



KnoDoBe: Cognition beyond problem solving

David Hales – www.davidhales.com

University of Szeged, Hungary

20/01/14



FuturICT.hu

- Two year project (2013-2014)
- Joint Hungarian / ESF funding
- Involves researchers from several Hungarian institutions
- Work within the scope of the previous FuturICT FET Flagship proposal
- One line of work looks at novel incentives, economic models and design patterns for radically decentralised information systems
- Lead: Mark Jelasity, University of Szeged
- <http://www.futurict.szte.hu/en>

New research directions in agency

- Workshop at Imperial College, London (May 2013)
- Brainstormed new research directions for computational notions of agency.
- 30 researchers working in agency mainly from the UK.
- Funded by the Non-equilibrium Social Science (NESS) FET Coordination Action.
- <http://www.nessnet.eu/>
- Draft towards a roadmap for NESS. Public deliverable will appear on website soon. If you want a copy sent to you then e-mail me.
- Here I present some potentially relevant ideas to the KnoDoBe call.
- I only found out about this call from the work program. I have not been involved in any prior inputs for formulating this call.

Outcome of the workshop

- We identified the following timely and important future application domains in which agent research could have significant impacts:
 - Financial and political stability
 - Environmental sustainability
 - Ethical and legal frameworks
- We note each of these domains evidences *collective action problems involving multiple agencies situated in complex dynamic global networks*.
- Based on this we proposed eight novel future agent research areas encompassing aspects of:
 - ICT engineering
 - Policy design
 - Legal and ethical frameworks
- Today I will overview just three.

Artificial Morality

- Semi-autonomous computational artifacts form parts of social systems.
- Consider automated trading algorithms or lethal autonomous robots.
- Also semi-autonomous surveillance systems.
- How to productively program these artifacts?
- Moral / Ethical frameworks could be used as a basis for design patterns.
- This could improve performance, predictability and trust.
- Already distributed systems use primitive forms of “moral code”.
- For example – tit-for-tat reciprocity in Bittorrent – the golden rule
- How can moral design patterns be created, represented and used in cyber-social systems?
- How do these relate to law?

Rationality 2.0

- Understanding and enabling social systems requires some model of agent action.
- Rationality can provide such a model yet it has often been narrowly defined as individual utility optimization (in economics for example).
- Broader notions of rationality, both of people and embodied in artifacts could focus on:
 - Comparative rationality: specify, compare and relate different forms.
 - Collective and individual rationality: how can collective or social rationality support individual rationality and vice versa.
 - Evolution of rationality: when do evolutionary processes produce rational outcomes (both individual and collective)
 - Rational history: what kinds of historical processes be viewed as rational?
 - Forming rational goals: how can agents form rational goals (or preferences)

Agent-Environment boundary

- The boundary between an agent and it's environment may be fluid.
- Agency may be projected by an observer onto many kinds of phenomena.
- The choices one makes identifying and bounding agency in a social or technological system has implications for the potential of that system.
- There are number of interesting topics that relate to this:
 - The extended agent mind: dynamic agent-environment boundaries (e.g. in robotics). Consider a robot building new actuators or sensors from materials in it's environment.
 - Inducing agency from data: automatically induce the previously unrecognized agents, their goals and action from big data? Conspiracy theories and imaginary friends.
 - Setting the agent-environment boundary: finding meaningful and useful boundaries for different kinds of social phenomena (riots, consumption patterns, financial markets). Can new boundaries reframe the problem? Suggest new solutions?





Thank you for your attention!

This work was partially supported by the European Union and the European Social Fund through project FuturICT.hu (grant no.: TAMOP-4.2.2.C-11/1/KONV-2012-0013).