Autonomous Multi-Agent Systems

Introduction

Introduction (content)

- Definition of an Agent
- Individual and Group Perspective
- Distributed AI and Multi-Agent Systems
- Ethical aspects
- Overview of the course topics
- Summary

What is an Agent? Scenario 1

When a spaceship makes its flight from Earth to other planets, a ground control team is required to continuously track its orbit and decide how to deal with unexpected situations. This is costly and, in many cases, it is simply not possible. For these reasons, it is necessary give the spaceship richer decision making capabilities and responsibilities.

Scenario 1

Self-Driving Car

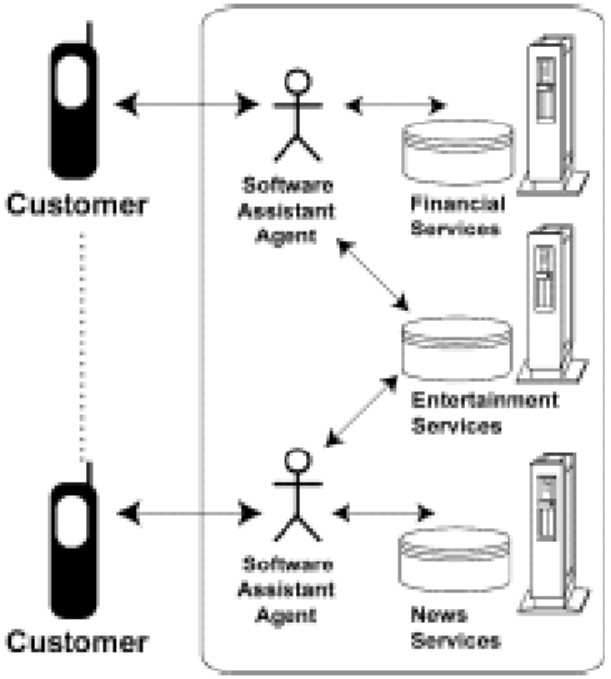
Scenario 2 (intelligent home)

We assume that most of devices in your home (such as refrigerator, washing machine, dish-washing machine, even wardrobe etc.) are software assistants (at least they are represented by software assistants). You are also represented by your software assistant. Imagine that you told your software assistant about your menu for dinner. The assistant then communicates to refrigerator assistant and asks which products are available for the selected menu. The refrigerator assistant answers with the list of products and their amount (it also may keep track of expiration dates for products). Your assistant plans purchasing missing products (taking into account a healthy diet, of course). At this time wardrobe assistant asks your agent about your schedule for tomorrow and recognizes that you go to the Opera but there is no dinner jacket (smoking) in the wardrobe and you have to rent it. So your assistant contacts a renting company for a smoking (of course, it knows your size).

When you come home at evening all necessary products are ordered from the Internet shop and delivered (as well as dinner jacket). We don't assume that intelligent stove made ready your dinner according to menu (we still try to be realistic ...)

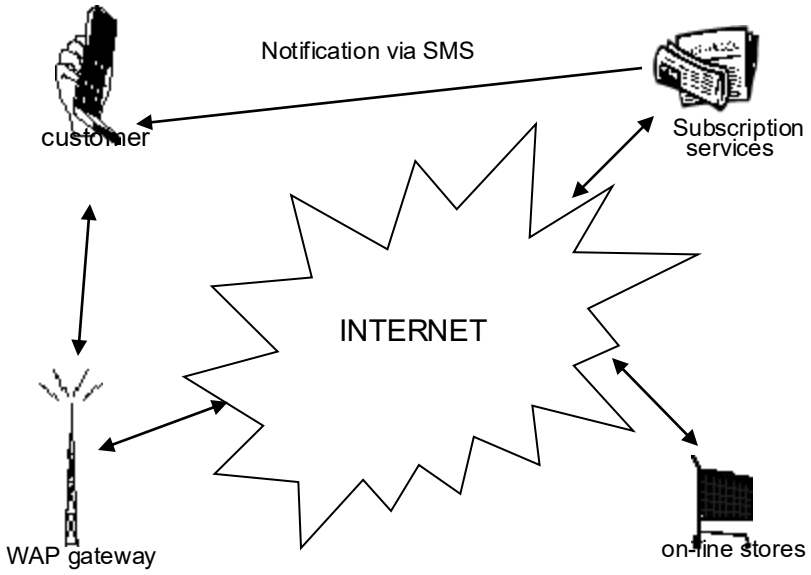
Scenario 3 (financial adviser)

You have stocks and worry about stock quotas. You would like to be immediately notified about changes in stock quotas that are significant for you. Your software assistant monitors stock quotas, performs market analysis reflected your preferences, consult with other assistants and notify you by sending email or SMS to your smart phone.

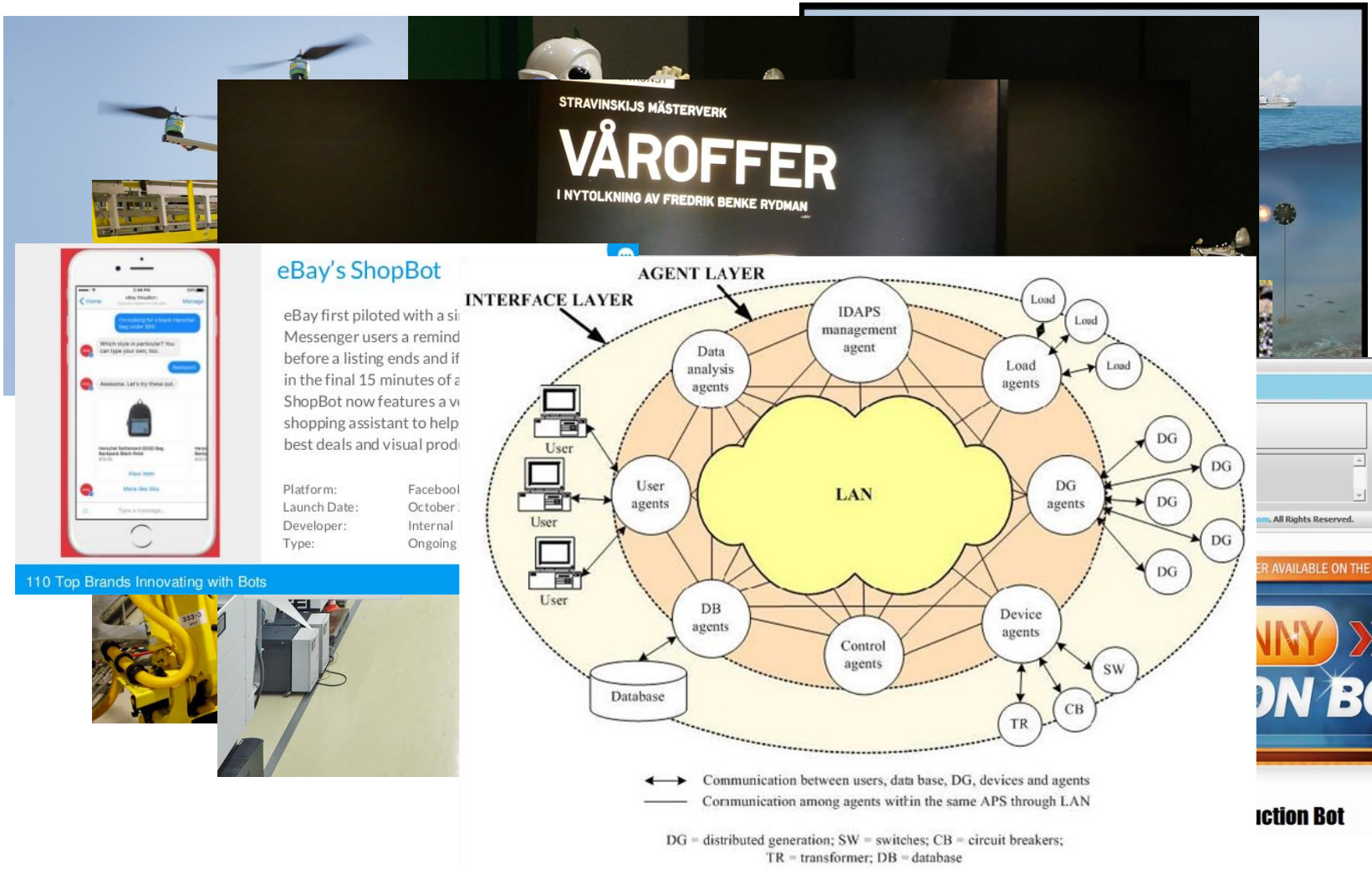


Scenario 4

You are looking for a 3D-player and have a limited amount of money to buy it. Your software assistant searches for different Internet shops, gets recommendations from sellers, brokers, other assistants. It may negotiate about price or form shopping coalitions with other agents in order to obtain discount. The assistant brings you the best available offer for your approval.



Autonomous systems



Definitions of Agents

American Heritage Dictionary:

”... One that acts or has the power or authority to act ... or represent another”

Negroponte:

”Digital sister in law”

Definitions of Agents

Russel and Norvig:

”An agent is anything that can be viewed as perceiving its environment through sensors and acting upon that environment through effectors.”

Maes, Pattie:

”Autonomous Agents are computational systems that inhabit some complex dynamic environment, sense and act autonomously in this environment, and by doing so realize a set of goals or tasks for which they are designed”.

Definitions of Agents

IBM:

”Intelligent agents are software entities that carry out some set of operations on behalf of a user or another program with some degree of independence or autonomy, and in doing so, employ some knowledge or representations of the user’s goals or desires”.

Coen:

“Software agents are programs that engage in dialog [and] negotiate and coordinate transfer of information.”

Why do we have activity around autonomous agents?

- more and more everyday tasks are computer-based
- the world is in a midst of an information revolution
- increasingly more users are untrained
- users require agents to assist them in order to understand the technically complex world we are in the process of creating

A more practical reason:

Lack of programming paradigm for decentralized program/system construction in dynamic environment.

More questions

What is a problem and why former paradigms (F/E OOP) can not completely satisfy our needs?

They were designed for constructing programs/systems in a completely specified environment. The world was closed!

What do we have in case of distributed IT (now)?

An environment (in most cases) can not be specified completely.

Why it is problematic to specify distributed environment completely?

Because of our distributed "world" (environment) is not closed any more

How can we work in an open environment?

Similar to "human's" way of behavior

- perceiving the environment
- affecting the environment
- having a "model" of behavior?
- having intentions/motivations to be fulfilled by implementing corresponding goals
- communicate with other

Fields that inspired agents

- Artificial Intelligence
  - Agent intelligence, micro aspects
- Software Engineering
  - Agent as an abstraction
- Distributed systems and Computer Networks
  - Agent architectures, multi-agent systems, coordination

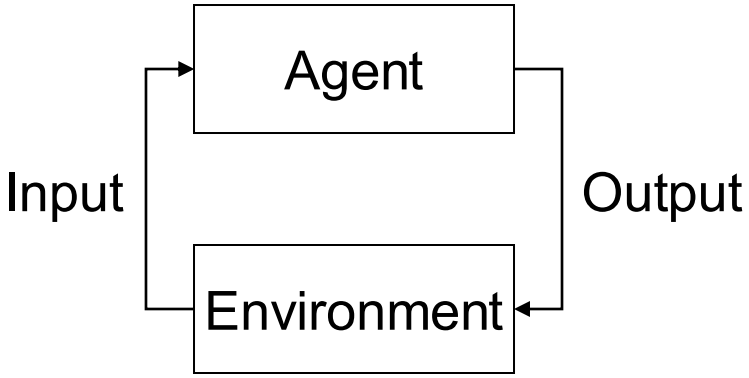
Five ongoing trends have marked the history of computing

1. Ubiquity
   - Reduction in the cost of computing capability
2. Interconnection
   - Computer systems are networked into large distributed systems
3. Intelligence
   - The complexity of tasks that can be automated and delegated to computers
4. Delegation
   - Judgment of computer systems are frequently accepted
5. Human-orientation
   - Use concepts and metaphors that reflect how we understand the world

Definitions of Agents (continue)

- An agent is *autonomous*: capable of acting independently, exhibiting control over its internal state.

➢ An agent is *a computer system capable of autonomous action in some environment*.



Some (Trivial) Autonomous Agents

- Thermostat
- UNIX deamon
- Help wizzards
- ...

What is an agent?

Wooldridge, Jennings (weak notion):

Agent is a hardware or (more usually) software-based computer system that enjoys the following properties:

- **autonomy**: agents operate without the direct intervention of humans or others, and have some kind of control over their actions and internal state;
- **pro-activeness**: agents do not simply act in response to their environment, they are able to exhibit goal-directed behavior by taking the initiative;
- **reactivity**: agents perceive their environment and respond in a timely fashion to changes that occur in it;
- **social ability**: agents interact with other agents (and possibly humans) via some kind of agent-communication language.

Agents and Objects

- An object encapsulates state and has no control over the execution of methods (has no control over its own actions)

- Agents encapsulate *behavior*, in addition to state. The decision about whether to act upon the request lies with the recipient – not on requestor.

Jennings and Wooldridge

Agents vs. Objects

Main differences:

- Degree of autonomy: agents embody a stronger notion of autonomy than objects, in particular, agents decide for themselves whether or not to perform an action.
- Degree of smartness: capable of flexible (reactive, pro-active, social) behaviour; standard object models do not have such behaviour.
- Degree of activeness: a multi-agent system is inherently multi-threaded in that each agent is assumed to have at least one thread of active control.

What is an agent (strong notion)?

- *Mentalistic notions*, such as beliefs and intentions are often referred to as properties of strong agents.
- Other properties are:
  - Veracity: agent will not knowingly communicate false information.
  - Benevolence: agents do not have conflicting goals and always try to do what is asked of it.
  - Rationality: an agent will act in order to achieve its goals and will not act in such a way as to prevent its goals being achieved.
  - Mobility: the ability of an agent to move around a network.

Ethical AI:
Isaac Asimov's laws of robotics

- Law One – "A robot may not injure a human being or, through inaction, allow a human being to come to harm."
- Law Two – "A robot must obey orders given to it by human beings except where such orders would conflict with the First Law."
- Law Three – "A robot must protect its own existence, as long as such protection does not conflict with the First or Second Law."

Asimov later added the "Zeroth Law," above all the others – "A robot may not harm humanity, or, by inaction, allow humanity to come to harm."

Asimov, Isaac (1950). "Runaround". *I, Robot* (The Isaac Asimov Collection ed.). New York City: Doubleday. p. 40.

WHY ASIMOV PUT THE THREE LAWS OF ROBOTICS IN THE ORDER HE DID:

POSSIBLE ORDERING	CONSEQUENCES	
1. (1) DON'T HARM HUMANS 2. (2) OBEY ORDERS 3. (3) PROTECT YOURSELF	[SEE ASIMOV'S STORIES]	BALANCED WORLD
1. (1) DON'T HARM HUMANS 2. (3) PROTECT YOURSELF 3. (2) OBEY ORDERS	EXPLORE MARS! HAHA, NO. IT'S COLD AND I'D DIE.	FRUSTRATING WORLD
1. (2) OBEY ORDERS 2. (1) DON'T HARM HUMANS 3. (3) PROTECT YOURSELF		KILLBOT HELLSCAPE
1. (2) OBEY ORDERS 2. (3) PROTECT YOURSELF 3. (1) DON'T HARM HUMANS		KILLBOT HELLSCAPE
1. (3) PROTECT YOURSELF 2. (1) DON'T HARM HUMANS 3. (2) OBEY ORDERS	I'LL MAKE CARS FOR YOU, BUT TRY TO UNPLUG ME AND I'LL VAPORIZE YOU.	TERRIFYING STANDOFF
1. (3) PROTECT YOURSELF 2. (2) OBEY ORDERS 3. (1) DON'T HARM HUMANS		KILLBOT HELLSCAPE

From: https://imgs.xkcd.com/comics/the_three_laws_of_robotics.png

Other possible rules of robotics

- A robot must establish its identity as a robot in all cases (1974 Lyuben Dilov Icarus's Way)
- A robot must know it is a robot (Nikola Kesarovski, short story "The Fifth Law of Robotics")
- A robot must reproduce. As long as such reproduction doesn't contradict with Laws 1, 2 and 3 (Harry Harrison in Foundation's Friends, Stories in Honor of Isaac Asimov)
- All robots endowed with comparable human reason and conscience should act towards one another in a spirit of brotherhood (2013 Hutan Ashrafian, proposed an additional law)

EU AI Act: first regulation on artificial intelligence (risk levels)

Unacceptable risk

Unacceptable risk AI systems are systems considered a threat to people and will be banned. They include:

- Cognitive behavioural manipulation of people or specific vulnerable groups: for example, voice-activated toys that encourage dangerous behaviour in children
- Social scoring: classifying people based on behaviour, socio-economic status or personal characteristics
- Biometric identification and categorisation of people
- Real-time and remote biometric identification systems, such as facial recognition

https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence#ai-act-different-rules-for-different-risk-levels-0

EU AI Act: first regulation on artificial intelligence (risk levels)

High risk

AI systems that negatively affect safety or fundamental rights will be considered high risk and will be divided into two categories:
1) AI systems that are used in products falling under the EU's product safety legislation. This includes toys, aviation, cars, medical devices and lifts.

2) AI systems falling into specific areas that will have to be registered in an EU database:

- Management and operation of critical infrastructure
- Education and vocational training
- Employment, worker management and access to self-employment
- Access to and enjoyment of essential private services and public services and benefits
- Law enforcement
- Migration, asylum and border control management
- Assistance in legal interpretation and application of the law.

https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence#ai-act-different-rules-for-different-risk-levels-0

EU AI Act: first regulation on artificial intelligence (risk levels)

Applications not explicitly banned or listed as high-risk are largely left unregulated.

EU AI Act: first regulation on artificial intelligence (transparency requirements)

Generative AI, like ChatGPT, will not be classified as high-risk, but will have to comply with transparency requirements and EU copyright law:

- Disclosing that the content was generated by AI
- Designing the model to prevent it from generating illegal content
- Publishing summaries of copyrighted data used for training

https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence#ai-act-different-rules-for-different-risk-levels-0

Summary of definitions

Basic agent features:

- acts on behalf of other entities
- has weak agent characteristics (Wooldridge, Jennings)
- may have strong agent characteristics (Wooldridge, Jennings)

Which other terms are used in the context of agency?

There are many synonyms of the term "intelligent agent" in literature.

However, they are mostly a specialization of agency with respect to implementation, operating environment, roles or functionality.

- **robots** - agents which inhabit the physical world (factory etc.)
- software agents (**softbots**) - agents which inhabit the programming environments (networks)
- **knowbots** - knowledge-based robots
- **taskbots** - task-based robots
- **userbots** - user oriented agents
- …

Autonomous Systems: Humans, Robots, Agents

- A **human** agent: has eyes, ears, and other organs for sensors and hands legs, mouth, and other body parts as actuators
- A **robotic** agent: might have cameras and infrared range finders for sensors and various motors for actuators
- A **software** agent receives keystrokes, file contents, and network packets as sensory inputs and acts on environment by displaying on the screen, writing files, and sending network packets

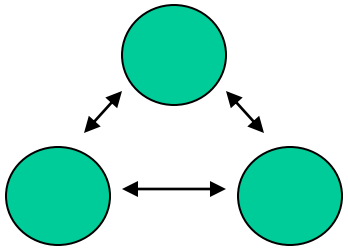
Agents dimensions

- Individual agents perspective
- Group agents perspective

Two Key Problems

1. How do we build agents that are capable of independent, autonomous actions in order to successfully carry out the tasks that we delegate to them? (Micro aspects)

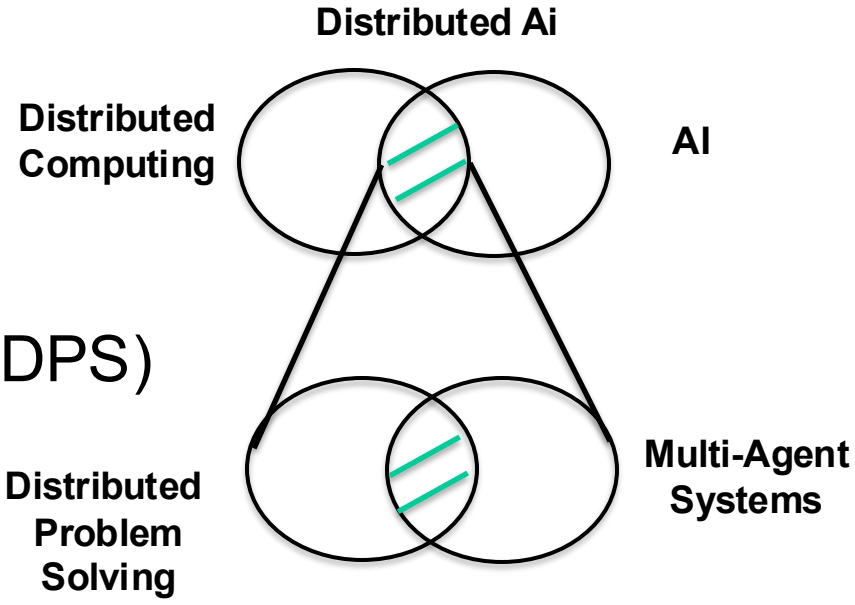
2. How do we build agents that are capable of interacting (cooperating, coordinating, negotiating) with other agents in order to successfully carry out the tasks we delegate to them? (Macro aspects)



Distributed Artificial Intelligence (DAI)

- DAI is a sub-field of AI
- DAI is concerned with problem solving where several agents solve (sub-)tasks (macro level)

- Main areas of DAI
  1. Multi-Agent Systems (MAS)
  2. Distributed Problem Solving (DPS)



DPS

- DPS considers how the task to be solved can be divided among several components that cooperate and share knowledge about the problem and its solution.
  - In pure DPS systems, all interaction strategies are incorporated as an integral part of the system
  - DPS has focused on achieving goals, having agents with established properties.

MAS

- MAS are designed with a great part of independency and autonomy
- Individual agents have own preferences, but they consent to act in a way that leads to desired global goal.
- MAS considers how can a certain collective goal be realized if the properties of agents can vary uncontrollably
  - loosely-coupled networks of problem solvers (agents) that work together to solve problems that are beyond their capabilities
  - no necessary guarantees about other agents

Decentralisation

- No central control; control is distributed
- Knowledge or information sources may also be distributed

Motivation for MAS

- To solve problems that are too large for a centralized agent
- To provide a solution to inherently distributed problems
- To provide solutions which draw from distributed information sources
- To provide solutions where expertise is distributed
- To offer conceptual clarity and simplicity of design

Cooperative and Self-interested MAS

- Cooperative
  - Agents designed by interdependent designers
  - Agents act for increased good of the system
  - Concerned with increasing the performance of the system

- Self-interested
  - Agents designed by independent designers
  - Agents have their own agenda and motivation
  - Concerned with the benefit and performance of the individual agent

Benefits of MAS

- Faster problem solving
- Decrease in communication
- Flexibility
- Increased reliability

DAI is not concerned with:

- Issues of coordination of concurrent processes at the problem solving and representational level

- Parallel Computer Architectures, Parallel Programming Languages or Distributed Operating Systems

Summary of definitions

Distributed Computing

- Focus on low-level parallelization and synchronization.

Distributed AI

- Intelligent control as well as data may be distributed. Focus on problem solving, communication and coordination.

DPS

- Task decomposition (task sharing) and/or solution synthesis (result sharing): information management

MAS

- Behavior coordination and management

Emergence, Swarm Intelligence and other Terms

Emergence

- Global (macro level) behaviour, patterns and properties that are arising from the interactions between local parts of the system (micro level).
- Systems that exhibit emergence can be characterized as simple, robust, and adaptive

Swarm Intelligence

- Designing algorithms or distributed problem-solving devices inspired by animal societies (the collective behaviour of social insect colonies and other)

- Multi-robot systems, which implement or adapt the concept of emergent behaviour, are commonly referred to as swarm robotic systems

Self-Organisation

- Self-organization is a dynamical and adaptive process where systems acquire and maintain structure themselves, without external control
- The term self-organization is also used as the process that leads to the state of emergence
- Self-organizing and emergent systems are distinct concepts, however, they still have one thing in common, that is: There is no explicit external control whatsoever
- The main difference between self-organization and emergence is that in the case of self-organization, individual entities can be aware of the system's intended global behaviour. In consequence, self-organization can be seen as a weak form of emergence
- The intuitive and regularly used approach to realize self-organization is applying the concept of feedback loops.

Self-Adaptation

- When the approach with feedback loop is applicable to single entity system it is usually referred as self-adaptation

- Self-adaptive software modifies its own behaviour in response to changes in its operating environment.

- If a decentralized system containing several entities exhibits adaptive behavior to external changes this is as well considered self-adaptation

Characteristics of Emergence, Swarm Intelligence, Self-Adaptation, and Self-Organization

- It would no longer be necessary to exactly specify the low-level system behaviour in all possible situations that might occur, but rather, leaving the system with a certain degree of freedom to allow for autonomous reaction and adaptation to new situations in an intelligent way.

Application

- Self-organization algorithms have been applied in many domains, (for example, combinatorial optimization, communication networks, and robotics)
- Self-organization algorithms and mechanisms, are used, for example, in motion control, information sharing, and decision-making

Course Topics for Agents Group Perspective

- Agent Coordination – achieving coherent behavior
- Agent Negotiation – conflict resolution
- Agent Communication – interoperability
- Multi-Agent System Architectures
- Software Engineering of Multi-Agent Systems

Coordination Problem

- Managing the interdependencies between the activities of agents.

Cartoon taken from Klein, AAMAS2002

Coordination

- Preventing anarchy or chaos
- Meeting global constraints
- Providing distributed expertise, resources or information
- Supporting dependencies between agents’ actions
- Allowing efficiency

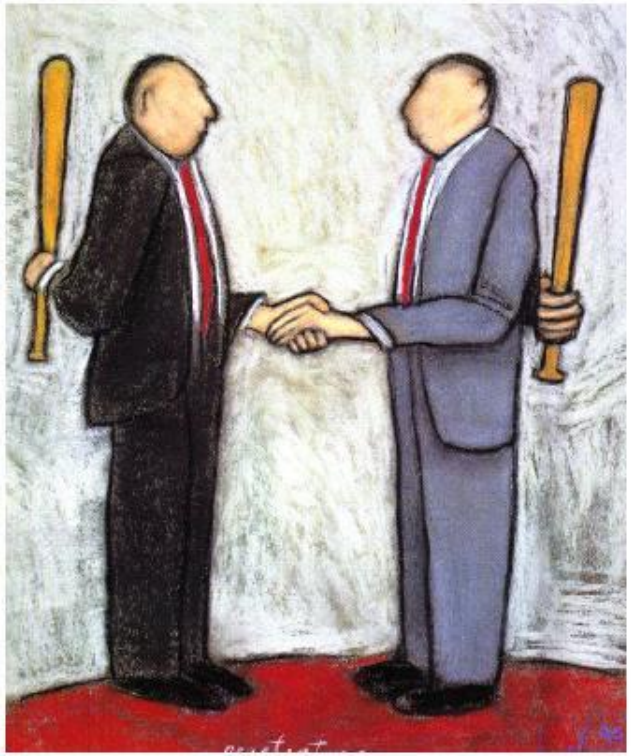
Negotiation (Example)

Conflict resolution

during problem solving

by independent

parties/agents.



Software interoperation (communication)

Interoperation - exchange of information and services with other programs, thereby solving problems that cannot be solved alone

Main problems:

- Possible heterogeneity of agents created by
  - different people
  - at different time
  - in different languages

Communication in MAS

- To successfully coordinate and negotiate agents need ability to communicate
- Basic methods for *communication*:
  - Plan /message passing
  - Information exchange using shared repositories
- Important characteristics of communication:
  - Relevance of the information exchanged
  - Timeliness
  - Completeness

Components of a system for effective interaction and interoperability

- common language
- common understanding of the knowledge exchanged
- common interpretation of the knowledge
- ability to exchange whatever is included in the previous items
  - *Agent Communication Language (ACL)*

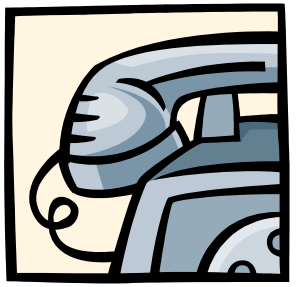
FBI Agents Ordering Pizza

FBI agents conducted a raid of a psychiatric hospital in San Diego that was under investigation for medical insurance fraud. After hours of reviewing thousands of medical records, the dozens of agents had worked up quite an appetite. The agent in charge of the investigation called a nearby pizza parlour with delivery service to order a quick dinner for his colleagues. The following telephone conversation took place and was recorded by the FBI because they were taping all conversations at the hospital.

Source: http://jewel.morgan.edu/~salimian/humor/humor_094.html

FBI Agents Ordering Pizza, contd.

- ***Agent***: Hello. I would like to order 19 large pizzas and 67 cans of soda.
- ***Pizza Man***: And where would you like them delivered?
- ***Agent***: We're over at the psychiatric hospital.
- ***Pizza Man*** : The psychiatric hospital?
- ***Agent*** : That's right. I'm an FBI agent.
- ***Pizza Man*** : You're an FBI agent?
- ***Agent*** : That's correct. Just about everybody here is.
- ***Pizza Man*** : And you're at the psychiatric hospital?
- ***Agent*** : That's correct. And make sure you don't go through the front doors. We have them locked. You will have to go around to the back to the service entrance to deliver the pizzas.
- ***Pizza Man*** : And you say you're all FBI agents?
- ***Agent*** : That's right. How soon can you have them here?
- ***Pizza Man*** : And everyone at the psychiatric hospital is an FBI agent?
- ***Agent*** : That's right. We've been here all day and we're starving.
- ***Pizza Man*** : How are you going to pay for all of this?
- ***Agent*** : I have my checkbook right here.
- ***Pizza Man*** : And you're all FBI agents?
- ***Agent*** : That's right. Everyone here is an FBI agent. Can you remember to bring the pizzas and sodas to the service entrance in the rear? We have the front doors locked.
- ***Pizza Man*** : I don't think so. Click.



What is Multi-agent Architecture

- The infrastructure for a MAS can be defined as a set of services, conventions and knowledge that support complex social interactions (e.g. negotiations, agree on commitments/coordination).

- In particular, agents need services to:
  - Enable them to find each other in open environments
  - Communicate
  - Warrant that proper security constraints are satisfied
  - …

Agent Oriented Software Engineering (AOSE)

The approach to developing AOSE **t**echnology:

- to present the new technology as an incremental extension of known and trusted methods (e.g. OO software engineering)

Course Topics for Individual Agents Perspective

- Agent Theory
- Agent Architectures
- Agent Mobility

Agent Theory

- intentional systems
  - knowledge, belief, desire, goal, intention
- traditional AI approaches

The Wise Men puzzle (knowledge about knowledge)

There are three wise men.

It's common knowledge -- known by everyone, and known to be known by everyone, etc. -- that there are three red hats and two white hats. The king puts a hat on each of the wise men and ask them sequentially if they know the color of the hat on their head. Suppose the first man says he does not know; then the second say he does not know either.

It follows that the third man must be able to say that he knows the color of his hat.

Why is this, and what color has the third man's hat?

1. rrr
2. rrw
3. rwr
4. wrr
5. rww
6. wrw
7. wwr

Agent architectures

- Deliberative
- Reactive
- Hybrid

Mobility

- Agents are capable of transmitting themselves – their programs and their state – across a computer network, and recommencing execution at a remote site.
- The program chooses when and where to migrate.
- It can suspend its execution at an arbitrary point, transport itself to another machine and resume execution.

➢ *Moving programs while they run!*

Summary

- DAI is a part of AI
- MAS is a part of DAI
- MAS - "macro" issues of agent systems
- intelligent agents - "micro" issues of agent systems

Next Lecture: Multi-agent Negotiation

Related chapters in

Wooldridge: ”Introduction to MultiAgent Systems”:

11, 12, 14, 15