Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence (2008)

Game Theory Pragmatics: A Challenge for AI

Yoav Shoham
Computer Science Department
Stanford University

Abstract

Game theory has been playing an increasingly visible role in computer science in general and AI in particular, most notably in the area of multiagent systems. I briefly list the areas where most of the action has been in the past decade or so. I then suggest that going forward, the most dramatic interaction between computer science and game theory – with a special role for AI – could be around what might be called *game theory pragmatics*.[1]

1 Introduction

Game theory has influenced many fields, from economics (historically its initial focus) to political science to biology, and many others. In recent years its presence in computer science has become impossible to ignore. It features routinely in the leading conferences and journals of AI, theory, certainly electronic commerce, as well as in networking and other areas of computer science. There are several reasons for this. One is application pull; the Internet calls for analysis and design of systems that span multiple entities with diverging information and interests. Game theory, for all its limitations, is by far the most developed theory of such interactions. The other is technology push; the mathematics and scientific mindset of game theory are similar to those of many computer scientists. Indeed, it is interesting to note that modern computer science and modern game theory in large measure originated at the same place and time, namely at Princeton under the leadership of John von Neumann.[2]

In this paper I would like to do two things. First, to very briefly identify the main areas of interaction between AI and game theory so far. Second, to point, in slightly greater length, to where the most interesting interaction yet may lie, an area which is still relatively under-explored.

2 Lessons from Kalai (1995)

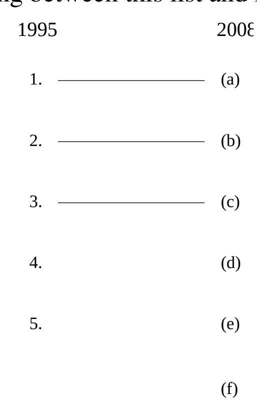
My departure point will be a thirteen-year-old survey paper by E. Kalai (Kalai 1995). Written by a game theorist with algorithmic and optimization sensibilities, and geared primarily towards computer scientists, the paper took stock of the interaction among game theory, operations research and computer science at the time. The paper points to the following areas:

1. Graphs in games.
2. The complexity of solving a game.
3. Multi-person operations research.
4. The complexity of playing a game.
5. Modelling bounded rationality.

The reason I start with this paper, beside the fact that it's interesting to start with the perspective of a non-computer-scientist, is the comparison with current CS-GT interaction, since both the match and the mismatch are instructive. When one looks at the interaction between CS and GT taking place today, broadly speaking one can identify the following foci of action:

(a) Compact game representations.
(b) Complexity of, and algorithms for, computing solution concepts.
(c) Algorithmic aspects of mechanism design.
(d) Game theoretic analysis inspired by specific applications.
(e) Multiagent learning.
(f) Logics of knowledge and belief and other logical aspects of games.

The crude mapping between this list and Kalai's as follows:



[1] This article is excerpted from (Shoham 2008).

[2] I thank Moshe Tennenholtz for this observation, which is especially true of GT and AI.

In a full version of this paper (Shoham 2008), I discuss the match up $(1 - a, 2 - b, 3 - c)$ as well as the currently active areas that were glossed over by Kalai (d, e, f). Here I will focus on the orphans on the other side, $(4, 5)$, discussed by Kalai but not yet been picked up as vigorously as they might, especially within AI.

3 Lessons from Linguistics

The field of linguistics distinguishes among syntax, semantics, and pragmatics. Syntax defines the form of language, semantics its meaning, and pragmatics its use. While the three interact in important ways, the distinction has proved very useful. I believe that game theory may do well to make similar distinctions, and that CS can help in the process. Just as in the case in linguistics, it is unlikely that game theory pragmatics will yield to unified clean theories as do some syntactic and semantic theories. But by the same token I expect game theory pragmatics to be as important to applying game theory in practice as pragmatics are to analyzing full human discourse, or understanding language by computers.

The distinction between the syntax and semantics of games is I think quite important. I feel that some of the disputes within game theory regarding the primacy of different game representations (for example, the strategic and extensive forms), suffer from the lack of this distinction. It is perhaps presumptuous for CS to intrude on this debate, except insofar as it lends logical insights, since I do think that logic is a useful lens through which to look at these issues (cf. (van Benthem 1998)). But perhaps indeed this is more the role of mathematical logic than of CS *per se*.

Where CS can truly lead the way is I think on the pragmatics of game theory. Game theory as we know it embodies radical idealizations, which include the infinite capacity of agents to reason and the infinite mutually-recursive modelling of agents. Backing off from these strong assumptions has proven challenging. A fairly thin strand of work under the heading of "bounded rationality" studies games played by automata (cf. (Rubinstein 1998)). This is an important area of research, and sometimes makes deep connections between the two fields. Early results, for example, showed that one of the well known pesky facts in game theory – namely that constant 'defection' is the only subgame-perfect equilibrium in the finitely-repeated prisoner's dilemma game – ceases to hold true if the players are finite automata with sufficiently few states (Neyman 1985; Papadimitriou & Yannakakis 1994). A more recent result shows that when players in a game are computer programs, one obtains phenomena akin to the Folk Theorem for repeated games (Tennenholtz 2004).

This connection between theoretical models of computation and game theory is I believe quite important and beautiful, but a fairly narrow interpretation of the term 'bounded rationality'. The term should perhaps should be reserved to describe a much broader research agenda, an agenda which may encourage more radical departures from the traditional view in game theory. Let me mention two directions that I think would be profitable (and hard) to pursue under this broader umbrella.

When one takes seriously the notion of agent's limited reasoning powers, it is not only some of the answers that begin to change; the questions themselves come into question. Consider the basic workhorses of game theory – the Nash equilibrium, and its various variants. They have so far served as the very basic analysis tool of strategic interactions. Questioning the role of equilibrium analysis will be viewed by some in GT as act of heresy, but real life suggests that perhaps we have no choice. For example, in the trading agent competition (TAC), Nash equilibrium of the game did not play a role in almost any participating program (Wellman, Greenwald, & Stone 2007), and this is certainly true of the more established chess and checkers competitions. It is premature to write off the Nash equilibrium as irrelevant. For example, one program (see chapter 8 of (Wellman, Greenwald, & Stone 2007)) did in fact make use of what can be viewed as approximate empirical NE. Another striking example is the computation of equilibria in a simplified game tree by a top scoring program in a poker competition (Zinkevich *et al.* 2007). It could be argued that maxmin strategies, which coincide with equilibrium strategies in zero-sum games, do play an important pragmatic role. But certainly computation of either maxmin or equilibrium strategies in competitions has been the exception to the rule. The more common experience is that one spends the vast majority of the effort on traditional AI problems such as designing a good heuristic function, searching, and planning. Only a little time – albeit, important time – is spent reasoning about the opponent. The impact of such pragmatic considerations on game theory can be dramatic. Rather than start from very strong idealizing assumptions and awkwardly try to back off from them, it may prove more useful and/or accurate to start from assumptions of rather limited reasoning and mutual modeling, and judiciously add those as is appropriate for the situation being modeled. What exact incremental modeling approach will win out is yet to be seen, but the payoff for both CS and game theory can be substantial.

The second direction is radical in a different way. Game theory adopts a fairly terse vocabulary, inheriting it from decision theory and the foundations of statistics.[3] In particular, agents have "strategies" which have minimal structure, and motivations which are encapsulated in a simple real-valued utility function (which in fact carries even less information than is suggested by the use of numbers, since the theory is unchanged by any positive affine transformation of the numbers). In real life, and in computer programs attempting to behave intelligently, we find use for a much broader vocabulary. Agents are *able* to take certain actions and not others, have *desires*, *goals* and *intentions* (the belief-desire-intention combination giving rise to the pun 'BDI agent architecture'), make *plans*, and so on. Apparently these abstract notions are useful to both effect intelligent behavior and reason about it. Philosophers have written about it (e.g.,

[3] Parenthetically it can be remarked that Savage's setting (Savage 1954) on which the modern Bayesian framework is based does not have an obvious extension to the multi-agent case. However this is not the focus of the point I'm making here.

(Bratman 1987)), and there have been attempts – albeit preliminary ones – to formalize these intuition (starting with (Cohen & Levesque 1990)). Some in AI have advocated embracing an even broader vocabulary of emotions (e.g., the recent provocative if informal (Minsky 2007)). Is game theory missing out by not considering these concepts?

Acknowledgements

I thank Alon Altman, Joe Halpern, Sam Ieong, Daphne Koller, Mike Wellman and anonymous referees for useful help and suggestions. Of course all opinions, errors, and erroneous opinions are mine alone. This work has been supported by NSF grant ITR-0205633.

References

Bratman, M. E. 1987. *Intention, Plans, and Practical Reason*. CSLI Publications, Stanford University.

Cohen, P. R., and Levesque, H. J. 1990. Intention is choice with commitment. *Artificial Intelligence* 42(2-3):213–261.

Kalai, E. 1995. Games, computers, and O.R. In *ACM/SIAM Symposium on Discrete Algorithms*.

Minsky, M. 2007. *The Emotion Machine: Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind*. Simon and Shuster.

Neyman, A. 1985. Bounded complexity justifies cooperation in finitely repeated prisoner's dilemma. *Economic Letters* 227–229.

Papadimitriou, C. H., and Yannakakis, M. 1994. On bounded rationality and computational complexity. In *STOC: Proceedings of the Symposium on the Theory of Computing*, 726–733.

Rubinstein, A. 1998. *Modeling Bounded Rationality*. MIT Press.

Savage, L. J. 1954. *The Foundations of Statistics*. New York: John Wiley and Sons. (Second Edition, Dover Press, 1972).

Shoham, Y. 2008. Computer science and game theory. *Communications of the ACM*.

Tennenholtz, M. 2004. Program equilibrium. *Games and Economic Behavior* 49:363–373.

van Benthem, J. 1998. When are two games the same? In *LOFT-III*. (ILLC preprint, 1999).

Wellman, M. P.; Greenwald, A.; and Stone, P. 2007. *Autonomous Bidding Agents: Strategies and Lessons from the Trading Agent Competition*. MIT Press.

Zinkevich, M.; Bowling, M.; ; and Burch, N. 2007. A new algorithm for generating equilibria in massive zero-sum games. In *Proceedings of the Twenty-Second Conference on Artificial Intelligence (AAAI)*, 788–793.

About the author

Yoav Shoham is Professor of Computer Science at Stanford University, where he has been since 1987. He is a Fellow of AAAI, Charter Member of the Game Theory Society, and recipient of the 2008 ACM-SIGART Autonomous Agents Research Award. His latest three books are:

- *Multiagent Systems: Algorithmic, Game Theoretic and Logical Foundations.* Y. Shoham and K. Leyton-Brown. Cambridge University Press, 2008.
- *Essentials of Game Theory: A Concise and Rigorous Introduction for Scientists and Engineers.* K. Leyton-Brown and Y. Shoham. Morgan Claypool Publishers, 2008.
- *Combinatorial Auctions.* P. Cramton, Y. Shoham, and R. Steinberg (Editors). MIT Press, 2006.