Intelligent Management of Data Driven Simulations to Support Model Building in Social Sciences

Catriona Kennedy and Georgios Theodoropoulos
School of Computer Science
University of Birmingham UK

Dynamic Data Driven Application Systems - DDDAS 2006
International Conference on Computational Science 2006 (ICCS 2006)
University of Reading, UK
May 28-31, 2006

Overview

- Part 1: Role of AI in DDDAS management
  - Symbiotic Simulation and cognition
  - Conceptual Architecture
  - Autonomous agent vs Assistant agent
- Part 2: DDDAS and social sciences
  - Why AI in social sciences?
  - How can DDDAS be applied to social systems?
  - A case study

Symbiotic Simulation and Cognition

- Cognition is a process of anticipation, continual update and revision
- Internal simulation can be used to anticipate events and generate "what-if" scenarios;
- Adaptation and model revision (or even ontology revision) may happen as a result of interaction with reality. This is the "data-driven" aspect.
- For example:
  - an agent expects to see an object on the table (e.g. cup);
  - expectancy causes direction of sensors and focus of attention;
  - reality of the object may be slightly different from expectancy (e.g. cup is stuck to the table)
  - new data leads to further questions and what-if scenarios (e.g. is the table sticky?)

What Simulation Architecture would AI fit into?

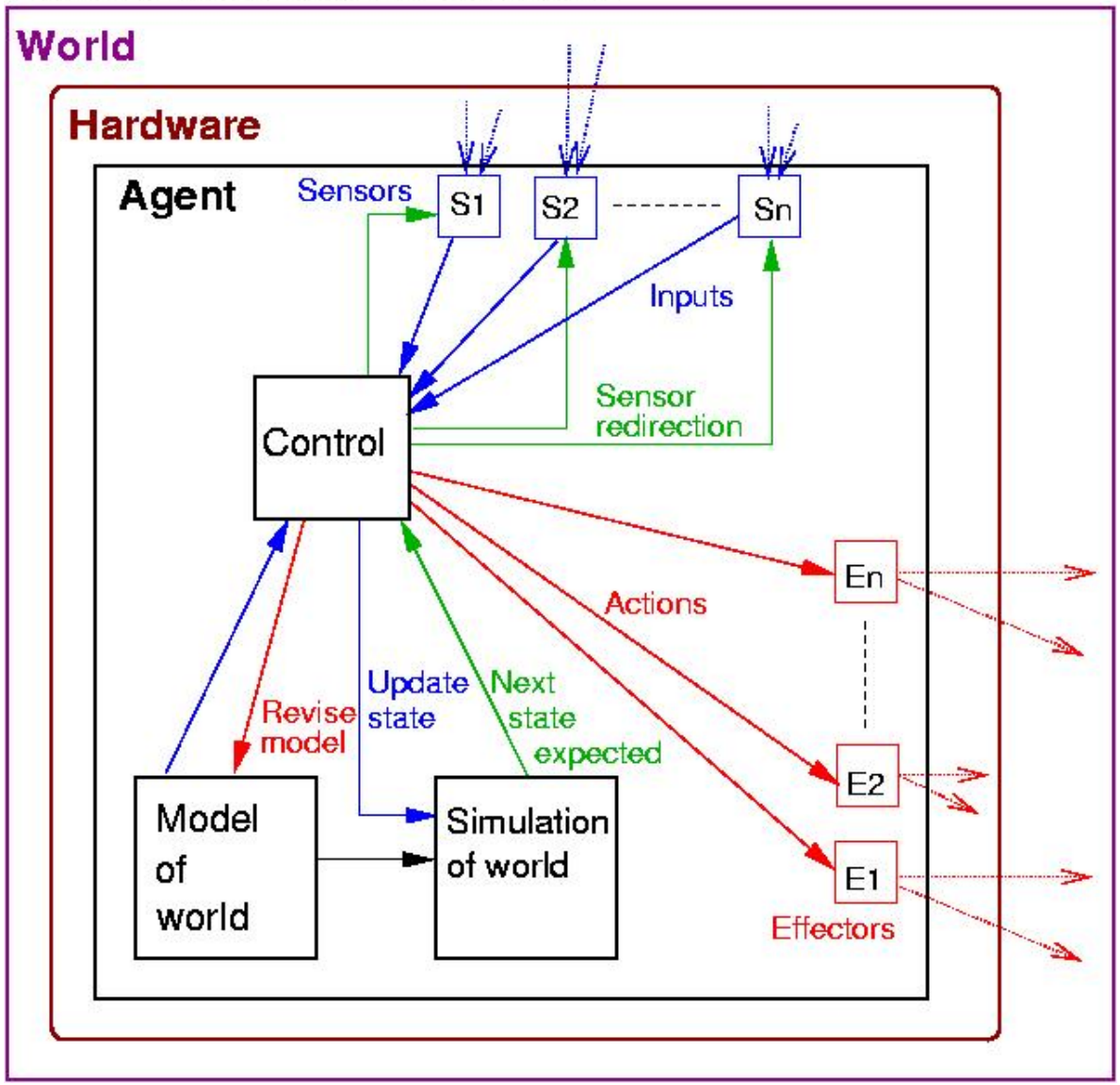
Autonomous agent:

- Simulation is internal and serves the survival of the agent (e.g. predict failure states of its hardware);
- Model on which simulation is based could be revised without human intervention as a result of agent adaptation;
- Agent is situated: it has direct control over sensors and effectors - making exploration easy.

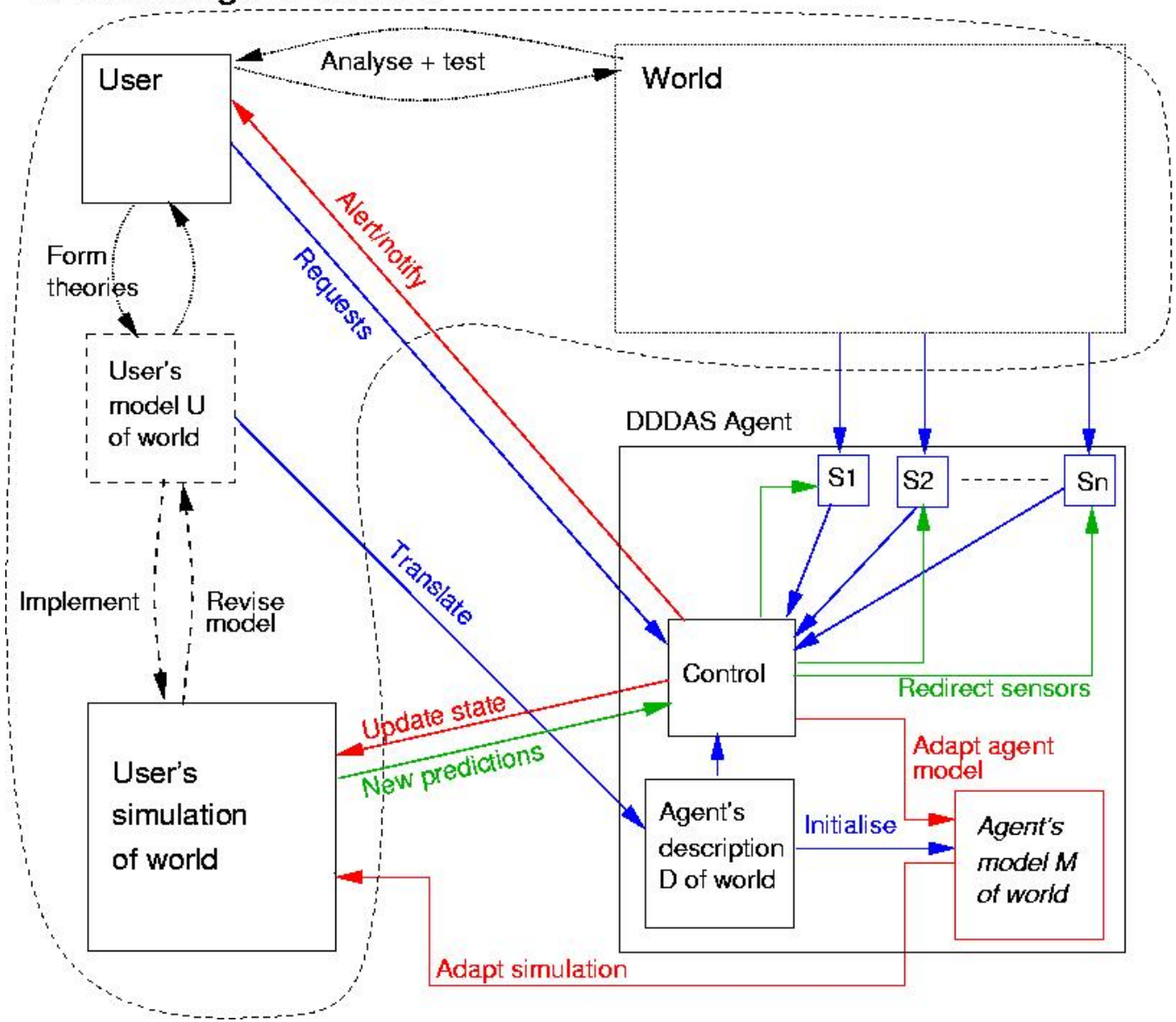
Assistant agent:

- Simulation is external and helps with predictions or "what-if" scenarios for human scientists or decision-makers;
- The model is based on a theory of the world (e.g. climate, society) and is *not* expected to be revised without human interaction;
- Agent is *not* situated: "sensors" are interfaces to software tools such as database query, data mining etc. (no "effectors")

Autonomous agent scenario



Assistant agent scenario



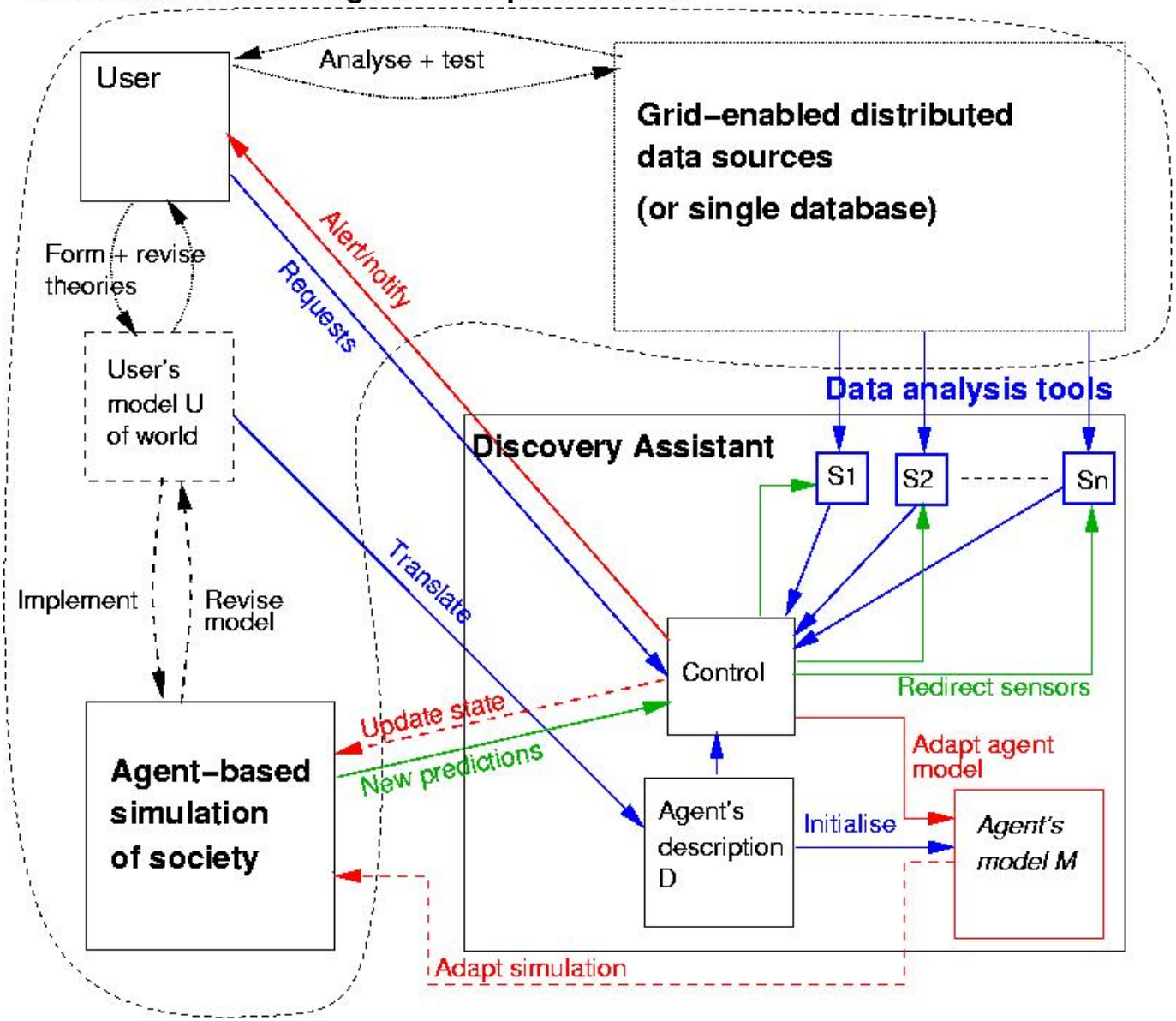
What is being simulated?

- An actual observed system (an instance):
  - simulation states are expected measurements of the real system;
  - direct update from measurements to simulation;
  - simulation may run concurrently with observed system
- A class of systems with similar properties (more likely in social sciences):
  - simulation states are abstract states applicable to this class;
  - model may be revised as a result of generalisations from data;
  - data may be collected from multiple instances of similar systems.

AI in Social Sciences

- Policy decision making requires understanding of a complex system;
- Data-driven simulation can assist with model building and revision;
- Agent-based simulations can predict future states, given current state - or more about present state, given partial state;
  - Also possible to run "what-if" scenarios for policy actions: start with a hypothetical state;
  - Simulated "agents" can represent individuals, groups, organisations etc.
- Note that there are TWO kinds of "agent":
  - the software agent that is building and testing the model (by data-driven simulation)
  - the agents in the world that are BEING MODELLED (which may include other software agents).

Modified Assistant Agent – not pure DDDAS



The AIMSS Project

- Funded by the UK e-Social Science Programme
- A collaboration between Computer Science and Public Policy (The Institute of Local Government Studies and the Centre for Urban and Regional Studies)
- Aim: To explore the feasibility of DDDAS for the prediction of complex Public Policy outcomes
  - Housing Policy scenarios
- In the longer term the project is oriented towards diagnostic interventions and resource management problems in public policy
- *http://www.cs.bham.ac.uk/research/projects/aimss*

Case Study: Housing Policy

- *Problem 1:*
  - Current housing market models are too simplistic;
  - assumptions may not hold in all scenarios
- *Problem 2:*
  - Understanding micro-level behaviour is a multi-dimensional problem;
  - incomplete data;
  - expensive data acquisition;
  - need assistance in determining what kinds of micro-level data are significant for policy goals.

Proposed Solution

- Use agent-based simulation to represent residents making decisions on whether to move house and where.
- Predicted states of the social simulation can be tested by analysing the data;
- The "predictions" are states that would be expected to exist now if the model's assumptions are true;
- Data analysis and mining tools are used to inquire whether the predicted state is actually true;
- DDDAS aspects are as follows:
  - *If insufficient data is available, the discovery agent can suggest new kinds of data that are required in future surveys (Problem 1);*
  - *Persistent discrepancies between the simulation predictions and the results of the data analysis prompts the discovery agent to suggest model revisions (Problem 2)*

Challenges

- Need for suitable ontologies for Social Science data sets
  - What is the meta-data D and how does the Discovery Assistant use it?
- How to specify what is *important* with regard to the policy goals; (e.g. affordable housing).
- Need for suitable data mining and analysis tools.
  - Need for fusion and summarisation;
- Semantic Grounding (longer term): semantics should be developed by interaction with the world and learning (exploration and adaptation)
  - BUT: the discovery assistance architecture may already have the potential to solve this - due to the symbiotic nature of the interaction between simulation and reality

Summary

- DDDAS and in particular symbiotic simulation may be used by a cognitive agent;
- Two kinds of scenario in which an agent can manage a simulation: autonomous agent and assistant agent;
- Modification of assistance scenario: Discovery assistance ine-Social Science: instead of pure “data-driven” simulation, provide interface to data analysis and mining tools;
- Semantic grounding is a fundamental problem BUT the symbiotic simulation architecture may itself provide a way of solving it.