Multi-Agent Systems

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 2020
  - 6 ECTS
- Teachers:
  - Lecturer: Eric Pauwels (eric.pauwels@cwi.nl)
  - TAs: Andrei Udriste, Thomas Bellucci, Zachary Nudelman, Xinyu Hu, Daniel van der Riet
- Recommended prior knowledge:
  - Elementary calculus and probability, basic programming

Contact Hours

- Period:
  - Nov - Dec (7 weeks)
- Plenary lectures:
  - Tuesday: 11.00 - 12.45
  - Thursday: 15.30 - 17.15
- Lab sessions:
  - Mon 15.30 - 17.00 and   17.30 - 19.00
- Final exam:
  - Tuesday, Dec 21 (12h)

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 3 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

Motivating the introduction of *Agents*

- Five ongoing trends within computing:
  - *ubiquity*
    (processing power is everywhere)
  - *interconnection*
    (computer systems are networked and distributed)
  - *intelligence*
    (computers solve increasingly complex tasks)
  - *delegation*
    (we transfer more and more control to computers)
  - *human-orientation*
    (machines are increasingly being treated as humans)

Progression in abstraction…

- Programming has progressed through:
  - machine code;
  - assembly language;
  - machine-independent programming languages;
  - sub-routines;
  - procedures & functions;
  - abstract data types;
  - objects;

to *agents*.

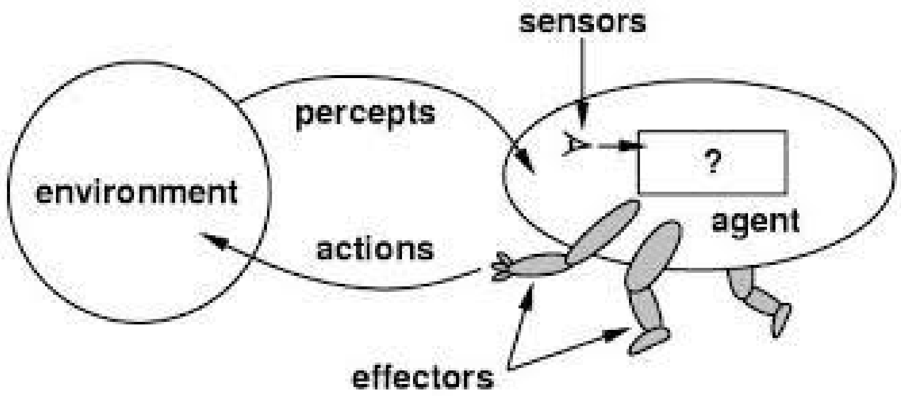
Where does it bring us?

- Delegation and Intelligence imply the need to build computer systems that can act effectively *on our behalf*
- This implies:
  - The ability of computer systems to act *independently (autonomously)*
  - The ability of computer systems to act in a way that *represents our best interests* while interacting with other humans or systems

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*

delegated: 委派的

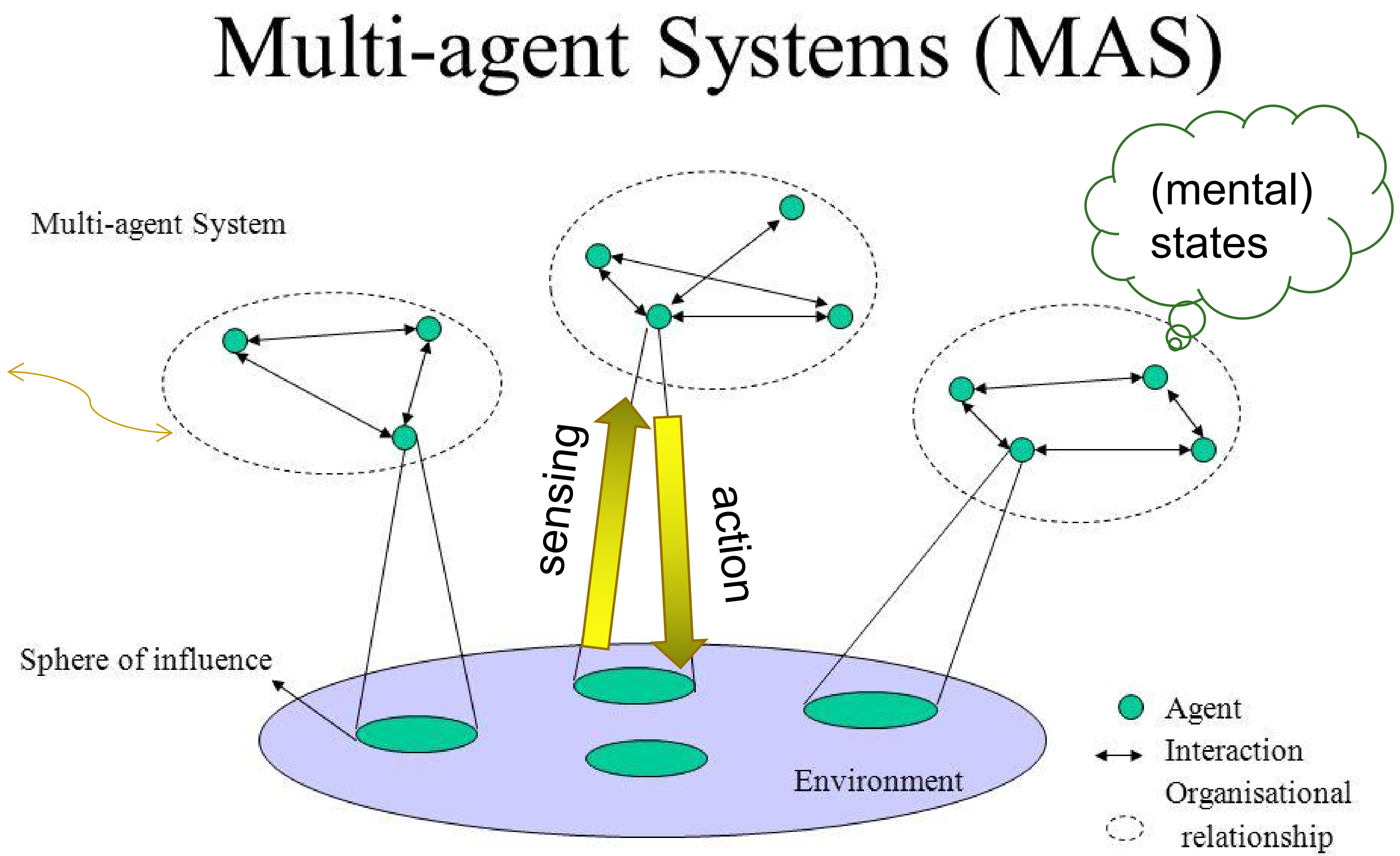


- Note: autonomy is a *spectrum*!

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*
- 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 coordination**
  - Cooperation: coordination among non-antagonistic agents
  - Negotiation: coordination among self-interested agents

Multiagent Systems

- Typical scientific questions addressed:
  - How can *cooperation* emerge in societies of self-interested agents?
  - 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?

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;

Multiagent Systems is Interdisciplinary

- The field of Multi-Agent Systems is influenced and inspired by *many other fields*:
  - Economics
  - Game Theory
  - Philosophy and Logic
  - Mathematics (e.g. optimal control)
  - Ecology
  - Social Sciences

- This can be both a *strength* and a *weakness*
- This has analogies with *Artificial Intelligence* itself

Overview of Today's Lecture

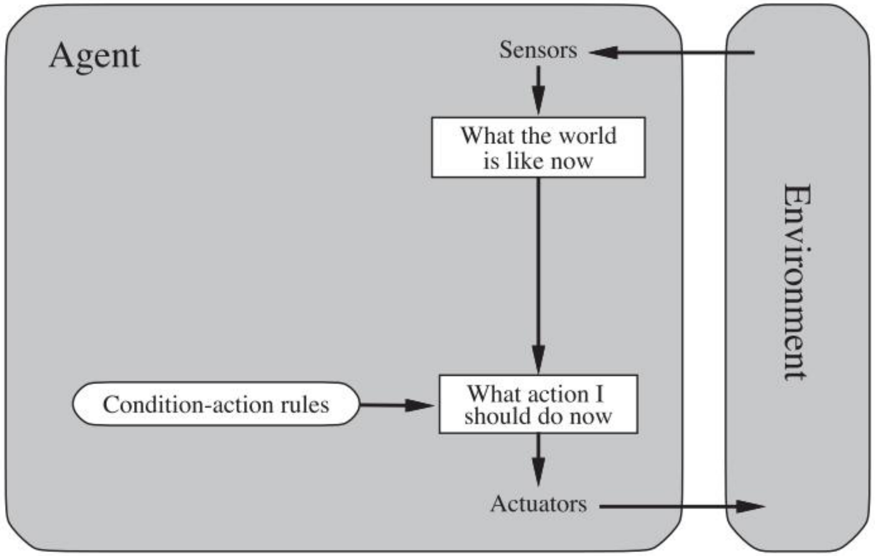
1. Course Organisation
2. Agents and Multi-Agent Systems
3. Agent Types
4. Environments
5. Topics discussed in this Course

Classification of Agents

- Intelligent agents are grouped in to five classes based on their degree of perceived intelligence and capability.
- ✓ 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:
  - Simple: ignore history
  - Partial observability



Type 2: Model-Based Reflex Agent

- A model-based agent can handle a partially observable environment.
- Its current state is stored inside the agent maintaining some kind of structure which describes the part of the world which cannot be seen.
- This knowledge about "how the world evolves" is called a model of the world, hence the name "model-based agent".

Type 3: Goal-Based Agent

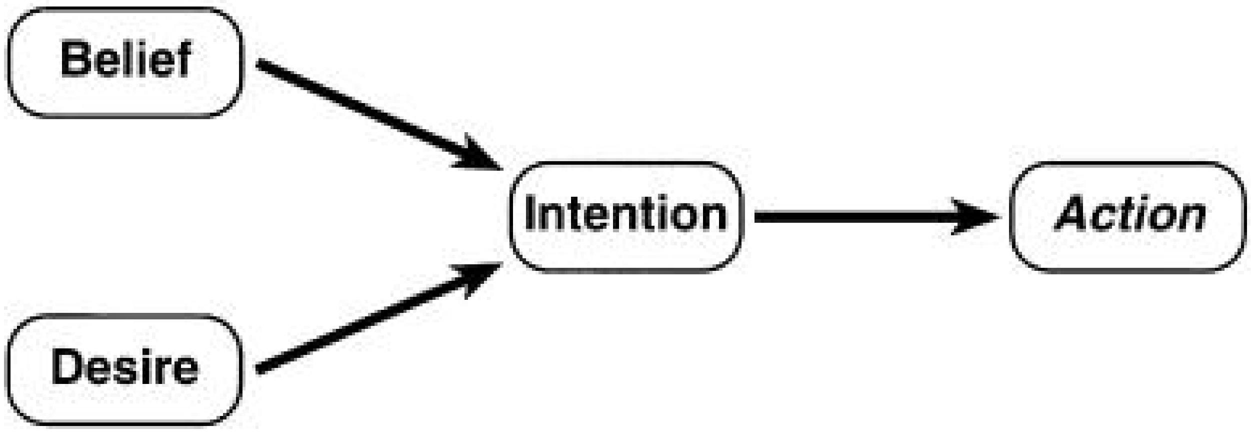
- Goal-based agents further expand on the capabilities of the model-based agents, by using "goal" information.

- Goal information describes situations that are desirable. This allows the agent a way to choose among multiple possibilities, selecting the one which reaches a goal state.

- Search and planning are the subfields of artificial intelligence devoted to finding action sequences that achieve the agent's goals.

Agents as Intentional Systems: BDI model

- **Beliefs:** informational state of agent (incl. inference rules)
- **Desires:** motivational state of agent: i.e. what he would like to accomplish;
- **Intentions:** deliberative state of agent: desires to which agent has committed!
- **Plans:** sequence of actions to achieve intentions



Type 4: Utility-Based Agent

- Goal-based agents only distinguish between goal states and non-goal states.
- It is possible to define a measure of how desirable a particular state is. This measure can be obtained through the use of a *utility function* which maps a state to a measure of the utility of the state.
- A more general performance measure should allow a comparison of different world states according to exactly how happy they would make the agent. The term utility, can be used to describe how "happy" the agent is.

Intelligent Agents

- An *intelligent agent* is a computer system capable of *flexible* autonomous action in some environment
- By *flexible*, we mean:

1. *reactive*

   (able to receive information from environment and respond)

2. *pro-active*

   (able to take initiative to work systematically towards goals)

3. *social*

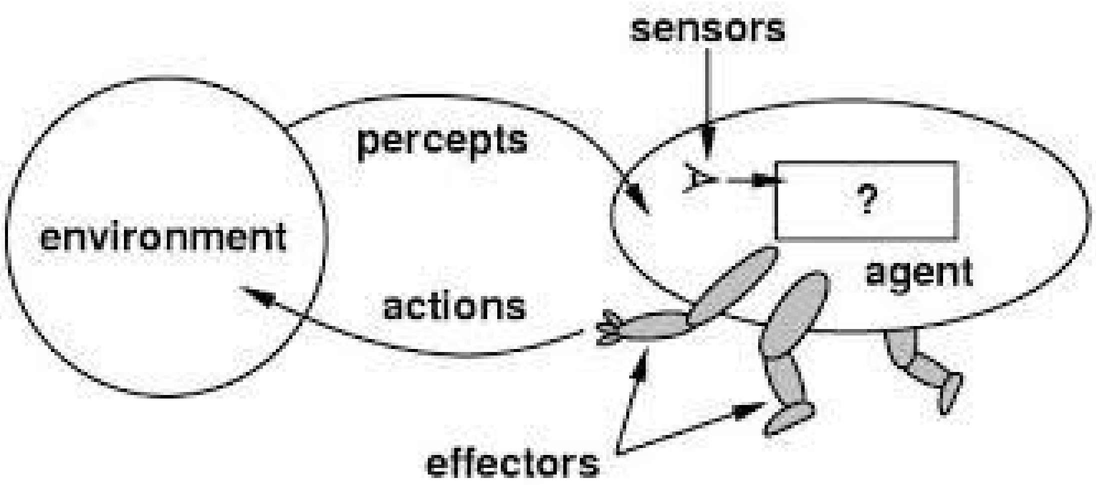
   (able to communicate, coordinate, negotiate and cooperate)

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*

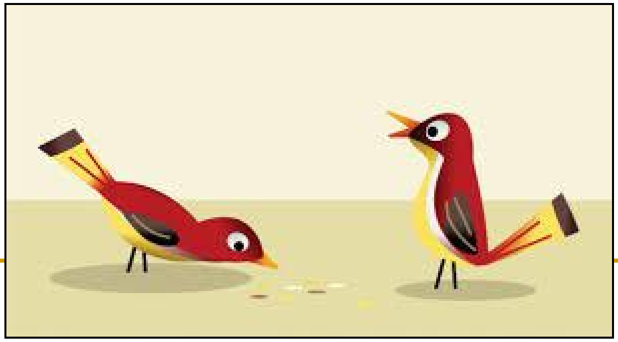
- An *accessible* environment is one in which the agent can obtain complete, accurate, up-to-date information about the environment's state
- Most moderately complex environments (including, for example, the everyday physical world and the Internet) are inaccessible
- The more accessible an environment 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

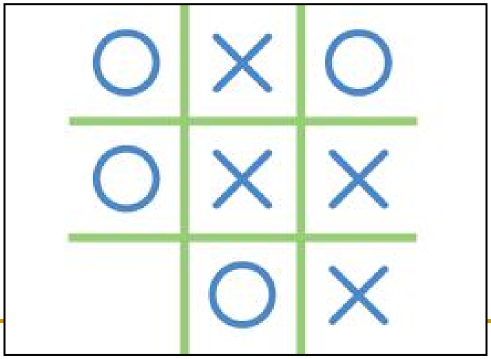
Environments - *Static* vs. *dynamic*

- 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)



Environments - *Discrete* vs. *continuous*

- An environment is *discrete* if there are a fixed, finite number of actions and percepts in it
- Continuous environments have a certain level of mismatch with computer systems
- Discrete environments could *in principle* be handled by a kind of “lookup table”



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:

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

Game Theory:

Strategic Thinking for Rational Agents

- Prisoners' dilemma:
- *Two members of a criminal gang are arrested and imprisoned, and they cannot communicate with one another. The prosecutors lack sufficient evidence to convict the pair on the principal charge of robbery, but they have enough to convict both on a lesser charge of trespassing. Each prisoner is given the opportunity to betray the other: if he does so while the other remains silent, he will go free while the other will feel the full force of the law (12 months in jail). If they both confess, the will both spend 8 months in jail, but if they both refuse to cooperate, they will both be convicted for the lesser charge of trespassing (which carries a 1 month jail-time).*
- What should they do? What will the outcome be?

Exploration versus exploitation

- You're hungry and want to dine at a restaurant. What should you do?
- Go to your favourite place which you know you like (exploiting your current knowledge)?
- Try a new place, which might be even better, or maybe....much worse (exploration)?

Reinforcement learning

- Unsupervised learning (e.g. clustering)
- Supervised learning
- Reinforcement learning: how to learn from interaction?
  - Cooking a complex dish: no feedback on every step or ingredient, but feedback on final result;

Summary

- Five trends in computing
  - ubiquity, interconnection, intelligence, delegation, human-orientation
- 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.

Properties of Autonomous Systems

- They live in some *environment*
- They *observe* this environment
- They maintain *knowledge* about the environment
- They make *decisions* about what to do
- They *act* in the environment
- They *communicate* with other systems
- They *coordinate* with other systems
- They *negotiate* with other systems