Drama theory as the behavioural rationale in agent-based models

J.W. Bryant
Sheffield Hallam University & Warwick Business School, UK

Abstract

Agent-based modelling has recently developed as an important approach to understanding the collective behaviour of individuals interacting, especially in organised and disorganised groups such as organisations, markets or crowds. It has been used as the basis of simulations to explore the dynamics of such systems, and so to predict the implications of deliberate interventions or chance events. However, agent-based models, which represent in an explicit manner the responses of individual decision-makers to the actions of others, can only be as good as the models of human behaviour upon which they are founded. This paper argues that the theory of dramatic resolution provided by drama theory provides suitable underpinning for this aspect of agent-based modelling and indicates how, through simulation, characteristic patterns of confrontation management could be exposed and assessed.

1. Agent-based Modelling

The collective behaviour of autonomous individuals frequently displays characteristics that are both systematic and unexpected. So, for instance Braess's Paradox – that adding new lanes to a highway may actually exacerbate traffic congestion – is just one of many counterintuitive phenomena resulting from the interdependent choices of independent parties. In a useful review, Bonabeau (2002) gives the example of Nasdaq's proposal to reduce the tick size of its stock market trading in an attempt to drive down the spread between bid and ask prices, to demonstrate how agent-based modelling can productively address the challenge of accounting for emergent behaviour. The approach is based upon creating a computer simulation that represents a situation at the level of the individual decision-maker, and which then permits these 'agents' to interact with each other within a context that is shaped by the plans or prospective policy whose implications are being examined.

Agent based modelling has been applied to a wide range of questions about market behaviour from studies of supermarket shoppers to the actions of stock market investors. However it has also been used as an investigative tool to study ways of improving staff motivation, to assess R&D strategy – focusing on the so-called 'selfish team syndrome' wherein an emotionally-committed group try to save a project that should rationally be terminated – to predict the outcome of new employee recruitment policies and to quantify the operational risk of a bank's asset management business. However in all these applications its reliance upon the inbuilt model of human behaviour immediately becomes apparent. As Bonabeau acknowledges, 'an agent-based model will only be as accurate as the assumptions and data that went into it', and this is a criticism that is repeatedly levelled at the approach. It is contended here that drama theory offers a more robust and realistic framework for representing human behaviour within agent-based models, especially when the nature of the interactions between agents demands a collaborative response.

2. Exploring co-operation and conflict

Human society is surprisingly co-operative. That it should be so, rather than exhibiting the victory of self-interest and selfish behaviour that economists or behavioural scientists might have us accept, has preoccupied philosophers since Thomas Hobbes, and is a subject of much recent scrutiny, a useful popular survey of which has been given by Ridley (1996). Among these attempts to explore the logic of co-operation, Axelrod's path-breaking 'tournament' (Axelrod, 1984) that pitted different algorithms – effectively differently programmed agents – against each other in repeated games of Prisoner's Dilemma is probably the best known. The success (in terms of the utility gained in 200 plays) of 'Tit-for-Tat' as a strategy was widely publicised; and this popularisation also attracted widespread criticism from game theorists. As later work (including some explicitly rooted in agent-based approaches) has shown (e.g. Danielson, 2002), although reciprocity is necessary to engender co-operative behaviour, it

is not sufficient to ensure sustained co-operation, because co-operators can be exploited by those with whom they have struck up apparently reciprocal relationships. However, notwithstanding the controversial nature of the results, these experiments demonstrated the flexibility of agent-based computational modelling for understanding competitive and co-operative behaviours.

The rationality mechanism with which agents have been supplied in modelling exercises has usually been based in part upon prediction of others' actions and in part upon a simple learning process that draws upon past experiences. However, as Elliott and Kiel (2002) have recognised, 'agents clearly need to be imbued with greater realism in terms of the range of behaviours'. In particular they highlight a need for greater focus upon cognitive dynamics at the level of the individual. It is contended in this paper that this is precisely the deficiency that drama theory can fill.

3. Collaborative Style

It is an everyday observation that individuals vary in their ability to collaborate with others. Furthermore, a similar evaluation can be applied to groups and organisations. Indeed the development of collaborative capability is widely recognised as an essential strength for businesses and public sector agencies of all types in an increasingly 'joined-up' world based upon fluid inter-organisational relationships and co-operative ventures. A number of studies (Hoffmann, 2002; Vantage Partners, 2002) have attempted to identify the key characteristics of successful collaborators and to suggest routes for building capacity to work in this way.

There is a considerable literature that addresses individual style in conflict management, and specific diagnostic tools such as the Thomas-Kilmann Conflict Mode Instrument (Kilmann and Thomas, 1974 & 1977). The latter suggests five basic conflict management styles - competing, collaborating, compromising, avoiding and accommodating - depending upon whether a party's own and whether others' demands are met. Although all styles are appropriate under some circumstances, people appear preferentially to use just one or two styles regardless of the situation. Since other studies (Volkema and Bergmann, 1995) have demonstrated a good relationship between these styles and specific behaviours in actual conflicts, it looks as though they may be used to anticipate as well as to explain an individual's behaviour.

Building upon this observation, the present author has piloted a diagnostic tool (Bryant, 2003a) that is intended to capture the characteristics of collaborative style. Significantly this is based upon the characteristic dilemmas of conflict and co-operation that are identified by drama theory (Bryant, 2003b). Briefly, drama theory depicts human interaction as involving 'characters' each seeking simultaneously to have others adopt their 'positions' (their advocated solutions) in collaborative situations. They each exert leverage to obtain this outcome by proclaiming 'fallback actions' which they may adopt to encourage or force others to comply, but these and the reliability of any agreements made are all susceptible to credibility and doubt. The dilemmas represent the challenges that each party seeks to overcome either to manage conflict and establish a shared solution, or to manage cooperation and implement joint action. The diagnostic tool assumed that characters might handle the dilemmas they faced in characteristic and repeatable ways, for example, by resigning their position and conceding to others, or by stiffening their resolve and becoming aggressively hostile, and that similar behaviours might be found across different interactions with another party or even across a character's various relationships. However there is presently insufficient evidence available to draw reliable conclusions from this preliminary work.

4. Interacting agents

The notion of collaborative style as something that can potentially be accessed and described in terms of a character's mode of handling drama-theoretic dilemmas suggests a way in which agent-based simulations of complex systems may be modelled. By having agents encounter others with whom they may or may not choose to co-operate, the emergent behaviour of such situations can be tracked. This is exactly the direction taken in the earlier simulations of collaboration described above. However the present paper suggests significantly elaborating this approach by using a finer characterisation of interactive behaviour. The way that this is done will be described in the remainder of this section.

Drama theory recognises six generic dilemmas. Four of these are dilemmas of conflict:

> *Threat Dilemma*: A character would rather not carry its fallback action, as it can see something that it can do unilaterally that will deliver it an even better outcome - so how can I use its threat to pressure others?

Rejection Dilemma: A character would rather adopt another character's solution than carry out its fallback action - so how can it persuade others to its solution?

Positioning Dilemma: A character would rather adopt another character's solution than proceed with its own proposal – so how can it convincingly advocate its own solution?

Persuasion Dilemma: A character sees that another character would prefer to carry out its fallback action that adopt its (the first character's) proposal - so how can it attract others ton its solution?

The other two are dilemmas of collaboration:

Co-operation Dilemma: A character can gain an improvement by unilaterally moving from an agreed solution - so why should others believe that it will maintain their agreement?

Trust Dilemma: A character recognises that another character can gain an improvement by unilaterally moving from their agreed solution - so why should it believe that this temptation will be ignored?

Each dilemma can be resolved by a character changing its preferences between the two outcomes to which the dilemma relates, but such changes bear an emotional 'cost'. So, for example, a Rejection Dilemma can be eliminated by a character strengthening its own belief in the worth of the fallback action (e.g. the media whipping up patriotic fervour in favour of a 'just' war), but this would be accompanied by an air of reluctant aggression regarding the 'inevitable' conflict. Clearly the necessary preference change can always be accomplished by one outcome becoming seen as more desirable than previously, or the other being seen as less desirable. So, to pursue the example of a Rejection Dilemma, this can be removed by coming to see the other character's solution as having less merit than previously appeared or contrastingly as coming to see the Fallback as having more to commend it than at first thought. For some of the dilemmas these routes correspond roughly to becoming more conciliatory or more confrontational.

These observations about dilemma elimination provide the basis for the rules to be used in an agent-based model. To demonstrate this, consider the following simple implementation of a simulation:

1. Establish a population of autonomous agents. The members of this population each have a characteristic way of resolving each of the six drama theoretic dilemmas they encounter. Since, as suggested above, there are two broad means of resolving each dilemma, then there are 2^6 (= 64) different patterns which an agent may possess. Within the simulation randomly select one of these patterns for each agent.

2. Cause the agents to encounter each other. For simplicity, suppose each encounter involves any two agents. Further, assume that within each encounter, each party has a single choice of action (i.e. can do an option, or not do that option). Then there are 4 possible outcomes of an encounter. Each agent will have 4! (=24) possible distinguishable rankings of these outcomes. Then there are 24^2 (=576) possible combinations of preference rankings across the two characters. Within the encounter (say between agents 'A' and 'B'), any of the 4 possible outcomes could be A's Position; there are also 4 possible candidate outcomes to be B's Position, and 4 possibilities for the Fallback. There are therefore potentially 576 x 4^3 (=36864) dramas, since a drama is defined in terms of the characters' Positions, the Fallback and the characters' ranking of the outcomes. Within the simulation randomly select one of these dramas.

3. Any drama will present a predictable set of dilemmas to each agent, these depending upon its features as explained above. Broadly speaking they will attempt first to resolve any Threat Dilemma present (to establish the Fallback); then to resolve the other dilemmas of conflict; and lastly to address the dilemmas of collaboration. Now the agents each possess a characteristic way of handling each of these dilemmas. Within the simulation, implement each agent's mode of dilemma elimination for the dilemmas encountered, considering the dilemmas in the order just indicated. As each dilemma is eliminated (in the manner that the agent's 'personality' suggests) new dilemmas may be created, either for itself or for others. These are then subject to the same process of elimination until the situation involving the pair of agents is resolved.

4. Drama theory hypothesises that dilemmas are 'uncomfortable' for characters and that they will seek to eliminate those they face (though they may seek to create dilemmas for other characters). Since the number and mix of dilemmas with which agents in the simulation are confronted is assumed to be independent of their characteristics, then the 'success' of an agent's strategy can be measured in terms of the ease with which it manages to eliminate its dilemmas. By maintaining a record of the dilemmas generated and resolved over a large number of simulated encounters, an evaluation of the different patterns of dilemma management can be obtained.

The protocol that has just been described is analogous to the experimentation in Axelrod's tournaments insofar as it is designed to highlight the features of 'successful' characters. Clearly it could be further developed in a number of ways, notably:

- The development of relationships through repeated encounters between the same agents, rather than the examination of encounters between randomly chosen pairs of agents might be tested.
- It might no longer be assumed that there is independence between the characteristics of agents and the dramas in which they find themselves. After all, it is plausible, for instance, that pacific agents are liable to find themselves confronted by Persuasion Dilemmas.
- Agents might be expected to 'learn' from their encounters. Rather than persisting in a mode of behaviour that repeatedly creates dilemmas, an agent might tend to shift to a 'better' mode that lessens this discomfort.

From these and other enhancements of the approach it should be possible better to understand how a particular pattern of dilemma management can underwrite a successful interactor.

5. Discussion

This paper has introduced a framework for simulating the behaviour of parties encountering each other in a situation in which either conflict or co-operation may emerge. The novel feature of this model and of the simulation experiment that is suggested is its dependence upon drama theory to define the nature of the confrontation and to provide a means of describing the characteristic ways in which autonomous decision-makers might handle the challenges that are presented. It is a relatively simple model and could clearly be enhanced in a number of ways. Nevertheless, as it is presented here, the model should still generate insights regarding the success or otherwise of different generic stances.

An agent-based approach has the potential to tell us much about the emergent properties of systems involving many agents, as well as about the worth of alternative strategies in multi-party situations. However such simulations also offer a means of investigating the practical implications of drama theory itself, and of beginning to explore how the learning that each agent gains from each dramatic resolution in which it is involved can be incorporated into a wider theory that takes a systemic view of interactions both through time and across a group.

REFERENCES

AXELROD, R. 1984. The Evolution of Co-operation. New York: Basic Books

BONABEAU, E. 2002. Predicting the Unpredictable, *Harvard Business Review*, **March 2002**, 5-11.

BRYANT, J. 2003a. Diagnosing collaborative style. In P. Hibbert (ed.) *Co-creating Emergent Insight: multi-organisational partnerships, alliances and networks*, pp.20-26. Glasgow: University of Strathclyde.

BRYANT, J. 2003b. The Six Dilemmas of Collaboration: inter-organisational relationships as drama. Chichester: Wiley.

DANIELSON, P. 2002. Competition among co-operators: altruism and reciprocity. *Proc. Nat. Acad. Sci. USA*, **99, Suppl. 3**, 7237-7242.

ELLIOTT, E and KIEL, L.D. 2002. Exploring cooperation and competition using agent-based modeling. *Proc. Nat. Acad. Sci. USA*, **99, Suppl. 3**, 7193-7194.

HOFFMANN, W. 2002. Emergence and institutionalisation of alliance management capability: empirical evidence from a cross-sectional and a longitudinal study. Paper to *Strategic Management Society Conference*. Vancouver: SMS.

KILMANN, R.H. and THOMAS, K.W. 1974. Thomas-Kilmann Conflict Mode Instrument. Palo Alto, CA: Consulting Psychologists Press.

KILMANN, R.H. and THOMAS, K.W. 1977. Developing a forced-choice measure of conflict-handling behaviour: the mode instrument. *Educational and Psychological Measurement*, **37**, 309-325.

RIDLEY, M. 1996. The Origins of Virtue. London: Viking

VANTAGE PARTNERS. 2002. *Making Partnerships Work: institutionalising relationship management capability*. Cambridge MA: Vantage Partners.

VOLKEMA, R. and BERGMANN, T.J. 1995. Conflict styles as indicators of behavioural patterns in interpersonal conflicts. *Journal of Social Psychology*, **135**, 5-15.