

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

Catriona Kennedy and Georgios Theodoropoulos

School of Computer Science, University of Birmingham, UK  
{cmk, gkt}@cs.bham.ac.uk

**Abstract.** Artificial intelligence (AI) can contribute to the management of a data driven simulation system, in particular with regard to adaptive selection of data and refinement of the model on which the simulation is based. We consider two different classes of intelligent agent that can control a data driven simulation: (a) an autonomous agent using internal simulation to test and refine a model of its environment and (b) an assistant agent managing a data-driven simulation to help humans understand a complex system (assisted model-building). In the first case the agent is situated in its environment and can use its own sensors to explore the data sources. In the second case, the agent has much less independent access to data and may have limited capability to refine the model on which the simulation is based. This is particularly true if the data contains subjective statements about the human view of the world, such as in the social sciences.

For complex systems involving human actors, we propose an architecture in which assistant agents cooperate with autonomous agents to build a more complete and reliable picture of the observed system.

**Keywords:** agent, cognition, decision support, fault-tolerance, simulation, social sciences.

## 1 Introduction

In the physical sciences, “dynamic data-driven application simulation” (DDDAS) is a method where data from a physical system is absorbed into a simulation of the system[1]. If DDDAS is applied to an artificial system, the simulation may influence the real physical system (for example, to optimise or adapt it). This is called “symbiotic simulation” because of the mutual benefits of the simulation and the physical system on each other. Examples include semiconductor component testing [2].

The management of a symbiotic simulation has many similarities to a cognitive process. In natural cognitive systems, anticipation is used to direct perception and focus attention on a particular object (e.g. a cup on the table). The reality of the object can modify the further expectancy which may in turn direct attention to further objects or data sources that would not have been anticipated initially (e.g. if the cup is cracked or stuck to the table).

In a symbiotic simulation, predictions can be used as a basis for action on the observed system, just as anticipation does in a cognitive system. If direct action is not appropriate (e.g. because the observed system is not artificial) the predictions can be used to focus on relevant data sources to be assimilated. We will call an agent that controls such a process a “DDDAS agent”.

### 1.1 Autonomous Agents

An autonomous agent is any robotic or software agent that must continue operation without human intervention. It can use symbiotic simulation to predict states of its environment (or its own components) and to adapt to its environment. We will call this kind of DDDAS agent an “autonomous DDDAS agent” and is a subset of general autonomous agents. Its architecture is shown in Figure 1. Sensors and effectors in the diagram are schematic and may involve software as well as hardware. The predicted state can determine what kind of data is important for subsequent absorption and this requires direction of sensors. Sensor and effector activation can involve a complex translation between high level directions (thick arrows) and the actual low-level measuring or making adjustments multiple thin arrows). Similarly the absorption of data (agent input arrows from sensors) can involve a non-trivial fusion and summarisation process.

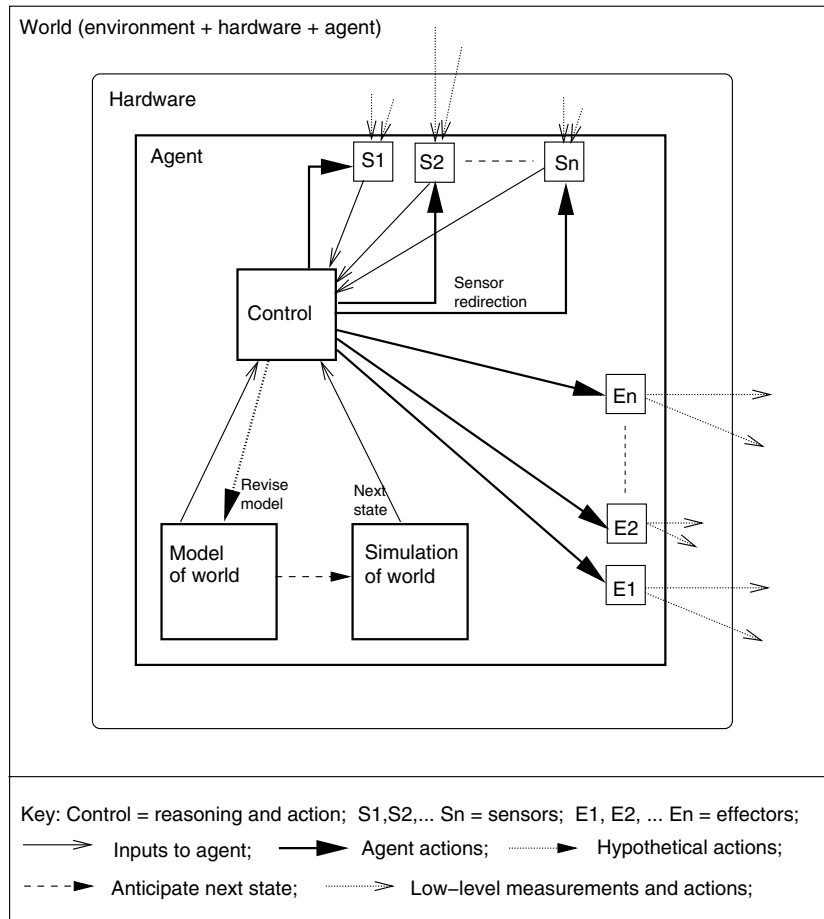


Fig. 1. An autonomous agent using internal simulation and some DDDAS capability

The internal simulation can be just a direct application of the model rules to the current state of the world to predict the next state. Control of sensors for data selection and the subsequent effect on the simulation may be regulated in different ways:

1. Evaluation-directed: if a predicted event is negative (e.g. if a component is predicted to fail), collect more data on the current state of the component; apply more detailed simulations based on the new data;
2. Uncertainty-directed: focus data collection on those areas where there is uncertainty (used in [3] and [4]).
3. Anomaly-directed: E.g. if the sensors indicate much less energy is available than the model predicted then use more sensors to collect data (as sensors may be faulty) and focus the simulation as in (2). If an anomaly persists, it may be the basis for model revision.

Model revision in a fully autonomous agent is ambitious (because of no human intervention). Therefore this action is labelled hypothetical in the diagram.

## 1.2 Assistant Agents

In contrast to the autonomous agent scenario, Figure 2 shows a scenario in which the DDDAS agent assists with a scientific process. Models and simulations have been developed separately for human understanding (labelled  $U$  in the diagram). The purpose of the simulation is primarily to help the scientist or other end-user. The simulation is an external application that the DDDAS agent interacts with.

If the agent is to select the relevant data for absorption into the simulation, it must have some description of what is in the simulation and a representation of what the goals and priorities are. The description, which we can call  $D$ , can be a representation of the main entities and relations in the original model ( $U$ ) in a form suitable for agent reasoning (e.g. it could be a set of rules or causal links). It may just cover a subset of  $U$ .

The DDDAS agent may also develop its own internal model of the world by adaptation. This is labelled  $M$  in the diagram and is an optional enhancement drawn in dotted lines.  $D$  can be used to initialise  $M$ . Model revision (of  $U$ ) has to be done by interacting with the human user's understanding of the simulation.  $M$  may be used to suggest revisions.

The simulation being managed by the DDDAS agent may include other "agents". They can be natural or artificial systems (e.g. humans, or software). Hence two kinds of "agent" exist: (a) the software DDDAS agent which manages the simulation and (b) the simulated agents, which represent real actors in the observed system. In a social world, examples include individuals, organisations etc. In an artificial world, examples include other software entities.

## 2 Application of DDDAS to the Social Sciences

Agent-based simulations may be used to model social systems and to assist decision-makers. Existing work includes geographical decision support systems (e.g. [5]) and fire evacuation [6]. A simulation can predict the effects of candidate policies or proposed interventions (or simply the effect of doing nothing). We assume that (a) a set of minimal

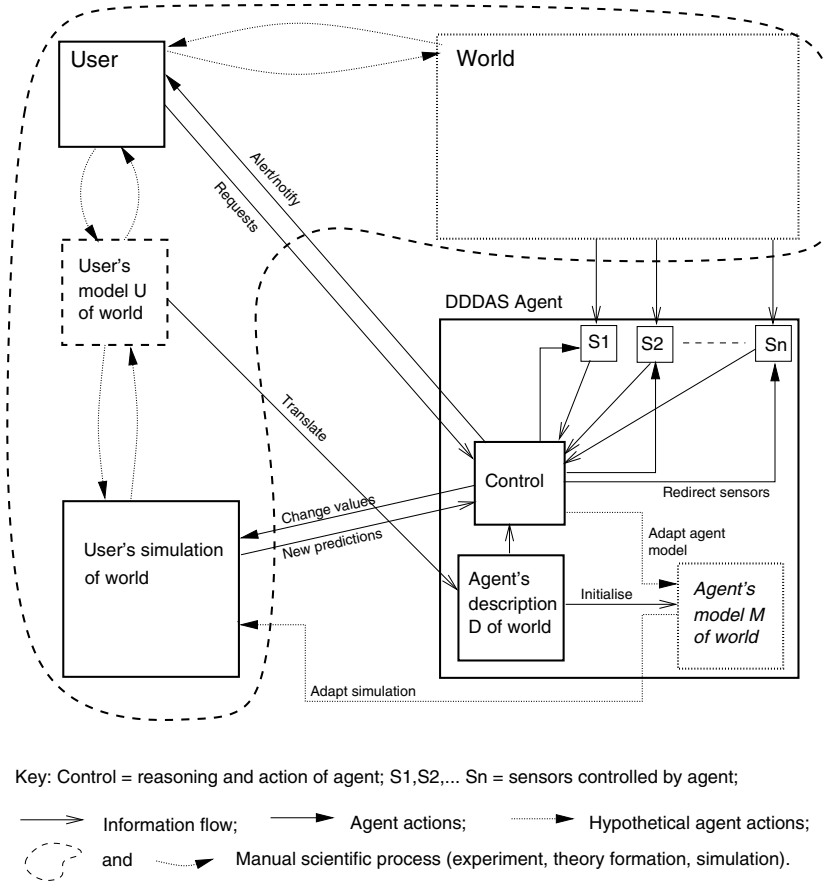


Fig. 2. A DDDAS agent assisting with modelling and simulation for a human user

requirements have to be met (e.g. relating to environment, health, crime-prevention) and (b) specific needs of participating agents have to be satisfied, with necessity for compromise when conflicts exist. The goals and priorities specified in *D* in Figure 2 are based on these requirements.

Introducing DDDAS into such a simulation leads to the possibility of validating its predictions “online” by continually comparing them with data from the real system. Unexpected features in the observed system can affect the simulation directly and possibly contribute to a revision of the theory that might not have been discovered otherwise.

### 2.1 How Can DDDAS Be Applied to Social Systems?

For simplicity we assume that a social simulation represents the evolution of a single observed system, not a class of systems. Thus the states of the observed system can fit directly into the simulation states (i.e. agent behaviour in the simulation can be checked against human agent behaviour in the observed system directly). The two systems could

run in parallel (e.g. passengers moving through an airport) or the simulation could be adjusted as historical data becomes available (e.g. housing decisions and mobility in a geographical area). It may also be possible to apply DDDAS to a simulation of a “typical” system, but the data would have to be selected carefully and generalised before being absorbed into the simulation.

The data selection and direction of the simulation can be regulated in the same way as in the physical sciences except that real-time data is not so easily available. The “sensors” will mostly involve database query and data mining tools.

For “evaluation-directed” selection, a negative prediction is a threat to the minimum requirements (e.g. health risk) or an unresolved conflict situation (e.g. leading to violence). (Section 2, (a) and (b)). The DDDAS agent may suggest certain kinds of data that have to be collected (e.g. what kind of behaviours tend to be associated with such events in reality?). It may also redirect the data collection autonomously using interfaces to databases etc. This is expected to be an interactive process.

### 3 Semantic Grounding

The DDDAS agent may use symbolic reasoning to interpret the simulation states, select appropriate data for absorption and to suggest model revisions. For example, non-monotonic reasoning involves making deductions using tentative assumptions and when necessary making subsequent revisions to these assumptions as new information becomes available (See e.g. [7], Ch. 7). We can apply this to Figure 2 by making updates to  $M$  and  $D$  interactively.

However, the AI system treats the symbols in  $D$  as formal patterns only, and serious errors in the initial version of  $D$  may not be detected. This problem has been called “symbol grounding” [8, 9]. We use the broader term “semantic grounding”, to refer to the checking of the validity of a model by independently interacting with the world and if necessary developing new concepts. Interaction with the world does not have to be physical. The important issue is the “data-driven” nature of the concept revision and its potential to fit into a DDDAS architecture.

#### 3.1 Agent Architectures to Support Grounding

Architectures already exist in AI which help to connect symbolic models with independent experience of a physical environment. For example, a hybrid agent architecture such as that in [10] may be appropriate. A hybrid architecture is one that integrates the symbolic tradition of AI with newer behaviour-based approaches such as that introduced by Brookes [11]. Behaviour-based approaches are “data-driven” because the result of the learning process is determined largely by low-level features in the environment and less by any pre-defined knowledge in an ontology. Recent work in concept formation involving symbol grounding includes [12, 13].

An ongoing research question is the degree to which the data-driven layer of a hybrid architecture can “interrupt” or influence the high level reasoning (e.g. if a dangerous situation is detected). Similarly, the degree to which the learning process is data-driven or top-down is important.

### 3.2 Multiple Ontologies

An alternative method of reducing the brittleness of a symbolic system without actually introducing “grounding” is to use multiple ontologies to introduce fault-tolerance. Multiple ontologies can co-exist as different viewpoints or descriptions of the observed social system. For example, in one description, agents could be modelled with objectively determined states and actions; another might emphasise the beliefs and affective states of agents.

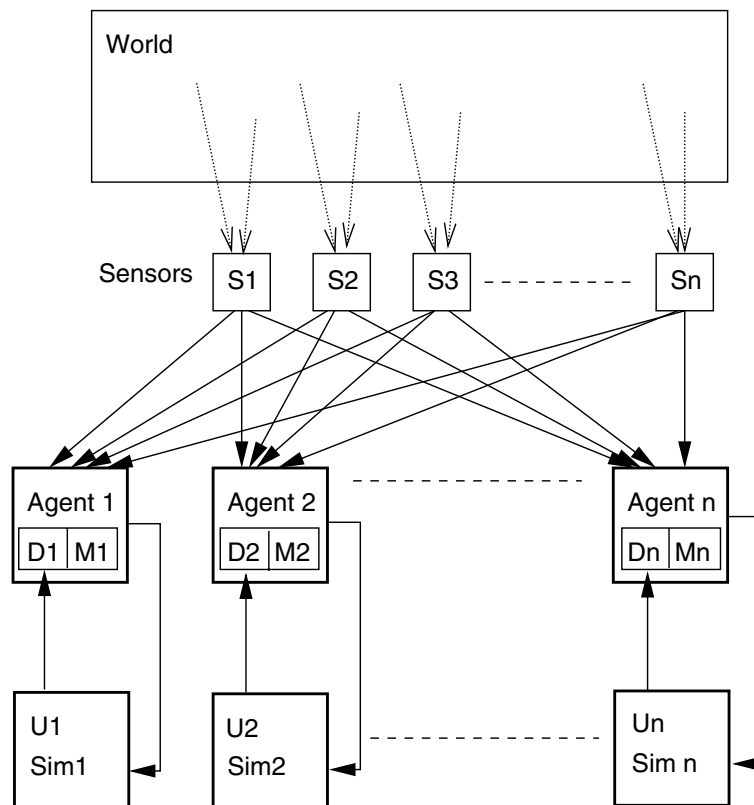


Fig. 3. Multiple agents, each representing a description of the world

Figure 3 shows a configuration involving multiple DDDAS agents, each controlling a simulation  $Sim_i$ . Each agent is a copy of the “assistant” type agent introduced earlier in Figure 2 (but with not all arrows and components shown). Each user-defined model  $U_i$  is translated into corresponding descriptions  $D_i$  for each agent, which may also have its own revisable model  $M_i$ . Any major disagreement between agents (and simulations) can be detected and investigated.

#### 4 Integrating Assistant Agents with Autonomous Agents

To implement the hybrid architecture of Section 3.1 in a science assistance scenario, the assistant agent in Figure 2 may be integrated with autonomous agents that can act as “sensing agents”. Similarly, in Figure 3, the “sensors” could have their own autonomous exploration and adaptation capability. However, this would be overridden by high level directives if necessary. Conversely, there may be “alarm” situations where the sensing agent can alert or even interrupt the assistant agent. Although the sensing agents are only semi-autonomous, their architecture can be similar to that of Figure 1 and they may even have their own DDDAS systems.

In most social science scenarios the interaction with the world requires human intermediaries and indirect access via speech acts, meaning that semantic grounding becomes more difficult than in a physical science assistant agent. However, the distinction between the detection of speech acts and the use of sensors is not sharp, since many data sources that could be called “sensors” are actually “event detectors” (i.e. they “say” that an event has occurred). In general, access becomes more indirect (and the concepts less “grounded”) the more the assistant agent relies on information that is pre-processed by another agent and the less control it has over it.

#### 5 Summary and Conclusions

DDDAS has the potential to improve the reliability of simulations used for decision support systems and can also assist with knowledge discovery and creativity in the social science domain. To reduce the brittleness associated with “semantically ungrounded” concepts in social simulations, we can conclude that the following is important:

1. *Autonomous learning and adaptation* is required, involving independent interaction with the world (i.e. data sources) in order to check the validity of a model and to revise it as necessary.
2. *Multiple ontologies* representing alternative descriptions of the world are advantageous. They can be the basis for different models, which can generate their own simulations. Potential problems can be detected if there are significant disagreements between model predictions.
3. *Cooperation between heterogenous agents* acting in different domains and levels of abstraction is important in order to exploit diverse sources of information that can be connected together. This increases the fault-tolerance of the system. The DDDAS agent itself is a service that is embedded in a wider Grid infrastructure and should exchange information with agents managing other services.

#### Acknowledgements

This research is supported by the Economic and Social Research Council as an e-Social Science feasibility study.

## References

1. Darema, F.: Grid Computing and Beyond: The Context of Dynamic Data Driven Applications Systems. Proceedings of the IEEE: Special Issue on Grid Computing **93** (2005) 692–697
2. Low, M.Y.H., Lye, K.W., Lendermann, P., Turner, S.J., Chim, R.T.W., Leo, S.H.: An Agent-based Approach for Managing Symbiotic Simulation of Semiconductor Assembly and Test Operation. In: Fourth International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS 2005), Utrecht, The Netherlands (2005)
3. Plale, B., Gannon, D., Reed, D., Graves, S., Droegemeier, K., Wilhelmson, B., Ramamurthy, M.: Towards Dynamically Adaptive Weather Analysis and Forecasting in LEAD. In: Workshop on Dynamic Data Driven Application Systems at the International Conference on Computational Science (ICCS 2005), Atlanta, USA (2005)
4. Patrikalakis, N., McCarthy, J., Robinson, A., Schmidt, H., Evangelinos, C., Haley, P., Lalis, S., Lermustaux, P., Tian, R., Leslie, W., Cho, W.: Towards a dynamic data driven system for rapid adaptive interdisciplinary ocean forecasting. In Darema, F., ed.: *Dynamic Data-Driven Application Systems*. Kluwer Academic Publishers, Amsterdam (2004)
5. Birkin, M., Dew, P., Macfarland, O., Hodrien, J.: HYDRA: A prototype grid-enabled spatial decision support system. In: First International Conference on e-Social Science, Manchester, UK (2005)
6. Chaturvedi, R., Filatyev, S., Gore, J., Mellema, A.A.: Integrating Fire, Structure and Agent Models. In: Workshop on Dynamic Data Driven Application Systems at the International Conference on Computational Science (ICCS 2005), Atlanta, USA (2005)
7. Rich, E., Knight, K.: Artificial Intelligence. McGraw-Hill Higher Education (1990)
8. Harnad, S.: The symbol grounding problem. *Physica D* **42** (1990) 335–346
9. Edmonds, B., Moss, S.: From KISS to KIDS - an anti-simplistic modelling approach. In: Joint Workshop on Multi-Agent and Multi-Agent-Based Simulation (MAMABS 2004) at the 3rd Conference on Autonomous Agents and Multi-Agent Systems (AAMAS-2004), Columbia University, New York City (2004)
10. Sloman, A., Scheutz, M.: A Framework for Comparing Agent Architectures. In: Proceedings of UKCI'02, UK Workshop on Computational Intelligence, Birmingham, UK (2002)
11. Brooks, R.A.: A Robust Layered Control System For A Mobile Robot. *IEEE Journal Of Robotics And Automation* **RA-2** (1986) 14–23
12. Roy, D., Pentland, A.: Learning words from sights and sounds: A computational model. *Cognitive Science* **26** (2002) 113–146
13. Gorniak, P., Roy, D.: Grounded semantic composition for visual scenes. *Journal of Artificial Intelligence Research* **21** (2004)