Collective Intelligence
A Multiagent Systems Perspective

Michael Rovatsos
mrovatso@inf.ed.ac.uk

Collective Intelligence

○ In theory: cooperation, competition, and coordination produce some **emergent** super-individual capabilities that transcend those of the individual

○ In practice (currently): crowdsourcing, folksonomies, social computation, human-based computation

○ **Emergence:** "you throw the pieces of a watch on a table and suddenly, there's a watch!"

○ Tempting: "emergent macro-scale phenomena will do our work for us". But how do you **design** for intelligence to emerge? (And what do you do when it does?)

Agents

○ Ambiguous, overused, overhyped term - let's just assume it means something like "AI with an integration focus"

○ Single-agent systems: integration of capabilities in one embodied, situated entity operating on behalf of a user

○ Multiagent systems: integration of several agents (esp with different goals) in the same environment

○ The "different goals" bit is the **genuinely** new element

Data

- Petascale data processing is the name of the game today, massive application benefits

- Data is not "intelligent" in itself, it is human interpretation that makes it useful

- Analysing data relies on human expert, this part often not explicit in systems design

- Effectively, the intelligent agents in the "data-driven" world of computing are still human experts

Design

- Human (and machine) intelligence as reflected in data says nothing about what to do with it

- Describing purposeful computation and aiding it with automated support is missing

- Top-down perspective of specification and reasoning vs. bottom-up perspective of data-driven computation

- Agents provide right technology for this, but largely disconnected from "data world"

Incentives

- The missing link between data and behaviour toward it – the key to design for collective intelligence

- Economics based models currently offer the only rigorous framework (game theory etc)

- These models don't account for many types of human behaviour (esp. in constantly changing Web world)

- Agents: modelling systems with potential conflict of interest, simulating human behaviour

Semantics

- **Meaning = Cognition + Incentive**
    - Human intelligence of individuals interprets information with a purpose, and produces knowledge

- Evolution of collective intelligence mediated by configuration of different participants' incentives and interpretations

- Emergent collective knowledge reflected in evolution of semantics of symbols used

- "Evolving Symbol Systems Hypothesis": Incentive-aware rules for manipulating semantics sufficient to achieve collective intelligence

Challenges

- Multi-level modelling and evolution of representations, linking "deep" to "shallow" reasoning

- Reasoning about incentives and behaviour, but engaging with humans to learn about these

- Lack of automation support for users and application designers: tapping into new data is useful, but not enough

(My own) Examples

○ Computing solutions to interaction problems from high-level problem specifications

○ Mining opponent and context models from executions of agent interaction protocols

○ Collaborative and competitive methods for distributed machine learning

○ Developing executable declarative languages for implementing social reasoning in agents