Higher-order social cognition in rock-paper-scissors: A simulation study (Extended Abstract)

Harmen de Weerd, Rineke Verbrugge, Bart Verheij
Department of Artificial Intelligence, University of Groningen
{hdeweerd,rineke,b.verheij}@ai.rug.nl

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence

General Terms

Algorithms

Keywords

iterated reasoning, repeated games, theory of mind, opponent modeling, agent-based simulation

1. INTRODUCTION

In settings where multiple strategic agents perform actions that influence each other's decision-making process, it is often necessary to accurately predict the behaviour of others in order to respond appropriately [9]. One option to do so is by modeling an opponent explicitly, e.g. through dynamic epistemic logic [13], Interactive POMDPs [6], multi-agent influence diagrams [7], or iterated best-response models such as cognitive hierarchy models [2] and level-n theory [1]. These models allow for recursive modeling of an opponent, by modeling the opponent as an opponent-modeling agent itself, creating increasingly complicated models to predict the actions of increasingly sophisticated opponents.

In humans, the ability to predict the actions of others by explicitly attributing to them unobservable mental content, such as beliefs, desires, and intentions, is known as *theory of mind* [10] or *social cognition*. Experiments in which humans play games show evidence that humans use theory of mind recursively in their decision-making process [8]. For example, when asked to search for a hidden object in one of four boxes, three of which are labeled 'A' and one of them 'B', participants tend to ignore the box labeled 'B', using their nested belief that a hider would believe that a seeker would consider the most obvious place to search for a hidden object to be the box labeled 'B' [4]. Whether any non-human species makes use of theory of mind is a controversial matter [3, 11], and although recursive opponent modeling could continue indefinitely, humans only use higher-order theory of mind (i.e. recursive theory of mind) up to a certain point [14]. In an evolutionary sense, the costs of using higher orders of theory of mind may outweigh the benefits.

Appears in: *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2012)*, Conitzer, Winikoff, Padgham, and van der Hoek (eds.), 4-8 June 2012, Valencia, Spain.

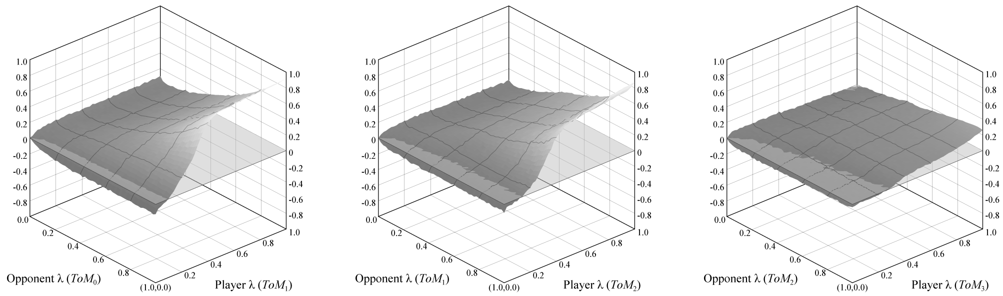
Agent-based modeling has proven to be a useful research tool to investigate how behavioural patterns emerge from interactions between individuals. We have used this approach to investigate the effectiveness of recursive opponent modeling in the setting of a specific game called Limited Bidding [5]. The results in this setting suggest that there may be a limit to the advantage that can be obtained through recursive opponent modeling, but these limitations may also be caused by the specific game structure. Here, we apply the model presented in [5] to rock-paper-scissors (RPS), perhaps the most transparent non-trivial setting in which the role of theory of mind can be investigated. In addition to the standard RPS game, in which 'rock' defeats 'scissors', 'scissors' defeats 'paper', and 'paper' defeats 'rock', we also investigate two variations. Elemental RPS (ERPS) preserves the structure of RPS with a unique best response to each action, but confronts agents with a choice out of five available actions. In rock-paper-scissors-lizard-Spock (RPSLS), agents choose from five actions as well, but this setting differs from ERPS in that each action is defeated by two other actions (e.g. both 'lizard' and 'scissors' defeat 'paper').

2. SIMULATION THEORY OF MIND

In our approach, an agent tries to take advantage of regularities in his opponent's strategy by predicting her behaviour through simulation-theory of mind [9]. An agent takes the perspective of the opponent, and simulates his opponent's decision-making process by making the decision himself. Through the implicit assumption that the opponent's thought process can be accurately modeled by his own, the agent predicts that his opponent will make the same decision he would have made if the roles were reversed.

A zeroth-order theory of mind agent models patterns in his opponent's behaviour, but does not attribute any mental content to her. In contrast, a first-order theory of mind agent considers the possibility that his opponent is trying to win the game for herself, and that she reacts to the choices he makes. For example, suppose that the first-order theory of mind agent remembers that he previously played 'rock' against the opponent he is facing. He realizes that if the roles were reversed, and he would remember her to have played 'rock' before, the agent would play 'paper' more often. The first-order theory of mind agent has the ability to attribute this thought process to his opponent, and predict that she will play 'paper' more often. Given this prediction, the agent reasons that he should play 'scissors' more often.

A second-order theory of mind agent also models his opponent as a first-order theory of mind agent. He believes



(a) First-order (b) Second-order (c) Third-order

Figure 1: Average performance of theory of mind agents playing rock-paper-scissors against an opponent of a lower order.

that his opponent may be putting herself in his position. If the second-order theory of mind agent remembers his opponent to have played 'rock' in previous encounters, he would therefore believe her to predict that he will be playing 'paper' more often. As a result, the second-order theory of mind agent would predict his opponent to play 'scissors' more often, in which case he should play 'rock' more often himself.

3. RESULTS

Using the mathematical model from [5], we determined the performance of the theory of mind agents described in the previous section by placing them in competition. Figure 1 shows the results for the RPS game as a function of the learning speed of both the agent and his opponent. We find that first-order and second-order theory of mind agents clearly outperform opponents that are more limited in their ability to model others. A third-order theory of mind agent mostly outperforms a second-order theory of mind opponent as well, but only marginally.

Results of similar competitions in the ERPS variation suggest that these diminishing returns on higher orders of theory of mind found in RPS are not related to the number of actions available to the agents. When agents choose from five instead of three possible actions, performance of third-order theory of mind agents only improves when playing against opponents that are unable to make use of theory of mind, and play according to a stationary mixed strategy.

Compared to the results of RPS, performance of theory of mind agents in RPSLS is greatly reduced. This suggests that the effectiveness of theory of mind is dependent on the existence of a unique best response. One explanation for this low performance is that when an agent is indifferent between two actions, he chooses either one with equal probability. A slight asymmetry, such that one option is preferable over the other, may therefore benefit agents making use of higher-order theory of mind. Such asymmetries may create a focal point [12] for agents with a lower order of theory of mind, which may result in more predictable behaviour.

4. CONCLUSION

Our results in the RPS game are qualitatively equivalent to those obtained in the setting of Limited Bidding [5]. This shows that the results reported in [5] are not exclusive to the specific game setting studied there, but are generalizable to games of different designs. This provides further support for the hypothesis that first-order and second-order theory of mind provide a clear advantage over opponents of a lower order, while deeper levels of recursion help less [5].

5. ACKNOWLEDGMENTS

This work was supported by the Netherlands Organisation for Scientific Research (NWO) Vici grant NWO 277-80-001.

6. REFERENCES

[1] M. Bacharach and D. Stahl. Variable-frame level-n theory. *Game Econ. Behav.*, 32(2):220–246, 2000.

[2] C. Camerer, T. Ho, and J. Chong. A cognitive hierarchy model of games. *Q. J. Econ.*, 119(3):861–898, 2004.

[3] P. Carruthers. Meta-cognition in animals: A skeptical look. *Mind Lang.*, 23(1):58–89, 2008.

[4] V. Crawford and N. Iriberri. Fatal attraction: Salience, naïveté, and sophistication in experimental "Hide-and-Seek" games. *Am. Econ. Rev.*, 97(5):1731–1750, 2007.

[5] H. de Weerd and B. Verheij. The advantage of higher-order theory of mind in the game of limited bidding. In J. van Eijck and R. Verbrugge, editors, *Proc. Workshop Reason. About Other Minds*, CEUR Workshop Proceedings 751, pages 149–164, 2011.

[6] P. Gmytrasiewicz and P. Doshi. A framework for sequential planning in multiagent settings. *J. Artif. Intell. Res.*, 24(1):49–79, 2005.

[7] D. Koller and B. Milch. Multi-agent influence diagrams for representing and solving games. *Game Econ. Behav.*, 45(1):181–221, 2003.

[8] B. Meijering, H. van Rijn, N. Taatgen, and R. Verbrugge. I do know what you think I think: Second-order theory of mind in strategic games is not that difficult. In *Proc. 33rd Conf. Cogn. Sci. Soc.*, pages 2486–2491. Cognitive Science Society, 2011.

[9] S. Nichols and S. Stich. *Mindreading*. Oxford University Press, USA, 2003.

[10] D. Premack and G. Woodruff. Does the chimpanzee have a theory of mind? *Behav. Brain Sci.*, 1(4):515–526, 1978.

[11] M. Schmelz, J. Call, and M. Tomasello. Chimpanzees know that others make inferences. *PNAS*, 108(7):3077–3079, 2011.

[12] R. Sugden. A theory of focal points. *Econ. J.*, 105(430):533–550, 1995.

[13] W. van Ditmarsch, H. van der Hoek and B. Kooi. *Dynamic Epistemic Logic*. Springer, 2007.

[14] R. Verbrugge. Logic and social cognition: The facts matter, and so do computational models. *J. Philos. Logic*, 38:649–680, 2009.