ESSAI-2024
Self-Governing Multi-Agent Systems
L1/10: Introduction to SGMAS

Jeremy Pitt and **Asimina Mertzani**
Department of Electrical and Electronic Engineering
Imperial College London

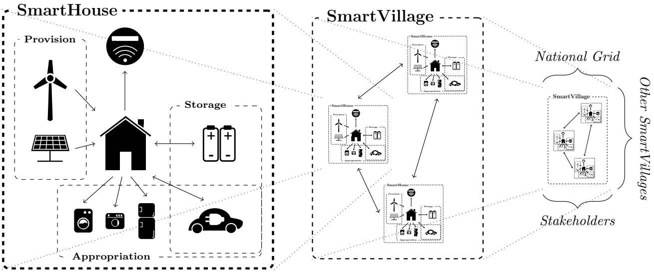
IMPERIAL

Aims and Objectives

- Aims
  - Through the use of motivating examples, introduce the concepts and building blocks for the study of self-governing multi-agent systems
    - Agents, multi-agent systems, self-organisation, socio-technical systems, social influence, learning, hybrid systems
- Objectives
  - Understand the overall research programme in the development of (next-generation) cyber-physical and socio-technical systems

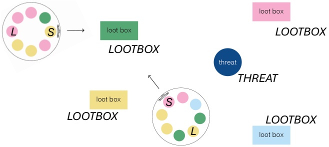
Example: 'Real World'

- Community Energy Systems



- Issues of agency, self-organisation and governance
  - Delegated – Cyber-physical: humans out of the loop
  - Programmed – Socio-technical (1): humans on the loop
  - Interactive – Socio-technical (2): humans in the loop
  - Attentive – Social: humans are the loop

- The *Megabike* Scenario



- Multiple interacting 'games', repeated, repeated
  - Voluntary association, entrenchment of rules
  - Role assignment (direction, steering), admission and exclusion
  - Linear Public Goods game (literally, free-riding)
  - Resource distribution (appropriation from lootboxes)
  - Monitoring, sanctions and punishment
  - Collective risk dilemma (avoiding existential threat)
  - Cooperation and competition between *megabikes*

The Agent Abstraction

- Highest level of abstraction
  - Ownership
    - Delegated responsibility for task
    - "Who did what that affected me"
  - Intelligence
    - Sufficiently 'large' number of internal state to be **not** predictable (deterministic) state machine
    - Exhibit some form of 'intelligent' behaviour
  - Asynchrony
    - Asynchronous communication with other agents/users
- Lowest level of abstraction
  - Embedded software process
  - Encapsulates some notion of state
  - Communicates by message passing

A Simple Agent

- In Qu-Prolog (multi-threaded Prolog)

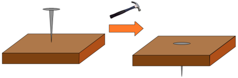
```
...
thread_fork( Name, launch(Name) ),
...

launch( Name ) :-
    thread_handle( Handle ),
    register(Name) ->> oracle:delphi,
    registered <<- _,
    make_agent_profile( Name, Handle ),
    agent_event_loop.

agent_event_loop :-
    Msg <<- From,
    process_message_a( Msg, From ),
    agent_event_loop.
```

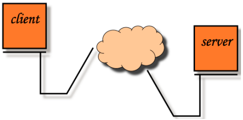
The Multi-Agent Abstraction

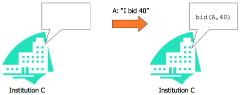
- Distributed (object-oriented) systems
  - physical distribution of data and methods
  - tightly coupled
  - location transparent
- Multi-agent systems
  - logical distribution of responsibility and control
  - loosely coupled
  - location significant
- Organization of individual and collective intelligence in context of:
  - differing functionality and diverse knowledge
  - cooperation, coordination and competition
  - planning and decision making wrt. individual and joint goals
- **Communication**

How To Communicate —

- — with physical objects
  - Change the state of the physical world
  - Given the ‘ideal’ physics (physical capability)
- — with software objects
  - Change the state of an object in the physical world
  - Given ‘appropriate’ programming language semantics and the semantics of the call
- — with words (as used by agents)
  - Change the state of the the *conventional* world
  - Given ‘validity’ of the of the action: *institutionalised power*



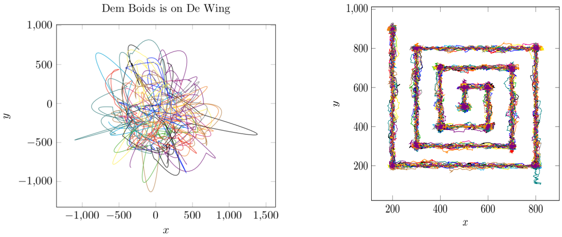




Institutionalised Power

- Searle: Speech Act Theory
  - X counts-as Y in context C
- Jones and Sergot (1996): Formal Characterisation of ...
- A standard feature of any norm-governed system whereby designated agents, acting in assigned roles, are empowered to create or modify facts of special significance conventionally agreed within the context of an institution
- This matters
  - Contract-net protocol and legal systems
  - Socio-technical systems and 'on behalf of'
- How To Do Things ... with intelligent agents
  - Specify which agent is empowered to do what actions
  - Compute how 'meaningful' or 'valid' actions change rules and roles (*If* certain conditions hold)
- This is **Self-Organisation**

- Complexity and Emergence
- Biological Systems
  - Structures: cellular automata
  - Patterns: animal fur
  - Movement: swarm behaviour, e.g. Boids
    - Separation: move to avoid over-crowding local flockmates
    - Cohesion: move towards average position of local flockmates
    - Alignment: move towards average heading of local flockmates
    - Synchrony: move towards average velocity of local flockmates

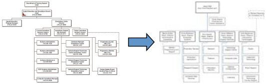


Properties

- These models:
  - Homogenous ‘agents’
  - Behaviour is hard-wired
  - Communication affects ‘cognitive’ state
  - No conventional rules
  - Emergence by complexity
  - Lightweight (points or particles)
- But we want:
  - Heterogenous agents
  - Behaviour is soft-wired (learning)
  - Communication affects ‘social reality’
  - Conventional rules (mutually agreed, mutable)
  - Emergence by complexity and plan
  - Heavyweight (separate threads, processes, machines, . . .)

Self-Organisation – Social Systems

- Focus on "socially constructed reality"
  - Teamwork
  - Business Process Re-engineering and Change Management
  - Role assignment
  - Legal proclamation



Organizational Adaptation

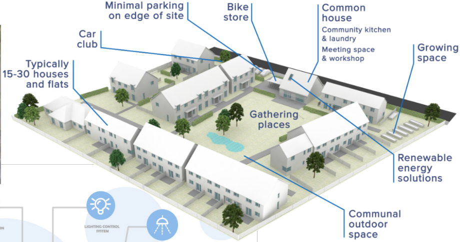
Role Assignment
(White smoke from the Sistene Chapel of St. Peter's Basilica announces a new Pope)

Legal Proclamation

Self-Governance

- == The self-determination of social arrangements
  - Social arrangements: the set of rules, roles, structures, procedures, policies, norms, conventions, contracts or laws that individuals in a group voluntarily agree to comply with, in order to hold each other accountable to that group
  - Self-determination: processes by which social arrangements are selected, modified and applied by those individuals who are affected by them
- Multi-Agent Systems
  - The individuals in the group
  - Embedded in a physical environment
  - Socially-constructed conceptual resources
- But there's more...

Another Example: Co-housing Community



Socio-Technical System $\Rightarrow$ NOT STATIC

How Do We Deal With This?

Let's define socio-technical systems...

Socio-Technical System

Committee	Cloud Systems	Smart Cities
		
Social System	Technical System	Socio-Technical System

Socio-Technical System - Problems

- Individuals have different (conflicting) objectives
- The population changes
- The environment changes
- The problems that they have to deal with change (or new problems appear)

Socio-Technical System - Requirement

- Requires **balancing out conflicting drivers** of the individuals and systems involved
- Requires **adapting individually and collectively** to the changes undergoing in each of those
- This requires **new rules, values, structures and mechanisms** (e.g. new social arrangements) for (self-)organising themselves

⌄

Need for Learning (to change, to adapt, ...)

Tools for Learning

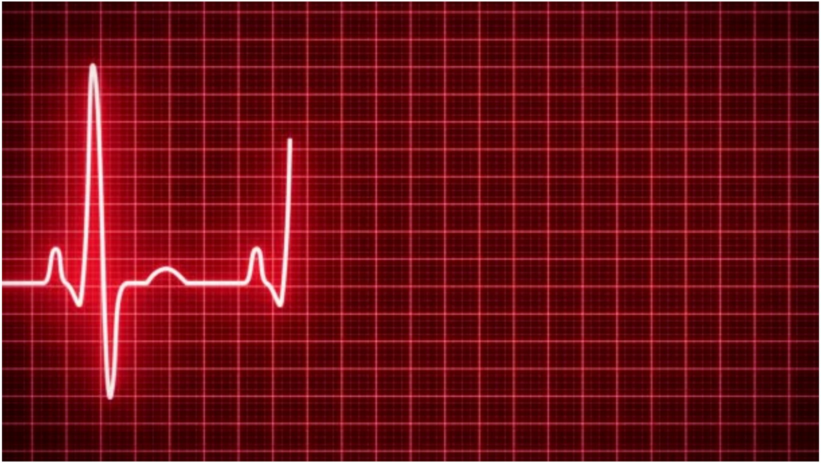
- Learning from humans (socially-inspired computing methodology)
- Learning for humans (machine learning – reinforcement learning)

Learning from/for Humans

- Learning from Humans:
  - Use a theory of **social influence** to enable a group to reach a collective agreement on a qualitative assessment (Mertzani et al., 2023).
  - Use a theory of **social influence** to enable a group to develop explanatory adequacy (Mertzani et al., 2022).
- Learning from and for Humans:
  - Draw inspiration from **psychoacoustics** and combine this with **reinforcement learning** to achieve the ethical self-regulation of a socio-technical system (Mertzani et al., 2024).
  - Use a **governance mechanism used in classical Athenian deliberative assemblies** and combine this with (multi-agent) **reinforcement learning** to enable a group form consensus (Mertzani et al., 2023).

But

The world is not linear!



Learn to Innovate

Especially during digital transformation humans and technology co-exist.

⌄

This causes unprecedented changes.

⌄

However, existing solutions might not be sufficient.

⌄

Therefore, to adapt we need to learn to systematically **innovate**.

How to Innovate?

We propose the **co-production between human and AI** to support innovation and move towards more **sustainable** self-governing multi-agent systems.

Summary

- Agents & Multi-Agent Systems
- Self-Organisation & Social Construction
- Socio-Technical System
- Learning & Innovation

$\Downarrow$

Self-Governing Multi-Agent Systems

- Cyber-physical systems
- Socio-technical systems
- Agent-based social simulation

Course Roadmap

- Lectures
  - Introduction (this lecture)
  - Sustainability
  - Distributive Justice
  - Knowledge Management
  - Constitutional Choice
  - Social Influence
  - Self-Regulation
  - Consensus
  - Innovation
  - Social Implications

⌄

"The principles and practices that enable autonomous agents to live together, better;
whether the agents are made of carbon, silicon, or both"