Proceedings of the 12th ANZSYS conference - Sustaining our social and natural capital
Katoomba, NSW Australia, 3rd-6th December, 2006

Agent Based Modelling of Real Issues: Acquisition of Data to Support AB Models of Social Phenomena

Susanne Tepe[1], Tim Haslett[2] & Charles Osborne[2]
[1]School of Applied Science, RMIT University, AUS
[2]Department of Management, Monash University, AUS

Agent Based Models (ABMs) are powerful models for describing the interactions between and among individual agents within a dynamic situation, such as in a community. In contrast to System Dynamics models, ABMs are based on rules describing agent behaviour. If modelled correctly, the ABM can mimic the behaviour observed in a community and new behaviours can emerge as a result of agents' interactions. But where does the modeller acquire the data to underpin and define the rules which can then be modelled? This paper describes and evaluates the advantages and short comings of various data acquisition techniques, with particular reference to social modelling. Gathering data via documented historical information, personal experience, group model building, qualitative analysis of public comment data, quantitative methods such as discrete choice questionnaires, mixed qualitative / quantitative methods such as the Cynefin technique and the use of ABM archetypes are examined. Recommendations are made for the use of appropriate techniques for different situations.

Introduction

Agent Based Modelling (ABM) can be a powerful tool for systems modellers. As opposed to modelling the stocks and flows as in Systems Dynamics (SD) models, ABM focuses on groups of agents and their interactions. Whereas SD models define the system 'in focus' from a big picture or top down point of view, ABM allows the system to emerge from the interaction of the agents, a more bottom up point of view. While both SD and AB models rely on the language of mathematics to represent their actions, SD generally uses mathematics to describe flows through the system based on quantitative data, while AB uses equations to define the rules of agent interaction. Both modelling types are capable of describing non-linear dynamics, with SD focusing on interacting balancing and reinforcing loops and ABM describing its rules for agent interaction with non-linear equations and IF / THEN statements. Essential to the function of SD models are feedback loops and delays in response. While these can happen in ABM, more focus is placed on agents undergoing 'state changes' (eg from a passive 'state' to an active 'state') as a result of parameters that influence them. While state changes and flows from stock to stock can be similarly defined by the dynamics of the model in both SD and ABM, ABM allows these dynamics to change as a result of recent history (Anderson & Johnson 1997; AnyLogic 2005; Axelrod 1997);(Axelrod & Tesfatsion 2005).

While SD and ABM are both useful techniques for the systems' modeller, they must be used to address appropriate issues. While applicable in other areas, SD works better for complicated interacting systems involving quantitative data while ABM is more appropriate for examination of issues where individuals interact, such as social phenomena. Table 1 summarises these observations.

Model formation

Despite these differences, all systems modelling require input from the environment in the form of data to formulate and validate the model. The data from the environment influence the model formation in different ways and at different times during the model formation. In order to describe this process more thoroughly, it is appropriate to describe the assumptions that underpin my description of model formation. Figure 1 shows a summary, very generic diagram of the model formation process.

As simplistically depicted in Figure 1, in establishing a systems model, the modeller starts with internal pictures of the 'world' being modelled, which Senge, et al (Senge et al. 1994) refer to as a Mental Model. Because at this point the Mental Model belongs to the modeller, Figure 1 refers to it as a Personal Mental Model (PMM). As the modeller interacts with and gathers data from the environment and reflects on this experience, their PMM evolves in clarity and the modeller's depth of understanding about this 'world' improves. The data drawn from the environment may take the form of discussions or interviews with other people concerned with the 'world' being examined. The modeller may acquire input from written documents or historical accounts concerning this 'world'. In addition, the modeller may discuss this 'world' with a group who in turn share their own versions of their mental models with each other and as a group they evolve a deeper understanding of this 'world'. As the modeller acquires this additional input, the PMM evolves into a richer understanding of this 'world', which Figure 1 refers to as an Evolved Mental Model.

At this point, the modeller may choose to cease keeping this model in their head and choose to record the model. While this can be done using any of the multitudes of systems modelling tools, in Figure 1 these models are differentiated by whether the data used to form the model is qualitative or quantitative. Generally, the modeller will first attempt to develop a qualitative model, such as a causal loop diagram (CLD). In the development of the CLD it is likely that PMM will continue to evolve, as indicated in Figure 1 by the feedback loop between the Qualitative Model and the Evolved PMM. When the modeller has a satisfactory CLD and if they have numerical data, they may choose to develop a quantitative model. Again, feedback from the model as it evolves will influence the PMM of the modeller.

Data acquisition

Given the above assumptions about the way that a systems model evolves, it is clear that data from the environment enters the model at various stages. Data from interviews with stakeholders, written documents and group discussions enrich the personal mental model of the modeller, which in turn can be recorded as a qualitative model. Numerical data can be used to convert the qualitative model into a quantitative model, such as a Systems Dynamics or an Agent Based model. But where does this information come from and how can it be used in Agent

	Systems Dynamics Modelling	**Agent Based Modelling**
Aim	To describe complex multifaceted interacting systems	To evolve system behaviour from the interaction of individual agents
Premise	System is defined by the modeller	Agents interact to define the system
Key components	Stocks and flows	Individual agents
Viewpoint	Top down / big picture / broad	Bottom up / individual
Outcomes	Behaviour over time	Emergent behaviour due to agent interaction
Role of mathematics	Equations describe flows	Equations describe rules of interaction
Role of history	Not influenced by history (each model run is usually a fresh run)	Agents influenced by their history (agents change to new states as a result of experience)
Non-linear dynamics	Due to balancing and reinforcing loops	Due to rules of interaction
Essential elements	• Feedback loops • Delays	• Agents undergo State changes • Groups of agents have a distribution of responses
Best areas of application	Complex systems with available quantitative data	Social issues where individual actors make decisions

Table 1 *A simplistic comparison between systems dynamics modelling and agent based modelling*

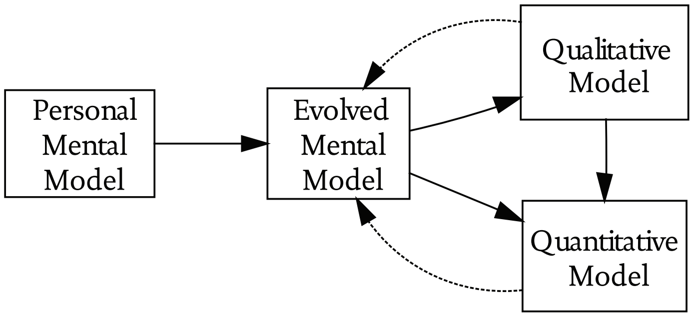

Figure 1 *Process of model development*

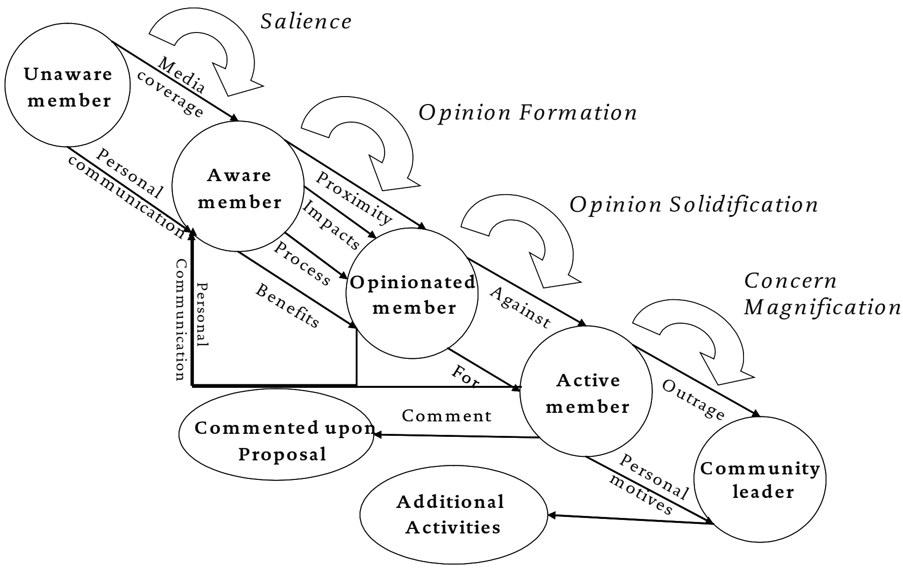

Figure 2 *State changes of community member agents*

Based Modelling of a social system?

Depending on the issue at hand, the modeller may or may not have personal experience with the issue or the 'system in focus.' As the modeller attempts to define and understand the issue, they will form a mental model of their idea of the system in focus and what influences it, who the stakeholders or actors are, what environment the system operates in, what system boundaries need to be defined, etc. Generally the modeller recognises that, as an individual, this is a shallow one-sided view of the issue and will desire to enhance their understanding by acquiring other viewpoints. At this point, the modeller may or may not choose to record their model as a qualitative model. Always at the back of the ABM modeller's mind is the question: how does this information relate to the rules of individual interaction? Who are the actors, what influences them, how do they interact with other actors and how can I represent this as mathematically expressed rules?

The modeller may decide to enhance their understanding of the issue through acquisition of other input. They may choose to read documents concerning the issue. In the case of these authors' research on community acceptance of hazardous waste facilities, there is a book written that describes a community who became outraged about a plan to establish a hazardous waste facility and their successful blocking of the proposed plan (Strangio 2001). This book was useful for providing additional insights to supplement the authors' experience, particularly concerning which groups should be construed as 'agents' and what parameters influence the agents at which stages of the process. For example, it became clear that an 'agent' community member could be represented in the variable states of 'unaware' of the issue, 'aware' of the issue, 'active' in their approach to the issue or 'community leader'. The parameters that influenced them included: the credibility of other community members who contacted the individual agent through word of mouth, the salience of media reporting and the potential impact of the development. Based on the mathematical relation between the parameters, the agents could undergo a state change from one state to the next. Figure 2 represents this set of state changes in a diagrammatical form.

In addition to deriving information from personal experience or from written material, a modeller can acquire information from the stakeholders in the issue. Implicit to this comment is that the modeller must recognise who the stakeholders are; this is usually as a result of previous personal experience or from case study input. Interviews with stakeholders can inform the personal mental model of the modeller in similar ways to reading a case study, but in a much more interactive way. The interviewee can correct any misunderstandings by the modeller and can specifically address topics of interest to the modeller. Interviews provide a broader range of viewpoints on the issue to supplement the modeller's own view. Because the stakeholder groups effectively become the 'agents' in the model, it would be best to have interview data from several individual stakeholders in each group to ensure that diversity of opinion is represented. The focus of the interviews should be about the parameters that influence the stakeholder / agent in changing from one state to the next. Consistency of view about the parameters is desired within the one agent group, while the range of opinion about the strength of influence of this parameter can be used to define the shape of the distribution of response. For example, all participants interviewed about what makes a person pay attention to a news article concerning the potential development of a hazardous waste facility (e.g. the news article's salience), indicated that geographical proximity of the proposed location to their own life was an important factor. In modelling terms, this data was interpreted as a relatively strong factor (so a numerically bigger value) with a narrow distribution of response.

Similar to interviews with individual stakeholders, a modeller can interview several member of one stakeholder group at one time, usually in a focus group type arrangement. This helps with ensuring that the breadth of stakeholder views are canvassed while also helping the stakeholders reflect on comments made by others in the group. This is not dissimilar to the group model building process used by SD modellers. Again, for ABM, the focus of the discussion should be based on the parameters that cause the stakeholder agents to change from one state to another, but obviously in a language that is consistent and sensitive to the issue at hand. In the example at hand, the government regulators were interviewed as a group, to get a consistent description of the process of review for development proposals.

All of the above methods for gathering information are qualitative in nature. They are very useful in forming and enhancing the mental model of the modeller, which in turn can be converted into CLD or ABM, but they do not give any numerical data about the mathematical relationships, the frequency of any events or the relative value of one influencing parameter compared to another.

At this point the modellers can test their mental model by using an ABM archetype. If the modellers recognise similarities between their issue and one already developed by a soft ware package, they can modify the existing software to reflect their own case. In the case of the authors' research, the AnyLogic software had a model of the Bass Diffusion Model, where prospective purchasers became actual purchasers due to influence of advertising and word of mouth (AnyLogic 2005). This model was able to be modified to reflect how community members could be influenced to become 'aware' of the issue and to become 'active' based on media reports and on contact with other community members. This archetype was then embellished to include more state changes and to incorporate feedback and reinforcing loops (see Figure 2). While it is possible to get the model to run and for community member agents to change state, the mathematical relationships and the values attributed to each parameter are contrived; there is no basis in reality for why a given parameter has a given value, only that those values make the model 'work'.

Numerical data for ABM modelling is more challenging to acquire than data for SD models. SD models frequently use existing data concerning levels of stocks and rates of flows from existing situations. However, for the authors' research, it has been very challenging to acquire data from existing sources. One opportunity has been to examine the public comment submissions required by the Environment Protection Authority; in determining if the hazardous facility is to go ahead, the EPA is legally required to accept public comment on the plans. The comments were analysed to determine the types or themes of issues of concern to the public. In addition, some numerical data were converted to ratios, eg the number of comments on a given theme compared with the total number of comments. While the prospect of gaining data in this manner is quite attractive, it has its own limitations. It is known to the EPA that public comment statements are heavily biased according to who submits comment, with people in favour of a development much less likely to comment than those opposed to it. In addition, simply counting the number of comments does not provide any insight into who made the comment; intuitively a comment made on behalf of a group of people should have more 'weight' than a comment made by an individual.

If existing numerical data do not prove to be useful, de novo data can be acquired through the use of questionnaires. There are many, many forms that a questionnaire can take, all of which are influenced by the mental model of the modeller. A questionnaire that will produce usable data for ABM modelling must derive information about the variables that the modeller believes influence state changes in the agents while producing frequency and relativity data. Choice based questionnaires, such as Discrete Choice or Contingent Choice methods where the respondent must state a preference for a given choice within a hypothetical scenario or as a real choice are very useful for this purpose (Train 2002). In addition, methods that generate Bayesian distributions and probability statistics can also be used in ABM (AnyLogic

2005).

An interesting method with possible application to ABM is the Cynefin technique (Snowden 2000) of narrative analysis. In this technique, qualitative data are gathered by asking respondents questions that generate stories. The respondents are asked to index their own stories against the characters (agents) and themes (parameters) which appear in their story. While this technique has not been used by the authors, it appears to provide an interesting blend of quantitative data which can be enriched by the narrative analysis. It allows for model development as well as detection of 'surprises' from respondents which can add to the mental model of the modeller.

Discussion

Agent Based Modelling can be a very powerful technique for modelling interactions between groups of agents in order to observe if relatively simple rules can cause unexpected behaviours to arise. But if ABM is to be more than just entertainment, it must be able to reflect real life experiences, and thus must be able to take data from experience and incorporate it into the model.

Qualitative inputs are essential in clarifying the mental model that drives the ABM, however quantitative data must be acquired eventually if the model is to meaningly reflect the life experience. Qualitative data are probably easier to acquire; virtually every stakeholder has an opinion that can contribute to the development of the modeller's mental model. Quantitative data are more difficult to acquire, primarily because it is unlikely that existing data will be in a format to be usable to drive the rules of agent interaction and the data will need to be acquired de novo.

The use of ABM archetypes is a feasible stepping stone between qualitative and quantitative models; however, the software suppliers will need to extend the availability of archetypes and make the process for modification more transparent.

References

Anderson, V & Johnson, L 1997, *Systems Thinking Basics: From Concepts to Causal Loops*, Pegasus Communications, Inc., Waltham, MA. ISBN 1-883823-12-9

AnyLogic 2005, *Standard Training Handout Materials*, XJ Technologies. www.xjtek.com

Axelrod, R 1997, 'Advancing the Art of Simulation in the Social Sciences', in ReH Conte, Rainer (ed.); Terna, Pietro (ed.) (ed.), *Simulating social phenomena.*, Springer, Berlin, pp. 21-40. ISBN: 3-540-63329-4

Axelrod, R & Tesfatsion, L 2005, 'A Guide for Newcomers to Agent_Based Modeling in the Social Sciences', in LTaKL Judd (ed.), *Handbook of Computational Economics, Vol. 2: Agent-Based Computational Economics*, North-Holland, Amsterdam, the Netherlands, vol. 2. ISBN 0444898573

Senge, P, Roberts, C, Ross, R, Smith, B & Kleiner, A 1994, *The Fifth Discipline Fieldbook: Strategies and Tools for Building a Learning Organisation*, Nicholas Brealey Publishing, London. ISBN 1-85788-060-9

Snowden, DJ 2000, 'The Art and Science of Story or "Are you sitting uncomfortably?"', *Business Information Review*, vol. 17, no. 3, pp. 147-56. ISSN: 0266-3821

Strangio, P 2001, *No Toxic Dump! A triumph for grassroots democracy and environmental justice*, Pluto Press, Annadale, NSW ISBN 1 86403 179 4.

Train, K 2002, *Discrete Choice Methods with Simulation: Cambridge University Press*, Cambridge University Press. ISBN: 0-521-81696-3.