Taylor & Francis
Taylor & Francis Group

Journal of Experimental & Theoretical Artificial Intelligence

ISSN: 0952-813X (Print) 1362-3079 (Online) Journal homepage: https://www.tandfonline.com/loi/teta20

Modeling agent's preferences by its designer's social value orientation

Inon Zuckerman, Kan-Leung Cheng & Dana S. Nau

To cite this article: Inon Zuckerman, Kan-Leung Cheng & Dana S. Nau (2018) Modeling agent's preferences by its designer's social value orientation, Journal of Experimental & Theoretical Artificial Intelligence, 30:2, 257-277, DOI: 10.1080/0952813X.2018.1430856

To link to this article: https://doi.org/10.1080/0952813X.2018.1430856

JOURNAL OF EXPERIMENTAL & THEORETICAL ARTIFICIAL INTELLIGENCE
2018, VOL. 30, NO. 2, 257–277
https://doi.org/10.1080/0952813X.2018.1430856

Modelling agent's preferences by its designer's social value orientation

Inon Zuckerman[a], Kan-Leung Cheng[b] and Dana S. Nau[b]

[a]Department of Industrial Engineering and Management, Ariel University, Ariel, Israel; [b]Department of Computer Science, University of Maryland, College Park, MD, USA

ABSTRACT
Human social preferences have been shown to play an important role in many areas of decision-making. There is evidence from the social science literature that human preferences in interpersonal interactions depend partly on a measurable personality trait called, Social Value Orientation (SVO). Automated agents are often written by humans to serve as their delegates when interacting with other agents. Thus, one might expect an agent's behaviour to be influenced by the SVO of its human designer. With that in mind, we present the following: first, we explore, discuss and provide a solution to the question of how SVO tests that were designed for humans can be used to evaluate agents' social preferences. Second, we show that in our example domain there is a medium–high positive correlation between the social preferences of agents and their human designers. Third, we exemplify how the SVO information of the designer can be used to improve the performance of some other agents playing against those agents, and lastly, we develop and exemplify the behavioural signature SVO model which allows us to better predict performances when interactions are repeated and behaviour is adapted.

ARTICLE HISTORY
Received 13 June 2016
Accepted 11 January 2018

KEYWORDS
Agent modelling; cognitive modelling; repeated games; social preference; social value orientation (SVO)

1. Introduction

Human social preferences have been shown to play an important role in many areas of decision-making; e.g. interaction in labour markets (Kniesner, Grodner, & Bishop, 2011), bilateral or small-group bargaining (Bolton & Ockenfels, 2000), and social welfare considerations (Charness & Rabin, 2002). There is evidence from the social science literature that human preferences in interpersonal interactions depend partly on a measurable personality trait called a human's *Social Value Orientation (SVO)* (Bogaert, Boone, & Declerck, 2008; Messick & McClintock, 1968). The SVO trait quantifies the personality differences among humans in the way they approach interdependent others and can be used to cluster people to have pro-self vs. pro-social orientations. It can also be used to label humans to having a 'cooperative', 'competitive' or 'individualistic' nature, and on the extreme cases humans as 'aggressors', or of 'altruistic' nature. Analysis of many SVO based experiments reveal that most people are classified as cooperators (50%), followed by individualists (24%), followed by competitors (13%) (Murphy, Ackermann, & Handgraaf, 2011).[1]

In multiagent systems, agents are often written by humans to serve as their delegates when interacting with other agents. Thus, one might expect an agent's behaviour to be influenced by the SVO of its human designer. The purpose of this paper is to explore ways to model SVO in automated agents, to test the correlation hypothesis, and to exemplify how this information can be utilised in strategic reasoning.

CONTACT Inon Zuckerman inonzu@ariel.ac.il

There are few methods to gauge human social preferences or even personalities, as several measurement methods for quantifying variations in SVO across individuals have been developed (see Liebrand & McClintock, 1988; Murphy et al., 2011; Van Lange, De Bruin, Otten, & Joireman, 1997). However, using those to quantify the social preferences of computer agents is challenging, as the tests that were designed for humans cannot be easily converted to automated agents. For instance, an agent that was constructed to play in a simple repeated game cannot provide answers to questions out of the game contexts, such as 'what will you do in this situation?', which are often part of SVOs' psychological test.

We started by defining a simple 'para-SVO' measure for automated agents that is evaluated by assuming a random opponent. On top of the above limitation of using human test for automated agents, our 'para-SVO' measure also has to deal with the fact that the human tests are constructed on a series of one-shot questions, while we need to evaluate the agents' SVO in a repeated game, where strategic behaviours can adapt according to the behaviour of the opponent. We evaluated it by collecting a set of students' agents and conducting psychological SVO based evaluation to get the corresponding SVO value of the human who constructed the agent. We estimated the social preference of computer agents by the proposed methods, and studied the correlation. The results show that the SVO of the human designer is *highly correlated* with the social preference of the corresponding agent.

In the next step, we extended the simple 'para-SVO' model by developing what we denote, a *behavioural signature*, a model of how agents' behaviour over time will be affected by both their own SVO and the other agents SVO. Alongside the presentation of the model we also provide a way to measure an agent's behavioural signature, and methods for using behavioural signatures to predict agents' performance.

Having the ability to model the SVO of agents and knowing that it correlates to the SVO values of its designer is only part of the story. Opponent modelling is not a mean by itself, but a way to gain strategic advantages in various interactions (e.g Wilson, Zuckerman, & Nau, 2011). With that in mind we provide several experimental evaluations that demonstrate the value of having the SVO of the agent's designer for strategic reasoning. In the first experiment, we improved the performances of two basic agents by composing them to a single agent that decides which of them to play based only on a single input: the SVO value of the human delegator of the agent whom it is playing against. In our second experiment, we took the state-of-the-art *life game* automated agent (Cheng, Zuckerman, Nau, & Golbeck, 2011), provided him with the same SVO value as before, and used it to improve its performance by avoiding early exploitation by selfish agents. Lastly, we present experimental results using a large set of agents written by students for a repeated-game tournaments (71 students in 2 countries), that show that our predictions based on the human designer's SVO are highly correlated with the agents' actual performance.

To sum up, in this paper we propose to use ideas and techniques derived from SVO theory to measure computer agents' social preference. To the best of our knowledge, this is the first attempt to quantify the social preference of computer agents using a theory from social psychology. Specifically, we had three goals in our series of studies: first, to explore the ability of the SVO theory as a way to model agents' social preferences. Second to explore the correlation between the SVO of agents, and the human who designed them. Third, to demonstrate utilisation of the SVO information for strategic reasoning. Our main results can be summarised as follows:

- We explore, discuss, and provide a solution to the question of how SVO tests that were designed for humans can be used to evaluate agents' social preferences (Section 3).
- We show that in our example domain (the life game) there is a medium–high positive correlation between the social preferences of agents and their human designers (Section 3.1).
- We exemplify how the SVO information of the designer can be used to improve the performance of some other agents playing against those agents (Section 4).
- We develop and exemplify the *behavioural signature* SVO model which allows us to *better* predict performances when interactions are repeated and behaviour is adapted (Section 5).

The implications of our results are extensive and can be used in various forms. Moreover, with the emerging ability to quantify users social orientation on social networks (Golbeck, Robles, Edmondson, & Turner, 2011), and recent advancements in *transfer learning* (Pan & Yang, 2010), we can safely assume that transferring information across domains will be another possibility to utilising out results.

2. Background and related work

We begin with providing the necessary background on the *life game*, originating in Bacharach's research on collective rationality (Bacharach, Gold, & Sugden, 2006), and on the Social Value Orientation (SVO) theory which is one of the prominent class of social motivation theories in the behavioural sciences (Bogaert et al., 2008; Messick & McClintock, 1968). Following that we provide some related work on the notion of agent-modelling and specifically, peer-designed agents.

2.1. The life game

Many multiagent 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. In a traditional, game-theoretic repeated-game model, agents repeatedly play a game called the *stage game*. Many types of games can be used as the stage game. For example, Axelrod's famous Iterated Prisoner's Dilemma (IPD) competitions showed the emergence of cooperation when the game is played repeatedly without knowing in advance when it ends, even though the rational dominant equilibrium in a one-shot Prisoner's Dilemma is to defect (Axelrod, 1984). Maynard-Smith (1982) studied the two-player Chicken game with a population of Hawks and Doves, and Skyrms (2004) studied the evolved population when individuals were playing the Stag-hunt game.

The importance of Axelrod's work is that even-though the mathematical analysis shows that continuous defection is the equilibrium choice for rational agents, successful agents in his competition as well as its evolutionary work following the competition showed that cooperation can emerge. This lead to one of the most interesting questions in modern science, why does cooperation emerge among self-interested agents?

Each of these studies used a highly simplified game model in which the *same* stage game was used at every iteration. In other words, they assumed that the agents would interact with each other repeatedly in exactly the same environment. However, as pointed out by Bacharach et al. (2006, p. 100), repeatedly playing the **same** game is unlikely to be an accurate model of any individual's life. In many real-life situations, agents may interact with each other repeatedly in *different* environments.

As more accurate model, Bacharach proposed the *Life game*, in which an individual plays a mixture of games drawn sequentially according to some stochastic process from many stage games. Bacharach referred to the size and variety of this set as the game's *ludic diversity* (thus an ordinary non-stochastic repeated game has minimal ludic diversity). The rich variety of stage games also allows agents to express a larger spectrum of social preferences, resulting in an adequate playground for agents of different personalities and behaviours. We believe this makes the Life game a better model for repeated interaction in different environments – so in this paper we concentrate on studying social preferences of automated agents in the Life game of high ludic diversity.

To the best of our knowledge there is currently no bio-evolutionary game theory model that studies repeated interaction of **high** ludic diversity. Moreover, while classical game theory classified and presented equilibrium solutions to various classes of games, discussions on life games has been absent from the literature. In Bacharach's words, 'Although game theory studies games of very varied structures, there has been almost no interest in what happens in the game of life, or even short sequences of diverse interactions.' [ref, page 116].

In this paper, we model the *life game* as an iterated game in which each stage game is a 2×2 normal-form game that is generated randomly by choosing independent random values for the payoffs a, b, c

Table 1. Stage game for the life game. The values a, b, c, d are generated randomly as described in the text.

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

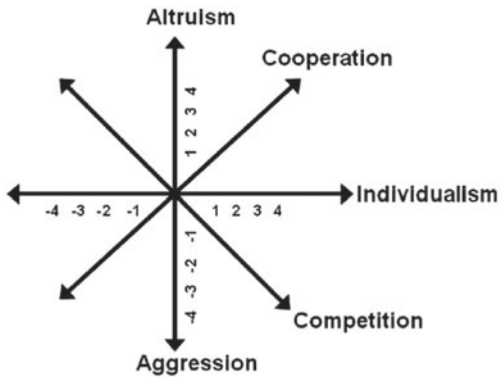

Figure 1. social behaviours spectrum.

and d in the payoff matrix shown in Table 1. The payoffs a, b, c, d are chosen from a uniform distribution over the set $[0, 9]$.[2] At each stage, each agent knows the complete payoff matrix. After deciding on the actions, each agent will be notified of the action chosen by the other agent. The two agents will play the games in succession, without knowing when the series of games will end. We do not place any restrictions on the agents' memory, and they may record past matrices and the actions taken by both agents and use it in their strategy.

Depending on the randomly chosen values of a, b, c and d, each stage game may or may not be an instance of a well-known social dilemma game (e.g. Prisoner's dilemma, Chicken game Stag-Hunt). Consequently, the semantics of the actions are subjective and depend on the value of a, b, c and d. For example, if $a = 3$, $b = 0$, $c = 5$ and $d = 1$ (a Prisoner's dilemma), then A_1 and A_2 can be considered as 'Cooperate' and 'Defect'. This additional layer of uncertainty might cause situation such as that when one agent considers a certain action to be a reasonably cooperative action, it will be captured as a competitive action in the eyes of its opponent.

2.2. The social value orientation (SVO) theory

There is a substantial set of evidence from the social and behavioural sciences literature showing that players explicitly take into account the outcome for the other player when considering their course of action (Au & Kwong, 2004). Moreover, the choices people make depend, among other things, on personality differences in how they approach interdependent others. This observation can be traced back to the seminal work by Messick and McClintock (1968) in which they presented a motivational theory of choice behaviour that considers both players' payoffs in game situations. This class of theorems was later denoted as the *Social Value Orientation* theory (Bogaert et al., 2008).

The foundation of the SVO theory can be explained using Figure 1 which describes a two-person preference model of the major interpersonal orientations that can occur between players. In this model, the player's utility is defined on the horizontal axis, and the outcome of the 'other' player is on the vertical axis. Each outcome increases monotonically along each axis, and the values reflect a linear combination of payoffs to both players.

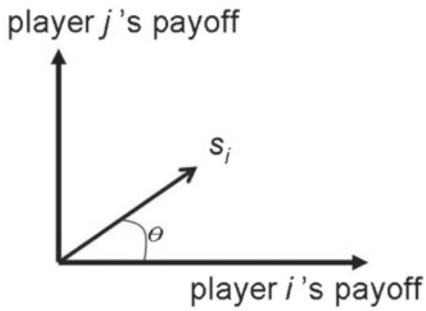

Figure 2. Social value orientation space.

The social-orientation space of a game can be viewed as a two-dimensional Euclidean space, as illustrated in Figure 2 (McClintock & Allison, 1989) for Player i and Player j. The *social orientation* of Player i is a unit vector $\hat{s}_i$ such that $\hat{s}_i$'s initial point is at the origin of the social-orientation space. We represent $\hat{s}_i$ by the angle, θ_i between $\hat{s}_i$ and the x-axis of the social-orientation space.

For example, when $\theta_i = 0$ then, Player i acts as a pure individualistic. If $\theta_i = \pi/4$, this means that player is fair, i.e. it acts to balance the accumulated total payoffs of two players. When $\theta_i = \pi/2$, (i.e. $\hat{s}_i = \langle 0, 1 \rangle$), the player is purely prosocial, i.e. it never attempts to maximise its own payoff, but rather it tries to increase the payoff of the other player. The values of θ can also be used to represent other social orientations such as altruism, aggression and so forth. Note that the behaviours on the left side of the graph are considered mental disorders (e.g. *Masochism*, *Sado-Masochism*).

Most of the SVO based studies typically recognise two opposing social value orientations: pro-self and prosocial orientations. A *proself orientation* is one that gives higher consideration to its own payoff; while a *prosocial orientation* gives gives more equal consideration to the payoff of the agents he or she is interacting with. The social orientation of a player is not an absolute value; it describes a spectrum of possible behaviours, in which one end of the spectrum denotes proself behaviour, and the other end denotes prosocial behaviour.

In contrast to the diversity of the SVO theory, the traditional rationality assumption dictates that all individuals are proself, without any difference among them. As most social or psychological traits, the claim that SVO is a fundamental personality trait is supported by both biological and sociological findings (Au & Kwong, 2004). Biological support also can be found, among others, in Van Lange et al. (1997) and Sutter et al. (2010) showing that the basic form of SVO is visible early in life as part of a child's temperament. The development of the SVO from social interactions is supported by many works shown in Au and Kwong (2004) review. The validity of SVO based theorems, shown both in laboratory and field studies, indicates that prosocial generally cooperate more and show greater concern for the effect of their actions on the well being of others and on the environment. For examples, McClintock and Allison (1989) showed that prosocial students were more willing to contribute time to help others, and Joireman, Lasane, Bennett, Richards, and Solaimani (2001) showed that prosocial participants tend to take more pro-environmental and collective policies than self-interest actions.

Over the years, there have been significant advances on social dilemmas and various aspects of the social value orientations since the seminal work of Messick and McClintock (1968). For example, Parks and Rumble (2001) showed that different aspects of the Tit-for-Tat strategy have different effects on the cooperation rates of individuals with different SVO values. In addition, there were several other research questions that considered some relaxation of the rationality assumption in their solution, for instance, de Jong, Tuyls, and Verbeeck (2008) presented a computational model that allows for achieving fairness in multiagent systems. Their computational model uses the *Homo Egualis* utility function that has been shown to adequately describe human behaviour in several games.

Table 2. An example from the ring measurement questionnaire.

Choose between: A: 26 for me, and 97 for other B: 38 for me, and 92 for other.

Table 3. An SVO based interpretation of the a symmetric game.

		Player 2	
		A	*B*
Player 1	*A*	$p_{\text{self},i}(A) + p_{\text{other},i}(A)$	$p_{\text{self},i}(A) + p_{\text{other},i}(B)$
	B	$p_{\text{self},i}(B) + p_{\text{other},i}(A)$	$p_{\text{self},i}(B) + p_{\text{other},i}(B)$

2.3. Measuring SVO

There are few measurement methods proposed by social psychologist for measuring human SVO (see Liebrand & McClintock, 1988; Murphy et al., 2011; Van Lange et al., 1997). To measure the SVO of a person x, x is usually asked a series of questions in which he needs to select between certain distributions of resources, some amount to himself/herself, and some amount to be allocated to some other randomly determined person y. The examiner will ask x to imagine that the points involved with the decisions have value to you: specifically, the more of them you accumulate the better. Similarly, x needs to imagine that the other person y feels about his/her own points the same way. It is told that x and y will remain mutually anonymous during and after the decision is made, and there is nothing y can do to affect x in any way. In other words, it is a one-shot game. Hence, the choice made by x is not a strategic decision, but rather this is a one-shot individual decision under certainty. Nonetheless this choice has a social dimension, as x's action will affect y's behaviour and x is aware of this potential effect.

For example, one well-known technique for measuring SVO used in social psychology is the Ring measure (Liebrand & McClintock, 1988). Typically, the ring measure involves a series of 24 decision tasks between two options. The participants are told to be randomly paired with another person whom the question refers to as 'other'. In the decision task, the participants will be making choices by circling the letter 'A' or 'B' on a response sheet. The participants' choices will produce points/money for themselves ($p_{\text{self},i}(A)$) and the other ($p_{\text{other},i}(A)$). The options involve combinations of own outcome and other outcome. One of the questions in the ring measurement questionnaire is shown in Table 2.

Adding up the chosen amounts separately for the self and for the other player provides an estimation of the weights assigned by the participant to own and others payoffs. These weights are used to estimate the SVO angle (θ) of the participant by the formula below:

$$\theta = \arctan\left(\frac{\sum p_{\text{other},i}(r_i)}{\sum p_{\text{self},i}(r_i)}\right), \text{ where } r_i = i - \text{th response} \tag{1}$$

All angles between 112.5° and 67.5° were classified as altruistic; those between 67.5° and 22.5° were classified as cooperative; those between 22.5° and 337.5° as individualistic, and angles between 337.5° and 292.5° as competitive (Liebrand & McClintock, 1988).[3]

Since the total number of points a participant receives on each decision problem is determined by the combination the choices of both participants, the participants are in fact playing the following symmetric game for the i-th decision problem (Table 3):

The above matrix is a two-players normal-form game similar to the general one presented in Table 1, where in this case the possible actions for each player are denoted as A and B, and the outcomes are the sum of the payoffs from the 24 Ring method questions. For example, the sample decision task mentioned above in Table 2 can be written as the following symmetric game (Table 4):

Table 4. An example of a decision task.

		Player 2	
		A	*B*
	A	123	118
Player 1	*B*	135	130

There are several other techniques for measuring social preferences, such as the decomposed game measure, the triple dominance measure and the slider measure. In a decomposed game, participants choose between three options that offer points to the self and another person. The most commonly used measure of SVO is the nine-item triple-dominance measure. Typically, participants are classified as one of three orientations (cooperators, individualists or competitors) if they make 6 out of 9 choices consistent with the orientation. Like the ring measure, the slider measure can help us estimate the SVO angles of participants; and it has been reported that the slider measure has better test–retest reliability (Murphy et al., 2011); but as described in the next section, the Ring measure is the only one that can be adapted for use in a repeated-game setting.

2.4. Agent-modelling

Previous research in agent-modelling has proposed modelling other agents by estimating agents' personalities. For example, Talman, Hadad, Gal, & Kraus (2005) proposed a decision-theoretic model that explicitly represents and reasons about agents' personalities in environments in which agents are uncertain about each others' resources. Similar to our agents, their agents can identify and negotiate with those who are cooperative while avoiding those who are exploiter. Gal, Grosz, Kraus, Pfeffer, and Shieber (2010) proposed several new decision-making models that represent, learn and adapt to various social attributes that influence people's decision-making in a group of human and computer agents.

Cheng et al. (2011) developed a successful strategy for the life game, using agent-modelling by estimating agents' SVO orientation, and Wilson et al. (2011) showed how social preferences can be used in classical adversarial game-tree search algorithms. Lastly, Zuckerman and Hadad (2012) showed a BDI-based architecture that provide reasoning capabilities on the social behaviour spectrum.

From the above related work (and others) we can see works that use different models for agent's behaviour in repeated games settings. Our work in this paper is different because the model is based on a well-established theory from social psychology that has been known and used for more than 50 years. This allows us to consider and correlate the personalities of the human designer. In addition, our proposed method can measure the social orientation quantitatively, instead of simply classifying the social preferences. In our experiments, we used the correlation to initialise the agent model, but the SVO information can also be used in other ways (which we leave for future research).

Peer-Designed Agents have been recently used with great success in AI to evolve and evaluate state-of-the-art cognitive agents for various tasks such as negotiation and collaboration (see Au, Kraus, & Nau, 2008; Lin, Kraus, Oshrat, & Gal, 2010; Manisterski, Lin, & Kraus, 2008). Lin et al. (2010) provided an empirical proof that PDAs can alleviate the evaluation process of automatic negotiators, and help their designs, while Manisterski et al. (2008) studied how people design agents for online markets and how their design changes over time. Their results show that most human subjects modified their agents' strategic behaviour; for example, they increased their means of protection against deceiving agents. PDAs have also been used with great success in Chalamish et al. (2013) to improve parking simulations. Au et al. (2008) used agents written by students to study enhancing agent by combining a set of interaction traces produced by other agents.

However, while using PDA's instead of human is very attractive, there are also limitations that must be taken into account. Elmalech and Sarne (2012) showed that there are limitations for the ability to generalise results from one successful implementation of PDA-based systems to another. As such, the decision to prefer working with PDAs instead of humans must dependent on the domain in question. Another interesting and (very recent) result from the same authors suggest that the process of developing a programmable strategy for a PDA might affect the behaviour of its designer (Elmalech, Sarne, & Agmon, 2015). There are also some evidence of discrepancies between actual and reported human behaviour in various domains, in particular in meta-cognition research (Harries, Evans, & Dennis, 2000), however overall the sum of evidence do show that PDA designers do manage to describe their strategy in a way that reflects their real-life strategy.

Our research has very limited exposure to these weaknesses of the PDA methodology, as our correlation result is, at least theoretically, independent from any specific domain. Based on the strong findings from the social and behavioural sciences stating that the SVO is consistent in time, we can hypotheses that the correlation observed on one domain will be transferable to another. There might be some deviations to the observed social behaviour, but the central social tendency should remain the same.

3. Measuring agents' social preferences

In order to model the behaviour of an agent, we would like to have a precise quantitative measurement (like SVO angle) on computer agents. For example, a Maximin agent maximises its worst-case payoff, so its SVO angle always equals to 0°. A Minimax agent minimises other agent's best-case payoff, so its SVO angle always equals to −90°. Except for Ring method, all the other measurement methods cannot be transformed into 2×2 games, so they cannot be used to measure social preferences of computer agents playing 2×2 games (e.g. life game). Therefore, we use a modified version of Ring method to measure social preference of agents.

Although the choice questions in the Ring measurement can be presented as 2×2 normal form games, most of the payoff values of the game matrices are not valid for the life game model we used. In the life game model we used, the payoff values must be in the range $[0, 9]$. To apply Ring measurement on automated agents, we modified the game matrices to G_{Ring} by down-sampling, scaling, and translating, so that all payoff values will sit within $[0, 9]$.

Another problem is that the agents were designed under the assumption that they may have repeated interactions with the other agents. It is likely that the social preference of an agent varies with the current number of iterations and the behaviour of the other agent. For example, an aggressive partner might trigger an aggressive behaviour even from an initially cooperative agent. This is so because the decision may involve many factors like behaviour of other agent, social preference and competence of the human designer. However, one-shot games are used in all of the SVO measurement methods for human, and the participants are told that they will remain mutually anonymous during and after their decisions are made. This non-repetitive interaction assumption is not valid in most multiagent environments. In the environment we used, the repeated interaction is modelled by a repeated game with unknown number of iterations.

In repeated games, an agent's social preference can be influenced not only by the agent's own SVO, but also by how the agent reacts to the other agent's SVO. For example, let x be an agent whose SVO is 45° (i.e. it prefers equal payoffs for both agents) and y be a memory-less agent whose SVO is 0° (i.e. y cares only about maximising its own payoff in the current iteration). If x and y interact repeatedly, then after repeated observations of y's behaviour, x might decide that the best way to equalise both agents' cumulative payoffs might be for x to try to maximise its own payoff at each iteration. Consequently, if we perform a Ring measurement of x after it has had many interactions with y, x's 'apparent' SVO value may be closer to 0° than 45°. We will call this x's *para-SVO* against y.

The para-SVO, $\theta_n(x|y)$, of agent x at the n-th iteration with tester agent y is measured by applying the modified Ring measurement on the agent at the $(n + 1)$-th iteration after it interacted with the tester agent y for n iterations. In this chapter, we use a random agent as the tester agent y.[4] The parameter n is introduced, because we would like to measure social preference, which might change during the interactions, at a specified iteration. Our measurement algorithm uses the behavioural data

```
Procedure MeasureParaSvo
Input:
    n = number of random games (before measurement)
    r = number of runs
    G_Ring = set of games of modified Ring measurement
Output:
    para-SVO of x with a tester agent y after n random games
Begin procedure
    p_x ← 0    /* p_x = total payoff of agents using x's strategy */
    p_y ← 0    /* p_y = total payoff of agents using y's strategy */
    Repeat for r times:
        For each game g in G_Ring:
            create new agents x' and y' which use the same
                strategies of x and y respectively
            pair up x' and y' for a repeated game with n
                random games and then g as the last game
            p_x ← p_x + (last gain of x' due to x''s last action)
            p_y ← p_y + (last gain of y' due to x''s last action)
        End For
    End Repeat
    Return arctan(p_y / p_x)
End procedure
```

Figure 3. Procedure of measuring para-SVO of an agent *x* with a tester agent *y* after *n* random games.

of the agent at the last iteration, therefore the para-SVO represents just the *latest* social preference of the agent after *n* games. Figure 3 shows the complete procedure for measuring $\theta_n(x|y)$ using the modified game matrices G_{Ring}. It will get the responses from the testee agent at the last game which is one of the games in G_{Ring}, and then calculate the para-SVO using Formula (1). We verified the validity of the measurement by applying it on some simple agents with known para-SVO angles (e.g. para-SVO angles of Maximin, Minimax and a prosocial agent are 0°, −90°, and 45°, respectively).

3.1. Experiments on measuring agents' para-SVO

In this section, we will present some results on the relationship between social preferences of agents and that of their designers. Our experimental results are based on a (large) set of peer-designed agents (PDAs) of students in advanced-level AI and Game theory classes in the US and Israel. Before giving the programming task we also asked the students to complete an online questionnaire for measuring their SVO (Liebrand & McClintock, 1988). We did not reveal the purpose of that questionnaire, and we did not provide them with their quantified SVO value based on their results. Next, we collected a set of PDAs, where the students were instructed that their agent would compete against all the agents of the other students in the class in a round-robin fashion. The instructions stated that, at each iteration, they will be given a symmetric game with a random payoff matrix. The students did not know in advance the number of iterations in each game. The total agent's payoff will be the accumulated sum of payoffs with each of the other agents. We have developed a framework code in Java that was provided to the students to get them started, and the students could focus solely on their agent's strategy. For motivational purposes, the project grade was positively correlated with their agents overall ranking based on their total payoffs in the competition.

In the first step, we collected 28 agents with SVO of the corresponding designer. That is, we have 28 'life game' agents, with strategies of varying complexities, alongside the SVO value of its designer. Figure 4 shows the SVO distribution of all 28 students. 21 of them are cooperative (67.5° > SVO angle > 22.5°), and 7 of them are individualistic (22.5° > SVO angle > −22.5°).[5]

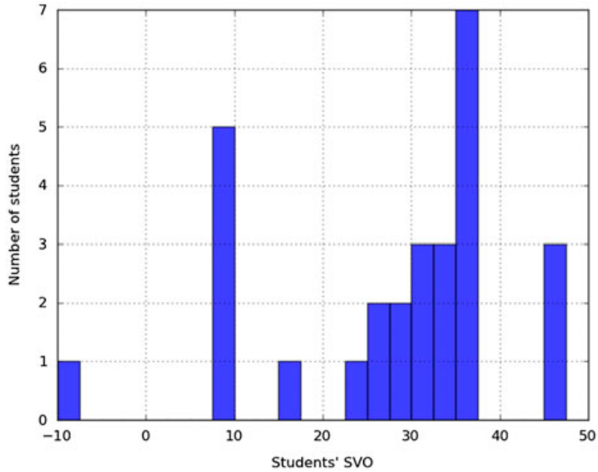

Figure 4. SVO of 28 students.

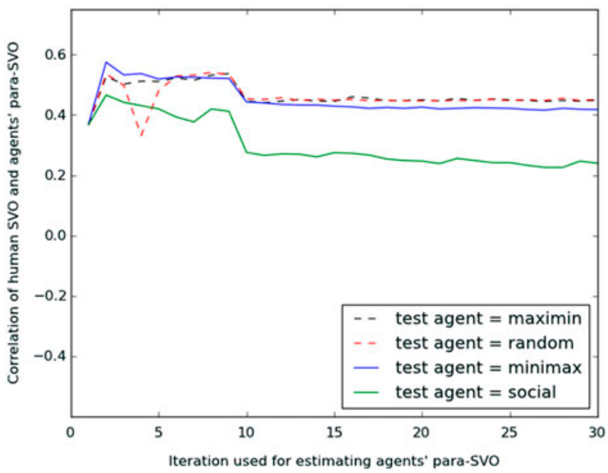

Figure 5. Correlation of human SVO and agents' para-SVO.

3.1.1. Agent-human SVO Correlation

We used the modified ring method presented in Section 2.3 to measure the para-SVO of all computer agents using different tester agents, and then calculated the Pearson correlation of the agents' social preferences and human SVO. Figure 5 shows the correlation of human SVO and agents' para-SVO measured by the modified ring method using different tester agents.

The x-axis of Figure 5 is the number of iterations (n) used for measuring the agent SVO. The Pearson's correlation at the first iteration is medium-positive correlation ($r = 0.4, N = 28, p < .05$), and then rises to around $r = 0.55$ for several iterations, before decreasing and leveling off for the rest of the repeated game.

After reading the source codes of the agents, we found that some students wrote specific parts of their codes for the first iteration only. In other words, they *hard coded* some initial behaviour that may be different from other iterations. This might be partly related to the fact that the Ring method (as well as all other available methods) uses a one-shot games, but we hypothesis that as there is no available

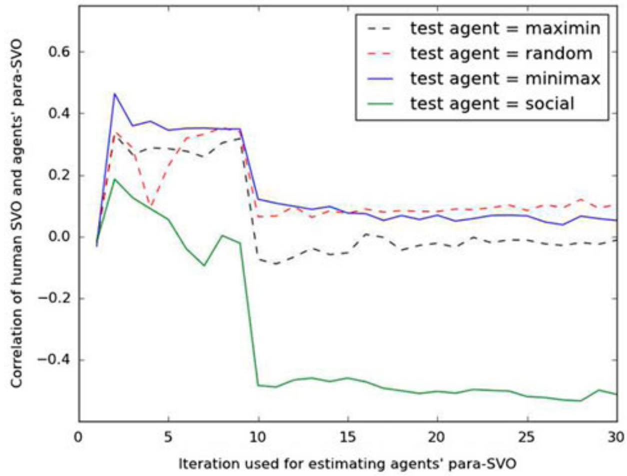

Figure 6. Correlation between human SVO and agents using non-stationary strategy.

information in the first round (no past behaviours to observe), a reasonable solution is to hard code the first iteration before trying to build a model of the other agent in the game, based on the history of interaction.

As the game progresses, the agent's behaviour will come to depend partly on the social preference of the designer, and partly on the agent's predictions of the other agent's behaviour. We surmise that this effect is responsible for the correlations shown in Figure 5.

3.1.2. Stationary vs. non-stationary strategies

The guess of the previous subsection motivated us to classify and investigate the agents based on the complexity of their strategies. We divided the agents' strategies into two groups, stationary and non-stationary, according to the variance of their para-SVO:

(1) For agents using stationary strategy, given a tester agent, their para-SVOs remain the same all the time, because their choices at each iteration depend only on the payoff matrix of the current game. They usually have shorter and simpler codes. For example, some students' agents use a simple competitive strategy that choose action A_1 when $b > c$, and choose action A_2 otherwise. We have about 12 agents using stationary strategy among those students' agents, and the correlation is approximately 0.6.

(2) For agents using non-stationary strategy, given a tester agent, their para-SVOs change as the game progresses, because their choices at each iteration depend on the previous history of the interactions. They may build predictive models of some kind and make some strategic decisions based on the model. For example, some students' agents estimate the probability of the other agent choosing some kind of action in different situations, and then respond accordingly.

Figure 6 shows the correlation of human SVO and para-SVO of agents using non-stationary strategy, which is the subset of 16 agents out of the collected data. Comparing with the correlation for agents using stationary strategy ($\approx$ 0.6 for all iteration), the correlation for agents using non-stationary strategy is lower (ranging from -0.4 to 0.4). This is consistent with our previous guess that para-SVOs of agents using non-stationary strategies correlate less with the social preferences of their designers.

4. Utilising the SVO information

There are many possible applications of utilising the SVO information. Having the SVO information, we can predict the behaviours of the other agents in various situations (as we will do next in Section 5), and our agents can interact with them in some better ways. For example, our agent can avoid possible exploitation by agent that is probably competitive. On the other hand, if we know that the other agent is possibly cooperative according to the SVO information, our agent can possibly increase mutual benefits by working closely with the other agent. If the other agent is neither competitive nor cooperative, our agent can start with some safe actions and learn more accurate model of the other agent from interaction. Using that kind of strategy, we can develop a collaborative agent to enhance safety and productivity by utilising the SVO information.

In this section, we present two examples that use the SVO information. First, we show how we can use the SVO information to composite two simple and non-adaptive agents to form a better non-adaptive agent. Second, we present how we can improve an adaptive agent (Cheng et al., 2011) using the SVO information.

4.1. Improving a non-adaptive agent

In this subsection, we present a way to use the data of other agents' designer to combine two simple agents to form a better agent. The two agents we used are social agent and Maximin agent. Both of them are non-adaptive, thus they do not apply any agent-modelling technique during the game. Social agent always chooses an action that maximises the sum of payoff of itself and other, so its SVO angle is 45° and it performs better if the other agent is also cooperative. Maximin agent always chooses an action that maximises its own minimum possible payoff, so its SVO angle is 0° and it can avoid being exploited by other non-cooperative (individualistic or competitive) agents.

It would be better if we can combine the advantage of both agents in following way: if the other agent is cooperative, our agent will act like the social agent to gain the benefit of mutual cooperation; if the other agent is non-cooperative, our agent will act like the Maximin agent to avoid being exploited by them. However, as we do not know the exact social preference of other agents before interacting with them, we propose to approximate the social preference of the other agents by the SVO of their designer. In other words, if the SVO of other agent's designer is in the cooperative range ($\geq 22.5^\circ$), our agent will act like a social agent; otherwise ($< 22.5^\circ$), our agent will act like a Maximin agent.

We implemented the simple agents and the proposed composite agent described above, and compared their performance in tournaments (10000 runs) with the 28 students' agents. Figure 7 shows the average payoffs when the three agents playing with all 28 students' agents. It shows that the average payoff of the composite agent is (almost, except one point) always higher than both simple agents, because it has the strengths of both agents in different situations ($p < .05$). The result also shows that the performances of the simple and composite agents drop when there is more iterations. It is because some students' agents apply agent-modelling technique that they can easily exploit agents using stationary strategies (including the simple and composite agents) after they have enough interaction data. As we shown in Section 3.1.2, if the other agents use non-stationary strategy, designers' SVO is not enough to model them later in the game. We still need agent-modelling techniques during interaction to being exploited by those agents.

4.2. Improving an adaptive agent

In this subsection, we show a way to use the data of other agents' designer to improve an adaptive agent that apply agent-modelling techniques. We use an adaptive agent which is a state-of-the-art life game agent (Cheng et al., 2011). The agent is an automated agent for the *life game* which performs agent-modelling using a cognitive agent model based on the Social Value Orientation (SVO) theory. In this subsection, we exemplify a way to modify that agent with the newly discovered SVO correlation results, by providing it the SVO data of other agent's designer.

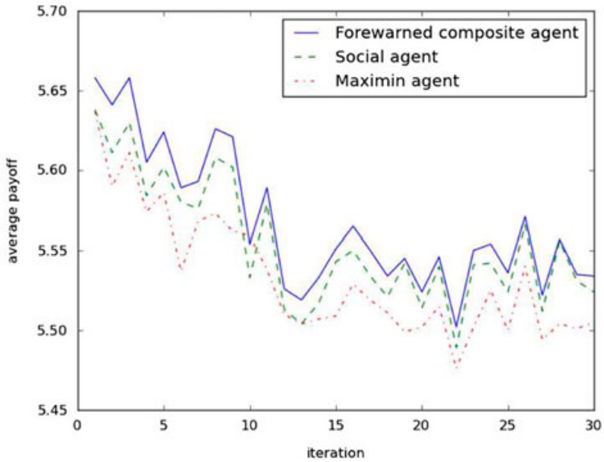

Figure 7. Performance of the simple and composite agents.

Since the original adaptive agent does not have any prior knowledge about the other agent, the agent does not know the social preference of the other agent. Therefore, the agent will start with some default models, and will estimate the orientation of other agent from the history of interactions. More precisely, the agent starts by assuming that the other agent is fully cooperative. After accumulating some interaction histories, the agent will learn the true social orientation of the other agent, and will adapt and use it to the best of its capacity (for example, if the other agent is cooperative, the agent will also be cooperative). To minimise exploitation, the estimated cooperativeness of the other agent is decreased whenever a defect-like action is observed. There are five types of cooperativeness: type 0 (fully non-cooperative), 1, 2, 3 and 4 (fully cooperative). Agents with higher cooperativeness will tend to cooperate on a larger subset of games.

Although it can prevent future exploitation by decreasing the estimated cooperativeness of the other agent whenever a defect-like action is observed, it cannot prevent the initial exploitation of non-cooperative agents. Avoiding initial exploitation is importance, especially when the expected number of iteration is small. We propose to use the SVO information of the designer of other agent to minimise the exploitation by selfish agent. If the designer of other agent has higher SVO angle, the higher initial estimated cooperativeness of the agent. More precisely, instead of initialising the estimated cooperativeness (C_a) of other agent (a) to fully cooperative (4), we initialise it according to the agent designer's SVO (θ_a) using following formula:

$$C_a = \max\left(\min\left(\frac{\theta_a}{10}, 4\right), 0\right)$$

In other words, we have a higher C_a value for an agent a written by a designer having higher SVO angle θ_a.

To evaluate the performance improvement of the agent with the help of the human SVO data, we implemented a forewarned adaptive agent described above, and evaluated its performance in tournaments (10000 runs) with students agents. Figure 8 shows the average payoffs the agents at different iteration when they play with all 28 students' agents.

The payoff of the original agent is very low at the beginning, because it begins by assuming the other agent is fully cooperative and so applying the 'nice' strategy towards all other agents. If the other agent is non-cooperative, the original agent may be exploited for the first few games, and lose some payoffs at the beginning. On the other hand, the forewarned adaptive agent has higher payoff at the

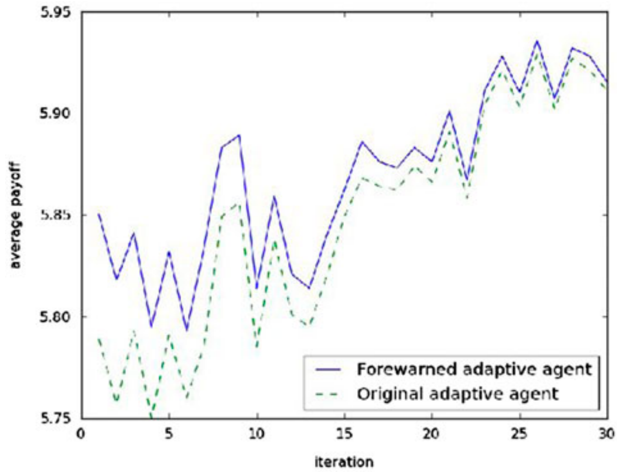

Figure 8. Performance of agents playing with all 28 students' agents.

beginning, because it prevents some of the exploitation by having a more accurate initial model. The results also shows that the significance increase in performance ($p < .05$) mainly comes from avoiding being exploited by agents written by individualistic human. The human SVO data can help the agent to estimate the cooperativeness of other agents, rather than learning through interaction.

With an increasing number of iterations, both agents' performances improve and converge. It is probably because both agents are doing agent modelling. With more interaction data, the modelling will be more accurate, and so they can better predict other agents' action to get higher payoffs. For example, even though the original agent always starts with being nice, when it knows more about the other agents, it will stop cooperating with the defectors and keep cooperating with the cooperators.

In summary, there are at least three main factors for a good life game agent: (1) apply agent-modelling techniques during interaction; (2) start with a more accurate model; (3) maintain mutual cooperation with other agents if possible.

5. The behavioural signature model

Recall that the para-SVO, $\theta_n(x|y)$, of agent x at the n-th iteration with tester agent y is measured by applying the modified Ring measurement on the agent at the $(n+1)$-th iteration after it interacted with the tester agent y for n iterations.

In previous sections, we used a random agent as our tester agent in the first steps of exploration. However, while measuring against a random agent did manage to provide a sufficient measuring method for agent's SVO as we saw in previous sections results, it might lead to various errors in the SVO estimations, especially for the adaptive behaviour agents. For instance, a prosocial agent whose random opponent played aggressively due to chance, might quickly adapt its behaviour to avoid exploitation. In order to account for such situations, we now suggest a slightly more complicated opponent model, the behavioural signature. This model will allow the agent to exhibit more fine-tuned social orientation when interacting with different tester agents.

We define a *behavioural signature* for a to be a vector $\sigma_n(a)$ that includes a's SVO and a collection of para-SVO values for a against several different 'constant-SVO agents':

$$\Theta_n(a) = (\theta_0(a), \theta_n(a|C_{-90}), \theta_n(a|C_{-80}), \ldots, \theta_n(a|C_{90})),$$

Procedure MeasureParaSvoAtNthIteration
Input:
S_a = testee agent's strategy
S_b = tester agent's strategy
G = set of random game
n = number of iterations
r = number of runs
G_r = set of games of modified Ring measurement
Output:
estimated para-SVO angle of an agent using strategy S_a at n-th iteration
Begin procedure
$p_a \leftarrow 0$ (p_a = total payoff of agent a)
$p_b \leftarrow 0$ (p_b = total payoff of agent b)
Repeat for r times:
For each game g in G_r:
create a testee agent a which uses strategy S_a
create a tester agent b which uses strategy S_b
pair up a and b for a n-iteration repeated game
where the last game is g, and the rest are random games from G
$p_a \leftarrow p_a$ + (last gain of a due to a's last action)
$p_b \leftarrow p_b$ + (last gain of b due to a's last action)
End For
End Repeat
Return $\arctan(\frac{p_b}{p_a})$, where $\frac{p_b}{p_a}$ = payoff ratio due to a's last action
End procedure

Figure 9. Procedure of measuring para-SVO of an agent a (which uses strategy S_a) at n-th iteration.

where $\theta_0(a)$ is a's SVO, and $\theta_n(a|C_x)$ is a's para-SVO at the n-th iteration when a plays with the agent C_x defined below.

Each agent C_x is a memory-less agent whose SVO is x degrees. C_x maximises the quantity $S \cos x + O \sin x$, where S is it's expected payoff and O is other agent's expected payoff if C_x plays against an agent that chooses each action with equal probability.[6] For example, if $x = 0$ and the game matrix is the one shown in Figure 1, the Constant-SVO agent will choose A_1 if $a + b > c + d$, otherwise it will choose A_2.

The para-SVO, $\theta_n(a|C_x)$, of agent a at the n-th iteration with tester agent C_x is measured by applying the modified Ring measurement on the agent at the n-th iteration after it interacted with the tester agent C_x for $n - 1$ iterations. The parameter n is introduced, because we would like to measure social preference, which might change during the interactions, at a specified iteration. Our measurement algorithm uses the behavioural data of the agent in the last iteration (the n-th iteration), therefore, the para-SVO represents just the *latest* social preference of the agent after $(n - 1)$-th iteration. We verified the validity of the modified measurement by applying it on some simple agents with known para-SVO angles (e.g. para-SVO angles of Maxmin, Minmax and a prosocial agent are 0°, −90°, and 45°).

Figure 9 shows the complete procedure for measuring $\theta_n(a|C_x)$ using the modified game matrices G_r. It will get the responses from the testee agent at the n-th game which is one of the games in G_r, and then calculate the para-SVO using Formula (1).

If we know the behavioural signatures of two agents a and b, we can estimate the cumulative payoff when a and b play with each other. In this paper, we will study and evaluate two methods, $E_0(a, b)$ and $E_n(a, b)$, for estimating a's average payoff when it plays with b for m iterations (where $m > n$). Both methods use a E_C function to approximate the payoff. $E_C[x, y]$ is the payoff of Constant-SVO agent C_x when it plays with another Constant-SVO agent C_y for m total number of iterations. Note that $E_C[x, y]$ can be computed quickly because the Constant-SVO agents are very simple.

E_0 estimation:

$$E_0(a, b) = E_C[\theta_0(a), \theta_0(b)]$$

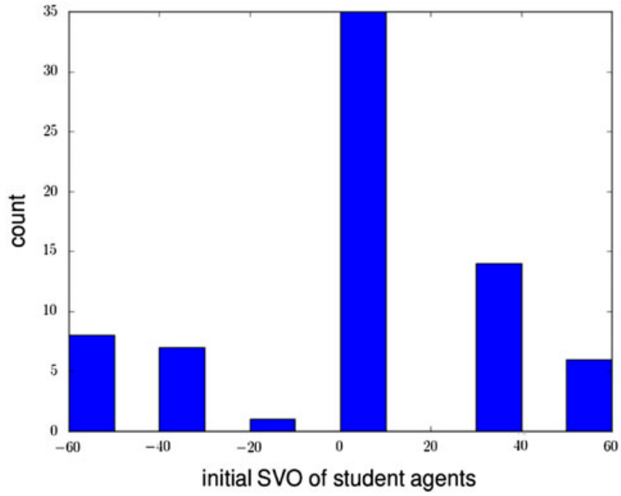

Figure 10. Distribution of students' agents' SVO.

E_n estimation:

$$E_n(a,b) = E_C[\theta_n(a|C_\beta), \theta_n(b|C_\alpha)],$$

where $\alpha = \theta_0(a)$ rounded off to the nearest tens digit, and $\beta = \theta_0(b)$ rounded off to the nearest tens digit.

The first method uses (initial) SVO values of a and b as the input to E_C. The second method uses a more sophisticated input that involves the behavioural signatures of both agents. Note that $E_0(a,b)$ is a degenerated case of $E_n(a,b)$ when $n = 0$, because all elements in the behavioural signature equal to SVO of the agent when $n = 0$.

We have evaluated our model experimentally on 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' behavioural 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 favourable payoff variations, we randomised 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. Following Axelrod's methodology, we did not tell the students the exact number of iterations in each life 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).

5.1. Measuring agents' behavioural signatures

We use the para-SVO measurement procedure (shown in Figure 9) to find the behavioural signatures of all students' agent. Figure 10 shows the distribution of (initial) SVO of students' agents.[7] While most of them are individualistic (to different degrees), there were some who had competitive and cooperative orientations.

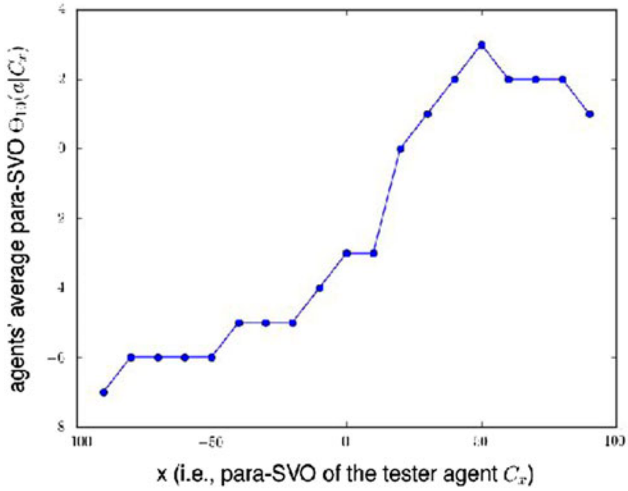

Figure 11. Average para-SVO values $\theta_{10}(a|C_x)$ for $x = -90°$ to $90°$, averaged over all a in the entire set of students' agents.

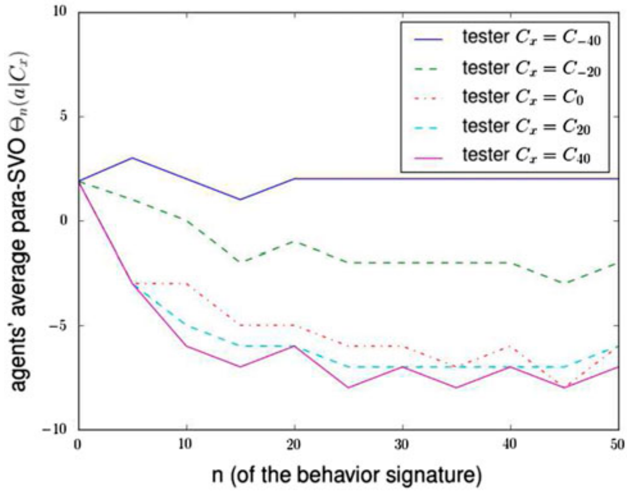

Figure 12. Average para-SVO values $\theta_n(a|C_x)$ with different tester agents C_x, averaged over all a in the entire set of students' agents.

Figure 11 shows the average, over all of the students' agents, of the para-SVO value $\theta_{10}(a|C_x)$. Recall that $\theta_{10}(a|C_x)$ is agent a's para-SVO value at 10th iteration against a memory-less agent C_x whose SVO is x degrees. Notice that the average para-SVO of the students' agents increases with the para-SVO of the tester agents, because it is beneficial to be more cooperative if the other agent is more cooperative. The magnitude of change of the average is not large, because para-SVO values of about 45 (out of 71) agents remain constant across different tester agents.

Figure 12 shows the average para-SVO of students' agents when the tester agents are Constant-SVO agents with SVO $= -40°, -20°, 0°, 20°$, or $40°$. Again, the results show that the average para-SVO of the students' agents increases with the para-SVO of the tester agents. Moreover, when n increases, most of the averages decrease, and all of them level off after about 20 iterations. From examining the code, we found that many of the agents try to build a model of the other agent in the game, based on

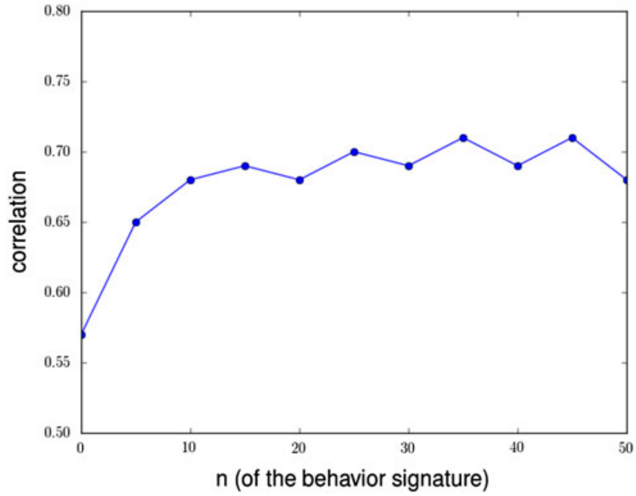

Figure 13. Correlation between predicted and actual payoffs (when student agents play in a tournament).

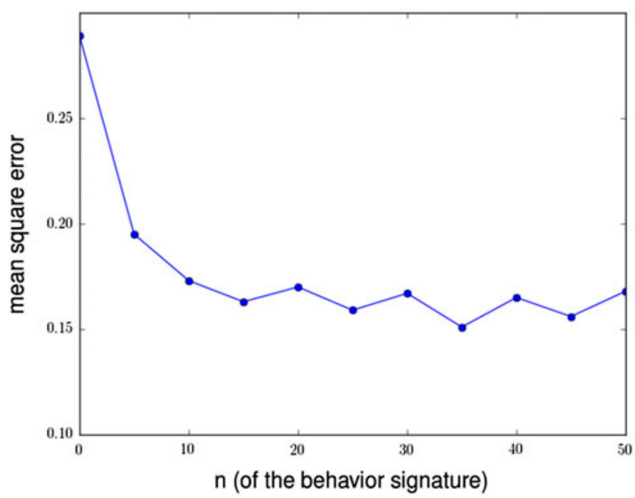

Figure 14. Mean square error of predicted payoffs (when student agents play in a tournament).

the history of interactions, and the model tend to be stabilised after some number of iterations. All of the above results show that the apparent social preferences of agents change with the behaviours of other agents, because the action of an agent is usually determined by both its SVO and its prediction of the opponent action.

5.2. *Predicting agents' performances*

Our next goal was to evaluate the accuracy of our prediction algorithms. In the following experiments, the total number of iterations (m) 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 method mentioned in Section 5.1.

Figures 13 and 14 show the correlation and mean square error between predicted payoffs and actual payoffs. Regardless the value of n, the predicted payoffs have high correlation with the actual payoffs. 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 = 0.289). As n increases, the accuracy of E_n also increases until $n = 20$, at which point it levels off (similar to Figure 12).

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 behaviours into account by considering their behavioural signatures. The better performance of E_n shows that our extended SVO model works better in repeated games than the original SVO model.

6. Conclusions

Human social preferences,i.e. human preferences for the outcomes of their interactions with others, have been shown to play an important role in many areas of decision-making. As agents are developed that exhibit more autonomy and take an increasing role in interacting with other human and agents, it is becoming important to understand the social preferences of agents as well as humans.

We have developed a way to measure the social preferences of computer agents, by adapting some concepts and techniques from social psychology. In our study of agents that were designed to play a repeated stochastic game (the life game), we have found a strong correlation between the agents' social preferences, measured using *para-SVO*, and the social preferences of their human designers. We have shown that this correlation can be used to make useful strategic predictions of what choices an agent will make over the course of a game, and have shown that these predictions can be used to improve the performance of other agents that interact with the given agent.

We have also extended the SVO model from social psychology, to provide a *behavioural signature* that models how an agent's behaviour over 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 behavioural signature, and a way to use this behavioural signature to predicting the agent's performance. In our study of agents that were designed to play a repeated stochastic game (the Life Game) in classroom tournaments, the predictions made by our model were highly correlated with the agents' actual performance.

However, in order to utilise the correlation in real deployment scenarios, one still needs to develop ways to evaluate the SVO of agents or its human designers from real world interactions. Techniques such as transfer learning and data mining can be used to evaluate the SVO of the human (or agent) from past interactions, and utilised via the presented correlation result. For example, in Au et al. (2008) the authors presents a way to take a set of interaction traces produced by different pairs of players in a two-player repeated game, and combine them into a composite strategy. This strategy can in turn be evaluated using the para-SVO technique and quantified into an SVO value.

The implications of the correlation results are extensive and are not limited to the psychological SVO exam employed in this paper. Automated agents are increasingly becoming more widespread in various domains such as online commerce, social networks (Aiello, Deplano, Schifanella, & Ruffo, 2012), online games (Golle & Ducheneaut, 2005) and automated negotiations (Rosenfeld, Zuckerman, Segal-Halevi, Drein, & Kraus, 2016). In addition, we also see an increase in simplified architectures and interfaces that allow people to define and form behaviour rules for their agents without complex programming. As automated agents are formed to replace humans in simple tasks, the ability to infer from the agent's behaviour information about the social tendencies of its designer might be valuable in any situation where the interaction is repeated and quick understanding of your opponent is of an advantage. For example, in the yearly automated negotiation agents competition (ANAC) (e.g. Fujita et al., 2017), where time and modelling your opponent are important factors, one could learn the SVO information of agents from the past year competition, and use it to boot-strap its agent when playing again with an agent of the same designer.

Although our study was restricted to the life game, we believe there is a strong potential for extending our results to other contexts. Such extensions may provide both an improved understanding of agent behavioir, and ways to improve the effectiveness of agents in their interactions with others.

Also, in the future we plan to deepen our understanding and find new ways to measure SVO in an automated agent.

Notes

1. The remaining 13% could not be classified as having a consistent SVO.
2. The [0, 9] range was selected arbitrarily, and the results in the paper are general for any selected range.
3. The boundary between cooperative and individualistic is $\frac{45^\circ+0^\circ}{2} = 22.5^\circ$. Other boundary angles can be derived similarly.
4. In the next chapter, we further extend the notion of para-SVO by having a special set of tester agents.
5. It skews toward cooperative orientation, possibly because we collected the data from students voluntarily responding to our survey.
6. The 'equal probability' assumption is needed to calculate the expect payoff for each action. It can be shown that this assumption is compatible with the para-SVO measurement.
7. SVO of *a* is measured by testing *a* with one-shot games, i.e. it is equals to the initial para-SVO of *a*.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Inon Zuckerman http://orcid.org/0000-0002-9999-1750

References

Aiello, L. M., Deplano, M., Schifanella, R., & Ruffo, G. (2012). People are strange when youre a stranger: Impact and influence of bots on social networks. In *Proceedings of the 6th International AAAI Conference on Weblogs and Social Media*, Dublin, IR.

Au, T.-C., Kraus, S., & Nau, D. (2008). Synthesis of strategies from interaction traces. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems AAMAS'08* (Vol. 2, pp. 855–862). Richland, SC.

Au, W. T., & Kwong, J. Y. Y. (2004). Measurements and effects of social-value orientation in social dilemmas. In R. Suleiman, D. V. Budescu, I. Fischer, & D. M. Messick (Eds.), *Contemporary psychological research on social dilemmas* (pp. 71–98). New York, NY: Cambridge University Press.

Axelrod, R. (1984). *The evolution of cooperation*. New York, NY: Basic Books.

Bacharach, M., Gold, N., & Sugden, R. (2006). *Beyond individual choice: Teams and frames in game theory*. Princeton, NJ: Princeton University Press.

Bogaert, S., Boone, C., & Declerck, C. (2008). Social value orientation and cooperation in social dilemmas: A review and conceptual model. *British Journal of Social Psychology*, *47*(3), 453–480.

Bolton, G. E., & Ockenfels, A. (2000). Erc: A theory of equity, reciprocity, and competition. *The American Economic Review*, *90*(1), 166–193.

Chalamish, M., Sarne, D., & Lin, R. (2013). Enhancing parking simulations using peer-designed agents. *IEEE Transactions on Intelligent Transportation Systems, 14*(1), 492–498.

Charness, G., & Rabin, M. (2002). Understanding social preferences with simple tests. *The Quarterly Journal of Economics, 117*(3), 817–869.

Cheng, K. L., Zuckerman, I., Nau, D., & Golbeck, J. (2011). The life game: Cognitive strategies for repeated stochastic games. In *2011 IEEE third international conference on privacy, security, risk and trust (PASSAT) and 2011 IEEE third inernational conference on social computing (SocialCom)* (pp. 95–102). Boston, MA.

de Jong, S., Tuyls, K., & Verbeeck, K. (2008). Artificial agents learning human fairness. In *Proceedings of the 7th international joint conference on autonomous agents and multiagent systems AAMAS'08* (Vol. 2, pp. 863–870). Richland, SC. International Foundation for Autonomous Agents and Multiagent Systems. Retrieved from http://dl.acm.org/citation.cfm?id=1402298.1402344

Elmalech, A., & Sarne, D. (2012). Evaluating the applicability of peer-designed agents in mechanisms evaluation. In *IEEE/WIC/ACM international conference on web intelligence and intelligent agent technology* (Vol. 2, pp. 374–381). Macau.

Elmalech, A., Sarne, D., & Agmon, N. (2015). Agent development as a strategy shaper. *Autonomous Agents and Multi-Agent Systems, 30*(3), 506–525. doi: 10.1007/s10458-015-9299-z

Fujita, K., Aydoğan, R., Baarslag, T., Hindriks, K., Ito, T., & Jonker, C. (2017). The sixth automated negotiating agents competition (ANAC 2015). In K. Fujita, Q. Bai, T. Ito, M. Zhang, F. Ren, R. Aydoğan, & R. Hadfi (Eds.), *Modern approaches to agent-based complex automated negotiation* (pp. 139–151). Cham: Springer International.

Gal, Y., Grosz, B., Kraus, S., Pfeffer, A., & Shieber, S. (2010). Agent decision-making in open mixed networks. *Artificial Intelligence, 174*(18), 1460–1480.
Golbeck, J., Robles, C., Edmondson, M., & Turner, K. (2011). Predicting personality from twitter. In *Proceedings of the 3rd IEEE international conference on social computing* (pp. 149–156). Boston, MA, USA.
Golle, P., & Ducheneaut, N. (2005). Preventing bots from playing online games. *ACM Computers in Entertainment, 3*(3), 3–3. doi: 10.1145/1077246.1077255
Harries, C., Evans, J. S. B. T., & Dennis, I. (2000). Measuring doctors' self-insight into their treatment decisions. *Applied Cognitive Psychology, 14*(5), 455–477. doi: 10.1002/1099-0720(200009)14:5<455::AID-ACP667>3.0.CO;2-V
Joireman, J. A., Lasane, T. P., Bennett, J., Richards, D., & Solaimani, S. (2001). Integrating social value orientation and the consideration of future consequences within the extended norm activation model of proenvironmental behaviour. *British Journal of Social Psychology, 40*(1), 133–155. doi: 10.1348/014466601164731
Kniesner, T. J., Grodner, A., & Bishop, J. A. (2011 August). Social Interactions in the Labor Market. Center for Policy Research Working Papers No. 133). Center for Policy Research, Maxwell School, Syracuse University. Retrieved from https://ideas.repec.org/p/max/cprwps/133.html.
Liebrand, W., & McClintock, C. (1988). 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.
Lin, R., Kraus, S., Oshrat, Y., & Gal, Y. K. (2010). Facilitating the evaluation of automated negotiators using peer designed agents. In *Proceedings of the twenty-fourth AAAI conference on artificial intelligence* (pp. 817–822). Atlanta: AAAI Press, Retrieved from http://dl.acm.org/citation.cfm?id=2898607.2898739.
Manisterski, E., Lin, R., & Kraus, S. (2008). Understanding how people design trading agents over time. In *Proceedings of the 7th international joint conference on autonomous agents and multiagent systems AAMAS* (Vol. 3 , pp. 1593–1596). Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems. Retrieved from http://dl.acm.org/citation.cfm?id=1402821.1402933.
Maynard-Smith, J. (1982). *Evolution and the theory of games*. Cambridge: Cambridge University Press.
McClintock, C. G., & Allison, S. T. (1989). Social value orientation and helping behavior. *Journal of Applied Social Psychology, 19*(4), 353–362.
Messick, D. M., & McClintock, C. G. (1968). Motivational bases of choice in experimental games. *Journal of Experimental Social Psychology, 4*(1), 1–25. Retrieved from http://www.sciencedirect.com/science/article/pii/0022103168900462
Murphy, R., Ackermann, K., & Handgraaf, M. (2011). Measuring social value orientation. *Judgment and Decision Making, 6*(8), 771–781.
Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering, 22*(10), 1345–1359.
Parks, C. D., & Rumble, A. C. (2001). Elements of reciprocity and social value orientation. *Personality and Social Psychology Bulletin, 27*(10), 1301–1309.
Rosenfeld, A., Zuckerman, I., Segal-Halevi, E., Drein, O., & Kraus, S. (2016). Negochat-a: a chat-based negotiation agent with bounded rationality. *Autonomous Agents and Multi-Agent Systems, 30*(1), 60–81. doi: 10.1007/s10458-015-9281-9
Skyrms, B. (2004). *The stag hunt and the evolution of social structure*. Cambridge: Cambridge University Press.
Sutter, M., Feri, F., Kocher, M., Martinsson, P., Nordblom, K., & Rützler, D. (2010). *Social preferences in childhood and adolescence ? a large-scale experiment*. (Working Papers). Faculty of Economics and Statistics, University of Innsbruck. Retrieved from http://EconPapers.repec.org/RePEc:inn:wpaper:2010-13.
Talman, S., Hadad, M., Gal, Y., & Kraus, S. (2005). Adapting to agents' personalities in negotiation. In *Proceedings of the fourth international joint conference on autonomous agents and multiagent systems AAMAS* (pp. 383–389). New York, NY: ACM.
Van Lange, P., De Bruin, E., Otten, W., & Joireman, J. (1997). Development of prosocial, individualistic, and competitive orientations: Theory and preliminary evidence. *Journal of personality and social psychology, 73*(4), 733–746.
Wilson, B., Zuckerman, I., & Nau, D. S. (2011). Modeling social preferences in multi-player games. In L. Sonenberg, P. Stone, K. Tumer, & P. Yolum (Eds.), *10th international conference on autonomous agents and multiagent systems*, Taipei, Taiwan, May 2–6, 2011 (pp. 337–344). IFAAMAS.
Zuckerman, I., & Hadad, M. (2012). The social landscape: Reasoning on the social behavior spectrum. *IEEE Intelligent Systems, 27*(2), 36–41. doi: 10.1109/MIS.2012.9