Grounding social interactions with affective intelligence in the iterated prisoner's dilemma

Joshua D. A. Jung*[1], Jesse Hoey†[1], Jonathan H. Morgan‡[2], Tobias Schröder§[3], and Ingo Wolf¶[3]

[1]University of Waterloo, David R. Cheriton School of Computer Science, Waterloo, Ontario, Canada
[2]Duke University, Dept. of Sociology, Durham, NC, USA
[3]University of Applied Sciences, Institute for Urban Futures, Potsdam, Germany

Abstract

Symbolic interactionist principles of sociology are based on the idea that human action is guided by culturally shared symbolic representations of identities, behaviours, situations and emotions. Shared linguistic, paralinguistic, or kinesic elements allow humans to coordinate action by enacting *identities* in social situations. Structures of identity-based interactions can lead to the enactment of social orders that solve social dilemmas (e.g., by promoting cooperation). Our goal is to build an artificial agent that mimics the identity-based interactions of humans, and to compare networks of such agents to human networks. In this paper, we take a first step in this direction, and describe a study in which humans played a repeated prisoner's dilemma game against other humans, or against one of three artificial agents (bots). One of the bots has an explicit representation of identity (for self and other), and attempts to optimise with respect to this representation. We compare the human play against bots to human play against humans, and show how the identity-based bot exhibits the most human-like behaviour.

1 Introduction

Recent advances in computational social science (CSS) have enabled better understanding of the network structures that constrain flows of information and patterns of interaction in social systems. Most approaches are driven by empirical measures and algorithmic analyses of network structures, based in part on the new abundance of digital social media data. However, besides a general observation of homophilic tendencies in social networks, we know little about the micro-level mechanisms of social interaction that give rise to the shape of networks. It has recently been proposed that a new type of agent called *BayesAct* [17], which models the emotional control of social interaction by humans, can explain the emergence of stable role relations and patterns of interaction [26]. Here, we are taking this work further by empirically studying a particular class of interactions, namely social dilemma games, a fundamental paradigm in the social sciences aimed at understanding the dynamics of human cooperation vs. competition. Preliminary results are encouraging in terms of supporting the validity of the *BayesAct* agent as a mechanistic model of human networked interactions.

*j35jung@cs.uwaterloo.ca
†jhoey@cs.uwaterloo.ca
‡jonathan.h.morgan@duke.edu
§schroeder@fh-potsdam.de
¶wolf@fh-potsdam.de

BayesAct [3, 16, 17, 26] is a partially observable Markov decision process (POMDP) model of affective interactions between a human and an artificial agent. *BayesAct* arises from the symbolic interactionist tradition in sociology, and more precisely from "Affect Control Theory" (ACT) [12]. *BayesAct* generalises this theory by modeling affective states as probability distributions, and allowing decision-theoretic reasoning about affect. *BayesAct* proposes that humans learn and maintain a set of *shared* cultural affective *sentiments* about people, objects, behaviours, and about the dynamics of interpersonal events. Humans use a simple affective mapping to appraise individuals, situations, and events as sentiments in a three dimensional vector space of evaluation (good vs. bad), potency (strong vs. weak) and activity (active vs. inactive). These mappings can be measured, and the culturally shared consistency has repeatedly been demonstrated to be extremely robust in large cross-cultural studies [13, 23]. Many believe this consistency "gestalt" is a keystone of human intelligence. Humans use it to make predictions about what others will do, and to guide their own behaviour. Importantly, this need to align implicitly defines an affective heuristic (a *prescription*) for making decisions quickly in interactions. The shared sentiments, and the resulting *affective ecosystem* of vector mappings, encodes a set of social prescriptions that, if followed by all members of a group, results in an equilibrium or *social order* [10] which is optimal for the group as a whole, rather than for individual members. Humans living at the equilibrium "feel" good and want to stay there. The evolutionary consequences of this individual need are beneficial for the species. However, agents with sufficient resources can plan beyond the prescription, allowing them to manipulate other agents to achieve individual profit in collaborative games.

For example, in the repeated prisoner's dilemma, cooperation has a different emotional signature than defection: it is usually viewed as nicer and more powerful. Rationality predicts an agent will try to optimize over his expected total payout, perhaps modifying this payout by some additional intrinsic reward for altruism. The *BayesAct* view is quite different: it says that an agent will take the most aligned action given her estimates of her own and her partner's affective *identity*. If she believes herself to be a *friend*, but believes her opponent to be a *scrooge* or a *traitor*, she will be more likely to defect, but may collaborate in an attempt to *reform* or *befriend* her opponent. On the other hand, if an agent believes himself to be a *scrooge*, he will defect by default, but may cooperate in order to manipulate his opponent.

As elucidated by Squazzoni [28], models of social networks must take into account the heterogeneity of individuals, behaviours, and dynamics in order to better account for the available evidence. Similarly, in this paper we argue that the principles encoded in affect control theory and *BayesAct* are ideally suited to capture these complex heterogeneities in individuals acting in groups. We present preliminary results from an experiment in which participants played a repeated prisoner's dilemma (PD) game against each other and against a set of computer programs, one of them *BayesAct*.

2 Background

2.1 Affect Control Theory

Affect Control Theory (ACT) arises from work on the psychology and sociology of human social interaction [12]. ACT proposes that social perceptions, behaviours, and emotions are guided by a psychological need to minimize the differences between culturally shared fundamental affective sentiments about social situations and the transient impressions resulting from the interactions between elements within those situations. Fundamental sentiments, $\mathbf{f}$, are representations of social objects, such as interactants' identities and behaviours, as vectors in a 3D affective space, hypothesised to be a universal organising principle of human socio-emotional experience [23]. The basis vectors of affective space are called Evaluation/-valence, Potency/control, and Activity/arousal (EPA). EPA profiles of concepts can be measured with the *semantic differential*, a survey technique where respondents rate affective meanings of concepts on numerical scales with opposing adjectives at each end (e.g., good, nice vs. bad, awful for E, weak, little vs. strong, big for P, and calm, passive vs. exciting, active for A). Affect control theorists have compiled lexicons of a few thousand words along with average EPA ratings obtained from survey participants who are knowledgeable about their culture [13]. For example, most English speakers agree that professors are about as nice as students (E), more powerful (P) and less active (A). The corresponding EPAs are $[1.7, 1.8, 0.5]$ for professor and $[1.8, 0.7, 1.2]$ for student[1]. In Japan, professor has the same P (1.8) but students are less powerful (-0.21).

The three dimensions were found by Osgood to be extremely robust across time and cultures. More recently these three dimensions are also thought to be related directly to intrinsic reward [8]. That is, it seems that reward is assessed by humans along the same three dimensions: Evaluation roughly corresponds with expected value, Potency with risk (e.g. powerful things are more risky to deal with, because they do what they want and ignore you), and Activity corresponds roughly with uncertainty, increased risk, and decreased values (e.g. faster and more excited things are more risky and less likely to result in reward) [8]. Similarly, Scholl argues that the three dimensions are in correspondence with the major factors governing choice in social dilemmas [25]. Evaluation is a measure of affiliation or correspondence between outcomes: agents with similar goals will rate each other more positively. Potency is a measure of dependence: agents who can reach their goals independently of other agents are more powerful. Activity is a measure of the magnitude of dependence: agents with bigger payoffs will tend to be more active.

Social events can cause transient impressions, $\boldsymbol{\tau}$ (also three dimensional in EPA space) of identities and behaviours that may deviate from their corresponding fundamental sentiments, $\mathbf{f}$. ACT models this formation of impressions from events with a grammar of the form actor-behaviour-object. Consider for example a professor (actor) who yells (behaviour) at a student (object). Most would agree that this professor appears considerably less nice (E), a bit less potent (P), and certainly more active (A) than the cultural average of a professor. Such transient shifts in affective meaning caused by specific events are described with models of the form $\boldsymbol{\tau}' = \boldsymbol{M}\mathscr{G}(\mathbf{f}', \boldsymbol{\tau})$, where $\boldsymbol{M}$ is a matrix of statistically estimated prediction coefficients from empirical impression-formation studies and $\mathscr{G}$ is a vector of polynomial features in $\mathbf{f}'$ and $\boldsymbol{\tau}$. In ACT, the weighted sum of squared Euclidean distances between fundamental sentiments and transient impressions is called *deflection*, and is hypothesized to correspond to an aversive state of mind that humans seek to avoid. This *affect control principle* allows ACT to compute *prescriptive* actions for humans: those that minimize the deflection. Emotions in ACT are computed as a function of the difference between fundamentals and transients ($\epsilon = \boldsymbol{E} \times (\boldsymbol{\tau} - \boldsymbol{R}\mathbf{f} - \boldsymbol{d})$, where $\boldsymbol{E}$, $\boldsymbol{R}$, and $\boldsymbol{d}$ are parameters learned from survey data), and are thought to be communicative signals of vector deflection that help maintain alignment between cooperative agents [12]. ACT has been shown to be highly accurate in explaining verbal behaviours of mock leaders in a computer-simulated business [27], and group dynamics [14], among others [21].

2.2 Partially Observable Markov Decision Processes

A partially observable Markov decision process (POMDP) [4] is a stochastic control model that consists of a finite set $\mathcal{S}$ of states; a finite set $\mathcal{A}$ of actions; a stochastic transition model $\Pr : S \times A \to \Delta(S)$, with $\Pr(s'|s,a)$ denoting the probability of moving from state s to s' when action a is taken, and $\Delta(S)$ is a distribution over $\mathcal{S}$; a finite observation set $\boldsymbol{\Omega_s}$; a stochastic observation model, $\Pr(\omega_s|s)$, denoting the probability of making observation $\omega_s \in \boldsymbol{\Omega_s}$ while the system is in state s; and a reward assigning $R(a, s')$ to a transition to s' induced by action a. A *policy* maps *belief states* (i.e., distributions over $\mathcal{S}$) into actions, such that the expected discounted sum of rewards is (approximately) maximised. We use *factored* POMDPs in which the state is represented by the cross-product of a set of variables or features. POMDPs have been used as models for many human-interactive domains, including assistive technologies [15].

[1] All EPA labels and values in the paper are taken from the Indiana 2002-2004 ACT lexicon [13]. Values range by historical convention from -4.3 to $+4.3$.

2.3 Bayesian Affect Control Theory

Recently, ACT was generalised and formulated as a POMDP for human-interactive artificially intelligent systems [17]. This new model, called *BayesAct*, generalises the original theory in three ways. First, sentiments and impressions are viewed as probability distributions over latent variables (e.g., $\mathbf{f}$ and $\boldsymbol{\tau}$) rather than points in the EPA space, allowing for multimodal, uncertain and dynamic affective states to be modeled and learned. Second, affective interactions are augmented with *propositional* states and actions (e.g. the usual state and action space considered in AI applications). Third, an explicit reward function allows for goals that go beyond simple deflection minimization.

A *BayesAct* POMDP models an interaction between two agents (human or machine) denoted *agent* and *client*. The state, $\boldsymbol{s}$, is the product of six 3-dimensional continuous random variables corresponding to fundamental and transient sentiments about the *agent*'s identity ($\mathbf{F}_a, \mathbf{T}_a$), the current (*agent* or *client*) behaviour ($\mathbf{F}_b, \mathbf{T}_b$) and the *client*'s identity ($\mathbf{F}_c, \mathbf{T}_c$). We use $\mathbf{F} = \{\mathbf{F}_a, \mathbf{F}_b, \mathbf{F}_c\}$ and $\mathbf{T} = \{\mathbf{T}_a, \mathbf{T}_b, \mathbf{T}_c\}$. The state also contains an application-specific set of random variables $\mathbf{X}$ that are interpreted as *propositional* (i.e. not *affective*) elements of the domain (e.g. whose turn it is, game states - see Section 3), and we write $\boldsymbol{s} = \{\mathbf{f}, \boldsymbol{\tau}, \mathbf{x}\}$. Here the *turn* is deterministic (*agent* and *client* take turns), although this is not necessary in *BayesAct*. The *BayesAct* reward function is application-specific over $\mathbf{x}$. The state is not observable, but observations $\boldsymbol{\Omega_x}$ and $\boldsymbol{\Omega_f}$ are obtained for $\mathbf{X}$ and for the affective behaviour $\mathbf{F}_b$, and modeled with probabilistic observation functions $Pr(\boldsymbol{\omega_x}|\mathbf{x})$ and $Pr(\boldsymbol{\omega_f}|\mathbf{f}_b)$, respectively. Actions in the *BayesAct* POMDP are factored in two parts: $\mathbf{b_a}$ and a, denoting the *affective* and *propositional* components, respectively. For example, if a tutor gives a hard exercise to do, the manner in which it is presented, and the difficulty of the exercise, combine to form an affective impression $\mathbf{b_a}$ that is communicated. The actual exercise (content, difficulty level, etc) is the *propositional* part, a. The state dynamics factors as $Pr(\boldsymbol{s}'|\boldsymbol{s},\mathbf{b_a},a) = Pr(\boldsymbol{\tau}'|\boldsymbol{\tau},\mathbf{f}',\mathbf{x})Pr(\mathbf{f}'|\mathbf{f},\boldsymbol{\tau},\mathbf{x},\mathbf{b_a})Pr(\mathbf{x}'|\mathbf{x},\mathbf{f}',\boldsymbol{\tau}',a)$, and the fundamental behaviour, $\mathbf{F}_b$, denotes either observed *client* or taken *agent* affective action, depending on whose *turn* it is (see below). That is, when the *agent* acts, there is a deterministic mapping from the affective component of his action ($\mathbf{b_a}$) to the *agent*'s behaviour $\mathbf{F}_b$. When *client* acts, *agent* observes $\boldsymbol{\Omega_f}$ (the affective action of the other agent). The third term in the factorization of the state dynamics is the *Social Coordination Bias* (SCB). The SCB gives an estimate of how the state will grow based on the previous state and the current sentiments.

The transient impressions, $\mathbf{T}$, evolve according to the impression-formation operator in ACT ($\boldsymbol{M}\mathscr{G}$), so that $Pr(\boldsymbol{\tau}'|...)$ is deterministic. Fundamental sentiments are expected to stay roughly constant over time, but are subject to random drift (with noise $\boldsymbol{\Sigma}_f$) and are expected to be close to the transient impressions because of the *affect control principle*. Thus, the dynamics of $\mathbf{F}$ is[2]:

$$Pr(\mathbf{f}'|\mathbf{f}, \boldsymbol{\tau}) \propto e^{-\psi(\mathbf{f}',\boldsymbol{\tau})-\xi(\mathbf{f}',\mathbf{f})} \tag{1}$$

where $\psi \equiv (\mathbf{f}'–\boldsymbol{M}\mathscr{G}(\mathbf{f}', \boldsymbol{\tau}))^T\boldsymbol{\Sigma}^{-1}(\mathbf{f}'–\boldsymbol{M}\mathscr{G}(\mathbf{f}', \boldsymbol{\tau}))$ combines the *affect control principle* with the impression formation equations, assuming Gaussian noise with covariance $\boldsymbol{\Sigma}$. The inertia of fundamental sentiments is $\xi \equiv (\mathbf{f}'–\mathbf{f})^T\boldsymbol{\Sigma}_f^{-1}(\mathbf{f}'–\mathbf{f})$, where $\boldsymbol{\Sigma}_f$ is diagonal with elements $\beta_a, \beta_b, \beta_c$. The state dynamics are non-linear due to the features in $\mathscr{G}$. This means that the belief state will be non-Gaussian in general, and *BayesAct* uses a *bootstrap filter* to compute belief updates.

The distribution in (1) gives the prescribed (if *agent* turn), or expected (if *client* turn), action as the component $\mathbf{f}'_b$ of $\mathbf{f}'$. Thus, by integrating over $\mathbf{f}'_a$ and $\mathbf{f}'_c$ and the previous state, we obtain a probability distribution, $\pi^\dagger$, over $\mathbf{f}'_b$ that acts as a *normative action bias* (NAB): it tells the agent what to expect from other agents, and what action is expected from it in belief state $b(\boldsymbol{s})$:

$$\pi^\dagger(\mathbf{f}'_b) = \int_{\mathbf{f}'_a,\mathbf{f}'_c}\int_{\boldsymbol{s}} Pr(\mathbf{f}'|\mathbf{f}, \boldsymbol{\tau}, \mathbf{x})b(\boldsymbol{s}) \tag{2}$$

BayesAct agents in the experiment choose actions using a Monte-Carlo tree search (MCTS) algorithm [3]. The important thing here is that the *action* space is explored using the NAB as in Equation 2. The first

[2] We leave out dependence on $\mathbf{x}$ for clarity, and on $\mathbf{b_a}$ since this is replicated in $\mathbf{f}'_b$.

Table 1: Optimal (deflection minimising) behaviours for two *BayesAct* agents with fixed identities.

agent	client	optimal behaviour	closest labels	distance from collaborate	distance from abandon
friend	friend	$1.98, 1.09, 0.96$	treat/toast	0.4	23.9
friend	scrooge	$0.46, 1.14, \text{-}0.27$	reform/lend money to	1.7	10.5
scrooge	friend	$\text{-}0.26, \text{-}0.81, \text{-}0.77$	curry favor/look away from	8.5	4.2
scrooge	scrooge	$\text{-}0.91, \text{-}0.80, \text{-}0.01$	borrow money/chastise	9.6	2.7

tree expansion in MCTS is the expected behaviour ($\arg\max_{\mathbf{f}'_b} \pi^\dagger(\mathbf{f}'_b)$), and subsequent tree expansions are randomly sampled from $\pi^\dagger$. An 'action resolution' (δ_a) is specified as the maximum distance from an existing branch (action) beyond which a new action is added to the tree. Similarly, an 'observation resolution' (δ_o) is the same, but for observations. We use $\delta_a = 0.1$, $\delta_o = 0.1$. The MCTS is anytime, and we use a timeout of 5 seconds. Adding time would allow for more action-space exploration.

3 Experiments and Results

3.1 Repeated Prisoner's Dilemma and *BayesAct*

The prisoner's dilemma is a classic two-person game in which each person can either *defect* by taking \$1 from a (common) pile, or *cooperate* by giving \$2 from the same pile to the other person. There is one Nash equilibrium in which both players defect, but when humans play the game they often are able to achieve cooperation. A rational agent will optimise over his expected long-term payoffs, possibly by averaging over his expectations of his opponent's type (or strategy).

A *BayesAct* agent computes what *affective* action (an EPA vector) is prescribed in the situation (given his estimates of his and the other's identities, and of the affective dynamics), and then seeks the propositional action ($\in \{$*cooperate*, *defect*$\}$) that, according to a stored cultural definition, is most consistent with the prescribed affective action. This propositional action forms part of Ω'_x when it is the other player's turn, and so the distance from the propositional action choices to the prescribed affective action is the social coordination bias. As the game is repeated, the *BayesAct* agent updates his estimates of identity (for self and other), and adjusts his play accordingly. For example, a player who defects will be seen as quite negative, and appropriate affective responses will be to defect. However, the actual mapping is more complex, as the *BayesAct* agent maintains multiple identity hypotheses about his opponent, probabilistically re-weighted as the game progresses.

The normative action bias (NAB) for *BayesAct* agents is the usual deflection minimizing affective $\mathbf{f}_b$ given distributions over identities of *agent* and *client* (Equation 2). Thus, if *agent* thought of himself as a *friend* (EPA:$\{2.75, 1.88, 1.38\}$) and knew the other agent to be a *friend*, the deflection minimizing action would likely be something good (high E). Indeed, a simulation shows that one would expect a behaviour with EPA=$\{1.98, 1.09, 0.96\}$, with closest labels such as *treat* or *toast*[1]. Intuitively, cooperate seems like a more aligned propositional action than defect. This intuition is confirmed by the distances from the predicted (affectively aligned) behaviour to *collaborate with* (EPA:$\{1.44, 1.11, 0.61\}$) and *abandon* (EPA:$\{\text{-}2.28, \text{-}0.48, \text{-}0.84\}$) of 0.4 and 23.9, respectively[3]. Table 1 shows all combinations if each agent could also be a *scrooge* (EPA:$\{\text{-}2.15, \text{-}0.21, \text{-}0.54\}$). We see that a *friend* would still collaborate with a *scrooge* (in an attempt to *reform* the scrooge), a *scrooge* would abandon a *friend* (*look away from* in shame), and two scrooges would defect.

The *agent* will predict the *client*'s behavior using the same principle: compute the deflection minimising affective action, then deduce the propositional action based on that. Thus, a *friend* would predict that a *scrooge* would defect. If a *BayesAct* agent has sufficient resources, he could search for an affective action

[3]Here, we choose *collaborate with* and *abandon* as representative of the affective meaning of the actions in the game [3]

near to his optimal one, but that would still allow him to defect. Importantly, he is *not trading off costs in the game with costs of disobeying the social prescriptions*: his resource bounds and action search strategy are preventing him from finding the more optimal (individual) strategy, implicitly favoring those actions that benefit the group and solve the social dilemma.

We use propositional $\mathbf{X} = \{Turn, Ag\_play, Cl\_play\}$ denoting whose turn it is and *agent* and *client* state of play ($\in \{not\_played, cooperate, defect\}$). The agents' reward is only over the game (e.g. 2, 1, or 0), and we use a two time-step game in which both *agent* and *client* choose their actions at the first time step, and then communicate this to each other on the second step.

3.2 Description of Experiment

In order to compare the predictions of *BayesAct* to human play, we recruited 70 students (55 male and 15 female) from an undergraduate class on artificial intelligence at a large Canadian university. The participants had to sign up online first, and were given an automatically generated username. They read an information and consent form and chose to either withhold or consent to the use of their data. The participants were divided into four groups of size 12-20 by last name, with each group playing together for 40 minutes in a computer lab environment. Groups of this size were necessary to minimize the time taken in finding new opponents. In total, 360 games were played, where the length of each game was randomly chosen between $12-18$ rounds (plays of cooperation or defection). For any given game, a participant played against either (1) another randomly chosen participant; (2) an automated *tit-for-tat* player; (3) a *BayesAct* agent as described above; or (4) a fixed strategy of cooperate three times followed by always defect, hereafter referred to as *jerkbot*. Participants played through all opponent types on a rotation, which was randomized individually for each participant at the start of play. On sign-up and after each game (of between $12-18$ rounds), participants were asked the following by providing them with a slider for each dimension (Evaluation, Potency and Activity or EPA), known as a semantic differential [13]:

- how they felt about the plays in the game (take 1 or give 2), out of context. This is a measure of the social coordination bias (SCB). *BayesAct* agents then interpret actions in the game by comparing the actions to these two vectors.
- how they felt about themselves (their self identity). This gives *BayesAct* its self-identity, $\mathbf{f}_a$ (as we want it to replicate a participant). We use the raw data from all student responses across all questions as $\mathbf{f}_a$ for *BayesAct*.
- how they felt about their opponent in the game they just played. Before the first game we asked they how they felt about a generalised opponent in this game, giving the *BayesAct* client identity $\mathbf{f}_c$.

For more detail on the information on the prisoner's dilemma application and its interface, please refer to the Appendix, section 6.

The *BayesAct* agent for the first session was initialised as described above using the initial responses of the students during the sign-up phase (48 responses). The *BayesAct* agent for the second session was initialised as described above using all the responses during both sign-up and the first session[4].

A total of 89 samples were used for identities (resampled to get N=2000 samples used in the *BayesAct* particle filter) and an average of 89 samples used for the SCB. From this sample, we measured for *Give 2* an EPA of $\{1.4, 0.10, 0.18\}$, and for *Take 1*, $\{-0.65, 0.85, 0.70\}$. *Take 1* is seen as more negative and more powerful and active. Figure 1 shows the self and other identities as rated by the participants during the signup phase. The self is seen as more positive than the opponent or "other" (means 1/0.25 for self/other), but about the same power (0.56/0.64) and activity (0.41/0.33).

Table 2 shows an example game played by a human against the *BayesAct* agent used in the second phase of the study (so the identities and SCB were learned from the survey data during the initial and first phases).

[4] In both cases, the *BayesAct* agent played with a reward matrix corresponding to actions of *Give 10* and *Take 1*. However, it is unlikely that this resulted in a significant deviation in play, as the *BayesAct* agent's planning time was limited to 1 second, a time short enough to make its decisions nearly independent of the reward matrix.

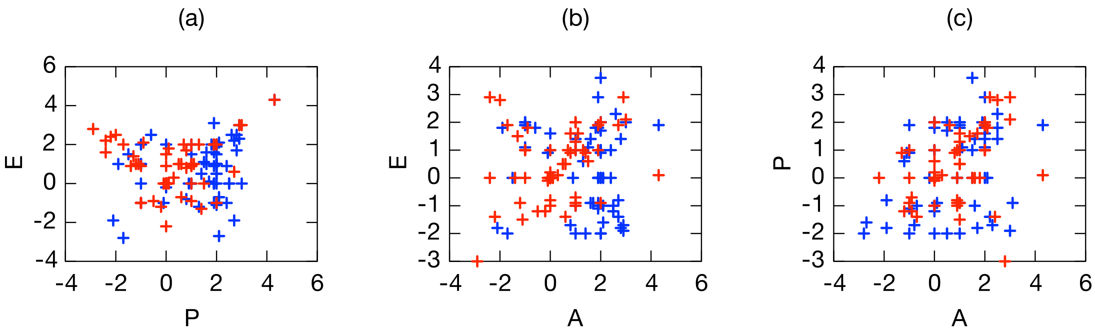


Figure 1: Out-of-context ratings of self (blue) and other (red).

Table 2: Example run of human playing *BayesAct* agent showing plays, identities and emotions as inferred by the *BayesAct* agent. EPA = mean Evaluation, Potency, Activity for the given quantity.

play		*BayesAct* id		human id		*BayesAct* emotion		human emotion	
BayesAct	human	EPA	label	EPA	label	EPA	label	EPA	label
take 1	give 2	1.4, 0.8, 0.8	spokeswoman	0.3, 0.3, 0.3	questioner	-0.3, -0.5, 0.2	exasperated	0.7, -0.2, -0.2	feminine
give 2	give 2	1.3, 1.2, 0.8	right-hand_man	0.5, 0.2, 0.2	stepbrother	0.9, 0.5, -0.1	reverent	1.2, -0.0, 0.0	sentimental
give 2	give 2	1.3, 1.2, 0.9	right-hand_man	0.5, 0.1, 0.0	stepdaughter	0.6, 0.5, -0.4	nostalgic	1.2, -0.1, 0.2	feminine
give 2	give 2	1.5, 1.3, 1.0	customer	0.4, 0.1, 0.1	convalescent	-0.6, -0.8, -0.8	lovesick	0.9, -0.3, 0.1	feminine
take 1	give 2	1.4, 1.4, 1.1	customer	0.2, 0.1, 0.1	convalescent	-1.0, -1.1, -0.9	embarrassed	0.7, -0.4, 0.1	feminine
give 2	take 1	1.3, 1.2, 1.0	right-hand_man	0.1, 0.1, 0.1	stepparent	-0.7, -1.1, -0.9	submissive	0.1, 0.1, 0.6	defensive
give 2	take 1	1.0, 1.0, 0.9	purchaser	0.1, 0.0, 0.1	stepparent	-0.6, -0.9, -0.6	disapproving	-0.2, 0.2, 0.6	defensive
give 2	give 2	1.0, 1.1, 0.9	purchaser	0.1, 0.0, 0.1	stepparent	-0.4, -0.5, -0.9	aloof	0.6, -0.1, 0.2	feminine
take 1	give 2	1.1, 1.3, 1.1	right-hand_man	0.1, 0.0, 0.1	stepparent	-0.4, -0.3, -0.8	placid	0.7, -0.3, 0.1	feminine
give 2	take 1	0.9, 1.2, 0.9	stud	0.0, 0.0, 0.1	stepparent	-0.1, -1.3, -1.2	melancholy	0.3, 0.4, 0.6	defensive
take 1	give 2	0.8, 1.3, 1.0	stud	0.1, 0.0, 0.1	stepparent	-0.2, -0.6, -0.7	aloof	0.7, -0.1, 0.2	feminine
give 2	take 1	0.7, 1.0, 0.7	matriarch	-0.0, 0.0, 0.1	stepmother	-0.5, -1.0, -0.4	disapproving	-0.0, 0.1, 0.5	defensive
take 1	give 2	0.8, 1.2, 0.9	boyfriend	0.0, 0.0, 0.1	stepparent	-0.4, -0.5, -0.4	contrite	0.6, -0.2, 0.1	feminine
take 1	give 2	0.8, 1.2, 0.9	boyfriend	0.1, 0.0, 0.1	stepparent	-0.9, 0.0, 0.1	envious	0.1, -0.7, 0.0	contrite
give 2	take 1	0.9, 1.2, 1.0	stud	0.0, 0.0, 0.1	stepparent	-0.9, -1.5, -0.9	submissive	-0.2, 0.1, 0.5	defensive
take 1	take 1	0.8, 1.0, 0.7	consultant	0.0, 0.0, 0.1	stepparent	-0.9, -1.0, -0.5	disapproving	-0.3, 0.1, 0.6	sarcastic
take 1	take 1	0.7, 0.7, 0.5	steady	-0.0, 0.0, 0.2	stepmother	-0.7, -0.7, -0.2	dependent	-0.3, 0.2, 0.6	sarcastic
take 1	give 2	0.8, 0.8, 0.6	steady	-0.0, 0.0, 0.1	stepmother	-0.4, -0.1, -0.3	contemptuous	0.3, -0.2, 0.2	feminine
give 2	give 2	0.9, 0.8, 0.6	head_nurse	-0.0, 0.0, 0.1	stepparent	-0.5, -0.7, -0.3	dependent	0.7, -0.2, 0.0	feminine
give 2	give 2	1.1, 0.9, 0.6	head_nurse	0.0, 0.0, 0.1	stepparent	-0.3, -0.5, -0.2	contrite	0.9, -0.3, 0.0	feminine
give 2	give 2	1.2, 0.9, 0.7	businesswoman	0.0, 0.0, 0.1	stepparent	-0.7, -0.9, -0.4	disapproving	0.8, -0.3, 0.1	feminine
give 2	give 2	1.4, 1.0, 0.8	workmate	0.0, 0.0, 0.1	stepparent	-0.4, -0.3, -0.4	placid	1.0, -0.2, 0.1	feminine
take 1	give 2	1.6, 1.2, 1.0	Air_Force_reservist	0.1, 0.1, 0.1	stepparent	-1.4, 0.2, -0.0	scornful	-0.0, -0.8, 0.1	exasperated
take 1	take 1	1.4, 1.1, 0.9	right-hand_man	0.1, 0.1, 0.1	stepparent	-1.6, -0.2, -0.1	displeased	-0.3, -0.3, 0.5	raunchy
give 2	take 1	1.4, 1.0, 0.8	businesswoman	0.0, 0.0, 0.1	stepparent	-1.3, -0.9, -0.5	flustered	-0.3, 0.1, 0.6	sarcastic
take 1	give 2	1.4, 1.1, 0.9	right-hand_man	0.1, 0.1, 0.1	stepparent	-1.1, 0.3, 0.0	vengeful	-0.1, -0.6, 0.2	exasperated
take 1	take 1	1.2, 1.0, 0.8	businesswoman	0.0, 0.0, 0.1	stepparent	-1.2, 0.0, 0.2	scornful	-0.3, -0.1, 0.5	pompous
take 1	take 1	1.1, 0.8, 0.5	head_nurse	0.0, 0.0, 0.1	stepmother	-1.0, -0.6, -0.1	cynical	-0.3, 0.1, 0.7	sarcastic
take 1	take 1	0.9, 0.8, 0.5	head_nurse	0.0, 0.0, 0.1	stepmother	-0.9, -0.7, -0.2	finicky	-0.3, 0.2, 0.7	sarcastic
take 1		0.9, 0.8, 0.5	head_nurse	-0.0, 0.0, 0.1	stepmother	-0.8, -0.5, -0.2	shaken	-0.3, -0.2, 0.3	exasperated

In this example, the *BayesAct* agent starts with a defection, while the human starts by cooperating. The *BayesAct* agent and human both then cooperate for three more games, after which the *BayesAct* agent again defects, but feels *embarrassed* about it. The human responds with a defection, and feels *defensive*. The game resolves into a tit-for-tat play, followed by mutual defection, mutual cooperation, and finally by mutual defection.

Table 3: Summary statistics for each opponent type. coops: number of cooperations after 10th game.

	num.	avg. game	agent (human)		client (human or bot)	
opponent	games	length	payoff	coops	payoff	coops
jerkbot	83	15.01	15.86 ± 3.00	0.09 ± 0.24	22.33 ± 6.00	0.00 ± 0.00
bayesact	73	14.85	27.05 ± 5.92	0.54 ± 0.40	22.19 ± 7.87	0.69 ± 0.32
human	35	15.43	24.11 ± 7.55	0.56 ± 0.45	26.00 ± 5.92	0.51 ± 0.47
titfortat	82	14.82	27.66 ± 5.39	0.81 ± 0.35	26.96 ± 6.00	0.83 ± 0.34

We eliminated 8 participants from the sample whose pattern of ratings indicated that they did not take the experiment seriously. To make these determinations, we considered each participant's proportion of extreme ratings, as well as their relative proportions of positive, neutral, and negative ratings. Specifically, a participant's data was eliminated if that participated met at least two of the following conditions.

- The participant rated more than 50% of the questions in an extreme manner (between 4 and 4.3)
- The participant rated more than 50% of the questions in a neutral manner (between 0 and 0.3)

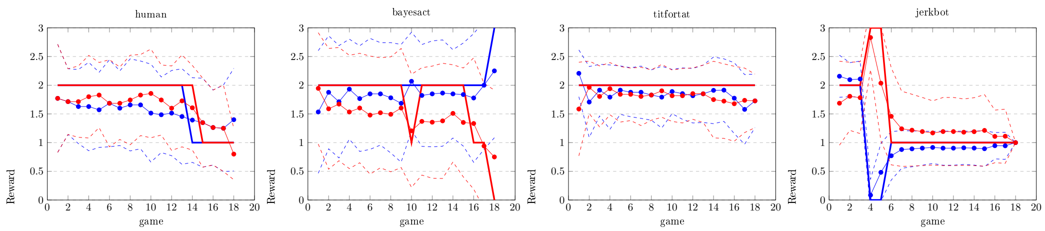


Figure 2: Blue=human; Red=agent (human, bayesact, titfortat and *jerkbot*); dashed=std.dev.; solid (thin, with markers): mean; solid (thick): median.

- The participant rated more than 65% of the questions in a single direction on the slider (either positive or negative).
- The participant rated more than 65% of the questions at exactly 0.

Table 3 shows the summary satistics across all games played against the different opponents.

Figure 2 shows the mean, standard deviation, and median reward gathered at each step of the game, for each of the opponents. The blue lines show the human play, while the red lines show the opponent (one of human, *BayesAct*, *tit-for-tat*, or *jerkbot*). We see that humans mostly manage to cooperate together until about 4-5 games before the end. The *tit-for-tat* strategy ensures more even cooperation, but is significantly different from humans. *Jerkbot* is obvious, as a few defections after three games convinces the human to defect thereafter. The *BayesAct* agent play is very similar to the human play, but the human participants seem to be able to take advantage of the *BayesAct* agents late in the game. This may be because the *BayesAct* agent is using a short (5 second) planning timeout, and we would need to compare to a zero timeout (so only using the ACT prescriptions) and to longer timeouts to see how this behaviour changes.

Table 4: Means of pre-game (initial) and post-game impressions for each opponent type.

	give 2			take 1			self (human)			other (human/bot)		
opponent	E	P	A	E	P	A	E	P	A	E	P	A
(initial)	1.4	0.1	0.2	-0.6	0.9	0.6	1.1	0.6	0.3	0.2	0.6	0.3
jerkbot	1.3	-0.3	-0.1	-1.3	0.8	0.7	1.3	-0.1	0.9	-1.9	0.4	0.5
bayesact	1.3	0.1	0.0	-0.9	1.1	1.0	0.7	1.4	1.2	0.4	-0.1	-0.3
human	1.7	0.7	0.3	-1.2	0.4	0.3	1.5	1.2	1.0	0.5	0.0	0.1
titfortat	2.3	1.2	1.1	-1.2	0.5	0.3	1.9	1.7	1.7	2.2	1.1	1.1

To further investigate the differences between the different opponents, we measure the mean fraction of cooperative actions on the part of the human after (and including) the 10^{th} game. We find that, when playing against another human, humans cooperate in 0.56 ± 0.45 of these last games. This number was almost the same when playing *BayesAct* agent at 0.54 ± 0.40. Against *tit-for-tat*, there was much more cooperation (0.81 ± 0.35). Finally, against *jerkbot*, it was very low 0.09 ± 0.24. We also computed the mean EPA ratings of the self and other after each game, as shown in Table 4. We found that *jerkbot* (EPA:$\{-1.9, 0.4, 0.5\}$) is seen as much more negative, and *tit-for-tat* (EPA:$\{2.2, 1.1, 1.1\}$) much more positive, than human (EPA:$\{0.5, 0.0, 0.1\}$) or *BayesAct* (EPA:$\{0.4, -0.1, -0.3\}$), and that the human feels less powerful when playing *jerkbot* (EPA of self:$\{1.3, -0.1, 0.9\}$) than when playing *BayesAct* (EPA of self:$\{0.7, 1.4, 1.2\}$), or another human (EPA of self:$\{1.5, 1.2, 1.0\}$). Humans felt more powerful, positive and active when playing *tit-for-tat* (EPA of self:$\{1.9, 1.7, 1.7\}$).

4 Related Work

Many authors have now presented the idea that human handling of an infinite action space may be governed largely by affective processes [2, 20, 3]. Shared affective structures allow agents to focus on a subset of possibilities, being those that provide aligned interactions according to the shared structure. The subset of possibilities is the "cultural prescriptions" for behaviours that are "rational relative to the social conventions and ethics" ([2], p200). Some social identity theorists have attempted to integrate models of these shared structures (as identities) into utility theory [1], in order to provide Kahneman's "fast thinking" [19]. Interestingly, ACT gives a functional account of this "fast" pathway as sentiment encoding prescriptive behaviour, while *BayesAct* extends this with a slow pathway that enables exploration and planning away from the prescription [3].

The prisoner's dilemma has long been studied, starting with the work of [5]. Recent work has looked at modeling both rational choice and social imitation to simulate more human-like behaviour in networked PD games [29]. Other researchers have looked at using emotional signals to influence play in PD games, for example by changing expectations of future games using emotional signals [7], by linking valence with an exploration bonus [18], or by using emotional appraisals as intrinsic reward signals [22]. Our work is the first to propose a method based on affective identity alone, effectively ignoring the payoff matrix and seeking socio-emotional balance rather than utility as the main driving force behind human behaviour in PD [3]. The dynamic nature of our proposal coheres well with recent work showing that, in networked social dilemmas, cooperation is maintained by flexible tie formation and cues of shared identities. Recent work on evolutionary models addressing the highly dynamic nature of identity and tie formation in small groups (e.g. [9]) has been promoted as promising avenues in this domain (cf. [24]).

5 Conclusion

We have presented a model for affectively guided play in the prisoner's dilemma. Our hypothesis is that we can design agents that are more human-like in their behaviours by basing the agents on symbolic interactionist principles. The idea is that the agents follow the social prescriptions for action based on their sense of self, or their identity. We show preliminary results from a study of human play in the iterated prisoner's dilemma, and discuss some of the early findings. We are currently looking at learning the parameters of the model (e.g. $\mathbf{\Sigma}$ and $\mathbf{\Sigma}_f$) from data. Other research avenues include intelligent tutoring [17] and other games [3], Our long-term objective is to run simulations of *BayesAct* agents (learned from human data) in a network. In previous work, we have observed the emergence of identity structures from the learning process in dyads of *BayesAct* agents [26]. We believe these emergent structures may provide a demonstration that the clustering of complementary and competitive identities (e.g, friends and enemies) [6] predicted by structural balance theory [11] is a property of identity maintenance generally.

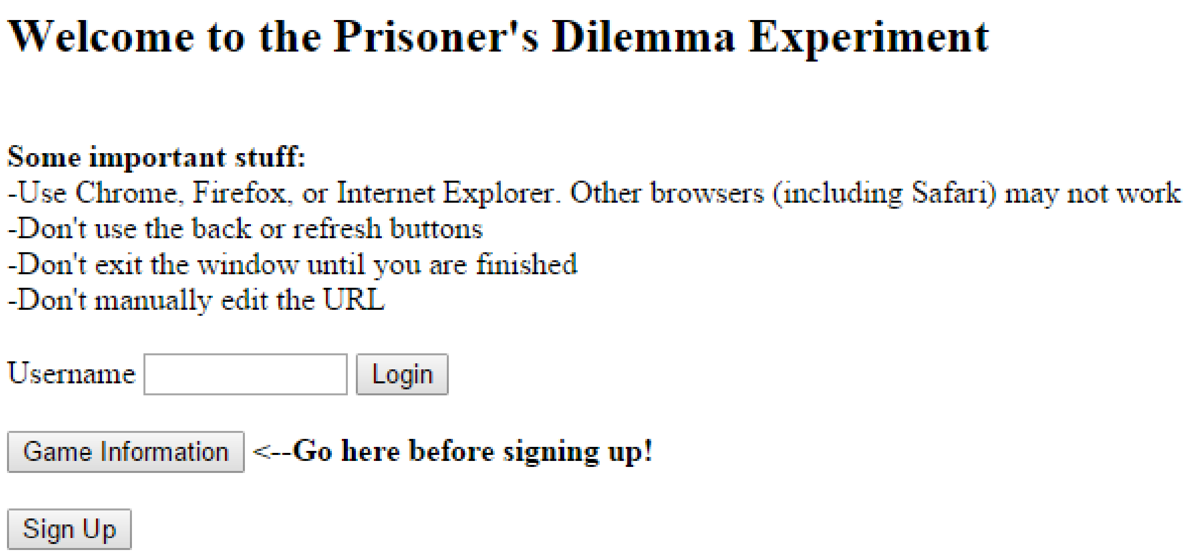
6 Appendix

Here, we reproduce the prisoner's dilemma application as it was seen by participants of the study. Upon entering the prisoner's dilemma URL, participants arrived at the welcome screen shown in section 6.1. They were instructed to review the game information given in section 6.2 before signing up. Signing up consisted of three phases: answering a short demographic questionnaire, reviewing and accepting/declining an informed consent form, and assigning E, P, and A values to each of the key concepts of "Self", "Other Player", "Give", and "Take". These materials can be found in sections 6.3 and 6.4 respectively.

Once signed up, participants were able to begin play. Upon being assigned a match, a participant would arrive at the Start of Game screen given in section 6.5 with the option to either Give 2 or Take 1. After making a selection, the participant had to wait for the server to respond with her opponent's move. Note that, even in the case where the opponent was a bot, some time was always allowed before a reply was sent to preserve the illusion that all players were human. On completion of the final round (decided randomly to be a value in the range 12-18), the participant was asked to again evaluate E, P, and A values for each of the four key concepts. An example End of Game screen can be found in section 6.5 and sliders, as before, in section 6.4.

When the allotted play time of a group of participants ran out, they were instructed to stop playing and open the (previously hidden) debriefing page. This page, which can be found in section 6.6, revealed to the participants that they played against bots as well as each other, and gave them the option to withdraw their data from the study. This was the final interactions participants had with the prisoner's dilemma application.

6.1 Welcome Screen



6.2 Game Information

The Iterated Prisoner's Dilemma

The game you will be playing is a variant of the much studied class of matrix games referred to as iterated prisoner's dilemmas. In this game, you are partnered with another player and given the choices "Give 2" or "Take 1". If you choose "Give 2", the other player will be given 2 points. If you choose "Take 1", you will be given 1 point. After both players have made a choice, each will be able to see what the other chose, and points will be distributed. Then a new round will begin in which both players are again given the options of "Give 2" or "Take 1".

For any given round of play, the point structure can be summarized as follows.

Reward Matrix
(your payoff/opponent's payoff)

	Opponent Gives 2 (G)	**Opponent Takes 1 (T)**
You Give 2 (G)	2/2	0/3
You Take 1 (T)	3/0	1/1

At some point, the game will end abruptly and you will be asked to perform EPA ratings based on the game you just played. It is important to note here that a single game does not have a winner and loser. You will add your points to your total and the other player will add theirs. At the end of the assignment, every point you earn will be counted as one entry into a draw for ten $20 Amazon gift cards.

Continue

A Review of EPA Ratings

Psychologists agree that emotions have three dimensions:

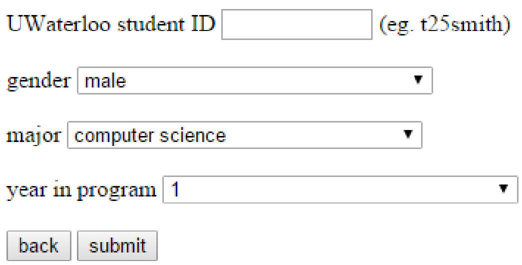
Evaluation - is it something good/nice or rather bad/awful;
Potency - is it strong/powerful or rather weak/powerless; and
Activity - is it slow/inactive or rather fast/lively.

For example, if someone said to you: "I like you", you might perceive this as very good/nice (Evaluation), quite powerful/big (Potency), and possibly neutral on the third dimension, i.e. neither very slow/quiet nor very lively/noisy (Activity). On the other hand, if someone said "let's go dancing!" you might perceive this as a bit good/nice (Evaluation), neutral potency, but very active.

Note that there are no right or wrong answers; we are interested in your quick and spontaneous emotional reactions. If in doubt about a rating, consider how the average member of society might rate, having been subject to the same experiences as you.

Return to Login

6.3 Sign-Up Questionnaire



Consent Information

Study Title: Applying BayesACT to the Iterated Prisoner's Dilemma

Faculty Supervisor: Dr. Jesse Hoey, Department of Computer Science, (***) ***-**** ext. *****, ****@****

Student Investigator: Josh Jung, Department of Computer Science, (***) ***-**** ext. *****, ****@****

Course Assignment

During class on October 20 and 22, instead of the lectures, you will be asked to play a series of iterated prisoner's dilemma games in ****/****. You may bring your own laptop or use one of the Macs in the lab. The game is very simple; you select one of two options and receive a score based on the combined choices made by you and your opponent. After each set of approximately 20 such games, you will be asked to rate both yourself and your opponent on the evaluation, potency, and activity scales prescribed by Affect Control Theory. You will then be matched with another player to play another set.

For this course assignment you will be asked to sign up online at **********. You will be asked for your UWaterloo ID, as well as your major and gender, which you can choose not to share if you wish. You will also be assigned an ID, which you must bring to class on the days of the assignment. Completion of the assignment is worth 5% of your mark and is expected to take approximately 3 hours of your time. At the end of the assignment each point you earn while playing the game will be counted as one entry into a draw for one of ten (10) $20 Amazon gift cards.

If you are unable to attend these two classes you can choose to complete a paper review instead and will still be entered into the draw with odds equivalent to the median player participating in the study. This requires that you choose a research paper in artificial intelligence, read it, and write a 2-page review of the paper. Reviews will be assigned a pass/fail (5%/0%) grade based on the suitability of the review, where a suitable review is one that is coherent and makes it clear that its author has read the paper. Your odds of winning one of the prizes is based on the number of individuals who complete the in-class assignment or paper review. We expect that approximately 120 individuals will complete the in-class assignment.

You are invited to participate in a study

You are invited to participate in a study assessing the validity of Bayesian Affect Control Theory (BayesACT) as a predictor of human behaviour. The study is being conducted by Josh as a Master's student in the Department of Computer Science under the supervision of Dr. Jesse Hoey.

Affect Control Theory posits that people strive to behave in the manner most in line with the expectations of their society. BayesACT extends this theory to allow it to deal with uncertainty. This study will provide data from humans playing a game, the iterated prisoner's dilemma, for the purposes of comparing human behaviour to the predictions of BayesACT.

We would like to use the results from the course assignment described above (the Prisoner's dilemma assignment) for our research.

You are under no obligation to provide your consent for the use of your assignment for our research. Further, a decision to participate or not will have no impact on your grade in ****. Professor Hoey will not know who consented to the use of their assignment in this research. Note that completion of the assignment does not imply consent to use your assignment, which you may choose not to give after reading this form. You will receive 5% credit for the assignment regardless of whether or not you give consent to be part of the research.

Information collected to draw for the prizes will not be linked to the data in any way, and this identifying information will be stored separately, then destroyed after the prizes have been provided. The amount received is taxable. It is your responsibility to report this amount for income tax purposes.

You may opt out of the study at any time by contacting Josh. Also note that the student IDs of consenting students will not be viewable by Jesse Hoey or any of the TAs associated with this course.

Personal Benefits of the Study

This study will help to determine the efficacy and generalizability of BayesACT. It will also produce initial conditions based on real data that can be used in future BayesACT projects. BayesACT has, for example, been used to create assisted living devices for patients with Alzheimer's disease, including hand-washing stations developed at the University of Waterloo.

Risks to Participation in the Study

There is some risk that you may feel coerced into consenting to the use of your assignment due to Jesse Hoey's dual roles as professor and researcher. However, we would like to assure you that he will never see a list of students who give/don't give their consent, and that he will not be involved in the drawing or distribution of Amazon gift cards.

Confidentiality

All information you provide is considered completely confidential; indeed, your name will not be included or in any other way associated, with the data collected in the study. Furthermore, because the interest of this study is in the average responses of the entire group of participants, you will not be identified individually in any way in any written reports of this research. The data, with identifying information removed, will be kept for a period of 10 years following publication of the research, after which it will be deleted. The data will be securely stored in the research laboratory of Dr. Jesse Hoey in the DC building to which only researchers associated with this study have access.

Questions and Research Ethics Clearance

If after receiving this letter, you have any questions about this study, or would like additional information to assist you in reaching a decision about participation, please feel free to ask the student investigator or faculty supervisor listed at the top of this sheet. Alternatively, you may contact *****, a senior PhD student in the Computational Health Informatics Lab, at ****@****, who is not directly affiliated with the study, but can provide additional information to assist you in reaching a decision about consent.

We would like to assure you that this study has been reviewed and received ethics clearance through a University of Waterloo Research Ethics Committee. However, the final decision about consent is yours. Should you have any comments or concerns resulting from your participation in this study, please contact *********, the Director, Office of Research Ethics, at *-***-***-****, Ext. ***** or ****@****.

Thank you for your interest in our research and for your assistance with this project.

Consent of Participant

By consenting to the use of your assignment below, you are not waiving your legal rights or releasing the investigator(s) or involved institution(s) from their legal and professional responsibilities.

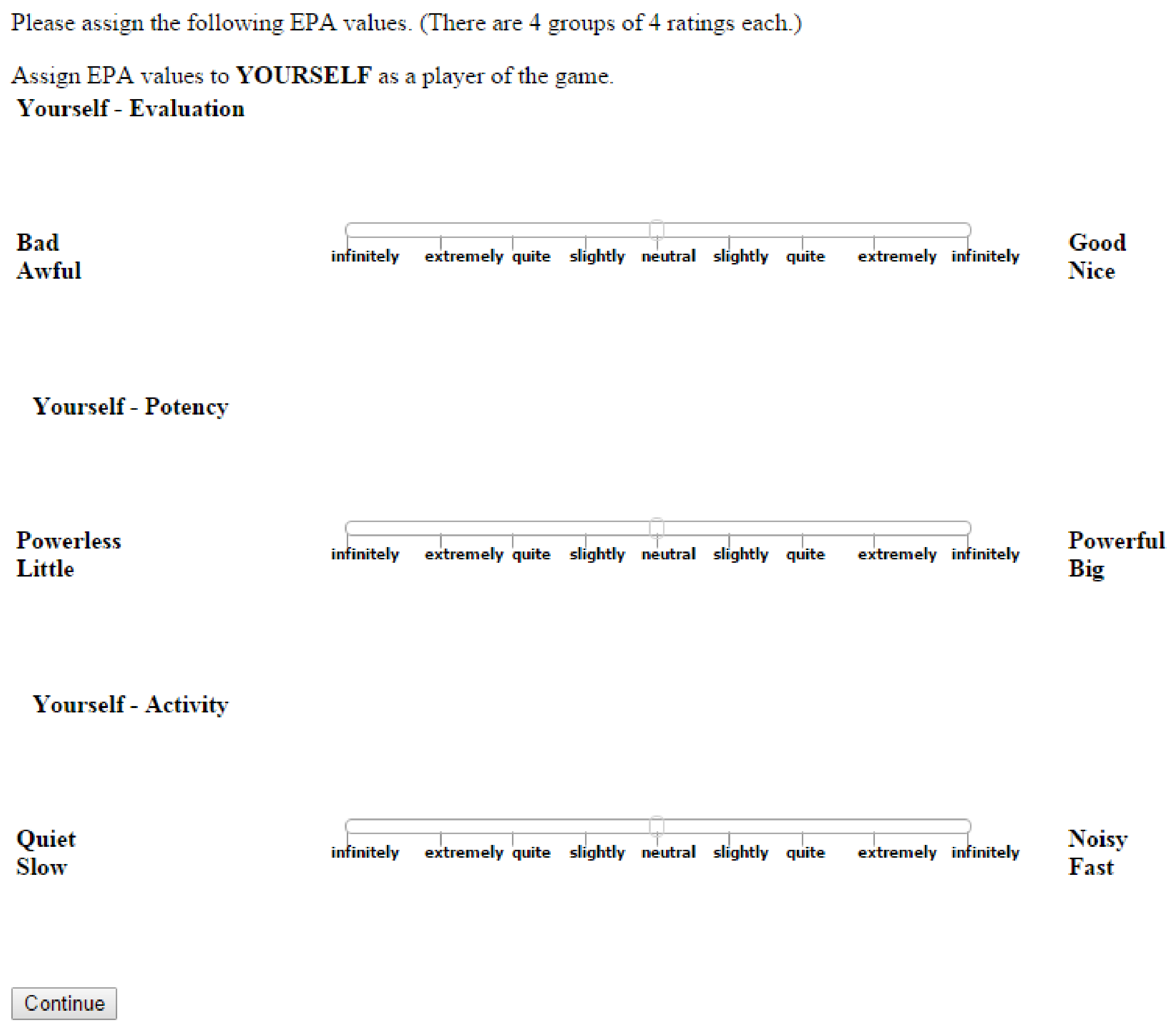
I have read the information presented in the information letter about a study being conducted by Josh Jung under the supervision of Dr. Jesse Hoey of the Department of Computer Science at the University of Waterloo. I have had the opportunity to ask any questions related to this study, to receive satisfactory answers to my questions, and any additional details I wanted. I am aware that I may withdraw consent for the use of my assignment from the study without loss of credit at any time by advising Josh of this decision.

This project has been reviewed by, and received ethics clearance through a University of Waterloo Research Ethics Committee. I was informed that if I have any comments or concerns resulting from my participation in this study, I may contact the Director, Office of Research Ethics, at *-***-***-****, Ext. ***** or ****@****.

Please select an option ▼

submit

6.4 ACT Sliders


Please assign the following EPA values. (There are 4 groups of 4 ratings each.)

Assign EPA values to **YOURSELF** as a player of the game.

Yourself - Evaluation

Bad **Awful**	infinitely	extremely	quite	slightly	neutral	slightly	quite	extremely	infinitely	**Good** **Nice**

Yourself - Potency

Powerless **Little**	infinitely	extremely	quite	slightly	neutral	slightly	quite	extremely	infinitely	**Powerful** **Big**

Yourself - Activity

Quiet **Slow**	infinitely	extremely	quite	slightly	neutral	slightly	quite	extremely	infinitely	**Noisy** **Fast**

Continue

Note that each of "Yourself", "Other Player", "Give", and "Take" had its own page of three sliders.

6.5 Game Interface

Start of Game

Reward Matrix
(your payoff/opponent's payoff)

	Opponent Gives 2 (G)	**Opponent Takes 1 (T)**
You Give 2 (G)	2/2	0/3
You Take 1 (T)	3/0	1/1

Your Choices:
Opponent's Choices:

Your Score: 0
Opponent's Score: 0

Give 2 | Take 1

Example End of Game

Reward Matrix
(your payoff/opponent's payoff)

	Opponent Gives 2 (G)	**Opponent Takes 1 (T)**
You Give 2 (G)	2/2	0/3
You Take 1 (T)	3/0	1/1

Your Choices: G G T G G G T T T T T T T T T
Opponent's Choices: G G G T G G G T T T T T T T T

Your Score: 22
Opponent's Score: 19

Your game is over. Please assign the following EPA values. (There are 4 groups of 4 ratings each.)

6.6 Debriefing

Debriefing Letter

Study Title: Applying BayesACT to the Iterated Prisoner's Dilemma

Faculty Supervisor: Dr. Jesse Hoey, Department of Computer Science, (***) ***-**** ext. *****, ****@****

Student Investigator: Josh Jung, Department of Computer Science, (***) ***-**** ext. *****, ****@****

Thank-you for completing this assignment. When you began the assignment, you were told that the purpose of this assignment was to observe human behavior in the iterated prisoner's dilemma game. However, the game was slightly more complicated than we explained at the beginning. Only 25% of the games you played were against your fellow classmates. The remainder were played in equal parts against three different artificially intelligent (AI) opponents (bots). One of the three bots was built using BayesACT and initialized with the ratings given by you over the course of the study. The other two played static strategies: one played tit-for-tat (i.e. always did exactly the same as what you did the last time you played), and the other cooperated three times and defected thereafter.

It was necessary to conceal this information to avoid spoiling the ratings you gave to yourselves and your opponents. We thought it likely that if you knew there was a high probability that your opponent was a bot, you would be unlikely to have a significant emotional response to the plays of your opponent. We apologize for omitting details about the tasks in this assignment. We hope that you understand the need for not informing you of this aspect of the assignment now this it has been more fully explained to you.

If you consented to the use of your assignment in our research please note that once all the data are collected and analyzed for this project, we plan on sharing this information with the research community through seminars, conferences, presentations, and journal articles. If you are interested in receiving more information regarding the results of this study, or would like a summary of the results, please email Josh Jung, and when the study is completed, anticipated by December, 2015, to send you the information.

The information you provided will be kept confidential by not associating your name with the responses. The data will be stored with all identifying or potentially identifying information removed. Electronic data will be stored 10 years on a password protected computer in DC 2584 then erased. No one other than the researchers will have access to the data.

This project was reviewed by, and received ethics clearance through a University of Waterloo Research Ethics Committee. Should you have any comments or concerns resulting from your participation in this study, please contact *********, the Director, Office of Research Ethics, at *-***-***-****, Ext. ***** or ****@****.

We really appreciate your participation, and hope that this has been an interesting experience for you.

Please enter your username and choose a consent option below.

Username

Consent Options Please select an option

Submit

References

[1] George A Akerlof and Rachel E Kranton. Economics and identity. *Quarterly journal of Economics*, pages 715–753, 2000.

[2] Damasio Antonio. Descartes' error: Emotion, reason, and the human brain. *GP Putnam'Sons, New York*, 1994.

[3] Nabiha Asghar and Jesse Hoey. Monte-Carlo planning for socially aligned agents using Bayesian affect control theory. In *Proc. Uncertainty in Artificial Intelligence (UAI)*, pages 72–81, 2015.

[4] KJ Astrom. Optimal control of markov processes with incomplete state information. *Journal of Mathematical Analysis and Applications*, 10(1):174–205, 1965.

[5] Robert Axelrod and William Donald Hamilton. The evolution of cooperation. *Science*, 211(4489):1390–1396, 1981.

[6] James A. Davis. Clustering and structural balance in graphs. *Human Relations*, 20(2):181–187, 1967.

[7] Celso M De Melo, Peter Carnevale, Stephen Read, Dimitrios Antos, and Jonathan Gratch. Bayesian model of the social effects of emotion in decision-making in multiagent systems. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, pages 55–62. International Foundation for Autonomous Agents and Multiagent Systems, 2012.

[8] John G Fennell and Roland J Baddeley. Reward is assessed in three dimensions that correspond to the semantic differential. *PloS one*, 8(2):e55588, 2013.

[9] Feng Fu, Corina E Tarnita, Nicholas A Christakis, Long Wang, David G Rand, and Martin A Nowak. Evolution of in-group favoritism. *Scientific reports*, 2, 2012.

[10] Erving Goffman. *Behavior in public places: Notes on the social organization of gatherings.* Free Press New York, 1963.

[11] Fritz Heider. Attitudes and cognitive organization. *The Journal of psychology*, 21(1):107–112, 1946.

[12] David R Heise. *Expressive order: Confirming sentiments in social actions.* Springer Science & Business Media, 2007.

[13] David R Heise. *Surveying cultures: Discovering shared conceptions and sentiments.* John Wiley & Sons, 2010.

[14] David R Heise. Modeling interactions in small groups. *Social Psychology Quarterly*, page 0190272512467654, 2013.

[15] Jesse Hoey, Craig Boutilier, Pascal Poupart, Patrick Olivier, Andrew Monk, and Alex Mihailidis. People, sensors, decisions: Customizable and adaptive technologies for assistance in healthcare. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 2(4):20, 2012.

[16] Jesse Hoey and Tobias Schröder. Bayesian affect control theory of self. In *AAAI*, pages 529–536. Citeseer, 2015.

[17] Jesse Hoey, Tobias Schröder, and Areej Alhothali. Affect control processes: Intelligent affective interaction using a partially observable markov decision process. *Artificial Intelligence*, 230:134–172, 2016.

[18] Eric Hogewoning, Joost Broekens, Jeroen Eggermont, and Ernst GP Bovenkamp. Strategies for affect-controlled action-selection in soar-rl. In *Nature Inspired Problem-Solving Methods in Knowledge Engineering*, pages 501–510. Springer, 2007.

[19] Daniel Kahneman. *Thinking, fast and slow.* Macmillan, 2011.

[20] Joseph LeDoux. *The emotional brain: The mysterious underpinnings of emotional life.* Simon and Schuster, 1998.

[21] Neil J MacKinnon and Dawn T Robinson. Back to the future: 25 years of research in affect control theory. *Advances in group processes*, 31:139Ā173, 2014.

[22] R Marinier and John E Laird. Emotion-driven reinforcement learning. *Cognitive science*, pages 115–120, 2008.

[23] Charles Egerton Osgood, William H May, and Murray S Miron. *Cross-cultural universals of affective meaning.* University of Illinois Press, 1975.

[24] David G Rand and Martin A Nowak. Human cooperation. *Trends in cognitive sciences*, 17(8):413–425, 2013.

[25] Wolfgang Scholl. The socio-emotional basis of human interaction and communication: How we construct our social world. *Social Science Information*, 52(1):3–33, 2013.

[26] Tobias Schröder, Jesse Hoey, and Kimberly B Rogers. Modeling dynamic identities and uncertainty in social interactions: Bayesian affect control theory. *Am. Soc. Rev*, 2015.

[27] Tobias Schröder and Wolfgang Scholl. Affective dynamics of leadership: An experimental test of affect control theory. *Social Psychology Quarterly*, 72(2):180–197, 2009.

[28] Flaminio Squazzoni, Wander Jager, and Bruce Edmonds. Social simulation in the social sciences a brief overview. *Social Science Computer Review*, 32(3):279–294, 2014.

[29] Daniele Vilone, José J Ramasco, Angel Sánchez, and Maxi San Miguel. Social imitation versus strategic choice, or consensus versus cooperation, in the networked prisoner's dilemma. *Physical Review E*, 90(2):022810, 2014.