Semantic Pluralism for Collaborative Model Development

Catriona Kennedy,
School of Computer Science, University of Birmingham, UK
cmk@cs.bham.ac.uk

Abstract

Ontology research in the social sciences has highlighted the need to represent differing conceptualisations of reality without specifying a single set of definitions as the correct one (semantic pluralism). Additionally, there is a need to allow some degree of "bottom-up" determination of definitions by exploiting new developments in Web 2.0 and text mining. In this discussion paper, we investigate whether a pluralist and bottom-up approach to semantics can also be applied to the specification of social simulation models. Ontologies are already being used to specify such simulations and the imposition of a single taxonomy can cause the same problems in modelling as it does for data classification. Taking the example of "deliberative democracy" as a case study, we propose a software infrastructure that can help to represent and combine differing viewpoints using multiple simulations.

Key words: folksonomy, machine learning, ontology, pluralism, simulation, text mining.

1 Introduction

Recent developments in ontology research in the social sciences has highlighted the need to take into account differing conceptualisations of reality without specifying a single set of definitions as the correct one (e.g. [10, 11]). Such co-existence of equally valid ideas is known as "pluralism". For example, poverty has multiple definitions. Additionally, there is a need to allow some degree of "bottom-up" determination of definitions by capturing the usage patterns of concepts in social groups, and in particular by exploiting Web 2.0 technologies used by communities such as *Flickr* and *del.icio.us* to classify web content. Since these concepts are often non-specialist and non-scientific, their informal usage patterns have been called "folksonomies" in contrast to a "taxonomy" which is an imposed top-down categorisation.

In this paper, we investigate whether a pluralist and bottom-up approach to semantics can also be applied to the specification of agent-based social simulations. Ontologies are already being used to specify such simulations and the imposition of a single taxonomy can cause the same problems in modelling as it does for data classification. We propose some building blocks for a software infrastructure that will help to represent and combine differing viewpoints using simulation in a public policy scenario.

The paper is organised as follows. The first part (Section 2) considers the requirements of researchers who are advising policy-makers involved in collaborative decision-making, giving particular consideration to "deliberative democracy". The second part (Sections 3 and 4) is a review of some technologies and methods that can assist with satisfying the requirements. This includes a review of the capabilities and limits of a software prototype developed as part of the AIMSS project[1] along with a general discussion of how to enhancing concept discovery and generate awareness of new policy options. Finally, section 5 proposes a method of incorporating these technologies to extend the current AIMSS architecture so that the requirements can be satisfied.

[1] `http://www.cs.bham.ac.uk/research/projects/aimss/`

2 Case Study: Deliberative Democracy

Public Policy research and decision-making is one area where it is necessary to understand the differing viewpoints and concerns of multiple stakeholders affected by a proposed policy implementation, e.g. minority ethnic groups, mothers with young children, the elderly etc. In addition, the allowance of "bottom-up" determination of policy options is a particular requirement of "deliberative democracy" where decisions are made by a process of conversations between representatives of different groups, and not just by "experts". An important feature of this kind of democracy is that it is not based on a predefined set of options from which participants must choose, but instead it provides an environment in which completely new options can emerge as a result of informed deliberation and the gaining of awareness of other points of view.

We identify two key requirements that researchers and decision-makers will want to meet:

1. Understanding policy options and their impact, which involves two levels:
   (a) determinining the impact of known actions and policies and to predict future impact of known policy options;
   (b) identifying new kinds of actions and policy interventions that might have a significant effect;
2. Representing pluralism, also on two levels:
   (a) understanding the perceptions, concerns and aspirations of different social groups.
   (b) identifying new ways of dividing society into groups, based on differing needs and attitudes as well as measures such as economic status, ethnic group etc. Different grouping schemas may exist in parallel, with overlaps between them.

Note that 1(a) and 2(a) can be satisfied without 1(b) and 2(b), simply by already having a list of identified groups and a list of predefined policy options. Requirements 1(b) and 2(b) are about *discovery* of new categories that were not obvious initially or which may have been excluded from consideration due to wrong assumptions. Similarly requirement 1(a) could be addressed on its own, if we assume that we already understand the concerns of those affected, or if we are interested in objective measures only.

The purpose of the software infrastructure is to *assist* with satisfying the above requirements. It is not a "black box" that outputs a result when given an input, but instead it needs to be integrated with the methodologies of social scientists in collecting and interpretating data as well as generating theories. The software does not replace the human judgement involved in data collection and interpretation, which is a normal part of the social sciences. However, it should enhance the research process. In the longer term, the aim is to also enhance deliberative democratic processes by improving communication and collaboration between participants beyond what would be possible without the technology.

Specifically, the potential users include the following:

- Researchers who advise policy-makers, along with other social scientists aiming to understand groups and power structures in society as well as democratic processes.
- Longer term: policy-makers and other participants in a deliberative democratic process, whose requirements match the above.

In a deliberative democracy, representatives from all identified groups should ideally participate in the understanding, decision-making and evaluation above (not just policy-makers and researchers). Requirement 2(b) implies a self-organising process if some groups define themselves or help to define others ("mutual bootstrapping").

2.1 Assumptions about Public Policy Scenarios

Even if we are addressing requirement 2(b) above, it is still necessary to start with some minimal concepts and classifications. We assume that the following three groups exist:

1. policy-makers and social science researchers who have to take into account different goals and values (e.g. optimise transport connections vs. housing regeneration);

2. decision-makers acting independently according to a single goal, whose decisions affect others (e.g. house builders, businesses);

3. occupants of the area (e.g. residents, shoppers, pedestrians) who will be affected by decisions.

Groups 2 and 3 may be further divided into subgroups according to their goals and perceptions. Subgroups in 1 and 2 may be aware of each other's viewpoints and can be involved in collaborative decisions. Subgroups in 3 are often not aware of each other's viewpoints or that of groups 1 and 2 and are not normally equal participants in collaborative decision-making, since they have limited power. However, policy-makers and researchers can consult occupants. This paper is mostly about capturing the differing viewpoints of occupants (raising awareness of them by groups 1 and 2) and also to increase communication and mutual understanding between occupant subgroups.

3 Simulation for prediction of policy impact

We now consider what technologies are available to help satisfy the requirements listed in Section 2. For requirement 1, given that groups and actions are already defined and their concerns are understood, how do we predict the effects of these pre-defined policies and actions? Agent-based simulations may be used to predict policy impact [3, 2] as well as to predict the effects of individual actions of agents on other agents.

An agent can be understood as a "sense-process-act" loop where actions are determined by internal processing as well as the sate of the environment. The processing can range in complexity from insect-like to human-like. An entity which is dynamic but does not sense or act (e.g. rain) would not normally be regarded as an agent, but instead as a dynamic part of the environment. In a social simulation, entities such as organisations, firms and governments can be represented as agents in addition to individuals.

3.1 AIMSS architecture: "data driven" simulation

Dynamic Data Driven Application Simlation (DDDAS) is explained in [9]. In the AIMSS project, we have interpreted this process as a semi-automated model revision loop, where the model is iteratively adjusted until its predictions are consistent with observations. The goal is to eventually apply this method to large scale realistic simulations in order to increase the reliability of their predictions and make it feasible to use them for decision support. In the short to medium term, the method is expected to be useful for researchers. The AIMSS architecture is described in detail in [14] and [13] but will be briefly summarised here.

Figure 1 is a schematic workflow for the AIMSS theory revision architecture. The figure can be divided into two processes: the first is a process of interpreting both the simulation predictions and the data content and determining whether they are consistent (arrows 1, 2 and 3 in the figure). The second involves adapting the simulation in the event of an inconsistency. In the current version of the prototype we have focused on the first process. This requires the following automated capabilities:

- Interpretation of simulation states (at regular intervals or on demand): this is not just about reading the current variable values, but about generating a *high level description* summarising patterns or trends.

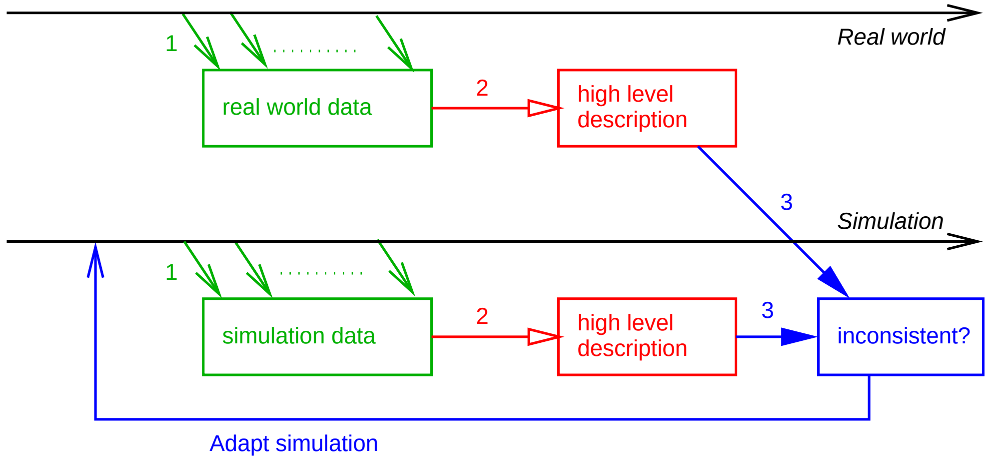

Figure 1: Data driven adaptation of a simulation

- Interpretation of real world data: the same methods are used for data interpretation as for simulation interpretation, except that the data is often more detailed and is usually noisy. Therefore pre-processing is required, which often involves the integration of data from multiple sources and the generation of higher level datasets that correspond to the simulation events.
- Consistency checking to determine whether the simulation states mostly agree with interpretations of data content.
- Re-direction and focusing of data collection in response to evaluation of simulation states or uncertainty in the data comparison.

While the simulation is running, a history of events is generated as a continual trace. Data mining is then applied to this synthetic dataset, resulting in a set of general logical statements about simulation trends and developments. The high level datasets generated from the raw data are also subjected to data mining and similar high level logical statements about data content are generated. Logic-based consistency-checking (based on satisfiability checking) is then applied to both sets of statements. This automated process may lead to the discovery of evidence in the data that refutes the model.

It is expected that the interpretation of the simulation and data content will be complemented by human visual inspection. Similarly the consistency checking may be overriden by the "common sense" judgements of the researcher. Visualisation and compact summarisation of the data are important technologies. Currently simulation adaptation is a purely manual process, but eventually this may also be augmented by machine learning or evolutionary programming techniques.

3.1.1 Role of ontologies

A simulation requires a definition of the entities that are in it: what kind of things are agents, what entities are in the environment and how are they related? In the AIMSS project we called such a definition an *ontology* in the broad sense in which the term is used in AI (not necessarily a formalised ontology as in Semantic Web services). This same ontology is used to generate different variants of a simulation, as well as to automatically interpret developments in the simulation during its execution, and to interpret the data used to find evidence to support or refute the simulation predictions. For the housing case study used in the AIMSS project, the ontology contained entities such as homes, regions, households and moves from one home to another. More details are in [14] and [13]. Although we used only an informal ontology in the AIMSS project (in XML), this may be encoded in a formal ontology language such as OWL and integrated with other formalised ontologies used in Semantic Web services. Interoperability is required here, since the kind of ontologies used to classify and retrieve data may not necessarily be suited to specifying of simulations.

A simulation also needs a specification of dynamic behaviour: how does the environment change and what are the behaviour rules and possible actions available to agents? We have called this a model (since it is similar to a physical model except that it uses discrete events and symbolic behaviour rules instead of equations).

3.1.2 Role of automated pattern recognition and reasoning

Automated pattern recognition includes data mining, text mining and other forms of "machine learning". Data mining is often classed as "machine learning" because the rules or structure that it discovers are generalisations which can be used to predict further data content. For example, if customers have been observed to buy one product whenever they buy another, this general pattern has been learned and can be used to predict future behaviour of customers. In the AIMSS architecture, data mining is used to extract and summarise data content. However, it could also be part of an automated learning process to predict new regularities. Such learning is an acquisition of new knowledge and could involve an automated extension or revision to the behaviour model or ontology. This is an ongoing area of research and will involve consultation with research on "belief revision", which is a traditional research area in AI (see e.g. [16], Ch. 7).

The automated reasoning component of the architecture currently involves consistency-checking between simulation content and data content using satisfiability checking. All of these automated capabilities are tools on a "work bench" which can complement visual inspection and manual revisions.

4 Representing Pluralism

Requirement 2 is about capturing and representing the different concerns of known groups. In particular, we are focusing on subgroups of group 3 (occupants). We assume for now that these subgroups are known.

In the AIMSS architecture, the theory revision process starts with an initial model, which is iteratively adjusted until it is consistent with the data (Figure 1). The model can be revised by an expert interpreting the data while being assisted by software components such as data mining and logic-based consistency-checking. However, there is no allowance for the representation of different views of the world (pluralism). Moreover, the actions of the system and the nature of the final model may depend considerably on the initial model. This means that we are arbitarily favouring one view of the world over alternatives.

This problem is related to the disadvantages of a single (formal) ontology to classify social science data, because it imposes a single classification hierarchy (a taxonomy) which does not allow for differing interpretations. This issue has been highlighted by [10] and [11].

There are two kinds of pluralism that need to be considered:

1. the involvement of stakeholders in the determination of concepts and ontologies on which the simulation is based, in particular those with limited power (group 3 in Section 2).

2. pluralism in social science *research*: this includes the implicit ontology used by the researcher (which is their theory of the world) and the methodology, e.g. quantitative or qualitative[15].

Although this paper is mostly about the first issue, the software infrastructure must support different social science methods and should allow their combination. In particular, the following subdivision of methodology is important:

- "theory-driven" (or "top down"): data collection and interpretation should be strongly directed by an existing theory along with implicit ontology (e.g. should the world be modelled as indivisible "households" or should it be power relations?)

- "data driven" (or bottom-up"): theories and concepts should as far as possible be formed during the collection and anaysis of data. Interpretation can only happen at a later stage, once some initial concepts have been formed. Similarly the form of data collection is expected to emphasise free-text responses.

We assume that these are not mutually exclusive and that a spectrum exists between them. Moreover, different stages in the research process or different kinds of research question may require an emphasis on one or the other.

We now look at available technologies that can help determine what concepts are used by others and to describe these concepts using an ontology. These may be represented in an ontology or behaviour model which can in turn be used to specify simulations.

4.1 Concept capture

To represent an occupant group's view of the world, it is necessary to identify what kind of things are important for them in everyday life. For example, this might be fear of crime due to poor street lighting, or it might be danger due to traffic or just too much concrete with too little green space. The key concepts used by these groups in describing their experience will be centred around their concerns.

We define a "concept" informally as a name which is used to refer to a class of entities, where the classes are discriminated by the actual usage of different names. The named entity may, for example, be an object in a picture, or a dynamic sequence of events (e.g. "travelling to work"). In one case the concept of a "street" may be divided into well-lit and badly lit, while in the other it may be subdivided into heavy traffic or light traffic.

A more formal definition of a concept is provided in the Formal Concept Analysis[2] literature (e.g. [1, 5]). The definition involves an explicit listing of attributes which members of a named class must have (which is called the *intent*) as well as a set of instances which have all these features (the *extent*).

Two broad methods are available for capturing concepts used in social groups:

- Folksonomies are classifications of online content which are determined by user-created tags (labels). In addition to popular examples such as *Flickr*, academic researchers can use *Bibsonomy* and *Connotea*.

- Text mining and other forms of machine learning: for example, some algorithms extract key terms from text (those that appear to be most important).

Folksonomies contribute to a bottom-up emergent classification instead of a top-down taxonomy. They may be combined with formal ontologies to produce classifications of tags [12]. Machine learning and data mining can discover organisation in the way that tags are used (e.g. [17]) and this in turn can contribute to an ontology of tags.

Text mining can be used to automatically generate formal ontologies or to extend or populate existing ontologies [8, 4]. This work uses the formal definition of a concept. A problem with generating ontologies in this way is that concepts used to describe experiences are "fuzzy" and difficult to represent formally. There may be some essential attributes of a concept that cannot be listed as a set of properties, since they are not linguistic. Instead, they may relate to visual impressions or moods, or a combination of visual experience, sounds and emotions (e.g. "a sense of community"). Therefore, a formal ontology is not expected to capture all details of such concepts and some degree of manual intervention in the generation and use of classifications is expected. In particular, the process of generating a behaviour model for a simulation is likely to be highly interactive as it will not just depend on an ontology, but will involve "common sense" knowledge. Manual intervention is expected to involve a balance between "top-down" expert knowledge and "bottom-up" approaches based on folksonomy.

Simulation and visualisation may help with the representation of subjective and imprecise concepts. A simulation requires a specification of dynamic behaviour of its entities along with an initial

[2]See for example, `http://www.upriss.org.uk/fca/fca.html`

configuration. Such specifications are not typically included within ontologies but they require additional domain expertise. Informal concepts may be used to partly specify this behaviour, even if they are not explicitly formalised in an ontology.

4.2 Some advantages of automated methods

Although they have many limitations, automated methods of pattern recognition, such as machine learning and data mining are valuable for discovering regularities in large amounts of data that would not be discovered otherwise due to its complexity. Such regularities can be used to automatically extend ontologies. In particular, machine learning could help to satisfy Requirement 2(b), which is the discovery of new ways of dividing society into groups. If we combine machine learning with tagging and free text, there may be no need to define groups in advance, since the demarcations can be discovered using data mining methods (e.g. clustering) according to answers given to questions. If Web 2.0 and free text are exploited, there may be no need to ask specific questions; the users can make their own decisions on how (or whether) to give an entity a label. Identified groups may be communities recognisable by their pattern of communication or they may be classes of individuals with similar attitudes who are not in any well-defined community.

Another potential advantage of machine learning (and hence text mining) is that it may overcome some limitations of both folksonomies and top-down research approaches. Both approaches may fail to recognise important categories due to culturally biased interpretations or (in the case of folksonomy) due to assumptions that one community can make about another. In such cases, it is important to ensure that the internal configurations of the machine learning algorithms do not implicitly reflect a biased set of preferences. The same advantages may apply to the automated logic-based consistency-checking which is part of the AIMSS architecture, although this is still at an early stage of development.

To conclude this section, we can identify two different "axes" for development of semantic definitions:

- top-down vs bottom-up;
- automated vs manual;

For both of these, it is a matter of finding the correct balance between the two and to provide the ability to *vary* the degree of each, depending on the dynamically changing research context. Note that "data-driven" simulation is not necessarily bottom up (i.e. not necessarily data-driven *research*). It means primarily that the theory should be modified to fit the data. However, the data interpretation may still be strongly influenced by the theory, even though the theory is evolving. In the next section we consider how to integrate concept capture and pluralism into an AIMSS type architecture and thus mitigate potential problems of relying on a single theory.

5 Extending the AIMSS Architecture

To incorporate pluralism, the AIMSS architecture may be extended in two ways:

1. Single initial model, with divergent model revision or ontology extension: this assumes one approach to social science research, but which requires representation of pluralism in society.
2. Multiple initial models, with multiple paths of revision: Initial models can represent theoretical and methodological pluralism in social science.

If we consider the first option, a schematic workflow for the iterative and collaborative development of models may look as follows:

1. Generate and revise initial model (composed of the following substeps):

(a) Generate a simulation using an initial minimal ontology and behavioural model.

(b) Invite representatives of different groups (participants) to interact with the simulation, and make comments or add tags; they should also have the opportunity to indicate things that are *not* being represented in the simulation but are of concern to them. Free text (and text-mining) are important at this stage.

(c) Revise the ontology and behavioural model, according to participants' tags and comments. The result should be a divergence of simulations, each emphasising different kinds of behaviour and entities in the environment and possibly different kinds of agents.

2. If appropriate, combine simulations together so that they share the same "reality" (as in for example, [7]). The result may be many simulations (see below), showing different views of the same reality. The views will correspond to entities that have been highlighted by different participants.

3. Invite participants to evaluate developments in each of the combined simulations and to interact with different views of them.

The result is a collection of different ontologies, each one corresponding to a class of simulations. For each ontology, variants of a simulation can exist due to differing configurations of dynamic behaviour, as well as other parameters. If we begin with a single model, the extended ontologies may all continue with the same top-level structure containing the initial concepts. The simplest case is when revision is limited to the creation of new subclasses of existing concepts or the creation of new relations between them. In such cases it should be possible to combine simulations. However, more radical revision may occur, where some of the original concepts are disputed. In this case, belief revision methods can be applied to revise ontologies, but it may not be meaningful to combine simulations based on these ontologies if they mutually exclude each other.

5.1 Combining simulations

Interactions between simulations based on different models is a current research area (e.g. [7, 6]). If there is a minimal fundamental ontology of agreed concepts, the different extensions to the ontology can be regarded as specialised models (e.g. a model of variable road traffic, or a model of changes in street lighting leading to different agent behaviour). We can also think of one model acting as an "environment" for another (e.g. physical models of fire propagation affecting the movement of agents in an agent-based simulation [7].

To show how this can work in the model revision workflow above we can imagine that a simulation S_1 is modified and renamed S_2 by adding a new kind of agent (indicated as important by a user) with its own set of actions which affect existing agents. The original simulation S_1 may be modelling people moving in a crowd in a shopping area. If we assume that road traffic was not modelled in S_1 but is included in the new simulation S_2 then the new actions happening in S_2 will also have to be translated into S_1 if they affect agents. For example, the existence of traffic may cause agents to follow different paths. If the simulations have a shared reality, the entities shared between them must be consistent (including the shoppers' movements). Therefore, in S_1 the agents will have to move in the same way as in S_2 but the cause of this is not explicitly represented. The agents will appear to be disrupted at random. Although this is a simple example, it may be applied to less obvious situations (for example, if households are modelled as single entities in one simulation and as differentiated multi-agent structures in another).

One possible way to implement this is to provide all simulations with a set of basic actions of the "environment" which can allow or deny any action by an agent or force unintended actions. Any new agents or environmental entities will then have their actions translated into those basic actions for those simulations that do not include the new agents. This can be included within a minimal objectively agreed ontology for all simulations which can be extended in divergent ways. Such translation processes may also be possible in cases where there are multiple models initially. In this case the "minimal" ontology will only contain those basic environmental actions.

5.2 Advantages of multiple simulations

An important consequence of having multiple simulations is that a participant can interact with all of them and obtain a view of *other groups'* perceptions, which may not have been possible using verbal communication only. Furthermore, if the simulations can be combined so that they interact, one simulation can act as "interference" for the other and force events that would not otherwise happen. This means that policy-makers and other decision-makers with power (groups 1 and 2 in Section 2) may be able to see the consequences of their proposed actions as side effects (i.e. "interference") on another group's perceived space, thus improving their understanding of the subjective experience of that group. Since this can lead to the discovery of actions that cause "negative interference", it may also help in the search for new policy options which can produce more helpful "interference".

Another advantage is that the different simulations may be subjected to a DDDAS workflow in parallel, effectively leading to many versions of Figure 1. If methodological pluralism is used, there is even more divergence between the initial models. For each model, the selection of data to be used as potential evidence can be influenced by the simulation content, which in turn depends on the ontology used to interpret simulation states. The end result is that consistency-checking is done in many different ways, not depending on an initial ontology.

6 Summary and Conclusion

In this paper, we have discussed technologies that can help to represent pluralism in a deliberative democracy scenario. Although many technical challenges remain, we proposed some building blocks for an infrastructure in which representatives of different communities can help to develop simulation models. Opportunities for tagging and free-text entry can enable participants to interact critically with a simulation, which in turn can lead to changes in the underlying model, producing multiple variants. Interaction with the new "democratically-determined" simulations may help participants to gain an insight into each other's concerns. In addition, simulations may be combined in a way that can help policy-makers to identify unexpected side-effects of proposed policies on the reality of an affected community. It is also possible that combining different realities in this way can lead to entirely new concepts, such as previously unexplored policy options.

Machine learning (including text mining and data mining) can also play a key role in the discovery of previously unrecognised concerns as well as self-defining social networks. Although machine learning has limitations in its capability to automaticaly generate rich representations, it has the potential to overcome some limitations of folksonomy that may result due to culturally biased assumptions made about other communities. Therefore, the infrastructure should support a balance between automation and interactivity, in the same way as it should support a balance between top-town and bottom-up semantics.

Acknowledgements

This research was supported by the Economic and Social Research Council.

References

[1] Bernhard Ganter, Gerd Stumme and Rudolf Wille (Eds.). *Formal Concept Analysis: Foundations and Applications*. Springer Verlag, Lecture Notes in Artificial Intelligence, no. 3626, 2005.

[2] Mark Birkin, Haibo Chen, Martin Clarke, Justin Keen, Phil Rees, and Jie Xu. MOSES: Modelling and Simulation for e-Social Science. In *First International Conference on e-Social Science*, Manchester, UK, June 2005.

[3] Mark Birkin, Pete Dew, Olga Macfarland, and John Hodrien. HYDRA: A prototype grid-enabled spatial decision support system. In *First International Conference on e-Social Science*, Manchester, UK, June 2005.

[4] Paul Buitelaar and Philipp Cimiano. Tutorial on Ontology Learning from Text (at EACL 2006), 2006. `http://www.aifb.uni-karlsruhe.de/WBS/pci/EACL_OL_Tutorial_06/`.

[5] Peter Burmeister. Formal Concept Analysis with ConImp: Introduction to the Basic Features, 2003. `http://www.mathematik.tu-darmstadt.de/~burmeister/ConImpIntro.pdf`.

[6] Alok Chaturvedi. A Society of Simulation Approach to Dynamic Integration of Simulations. In *Proceedings of the 2006 Winter Simulation Conference*, 2006.

[7] Alok Chaturvedi, Angela Mellema, Sergei Filatyev, and Jay Gore. DDDAS for Fire and Agent Evacuation Modeling of the Rhode Island Nightclub Fire. In *Workshop on Dynamic Data-Driven Applications Simulation at ICCS 2006, LNCS 3993*, pages 433–439, Reading, UK, May 2006. Springer-Verlag.

[8] Philipp Cimiano, Johanna Volker, and Rudi Studer. Ontologies on Demand? A Description of the State-of-the-Art, Applications, Challenges and Trends for Ontology Learning from Text. *Informationswissenschaft*, 57:315–320, 2006. `http://www.aifb.uni-karlsruhe.de/WBS/pci/Publications/iwp06.pdf`.

[9] Frederica Darema. Grid Computing and Beyond: The Context of Dynamic Data Driven Applications Systems. *Proceedings of the IEEE: Special Issue on Grid Computing*, 93(3):692–697, March 2005.

[10] Phil Edwards. Cloudbuilding: Semantic Imprecision in Social Science Data (and how to live with it). In *Second International Conference on e-Social Science*, Manchester, UK, June 2006.

[11] Abdelkader Gouaich. Semantic Grid Challenges for eResearch: the PolicyGrid View. In *Workshop on Semantic Grid, Second International Conference on e-Social Science*, Manchester, UK, June 2006.

[12] Thomas Gruber. Ontology of Folksonomy: A Mash-up of Apples and Oranges, 2006. `http://tomgruber.org/writing/ontology-of-folksonomy.htm`.

[13] Catriona Kennedy, Georgios Theodoropoulos, Ed Ferrari, Peter Lee, and Chris Skelcher. Towards an Automated Approach to Dynamic Interpretation of Simulations. Invited Paper. In *First Asia International Conference on Modelling and Simulation (AMS 2007)*, Phuket, Thailand, March 2007.

[14] Catriona Kennedy, Georgios Theodoropoulos, Volker Sorge, Ed Ferrari, Peter Lee, and Chris Skelcher. AIMSS: An Architecture for Data Driven Simulations in the Social Sciences. In *Workshop on Dynamic Data-Driven Applications Simulation at ICCS 2007*, Beijing, China, May 2007. Springer-Verlag.

[15] Wendy Olsen. Pluralism, Poverty and Sharecropping: Cultivating Open-mindedness in Development Studies. *Journal of Development Studies*, October 2006.

[16] Elaine Rich and Kevin Knight. *Artificial Intelligence*. McGraw-Hill Higher Education, 1990.

[17] Christoph Schmitz, Andreas Hotho, Robert Jaschke, and Gerd Stumme. Mining Association Rules in Folksonomies. In *Proc. IFCS 2006 Conference*, Ljubljana, July 2006.