Self-Awareness in Autonomic Systems
Layman Seminar
Awareness Slides Factory
Barcelona, September 2012
Content Outline

• Introductory examples
  – Robotics
  – Autonomic systems
    • Robot Swarm
    • Science Cloud
    • Cooperative E-Vehicles
  – Networks
    • Communication Networks
    • Power Networks

• Motivation
Content Outline (2)

• “Definition”/circumscription of
  – Autonomic Systems
  – Self-Awareness

• Summary of existing research projects
  – Ascens, EPiCS, Sapere, Recognition, CoCoRo, Symbrion

• Existing “self-aware” systems (sources of inspiration)
  – Nature
  – Man-made
Content Outline (3)

• Some related general properties
  – Different levels of awareness (quality)
  – Collectivity/Swarm/Distributedness
  – Perception (sensing data + context)
  – Internal Model

• Impact short/long-term
  – Safety
  – Sustainable Future
  – Philosophical

• Questions & Challenges
Introductory examples

ROBOTICS
Robots

• How hard can a robot push a person
  – How hard do you have to hit someone before ABH becomes GBH?
• [cartoon – Japanese underground in rush hour]
• Suppose we want to automate that
• [cartoon – R2D2 pushing a person]
• [cartoon – adding a 100 ton falling weight]
• What does R2D2 need to know
  – There is a rule
  – When is it OK to break the rule
  – Be aware of the physical environment
  – How the environment can change
  – “Ask not what the environment can do you for you, but what you can do for the environment” – proactivity
  – Act so as to fix/maintain some particular property of the environment or object in the environment
Introductory examples

AUTONOMIC SYSTEMS
Robot Swarm

• A swarm of robots is used for localizing and transporting objects to a goal location
  – e.g. for gardening in a large park or for cleaning a devastated or dangerous area
• In difficult environments with holes, hills, obstacles, . . . the robots have to cooperate
  – e.g. for transporting an object, or
  – Passing an obstacle or climbing a steep hill
Robot Swarm

• Each robot has only limited battery resources and should use as little energy as possible
• Each robot may fail
  – Parts may break done
  – Malicious behaviour may destroy a robot
• Partly failed robots may adapt their task and still provide some functionality
  – such as serving as a communication relay for normally working robots
Science Cloud

- Consists of a collection of notebooks, desktops, servers, or virtual machines
  - running a cloud platform/application
  - communicating over the Internet (IP protocol), forming a cloud
  - providing data storage and distributed application execution

- Every participant is
  - provider and possible user of resources
  - knows about
    - itself (properties set by developers),
    - its infrastructure (CPU load, available memory), and
    - other SCPis (acquired through the network)
Science Cloud

• The science cloud
  – is dynamically changing
    • Participants may dynamically join or leave the cloud or just disappear from the cloud
  – is fail-safe
    • Continues working if one or several nodes fail
  – provides load balancing
    • By parallelly executing applications if the load is high, but not before that.
  – aims at energy conservation
    • virtual machines are shut down or are taken out of the configuration if not required
Cooperative E-Vehicles

• In a few years the e-mobile cars of a big town will be able to communicate with
  – each other and the time tables of the users
  – traffic management servers,
  – battery loading stations,
  – parking lots, etc.

• In such an ensemble, the communicating entities and users may pursue different goals and plans
  – several users may share cars, but have different time tables
  – Loading stations have only limited capabilities; so cars may not be able to use the nearest station for changing the battery
Cooperative E-Vehicles

- Communication and cooperation between the entities of the ensemble leads to better Quality of Service w.r.t.
  - reliability
    - e.g. transport/delivery reliability, adherence to schedules, guarantee to reach the goal, recharging-in-time assurance
  - adaptability to changes
    - e.g. traffic flow, daily personal schedule of the driver
  - predictability of plans
    - confidence in reaching a desired location at a preferred time
Introductory examples

NETWORKS
Communication Networks

• People require to communicate
  – Everywhere
  – Every time
  – From whatever device

• Communication networks are pervasive and at the same time very complex

• They are subject to different problems
  – Congestions
  – Software failures
  – Hardware failures
  – Natural events (e.g., earthquakes and tsunamis)
Communication Networks – An Example

• On September 24 2012, the whole network of the University of Modena and Reggio Emilia, in Italy, was down all day long
• The technicians worked to resume it, but the problem was that there no track of the failure in the logs
• So, the technicians were completely unaware about the failure reasons
• Since the network was large and depended also on external providers, the work was hard and took two days to resume the network
Communication Networks - Control

- Human control is hard to be enacted
- Now, in case of problems, humans try to understand and repair the problem, or to reorganize the network
  - It takes a lot of time and resources
- Can a network try to prevent problems?
- Can a network try to repair itself?
- Can a network try to reorganize itself?
Power Networks

• Current power networks rely mainly on big companies, which generate and distribute energy
• The scenario is quickly changing:
  – Renewable energy (solar panels, wind turbines, …)
  – “Home-made” energy
  – Smart devices
• This opens to a lot of opportunities, but requires an appropriate management
Power Networks – A New Scenario

• People can produce its own energy
• People can sell energy they do not use
  – To their neighbors in a peer-to-peer fashion
• Renewable energy impacts positively on the environment
• Smart devices can help in controlling the energy consumption and in providing us with information
Power Networks - Renewable

- US Nationwide energy dispatch without (a) and with (b) renewable contribution

Power Networks –
The New Scenario’s Issues

• The new scenario introduces some peculiarities
  – The production is “distributed” among a possibly large number of producers (or “prosumers” if they consume energy)
  – The production is subject to external conditions (e.g., weather)
  – Smart devices are better than old ones but must be coordinated

• In general, we have a more dynamic and unpredictable scenario
Power Networks - Control

• But how this situation can be controlled?

• A human control
  – Is difficult (many parameters, autonomous entities, ...)
  – Can be not impartial (big companies are self-interested)

• Can a power network control itself?
Why awareness in ICT systems?

MOTIVATION
What is needed?

• In both cases, for networks’ self management/organization we need:
  – Mechanisms, which can enable the network to act on itself
  – Policies or goals, which leads the networks in taking decisions
Why do we need awareness in ICT systems?

- Huge explosion of the number of computers in our technological environment.
- Nearly all our everyday devices are acquiring computing power and are being networked.
- Being ubiquitous these devices are highly distributed and decentralized.
- Being highly decentralized they are very difficult to manage and maintain stable and functional.
Why do we need awareness in ICT systems?

• Because all our devices are acquiring computing power they can present various forms of intelligence and adaptability.

• Becoming intelligent they could improve the services they offer and offer more innovative and elaborate services.

• To become intelligent the devices have to perceive and understand what is happening in their environment and into themselves.

• They would also need forms of anticipation.
Why do we need awareness in ICT systems?

• Why do we (=citizens) need intelligent systems at all?
  – Cars
  – Smartphones
  – Computer networks, Internet
  – Domestic appliances and embedded environmental sensors and actuators
  – Networks of energy production
  – Infrastructures: roads, traffic lights, public transports, water supply, waste disposal, ...
  – Manufacturing & industrial processes
  – Etc.

• [include illustration]
“Definition”/Circumscription

AUTONOMIC SYSTEMS
Autonomic Systems

- Self-governing, can operate free of external direction
- May be a single entity, or a network of entities acting together to achieve a goal
- Often in complex, dynamic, uncertain environments
- Must balance alternative actions, and select best to achieve goals
- Examples:
  - Refer back to examples used before (pictures)
“Definition”/Circumscription

SELF-AWARENESS
Where does “self-awareness” come from?

- Term first appeared around the start of the 20th century.
- The capacity to become the object of one's own attention.
- The capacity for subjective experience.
Self-awareness

- Not an “on or off” capability!
- Private self-awareness: concerns internal information.
- Explicit self-awareness: concerns me as an object.
- Implicit self-awareness: concerns my subjective experience.
- Also: levels of self-awareness capabilities.
Existing Research Projects

ASCENS
ASCENS

- Self-aware, self-adaptive and self-expressive autonomic components, running within environments which are called “ensembles”.
- The goal is to build ensembles by combining the wide applicability of software engineering with the low management overhead and optimal utilization of resources promised by autonomic, adaptive, self-aware systems.
Existing Research Projects

COCORO
CoCoRo

• Aims to create an autonomous swarm of interacting, autonomous underwater vehicles (AUVs).

• A robotic system in such a complex and unpredictable environment faces new challenges that have not yet been solved.

• For this purpose, collective cognitive capabilities derived from animals (e.g. social insect societies) will be used underwater for the first time.

• Focal tasks are: ecological monitoring, searching, maintaining, exploring and harvesting resources in underwater habitats.
Existing Research Projects

EPICS
EPiCS

• Concerns so-called proprioceptive computing systems
  – Systems collect and maintain information about their state and progress
  – Self-awareness: use info to reason about their behavior
  – Self-expression: adapt behavior to changing circumstances

• Develops new hardware and software platforms

• Application domains:
  – High-end computers
  – Distributed smart cameras
  – Hypermusic
Existing Research Projects

RECOGNITION
RECOGNITION

• Project draws inspiration from human cognitive processes to achieve self-awareness
• Try to replicate core cognitive processes in computer systems:
  – e.g. inference, beliefs, similarity, and trust
  – embed them in ICT
• Application domain: internet content
  – Manage and acquire content in an effective manner by means of self-aware computing systems
Existing Research Projects

SAPERE
Self-aware Pervasive Service Ecosystems

• Computers for handling data and providing services are integrated into an “ecosystem”

• System is extended with
  – methods for data and situation identification
  – decentralized algorithms for spatial self-organization, self-composition, and self-management

• Thus, we obtain automated deployment and execution of services and for the management of contextual data items

• Questions addressed:
  – Can we invent and implement the “add-on” methods that extend the computer system with self-everything?
Existing Research Projects

SYMBRION
SYMBRION

Symbiotic Evolutionary Robot Organisms

• Hundreds of small cubic robots are built and deployed in an environment
• Robots sense each other and the environment and are capable of aggregating into “multi-cellular” organisms
• Aggregation and disaggregation is self-driven, depending on the circumstances: different environments, different tasks

• Questions addressed:
  – Can we build such robots and program the basic behaviors needed for appropriate (dis)aggregation?
  – Can we provide adaptive mechanisms that enable newly “born” organisms learn to operate (sense, move, act, …)?
Existing “Self-Aware” Systems (source of inspiration)

NATURE
Flocking

- Flocking is a kind of self-organising **motion coordination** behaviour of a herd of animals of similar size and body orientation, often moving in masse or migrating in the same direction and with a common group objective.

- Reynolds’s rules:
  1. Cohesion (Flocking centering)
  2. Separation (Obstacle avoidance)
  3. Alignment (Velocity and direction matching)

- Each individual adjusts its position, orientation and speed according to its nearest neighbours.
Ant Foraging

- Ant foraging is the activity where a set of ants collaborate to find food.
- Ants coordinate their behaviour to find the shortest path from the nest to the food.
- Ant colonies use a **stigmergic** communication means, i.e. ants modify the environment by depositing a chemical substance called **pheromone**.
- This pheromone drives the behaviour of other ants in the colony, pheromone concentrations being used to recruit other ants.
- Following the highest pheromone concentration, ants find the shortest path from the nest to the food, and adapt this path when obstacles appear or when food is depleted.
Quorum Sensing

- **Quorum Sensing** process is a type of intercellular signal used by bacteria to monitor **cell density**.
- According to the cell density, a bacteria can coordinate collaborative behaviours, such as coordinated attack.

- Examples:
  - **bioluminescent bacteria** (Vibrio Fischeri) found in some species of squids. These bacteria self-organise their behaviour to produce light only when the density of the bacteria is sufficiently high.
  - **Salmonella** (*Salmonella enterica*). *This bacteria wait until the* density is enough to overcome the body defenses (coordinated attack).
Chemotaxis

- **Chemotaxis** is the phenomenon in which single or multi-cellular organisms direct their movements according to certain chemicals gradient present in their environment.

- Examples in nature include:
  - leukocyte cells moving towards a region of a bacterial inflammation, or
  - bacteria migrating towards higher concentrations of nutrients.
Morphogenesis

- Morphogenesis is the biological process that cause an organism to develop its shape.
- In the biological morphogenetic process some cells create and modify molecules which diffuse, creating gradients of molecules.
- The spatial organisation of such gradients is the morphogenesis gradient, which is used by the cells to differentiate the role that they play inside the body, e.g. in order to produce cell differentiations.
- The behaviour of each cell depends on the relative spatial position of other cells.
Gossiping

• Human social behaviour linked to spreading rumours
• People add their own information to information received from other people, increasing their knowledge and spreading this knowledge further.
• When the process is repeated several times, people start to share the same knowledge that results from the sharing of the knowledge of different people.
Self-Aware Systems in Nature

The Immune System

- Collection of organs and vast numbers of cells responsible for maintaining your health
- Includes fighting bacteria, viruses, fungi
Awareness in the Immune System

- Recognizes to destroy only what is dangerous
- Bacteria in the gut not destroyed
- How does the immune system know to attack the harmful, and not the *self*?
- How does autoimmunity manifest?
- ... and why do most people not get sick when there are autoimmune cells in all of us?
Immune System Awareness

• Immune response is very heavily regulated.
  – Positive reinforcements to support what it is doing
  – Negative feedbacks to prevent potentially harmful actions

• Immune cells recognize pathogens

• Immune cells that recognize other immune cells
Immune System Operation

- Immune response manifests from many cell interactions, and the body signaling damage to itself
Immune System Awareness

- Immune system monitors itself, what it is doing, and whether that is causing harm or not.
- It is **aware** of its own actions, and their consequences.
- Based on this, it can change what it is doing.
- This awareness is...
  - Distributed: cells & organs throughout the body.
  - Decentralized and self-organizing: No one cell/organ dictates
  - Flexible: response can change over time
Existing “Self-Aware” Systems (source of inspiration)

MAN-MADE
The Internet

• The Internet
  – At a low level
    • Autonomous components
    • Local decision making, local communication, sensed local environment
    • Rules that can change over time
    • Packet transmission based on a notion of fairness
  – At a high level
    • “Winfield Test”
    • Anthropomorphise interaction
      – “Does the Internet know who I am?”
DIFFERENT LEVELS OF AWARENESS

Some related general properties
Levels of self-awareness

- The self-awareness capabilities of an individual can be described as being at one or more levels.

- These levels range from very simplistic capabilities to highly complex ones.
Neisser's levels of self-awareness

1. **Ecological self**
   - (Awareness of internal or external stimuli).

1. **Interpersonal self**
   - (Awareness of interactions with others).

1. **Extended self**
   - (Awareness of time: past and/or future).

1. **Private self**
   - (Awareness of own thoughts, feelings, intentions).

1. **Conceptual self**.
   - (Awareness of one's own self-awareness, possession of an abstract model of oneself).
Computational framing of the levels

- Ecological self → Stimulus awareness
- Interpersonal self → Interaction awareness
- Extended self → Time awareness
- Private self → Goal awareness
- Conceptual self → Meta-self-awareness
Some related general properties

COLLECTIVITY/SWARM/
DISTRIBUTEDNESS
“Natural” Complex Systems

- Emergence on multiple levels of self-organization

**complex systems:**

- a) a large number of elementary agents interacting locally
- b) simple individual behaviors creating a complex emergent collective behavior
- c) decentralized dynamics: no master blueprint or grand architect
“Natural” Complex Systems

All agent types: molecules, cells, animals, humans & tech

- The brain
- Organisms
- Ant trails
- Termites mounds
- Animal flocks
- Cities, population
- Internet, Web
- Markets, economy
- Social networks
Common Properties of Complex Systems

- **Emergence**
  - the system has properties that the elements do not have
  - these properties cannot be easily inferred or deduced
  - different properties can emerge from the same elements

- **Self-organization**
  - “order” of the system increases without external intervention
  - originates purely from interactions among the agents (possibly via environment)

- **Positive feedback, circularity**
  - creation of structure by amplification of fluctuations
    - ex: the media talk about what is currently talked about in the media

- **Decentralization**
  - the “invisible hand”: order without a leader
    - **distribution**: each agent carry a small piece of the global information
    - **ignorance**: agents don’t have explicit group-level knowledge/goals
    - **parallelism**: agents act simultaneously
Spontaneous Self-Organization of Human-Made Systems

• Burst to large scale: *de facto* complexification of ICT systems
  – ineluctable breakup into, and *proliferation* of, modules/components

→ trying to keep the lid on complexity won’t work in these systems:
  - cannot place every part anymore
  - cannot foresee every event anymore
  - cannot control every process anymore ... but do we still want to?
Spontaneous Self-Organization of Human Organizations

• Burst to large scale: *de facto* complexification of organizations, via *techno-social* networks
  – ubiquitous ICT capabilities connect people and infrastructure in unprecedented ways
  – giving rise to complex *techno-social* systems composed of a multitude of *human users* and *computing devices*
  – explosion in size and complexity in all domains of society:
    ▪ healthcare  ▪ energy & environment
    ▪ education  ▪ defense & security
    ▪ business   ▪ finance
  – large-scale systems have grown and reached unanticipated levels of complexity, beyond their components’ architects

→ impossible to assign every single participant a predetermined role
Regaining Control of Self-Organization

- ... by exporting models of natural CS to ICT: “(bio-)inspired” engineering

**CS Science:** observing and understanding "natural", spontaneous emergence (including human-caused)

<table>
<thead>
<tr>
<th>ex: neurons &amp; brain</th>
<th>ex: ant colonies / bird flocks</th>
<th>ex: genes &amp; evolution</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>biological neural models</strong></td>
<td><strong>trails, swarms / collective motion</strong></td>
<td><strong>laws of genetics</strong></td>
</tr>
<tr>
<td>binary neuron, linear synapse</td>
<td>move, deposit, follow “pheromone” / separation, alignment, cohesion (“boids”)</td>
<td>genetic program, binary code, mutation</td>
</tr>
<tr>
<td>artificial neural networks (ANNs) applied to machine learning &amp; classification</td>
<td>ant colony optimization (ACO) graph theoretic &amp; networking problems / particle swarm optimization (PSO) “flying over” solutions in high-D spaces</td>
<td>genetic algorithms (GAs), &amp; evolutionary computation for search &amp; optimization</td>
</tr>
</tbody>
</table>

**CS (ICT) Engineering:** creating and programming a new, artificial self-organization / emergence
Some related general properties

PERCEPTION
Awareness requires perception

- Perception is extracting the relevant information from the environment and from itself in order to be able to act appropriately.
- Perception is a difficult task as beings are surrounded by a lot of information and data.
- Perceiving the relevant information depends on the context and the purpose of a task.
- Thus there is a subtle interplay between awareness and perception.
Awareness requires perception

- Perception is a complicated process that requires appropriate sensing mechanisms.
- Perception can require forms of memory, knowledge and learning.
- Thus, perception can involve complicated forms of cognition.
- Awareness and perception allow producing appropriate attention.
Awareness requires perception

• Appropriate attention depends on what you are.
• Each type of intelligent machine and each individual machines can require different appropriate attention.
• Appropriate attention is complicated because it cannot be simply directly programmed it has to emerge from complex interactions between the individuals, their environment, the context, the tasks, their current states, their history, etc.
Some related general properties

INTERNAL MODEL
Internal Models

• A characteristic of all(?) self-aware systems is that they have internal models

• What is an *internal model*?
  – It is a mechanism for representing both the system itself and its current environment
    – example: a robot with a *simulation* of itself and its currently perceived environment, inside itself
  – The mechanism might be centralized (as in the example above), distributed, or emergent

To do: insert a picture
Internal Models

• Why do self-aware systems need internal models?
  – Because the self-aware system can run the internal model and therefore test what-if hypotheses*
    • what if I carry out action x..?
    • of several possible next actions, which should I choose?
  – Because an internal model (of itself) provides the self in self-aware

*Reference Dennett’s model of ‘generate and test’
Impact short-/long-term

SAFETY
Safety

• No system can have pre-determined responses to every eventuality in unpredictable environments
  • example: robots that have to interact with humans
    – therefore no system that works in unpredictable environments can be guaranteed to be safe
  – Self-awareness could provide a powerful solution to this fundamental problem
Safety

• How can a self-aware system be safer (than a system without self-awareness)?
  – Because a self-aware system with an internal model of itself and its environment could*
    1. *Represent* the currently perceived (unforeseen) situation in its internal model
    2. *Run* each possible next action in its internal model *(in a sense *imagine* each course of action)*
    3. *Evaluate* the *safety* of each action
    4. *Choose* the *safest* of those actions, and then actually carry out that action

*a major engineering challenge is to build a system that can do this quickly
Impact short-/long-term

SUSTAINABILITY
Sustainable Futures

• Make Critical infrastructure more adaptive
  – Royal Commission on Environmental Pollution
  – Tragedy of the Commons not inevitable

• Take into account
  – Social arrangements of citizens
  – Attributes of the infrastructure with which the interact
  – Context of institutions

• Adaptive Institutions
  – Individuals, ICT-enabled devices and institutions are deeply entangled
  – ICT devices can be equipped with social awareness and can participate in the collective endeavour
  – Out of the entanglement new structures can emerge
  – People retain the power to self-organisethese structures

• Computational Sustainability
  – There is a reason whtElinorOstrom won the Nobel Prize for Economic Science
    – empowering individuals with collective awareness
Impact short-/long-term

PHILOSOPHICAL
Philosophy

• The conception and implementation of self-aware systems might have philosophical implications
  – If self-aware systems are, in some way, models of living systems then could we gain insights into self-awareness in living systems by testing such models?
  – Is self-awareness the first step toward long-term goals of artificial theory-of-mind, and machine consciousness?
  – Could we gain *ontological* insights by asking questions such as, at what point does a self-aware system make the transition to a self-determining autonomous agent, i.e. ‘being’
Open Issues

QUESTIONS & CHALLENGES
Current research questions and challenges

- Dilemma of wishing to make our designed artefacts autonomous but not too much (safety).
- To have a metrics to measure properties related to awareness, autonomy.
- We do not know how to engineer self-organization and emergence.
- We do not know how to cope with autonomy and variability. Dilemma of system stability and reliability incorporating randomness and variability.
- How to design and implement self-aware systems?
- What kind of tools and methodology can we use here?
- Is it ethical to build self-aware systems?
- Can we build autonomic self-aware systems that behave in an ethical way? Related: legally correct behaviour, behaviour compliant with some set of rules and regulations.
- What makes known natural systems self-aware?
- Describing the scope of the future behaviour of a self-aware system.
Current research questions and challenges

- Predicting the behaviour of autonomic systems and their interactions with the environment.
- How to ensure safety and security of autonomic self-aware systems? How to differentiate malicious from benign behaviour?
- What does the system theory of autonomic self-aware systems look like?
- How to build an autonomic self-aware system that would last 100 years?
- To what extent can Big Data be treated as an autonomic self-aware system?
- Can you separate an autonomic self-aware system from its environment?
- In what sense is human and machine self-awareness different? What implications do these differences have on developing them?
- How can we draw inspiration from human self-awareness for designing machine self-awareness?
- How to do the second order design needed in autonomic self-aware systems?
- Will autonomic self-aware systems develop their own medical science?
- Goal: build an autonomic self-aware energy production system.
- Goal: build a smart city / computer network / communication network.
With thanks to

These slides were designed at the Awareness Slides Factory by

• Giacomo Cabri
• Rene Doursat
• Jose Luiz Fernandez Marquez
• Jose Halloy
• Peter Lewis
• Jeremy Pitt
• Mark Read
• Alan Winfield
• Martin Wirsing