
  - Expert systems are knowledge-based systems capable of problem solving in rich, complex domains
  - They can be intelligent, but they lack situatedness and usually don't cooperate with each other
- Agents vs. intentional systems
  - “The intentional stance”: ascribing mental attitudes to machines (beliefs, intentions, goals etc)
  - Not necessary when simpler model is available but may aid human understanding

18 / 20

## Summary

- Trends in computing, new challenges
- Fundamental issues of MAS research
- Relationships to other fields of research
- Autonomy: a difficult notion
- Environments for agents
- Advanced properties of intelligent agents
- What agents are and are not
- Next time: **Abstract Agent Architectures**

20 / 20