Multi-Agent Systems

Nov-Dec 2023

Lecturer: Eric Pauwels

eric.pauwels@cwi.nl

Overview of Today's Lecture

1. Course Organisation
2. Agents and Multi-Agent Systems
3. Agent Types
4. Environments
5. Topics discussed in this course

Practical Information

- General:
  - Period: Nov-Dec 2023
  - 6 ECTS
- Teachers:
  - Lecturer: Eric Pauwels (eric.pauwels@cwi.nl)
  - TAs: Andrzej Szczepura, Abdallah Al-Janabi, Yassin Ben Allal
- Recommended prior knowledge:
  - Elementary calculus and probability, basic programming
  - Schaum Outline Series (pdf online)

Contact Hours

- Period:
  - Nov - Dec (7 weeks)
- Plenary lectures:
  - Tuesday: 15.30 - 17.15
  - Thursday: 11.00 - 12.45
- Lab sessions:
  - Thu: 13:30-15:15, 15:30 - 17:15 and 17:30 - 19.15
- Final exam:
  - Tuesday, Dec 19 (18:30)   Emergo !!

Lab sessions & Homework assignments

- Week 2-6: Report on *homework assignment*
  - *Grade: Pass or Fail: (2 Fails = 0/1)*
  - *Deadline: next lab session;*
  - *Groups of 4 students: single report*
- *Week 7 : Individual homework, graded (4/4)*
  - *Deadline: to be announced;*
- Final exam: 5/5
- Total score: 10/10

Course Material

- Recommended Books: .
  - Y. Shoham & K. Leyton-Brown:
    MultiAgent Systems, Cambridge

    Online: http://www.masfoundations.org/mas.pdf
  - R. S. Sutton & A.G. Barto:
    Reinforcement Learning, MIT Press (2nd edition)
    - http://incompleteideas.net/book/bookdraft2017nov5.pdf
  - William Spaniel: Game Theory 101
    - (book + YouTube channel)

  You only need to know what was covered during lectures!

Course Material: Optional

N. Nisan, T. Roughgarden, E. Tardos, V. Vazirani:

Algorithmic Game Theory. Cambridge UP.

Solid, mathematical. Advanced.

A. Dixit, B. Nalebuff: Thinking Strategically. Norton.

Lots of context and background. Interesting and non-technical.

D. Fudenberg and J. Tirole: Game Theory. MIT Press

Solid, mathematical. Advanced.

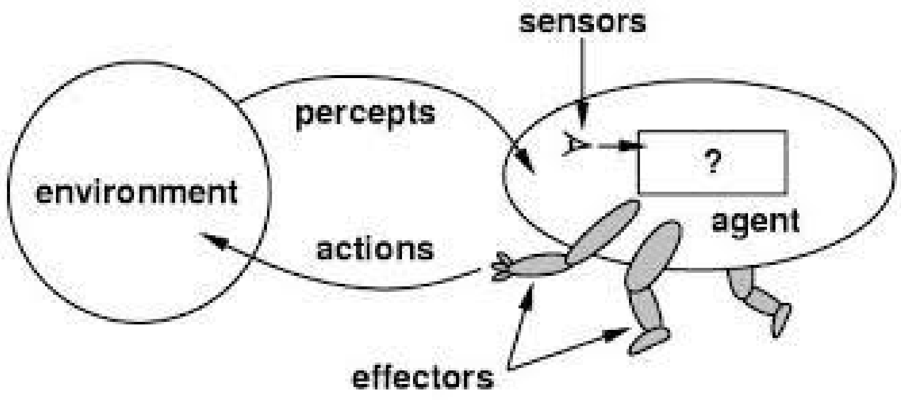
You only need to know what was covered during lectures!

Overview of Today's Lecture

1. Course Organisation
2. Agents and Multi-Agent Systems
3. Agent Types
4. Environments
5. Topics discussed in this Course

What is an Agent?

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 delegated *objectives*



- Note: autonomy is a *spectrum*!

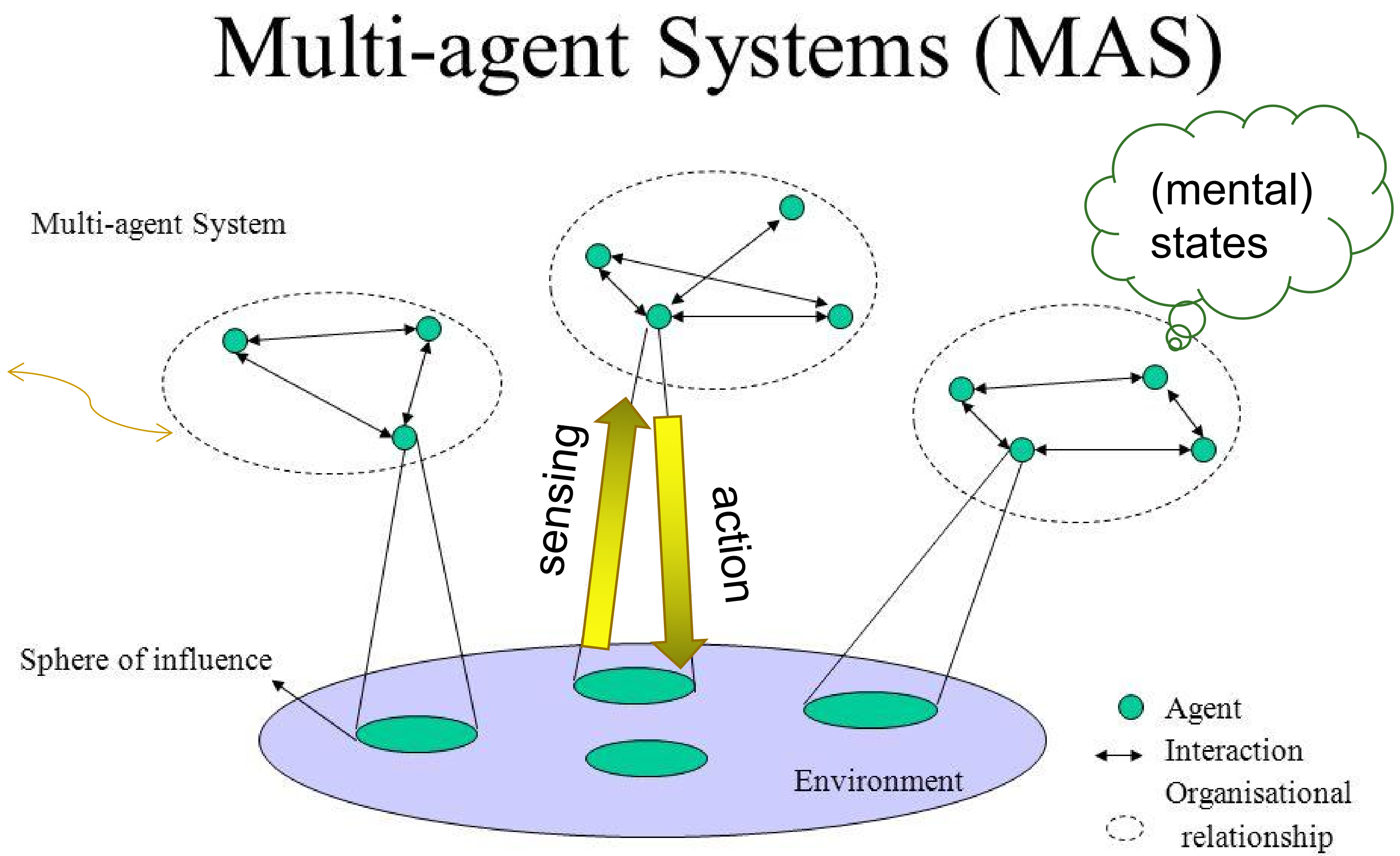
Properties of Agents

- Live in some *environment*
- *Observe* this environment
- Maintain *knowledge* about the environment
- Make *decisions* about what to do
- *Act* in the environment
- *Communicate* with other systems/agents
  - *Coordinate* with other agents (cooperative setting)
  - *Negotiate* with other agents (non-cooperative setting)

Multi-Agent Systems, a Definition

- A **Multi-Agent** System is one that consists of a number of agents that *interact (with each other and the environment)*
- In general, agents will have *different goals (often conflicting!)*
- To successfully interact, they will have to *learn, cooperate*, *coordinate*, and *negotiate*

Agents and Environment



Motivations for studying MAS

Techological:

- Growth of distributed, **networked computer** systems
  - (computers act more as **individuals** than parts)
- **Robustness**: no single point of failure
- **Scalable** and **flexible**:
  - adding new agents when needed
  - asynchronous, parallel processing
- **Development** and **reusability**
  - components developed independently (by specialists)

Application: Robotics

- Robots as Physical Agents (Embodiment)
  - Internet of Things (IoT)
  - Swarms of drones,
  - Fleet of autonomous vehicles
  - Physical internet

Motivations for studying MAS (2)

Scientific:

- Models for **interactivity** in (human) **societies**,
  - e.g. economics, social sciences

- Models for **emergence of cooperation**
  - Coordination: cooperation among non-antagonistic agents
  - Negotiation: coordination among self-interested agents

Multiagent Systems

- Typical scientific questions addressed:
  - What actions should agents take to optimize their rewards/utility?
  - How can self-interested agents learn from interaction with the environment and other agents to further their goals?
  - How can autonomous agents *coordinate* their activities so as to cooperatively achieve goals?
  - How can *cooperation* emerge in societies of self-interested agents?

MAS as Distributed AI (DAI)

- **AI :** Cognitive processes in individuals
  - Inspiration: neuro-science, behaviourism, ...
- **DAI:** Social processes in groups
  - Inspiration: social sciences, economics, ...

MAS as Distributed AI (DAI)

Basic question in DAI

- How and when should which agents interact (compete or collaborate) in order to achieve their design objectives?

Approaches:

- **Bottom-up:** given specific capabilities of individual agents, what collective behaviour will emerge?
- **Top-down:** Search for specific group-level rules (e.g. conventions, norms, etc.) that successfully constrain or guide behaviours at individual level;

Overview of Today's Lecture

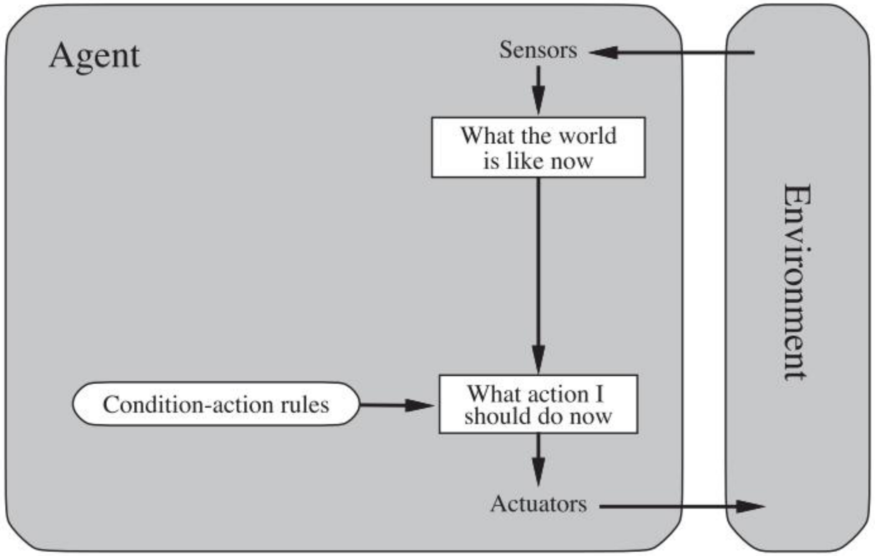
1. Course Organisation
2. Agents and Multi-Agent Systems
3. Agent Types
4. Environments
5. Topics discussed in this Course

Simple Typology for Agents

Intelligence in agents covers a spectrum:

- Reflex agents
  - Simple reflex agents
  - Model-based reflex agents
- Goal based agents
- Utility based agents
- Learning agents

Type 1: Simple Reflex Agent



- Reacts to environment
  - Percept ----> Action
    Based on simple
      if-then rules
    (condition-action)
- Properties:
  - No state: ignore history
  - Pre-computed rules
  - NO Partial observability

Type 2: Model-Based Reflex Agent

- Reflex agent with state
- Agent uses memory to store an internal representation of its world
- Internal model based percept history
- This internal model allows him to handle partially observable environment

Type 3: Goal-Based Agent

- Goal = desired outcome
- Goal-based (planning) agents act by reasoning about which actions to achieve the goal
- Less efficient, but more adaptive and flexible
- Search and planning: AI subfields concerned with finding sequences of actions to reach goal.

Terminal vs. instrumental goals

- **Terminal goals:** you want them bcs... *you want them!* -- stop asking stupid questions :-)
- **Instrumental goals:** you want them bcs they help you achieve your terminal goals
- **Convergent instrumental goals:** instrumental to wide range of terminal goals:
  - Self preservation and self-improvement
  - Goal preservation
  - Resource acquisition

Type 4: Utility-Based Agent

- Utility-based agents:
  - distinguish btw goal and non-goal states
- Utility-based agents use utility function
- Utility function
  - U(state) quantifies "happiness" (as real number)
  - Preferred world state has higher utility for agent
- Strategic Agents: Allows rational decisions in more situations
  - Evaluation of the tradeoffs among conflicting goals
  - Evaluation of competing goals

Type 5: Learning Agent

Four essential components

- Actor: responsible for selecting action in environment
- Critic: quantifies how well the agent is doing wrt. performance standard (e.g. utility).
- Learner: responsible for making improvements by learning from interactions.
- Explorer: responsible for suggesting new actions that will lead to novel and informative experiences.
-

Intelligent Agents

- An *intelligent agent* is a computer system capable of *flexible* autonomous action in some environment
- Autonomous: not pre-determined by designer
- By *flexible*, we mean:
  1. *reactive*

     (able to receive information from environment and respond)
  2. *pro-active*

     (able to reason and/or learn and work towards goals)
  3. *social*

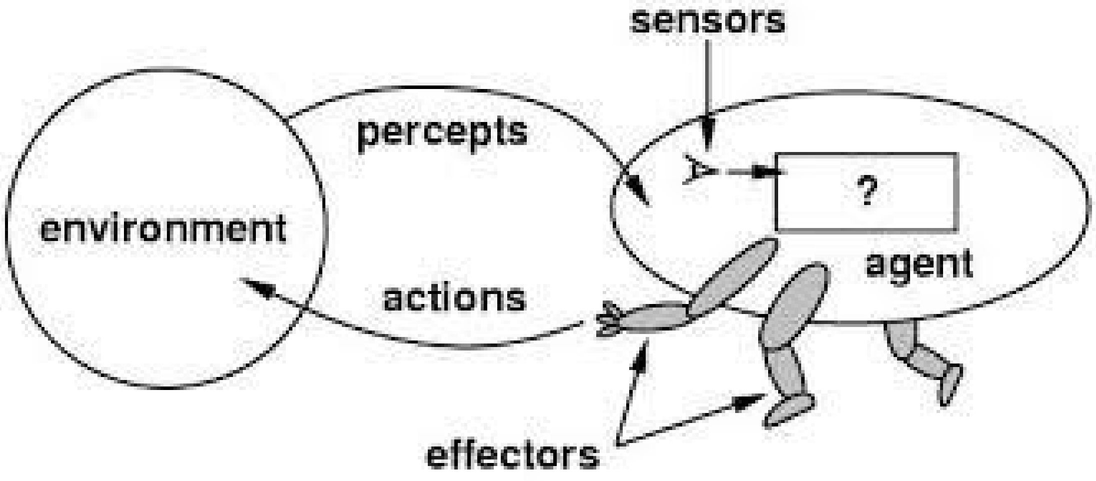
     (able to communicate, coordinate, negotiate and cooperate)

Overview of Today's Lecture

1. Course Organisation
2. Agents and Multi-Agent Systems
3. Agent Types
4. Environments
5. Topics discussed in this Course

Environments

- Agents act in/on environments



Environments - *Accessible* vs. *inaccessible*

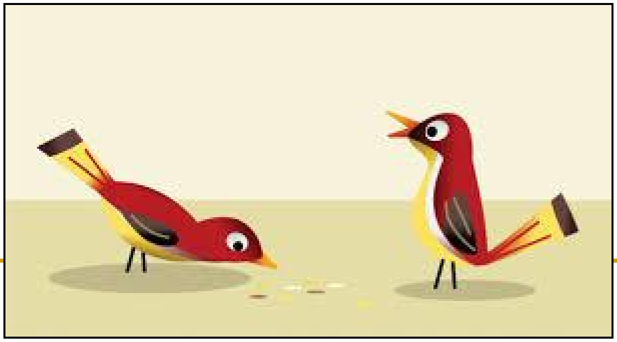
- An *accessible* environment: agent has complete, accurate, up-to-date information about the environment's state
- Most moderately complex environments (including, for example, the everyday physical world and the Internet) are inaccessible
- The more accessible an environment is, the simpler it is to build agents to operate in it

Environments - *Deterministic* vs. *nondeterministic*

- A *deterministic* environment is one in which any action has a single guaranteed effect - there is no uncertainty about the state that will result from performing an action
- The physical world can to all intents and purposes be regarded as non-deterministic
- Non-deterministic environments present greater problems for the agent designer

Environments - *Static* vs. *dynamic*

- A *static* environment is one that can be assumed to remain unchanged except by the performance of actions by the agent
- A dynamic environment is one that has other processes operating on it, and which hence changes in ways beyond the agent's control
- Other processes can interfere with the agent's actions (as in concurrent systems theory)



Overview of Today's Lecture

1. Course Organisation
2. Agents and Multi-Agent Systems
3. Agent Types
4. Environments
5. Topics discussed in this Course

Main topics in this course:

- **Strategic** Agents: Elementary Game Theory:
  - Multiple (stateless) competing agents,
  - Rational choice among multiple actions (max reward)
- Exploration versus Exploitation
- **Learning** Agents: Reinforcement Learning
  - Single agent learning to optimize reward from sequential interactions with environment;
  - Multi-agent Reinforcement Learning (MARL)
    - Intro and some pointers

Summary

- An agent is a computer system that is capable of *autonomous* action in some *environment*, in order to achieve its delegated *objectives*
- Agent properties: reactive, proactive, social
- A multi-agent system is a system that consists of a number of agents, which *interact* with one another and the environment.
- Three main topics in this course:
  - Strategic agents (game theory)
  - Exploration versus Exploitation
  - Learning agents