

---

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

**Georgios Theodoropoulos and Catriona Kennedy**

School of Computer Science  
University of Birmingham UK

PADS 2006

Singapore

May 24-26, 2006





# The AIMSS Project

## Adaptive Intelligent Model-Building for the Social Sciences - using Symbiotic Simulation

---

- Funded by the UK e-Social Science Programme
- A collaboration between Computer Science and Public Policy in Birmingham
  - The Institute of Local Government Studies
  - 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



# DDDAS Management

- Management of a DDDAS system involves the following decisions (among others):
  - What **new data** should be absorbed, **when** and **how**?  
**When** should attention be focused on a **particular data source** for absorption into the simulation?
  - **When** should newly **observed values** that are outside of the normal range **replace predicted ones** in the same context and **when** should there be some **revision of the model** on which the simulation is based, and **how**?
- Can the above be automated and how?



# Symbiotic Simulation and AI

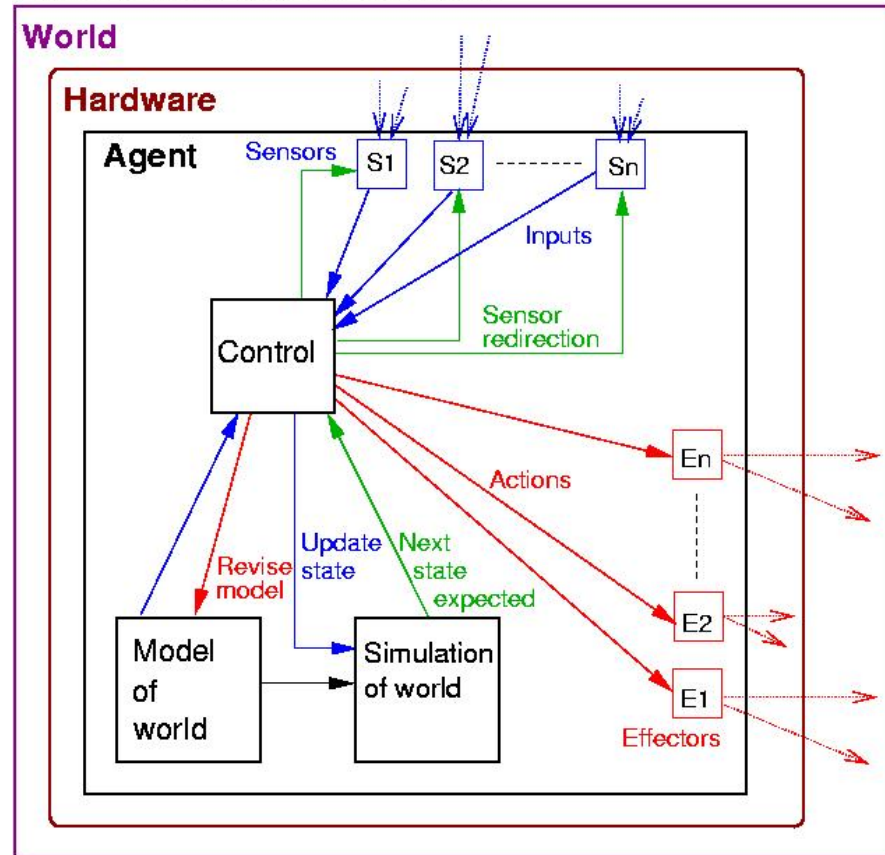
- **Cognition** is a process of anticipation, continual update and revision: predicting future states of the environment and continually adjusting them to new events is an important requirement of cognition.
- **Internal simulation** can be used as a form of “what-if” reasoning for an agent, to anticipate events and generate “what-if” scenarios;
- The degree and type of learning and revision may vary: from purely hardwired (little or no revision of the initial model) to fully “discovered”: **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?)
- Such an agent can also manage a DDDAS system so that it assimilates aspects of the reality and revises existing models as necessary



# Autonomous Agent

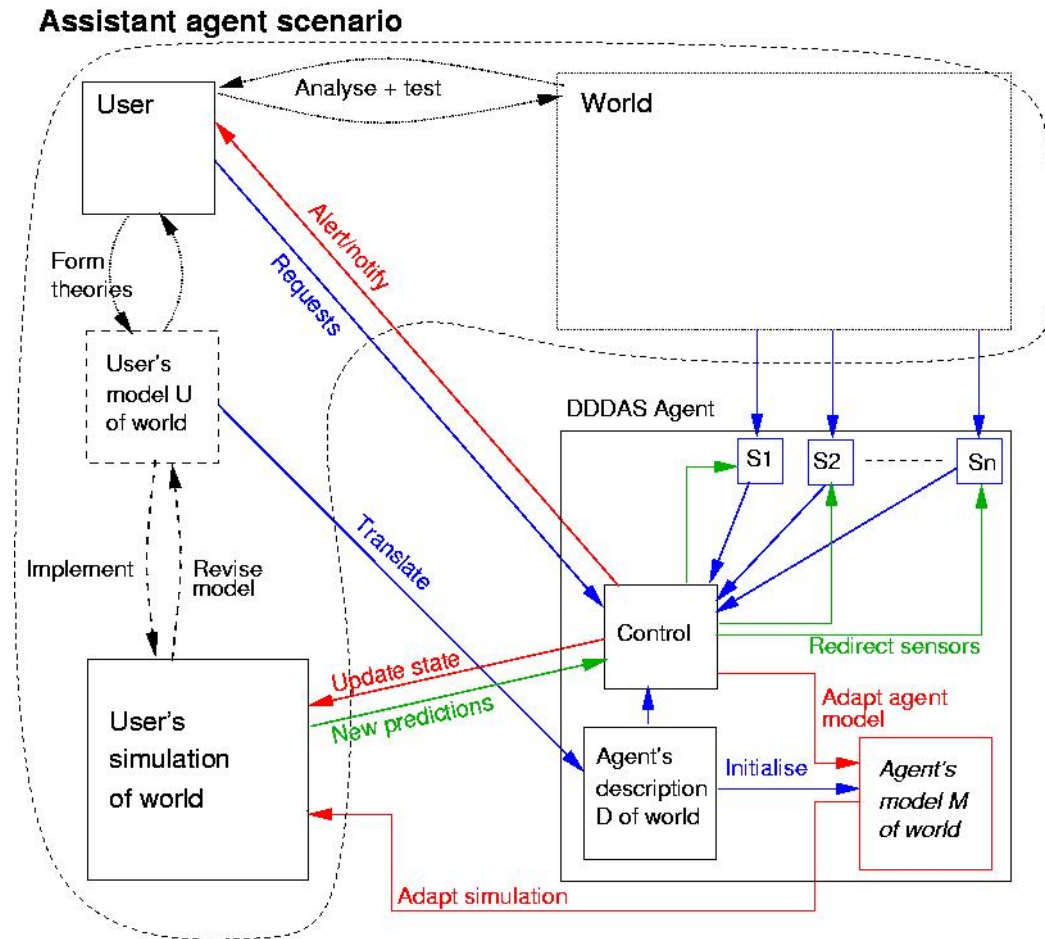
- Any robotic or software agent that must continue operation without human intervention
- Agent is **situated**: it has direct control over sensors and effectors - making exploration easy.
- Simulation is **internal** and serves the **survival of the agent** (e.g. predict failure states of its hardware)
- Model on which simulation is based could be **revised without human intervention** as a result of agent adaptation (very ambitious)
- Predicted states can determine importance of data and direct sensors:
  - Evaluation, uncertainty or anomaly direction

Autonomous agent scenario



# Assistant Agent

- Most DDDAS systems are used to help humans understand a complex system and make decisions
- 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”)



# 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.



# DDDAS in Social Sciences

- **Policy decision making** requires understanding of a complex system
- **Agent-based simulations** may be used to model social systems and can assist in the prediction of the effects of candidate policies or proposed interventions
  - can predict **future states, given current state** - or more about present state, given partial state
  - possible to run **“what-if” scenarios** for policy actions: start with a **hypothetical state**;
  - Simulated “agents” can represent individuals, groups, organisations etc.
- **Data-driven simulation** can assist with **model building and revision**



# 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.

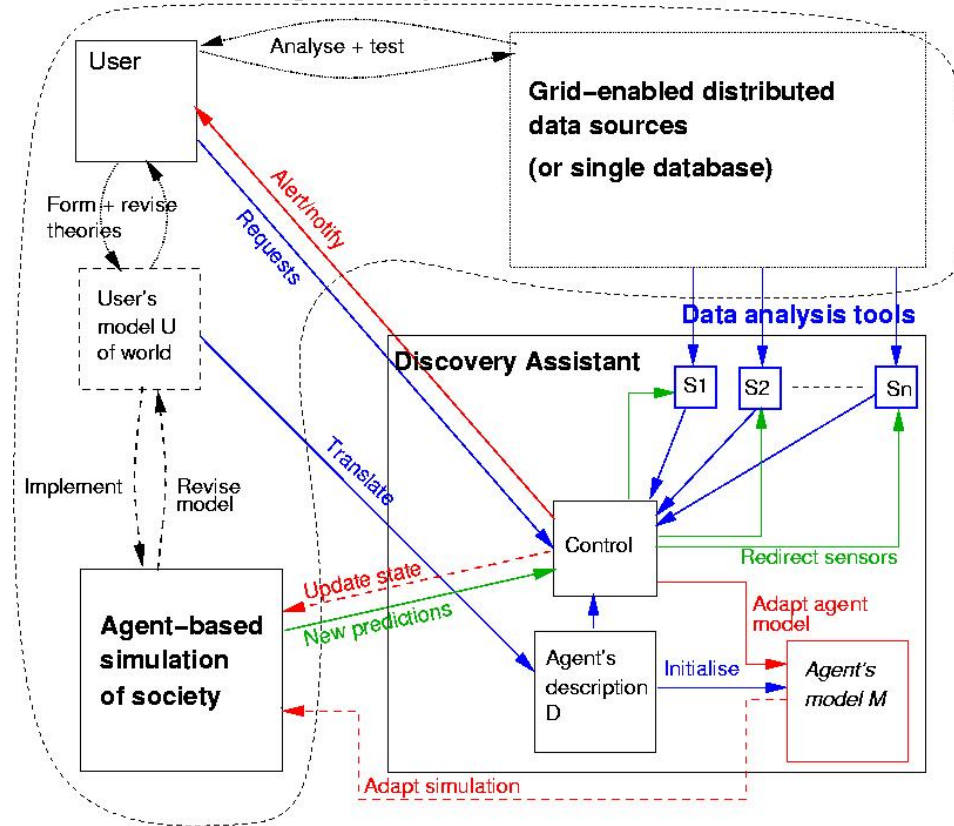


# Proposed Solution

- 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 (P2);
- **Persistent discrepancies** between the simulation predictions and the results of the data analysis prompts the discovery agent to **suggest model revisions** (P1)

Modified Assistant Agent – not pure DDDAS



# Challenges and future work

- Specification of  $D$ , generation of  $M$  and automated adaptation of  $U$
- Need for autonomous learning and adaptation: involving independent interaction with the world to check the validity of the model and to revise it as necessary
- Multiple ontologies representing alternative descriptions of the world are advantageous (problems can be detected if there are significant disagreements between model predictions)
- Cooperation between heterogeneous agents acting in different domains and levels of abstraction is important in order to exploit diverse sources of information
- In the e-Social Science scenario:
  - 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;



---

# More Info

*<http://www.cs.bham.ac.uk/research/projects/aimss>*

