ECAI 2014
T. Schaub et al. (Eds.)

doi:10.3233/978-1-61499-419-0-985

Predicting Agents' Behavior by Measuring their Social Preferences

Kan-Leung Cheng[1] and **Inon Zuckerman**[2] and **Dana Nau**[3] and **Jennifer Golbeck**[4]

Abstract. There are many situations in which two or more agents (e.g., human or computer decision makers) interact with each other repeatedly in settings that can be modeled as repeated stochastic games. In such situations, each agent's performance may depend greatly on how well it can predict the other agents' preferences and behavior. For use in making such predictions, we adapt and extend the Social Value Orientation (SVO) model from social psychology, which provides a way to measure an agent's preferences for both its own payoffs and those of the other agents.

The original SVO model was limited to one-shot games, and assumed that each individual's behavioral preferences remain constant over time—an assumption that is inadequate for repeated-game settings, where an agent's future behavior may depend not only on its SVO but also on its observations of the other agents' behavior. We extend the SVO model to take this into account. Our experimental evaluation, on several dozen agents that were written by students in classroom projects, show that our extended model works quite well.

1 Introduction

Many multi-agent domains involve human and computer decision makers that are engaged in repeated collaborative or competitive activities. Examples include online auctions, financial trading, and computer gaming. Repeated games are often viewed as an appropriate model for studying these kinds of repeated interactions between agents. Compared to one-shot games, repeated games are much more complex as they allow agent to adapt their behavior between the rounds. The relevant literature contains many demonstrations of how an agent's behavior can change as it develops a better understanding of the other agents' behavior [6, 1, 3, 8, 4].

In order to model the behavior of an agent and predict its performance, we adapt and extend a construct, Social Value Orientation (SVO), from social psychology [2]. SVO theory assumes that in interpersonal interactions, an individual's choices depend not only on his/her own payoffs but also on his/her preferences for the *other* individual's payoff, and that these preferences remain stable over time. SVO theory provides a way to measure these preferences, and experimental validations of these measurements on human subjects.

If a human writes an agent to act as the human's delegate in a multi-agent environment, one might expect the computerized agent to have social preferences as well. Knowing an agent's social preference would make it possible to make informed guesses about the agent's future actions.

[1] University of Maryland, College Park, USA, email: klcheng@cs.umd.edu
[2] Ariel University, Ariel, Israel, email: inonzu@ariel.ac.il
[3] University of Maryland, College Park, USA, email: nau@cs.umd.edu
[4] University of Maryland, College Park, USA, email: golbeck@umd.edu

A critical limitation of the SVO model is that it only looks at agent's preferences in one-shot games. This is inadequate for predicting agents' behaviors in the situations where the agents interact with each other repeatedly. An agent's actions may depend on both its SVO and its model of the other agent's behavior. To use the SVO model effectively in repeated games, it is necessary to extend the SVO model to take into account how an agent's behavior will change if it interacts repeatedly with various other kinds of agents.

Our contributions in this paper are as follows (for details, see [5]). First, we extend the SVO model by developing a *behavioral signature*, a model of how an agent's behavior over time will be affected by both its own SVO and the other agent's SVO. Second, we provide a way to measure an agent's behavioral signature, and methods for using behavioral signatures to predict agents' performance. Third, we present experimental results using agents that students wrote to compete in repeated-game tournaments. The experimental results show that our predictions are highly correlated with the agents' actual performance in tournament settings. This shows that our proposed model is an effective way to generalize SVO to situations where agents interact repeatedly.

2 Modeling Computer Agents

Even if an agent has a fixed SVO in general, its actions toward a specific agent may depend on its past experience of that agent. For example, an agent that is normally cooperative might behave aggressively toward another agent that behaved aggressively toward it in the past. SVO measurements do not capture this influence on an agent's behaviors, because SVO measurements are always done on one-shot games against an abstract opponent. This non-repetitive interaction assumption is not valid in most multi-agent environments. In the environment we used, the repeated interaction is modeled by a repeated game with unknown number of iterations. In order to model the behavior of an agent, we use a modified version of Ring method, a well-known technique for measuring SVO used in social psychology [7] to measure social preferences of agents.

In the technical report [5], we present a way to measure the social preferences of a computer agent after the agent played n iterations with a tester agent. We also define a *behavioral signature* for an agent x to be a vector $\Theta_n(x)$ of x's social preferences when x plays against nineteen different tester agents after n iterations. Each tester agent is a memoryless agent whose social preference is constant and unique. Therefore, behavioral signature of an agent takes into account how the agent's behavior will change if it interacts repeatedly with various other kinds of agents. For details, see the technical report.

If we know the behavioral signatures of two agents x and y, we can estimate the cumulative payoff when x and y play with each other. Below, we summarize the results of experiments with two methods,

2x2 symmetric game		Player 2	
		A_1	A_2
Player 1	A_1	(a,a)	(b,c)
	A_2	(c,b)	(d,d)

Figure 1. Stage game. The values a, b, c, d are generated randomly.

$E_0(x, y)$ and $E_n(x, y)$, of estimating x's average payoff when it plays with y for N iterations (where $N > n$). The first method, $E_0(x, y)$, only uses (initial) SVO of x and y. The second method, $E_n(x, y)$, uses the behavioral signatures of both agents. $E_0(x, y)$ is actually a degenerated case of $E_n(x, y)$ when $n = 0$, because the behavioral signature becomes initial social preferences of the agent when $n = 0$.

3 Experiments

For our experimental evaluation, we used a large collection of agents that were written by students in several advanced-level AI and Game Theory classes. In each case, the students wrote their agents to compete in a round-robin tournament among all the agents in their class. To attain a richer set of agents, the classes were held at two different universities in two different countries: one in the USA, and one in Israel.

Our experimental studies involved measuring the agents' behavioral signatures, playing round-robin tournaments among the entire set of agents, and comparing the agents' performance with the predictions made by our model. To eliminate random favorable payoff variations, we randomized the series of games, and used the same series between all agents in the population. The instructions stated that at each iteration, they will be given a symmetric game with a random payoff matrix of the form shown in Figure 1. We did not tell the students the exact number of iterations in each repeated game. The total agent's payoff will be the accumulated sum of payoffs with each of the other agents. For motivational purposes, the project grade was positively correlated with their agents overall ranking based on their total payoffs in the competition. Overall, we collected 71 agents (47 from the USA and 24 from Israel).

In the following experiment, the total number of iterations (N) is 100, and the number of runs is also 100. We predicted the average payoff of all possible games of any two students' agents (including playing with itself, i.e., 71×71 data points for each run), using the methods mentioned in Section 2.

Figure 2 shows mean square error between predicted payoffs and actual payoffs. Regardless the value of n, their mean square errors are low, comparing with the average payoff ≈ 5.5. When $n = 0$, the accuracy of E_n is good (mean square error = 0.284). As n increases, the accuracy of E_n also increases until $n = 20$, at which point it levels off.

When $n = 0$, E_n degenerates to E_0 which only considers the (initial) SVO value of the agents. When $n > 0$, E_n takes the agents' adaptive behaviors into account by considering their behavioral signatures. The better performance of E_n shows that our extended SVO model works better in repeated games than the original SVO model.

4 Conclusions

We have extended the SVO model from social psychology, to provide a *behavioral signature* that models how an agent's behavior over

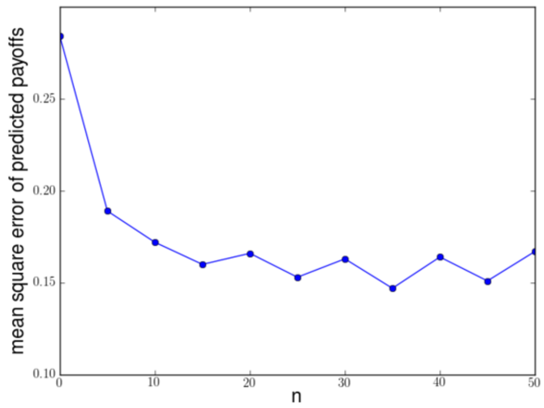

Figure 2. Mean square error of predicted payoffs (when student agents play in a tournament). n is the number of iterations before measuring an agent's behavioral signature.

multiple iterations will depend on both its own SVO and the SVO of the agent with which it interacts. We have provided a way to measure an agent's behavioral signature, and a way to use this behavioral signature to predicting the agent's performance. In our study of agents that were designed to play a repeated stochastic game in classroom tournaments, the predictions made by our model were highly correlated with the agents' actual performance.

ACKNOWLEDGEMENTS

This work has been supported in part by ARO grant W911NF1110344, ONR grant W911NF0810144, and AFOSR grant FA95501410020. The information in this paper does not necessarily reflect the position or policy of the funders, and no official endorsement should be inferred.

REFERENCES

[1] T.-C. Au and D. Nau, 'Is it accidental or intentional? a symbolic approach to the noisy iterated prisoners dilemma', *The iterated prisoners' dilemma: 20 years on*, **4**, 231, (2007).

[2] S. Bogaert, C. Boone, Declerck, and C. Declerck, 'Social value orientation and cooperation in social dilemmas: A review and conceptual model', *Brit. Jour. Social Psych.*, **47**(3), 453–480, (September 2008).

[3] Colin F Camerer, Teck-Hua Ho, and Juin-Kuan Chong, 'Sophisticated experience-weighted attraction learning and strategic teaching in repeated games', *Journal of Economic Theory*, **104**(1), 137–188, (2002).

[4] Ananish Chaudhuri, Sara Graziano, and Pushkar Maitra, 'Social learning and norms in a public goods experiment with inter-generational advice', *The Review of Economic Studies*, **73**(2), 357–380, (2006).

[5] Kan-Leung Cheng, Inon Zuckerman, Dana Nau, and Jennifer Golbeck. Behavioral signatures for predicting agents' behavior. `http://www.cs.umd.edu/~klcheng/tech_report_behavioral_sig.pdf`, 2014.

[6] D. Egnor, 'Iocaine powder explained', *ICGA Journal*, **23**(1), 33–35, (2000).

[7] W.B.G. Liebrand and C.G. McClintock, 'The ring measure of social values: A computerized procedure for assessing individual differences in information processing and social value orientation', *European journal of personality*, **2**(3), 217–230, (1988).

[8] Robert L Slonim, 'Competing against experienced and inexperienced players', *Experimental Economics*, **8**(1), 55–75, (2005).