

Game theoretic and decision theoretic agents

Simon Parsons and Michael J. Wooldridge,
Department of Electronic Engineering,
Queen Mary and Westfield College,
University of London,
London E1 4NS,
United Kingdom.

July 17, 2001

1 Introduction

In the last few years, increasing numbers of members of the agent community have been adopting techniques from game theory and decision theory. Broadly speaking, decision theory [15] is a means of analysing which of a series of options should be taken when it is uncertain exactly what the result of taking the option will be. Decision theory concentrates on identifying the “best” decision option, where the notion of “best” is allowed to have a number of different meanings, of which the most common is that which maximises the expected utility of the decision maker. Game theory [1] can be considered as a variant of decision theory in which the outcome of taking a particular decision is dependent upon the actions of another, usually an opponent which is trying to maximise its own benefit at the cost of the decision maker. Alternatively, game theory can be considered a mechanism for analysing games between two players in which each gets to choose a move from some limited set of options and, depending on what both have chosen, each receives a payout. Since the payout one player receives depends upon the move made by the other then, in order to maximise its payout, each player needs to take into account the likely move taken by its opponent. From this perspective decision theory can be considered to be the study of games played against nature, an opponent which does not look to gain the best payout, but rather acts randomly.

Given this brief description, it is no surprise to learn that many of the applications of game theory in agent systems has been to analyse multi-agent interactions, particularly those involving negotiation and co-ordination activities. Indeed, as far back as 1985 it was suggested as a means by which software agents can make deals [16]. Because making these deals does not require any communication, since both agents just compute the best outcome and know that the party they are dealing with will do the same, game theory was also suggested as a mechanism by which software agents can co-ordinate their actions without the need for communication [5] since all the agents involved could work out what the most mutually beneficial outcome of any interaction was. Similarly, decision theory has often been used to analyse the interactions between agents and their environments.

This increasing interest in game theory and decision theory in the agents community led us to believe that the time was ripe to hold a workshop which focused exactly on these matters. This was held in London on the 3rd July 1999 in conjunction with the 5th European Conference on Quantitative and Symbolic Approaches to Reasoning and Uncertainty [10], and comprised an invited talk and seven submitted papers. Broadly these split into two groups—those which focused on game theory and those which focussed on decision theory. All of the papers described here are available in the workshop proceedings, which is available on-line at <http://www2.elec.qmw.ac.uk/~sp/events/gtdt99/proceedings.html>.

2 Game theory

The workshop began with an invited talk by Nir Vulkan, entitled “Bargaining with deadlines”. This described Nir’s recent work with Tuomas Sandholm [17] which considered the behaviour of pairs of negotiating game-theoretic agents. The aspect of this work which sets it apart from previous efforts in this area,

for example [13], is that in Sandholm and Vulkan's work each agent has some deadline by which it has to complete the negotiation, and this deadline is known only to that agent. The main result of the work is to show that the agent with the earliest deadline is always at a disadvantage. The best strategy for the agent with the later deadline is to refuse any offers and wait until the other agent runs out of time and has to concede—any offer made by a rational opponent before the deadline will be at a worse price for the recipient than that offered at the deadline. Similarly, the best strategy for the agent with the earlier deadline is to wait until the deadline and then concede—the only offer that a rational opponent will accept before the deadline will be at a worse price for the agent making the offer than that it can obtain at the deadline. This result is particularly interesting in that it undermines the classic motivation for using agents as proxies in negotiations, which is that such agents can perform the calculations necessary to obtain rational (in a game theoretic sense) behaviour much quicker than humans can, thus allowing negotiations to be speedily concluded. If agents have to wait on long deadlines before reaching a resolution, this advantage cannot be realised.

As discussed by Guttman *et al.* [7] and Crabtree [3], one of the niches in which autonomous agents are rapidly proving their worth is electronic commerce. Here agents help to “grease the wheels” that must turn in order that goods and services can be bought and sold across the Internet. One class of these wheel-greasing agents are shopbots, agents which search the Internet on behalf of consumers, comparing prices across dozens of web sites. Shopbots thus help to cut consumers' costs, not just in the sense of allowing them to find the cheapest source of the good they require, but also in the more general sense of reducing the cost of obtaining optimal price and quality. Shopbots can also help to reduce the costs of suppliers, by reducing the cost of evaluating, updating and advertising prices, and thus have the potential to significantly affect the way markets operate. As a result Kephart and Greenwald [11] have investigated the impact of shopbots in single commodity markets, modelling the behaviour of both buyers and sellers using game theoretic techniques, and at the workshop presented a summary of their results.

Shopbots can be considered as operating on the side of buyers in an electronic market helping them to find the best deals. The complementary kind of agent which operates on the side of sellers are what have been called “pricebots” autonomous agents which fix the prices charged by a seller in order to secure the best price that sellers are prepared to pay. Tesauro and Kephart [19] have investigated how such price-setting agents can be made adaptive, in particular how they can make use of Q-learning [22], an approach which factors in the long-term expected reward for a given action taken in a given state. They considered a number of different model economies, including one in which buyers are assumed to make use of shopbots, and found that the Q-learning approach led, broadly speaking, to increased profits for sellers, in part because it reduced the effect of price wars (when sellers repeatedly undercut each other in an attempt to capture a bigger share of the market).

One way of viewing game theoretic analyses of negotiations between agents is as a process of resource allocation. Take the pricebot scenario discussed above. There we have a seller who has a large quantity of items for sale, and a set of buyers each of which has some money which it wants to trade for one of more items. Once a price is fixed, agents buy and sell. As a result, they reallocate the items and money amongst themselves in accordance with their relative preferences for those resources. Now, as described, the situation is relatively simple, but it rapidly becomes more complex when there are a number of different items on offer because of the need to trade-off the advantages of bundles of different types of items. It is this more complex kind of resource allocation problem which is the subject of Conen's paper [2]. As he points out, most attempts to consider the problem have considered the case in which the various items in a bundle have additive valuations—the value of the bundle is just the sum of the values of the components. In contrast, Conen provides a mechanism for resolving the resource allocation problem when the valuations combine in a range of different ways, including sub-additivity and positive synergy¹, which ensures that resources are used efficiently and agents end up with bundles which best match their requirements.

Conen's paper, as with that of Sandholm and Vulkan, can be considered to be an extension of classical game theoretic notions. This is also true of the work of Stirling and Goodrich [18]. In applications of game theory in multi-agent systems, agents are usually taken to maximise their expected utility. This approach, however, is not always practical—there are often bounds on computational resources which prevent the optimal solution being computed. As a result, there has been much interest in computing solutions under

¹He distinguishes between sub-additivity and negative synergy. Sub-additivity is when the sum of the values of the parts is greater than the value of the whole, and negative synergy is when the sum of the values of the parts is much greater than the value of the whole.

bounded rationality, that is approaches which aim to be rational in the sense of computing the solution with maximum expected utility, but which acknowledge bounds on their resources, and so relax one or more assumptions of the optimal approach. Stirling and Goodrich [18], consider the adaptation of Simon’s idea of satisficing—that is searching for an optimal solution until the cost of continuing the search outweighs the improvement in solution that further work will bring. They do this by introducing the notion of praxeic utility, a measure which explicitly models the resources consumed, and allows these to be balanced against the desire to obtain the best solution.

3 Decision theory

When constructing multi-agent systems, one major element in the design is the mechanism the agents use for making decisions about which actions to adopt in order to achieve their goals. What is usually required is as mechanism which ensures that the actions adopted lead to benefits for both individual agents, and the community of which they are a part. As Hogg and Jennings [8] argue, while it is possible to make use of classical decision theory to build agents which are *self interested* and ensure that their actions benefit them, such agents will not necessarily act in a way which benefits the whole community. To address this latter issue, Hogg and Jennings introduce a variant of classical decision theory in which the utility function of an individual agent explicitly includes a component which models the utilities of other agents. This then makes it possible to calculate the expected utility of a particular action for the community as a whole. Varying the weight given to the utility of other agents in this calculation makes it possible to model different types of agents, from the completely selfish to the totally selfless, and Hogg and Jennings give experimental results which suggest that, in certain circumstances, moving away from completely selfish agents can be beneficial for the community as a whole.

Gmytrasiewicz [6], while basically interested in solving the same problem as Hogg and Jennings, approaches it in a different way. This involves a more conventional use of decision theory as embodied in the use of Bayesian networks [14] and influence diagrams [9], in particular Koller and Pfeffer’s object-oriented extension [12]. When building a multi-agent system using Gmytrasiewicz’s approach, each agent is provided with a model of the world, which it uses to make decisions by maximising expected utility in the normal way. However, as part of its model of the world, each agent has a model of the decision making mechanism of the other agents in the system², and the behaviour of all the other agents is therefore factored into each agent’s decision making. Since each agent can independently compute the best course of action, as in [5] there is no need for communication between agents. This approach has been tested experimentally in a scenario in which the agents control anti-missile defences and the aim is to minimise the damage caused by incoming missiles (a scenario not fundamentally different to that of Hogg and Jennings where the agents are controlling fire-fighting equipment and the aim is to minimise fire damage). The results show that the approach not only out-performs human controllers, partly because human controllers tended to focus redundant effort on the most threatening attacker, but also out-performs teams of agents which individually maximise utility but without considering the decision making of the other agents.

Both Hogg and Jennings and Gmytrasiewicz are using the basic decision theoretic mechanism of expected utility unaltered, adapting it for the multi-agent case by extending the scope of the calculation to take other agents into account. As a result, both take utilities as primitive. In contrast, van der Torre and Weydert [20] start with an agent’s goals, and consider how an agent might reason about its goals and use these to define its utility function. Their approach is logic-based, and thus an extension of recent work on qualitative decision theory [4].

All the papers described so far in this section have been concerned with the use of decision theory alone, and thus deal with combinations of beliefs and utilities. In the case of both Hogg and Jennings and Gmytrasiewicz, these beliefs are distributed over states of affairs (roughly speaking conjunctions of propositions), while in the work of van der Torre and Weydert, the beliefs are taken over individual propositions. Thus, in neither case, is there much structure to the items that beliefs are distributed over. Vane and Lehner [21], on the other hand, build belief distributions over much more complex objects—in fact they deal with beliefs over games, in the sense of game theory. In essence, their hypergame framework allows a agent in a game theoretic setting to hedge its bets about what its opponent is doing. It does this by identifying a set of possible games, representing the possible behaviours that its opponent might engage in, building a

²The recursive nature of this kind of model is carefully considered in the paper.

probability distribution over these games, and evaluating the best moves by the usual maximum expected utility algorithm. The result is an elegant formalism which is a strict generalisation of both game theory and decision theory, and which looks to be a useful tool for building future generations of mixed game theoretic and decision theoretic agents.

4 The future

In addition to the usual nebulous outcome in terms of contacts made, information disseminated, and discussions partaken in, there are two very concrete of this workshop on Decision Theoretic and Game Theoretic Agents. The first is a special issue of the journal *Autonomous Agents and Multi-Agent Systems*, edited by the authors, which will contain revised and extended versions of the papers by Greenwald and Kephart, Gmytrasiewicz, Kephart and Tesauro, Stirling and Goodrich, and van der Torre and Weydert. The second outcome is another workshop on the same topic. This will be held in conjunction with the 4th International Conference on Multi-Agent Systems (ICMAS 2000), in Boston on 7th July 2000, and will be chaired by Simon Parsons and Piotr Gmytrasiewicz.

References

- [1] K. Binmore. *Fun and Games: A Text on Game Theory*. D. C. Heath and Company, Lexington, MA, 1992.
- [2] W. Conen. Economic co-ordination, bundled goods, and the impact of complementarities. In S. Parsons and M. J. Wooldridge, editors, *Proceedings of the First Workshop on Game Theoretic and Decision Theoretic Agents*, pages 1–12, 1999.
- [3] B. Crabtree. What chance software agents? *The Knowledge Engineering Review*, 13(2):131–136, 1998.
- [4] J. Doyle and R. H. Thomason. Background to qualitative decision theory. *AI Magazine*, 20(2):55–68, 1999.
- [5] M. R. Genesereth, M. L. Ginsberg, and J. S. Rosenschein. Cooperation without communication. In *Proceedings of the 5th National Conference on Artificial Intelligence*, pages 51–57, Los Altos, CA, 1986. Morgan Kaufmann.
- [6] P. Gmytrasiewicz. Implementing a decision-theoretic approach to game theory for socially competent agents. In S. Parsons and M. J. Wooldridge, editors, *Proceedings of the First Workshop on Game Theoretic and Decision Theoretic Agents*, pages 13–28, 1999.
- [7] R. Guttman, A. G. Moukas, and P. Maes. Agent-mediated electronic commerce. *The Knowledge Engineering Review*, 13(2):147–159, 1998.
- [8] L. Hogg and N. R. Jennings. Variable sociability in agent-based decision making. In S. Parsons and M. J. Wooldridge, editors, *Proceedings of the First Workshop on Game Theoretic and Decision Theoretic Agents*, pages 29–42, 1999.
- [9] R. A. Howard and J. E. Matheson. Influence diagrams. In R. A. Howard and J. E. Matheson, editors, *Readings on the Principles and Applications of Decision Analysis*, pages 719–762. Strategic Decisions Group, Menlo Park, CA, 1984.
- [10] A. Hunter and S. Parsons, editors. *Symbolic and Quantitative Approaches to Reasoning and Uncertainty*. Springer Verlag, Berlin, Germany, 1999.
- [11] J. O. Kephart and A. R. Greenwald. Shopbot economics. In S. Parsons and M. J. Wooldridge, editors, *Proceedings of the First Workshop on Game Theoretic and Decision Theoretic Agents*, pages 43–55, 1999.

- [12] D. Koller and A. Pfeffer. Object-oriented Bayesian networks. In D. Geiger and P. P. Shenoy, editors, *Proceedings of the 13th Conference on Uncertainty in Artificial Intelligence*, pages 302–313, San Francisco, CA, 1997. Morgan Kaufmann.
- [13] S. Kraus, J. Wilkenfeld, and G. Zlotkin. Multiagent negotiation under time constraints. *Artificial Intelligence*, 75:297–345, 1995.
- [14] J. Pearl. *Probabilistic reasoning in intelligent systems; networks of plausible inference*. Morgan Kaufmann, San Mateo, CA., 1988.
- [15] H. Raiffa. *Decision analysis: introductory lectures on choices under uncertainty*. Addison Wesley, Reading, MA., 1968.
- [16] J. S. Rosenschein and M. R. Genesereth. Deals among rational agents. In *Proceedings of the 9th International Joint Conference on Artificial Intelligence*, pages 91–99, Los Altos, CA, 1985. Morgan Kaufmann.
- [17] T. Sandholm and N. Vulkan. Bargaining with deadlines. In *Proceedings of the 16th National Conference on Artificial Intelligence*, pages 44–51, San Mateo, CA, 1999. AAAI Press/MIT Press.
- [18] W. Stirling and M. Goodrich. Satisficing equilibria: a non-classical theory of games and decisions. In S. Parsons and M. J. Wooldridge, editors, *Proceedings of the First Workshop on Game Theoretic and Decision Theoretic Agents*, pages 56–70, 1999.
- [19] G. Tesauro and J. O. Kephart. Pricing in agent economies using multi-agent q-learning. In S. Parsons and M. J. Wooldridge, editors, *Proceedings of the First Workshop on Game Theoretic and Decision Theoretic Agents*, pages 71–86, 1999.
- [20] L. van der Torre and E. Weydert. Goals, desires, utilities and preferences. In S. Parsons and M. J. Wooldridge, editors, *Proceedings of the First Workshop on Game Theoretic and Decision Theoretic Agents*, pages 87–102, 1999.
- [21] R. R. Vane and P. E. Lehner. Using hypergames to select plans in adversarial environments. In S. Parsons and M. J. Wooldridge, editors, *Proceedings of the First Workshop on Game Theoretic and Decision Theoretic Agents*, pages 103–111, 1999.
- [22] C. J. C. H. Watkins. *Learning from delayed rewards*. PhD thesis, Cambridge University, 1989.