Computational Affective Model of Human Action in a Social Dilemma

Joshua D. A. Jung and Jesse Hoey
David R. Cheriton School of Computer Science, University of Waterloo

keywords: Networked Prisoner's Dilemma; affect control theory; human decision making; emotion; modeling humans;

Affect Control Theory (ACT) is a powerful and general sociological model of human affective interaction [6]. It provides an empirically derived mathematical model of culturally shared sentiments as heuristic guides for human decision making [10]. BayesACT, a generalisation of ACT, combines affective reasoning with expected utility maximization (rationality) [8]. BayesACT allows for the creation of agents that are both emotionally guided and goal-directed. In this work, we simulate BayesACT agents in the Iterated Networked Prisoner's Dilemma (INPD), and we show that four out of five known properties of human play in the INPD [5] are replicated by these socio-affective agents. In contrast, previously used imitation-based agents are only able to replicate one of the five properties.

Cultural consensus on affective sentiments about words describing identities (e.g. *"doctor","mother"*) and behaviours (*e.g. "counsel", "comfort"*) has been measured through large-scale sociological surveys, and found to be reliably replicated across cultural groups and languages [10, 7]. Affective sentiments are measured across a large set of dimensions of meaning, but three significant factors are found to explain most of the variance observed. These factors are **E**valuation (roughly good vs. bad), **P**otency (roughly strong vs. weak), and **A**ctivity (roughly fast/loud vs. slow/quiet) [6]. These out-of-context sentiments are also found to combine in culturally consensual ways to form transient impressions when identities and behaviours are observed in social settings. That is, if E,P,A values are known for two interacting agent's identities, and for the behaviour performed, then culturally shared dynamics predict impressions that may differ from sentiments. The difference, called *deflection*, corresponds with the unexpectedness of the interaction. It is from this quantity that ACT draws its predictive power as a model of human behaviour, stating that people naturally act in a way that minimizes the deflection they create [6]. That is, a human's default action is that which aligns best with society's expectations. The dynamics equations in ACT, measured through cultural surveys, can therefore be used to predict behaviour. In BayesACT, a Monte-Carlo Tree Search (MCTS) method trades off this alignment bias with the maximization of expected utility. The MCTS is an anytime algorithm that smoothly shifts from societal expectations (affective alignment) to rational behaviour as a function of increased computational resources.

Grujić et al. [5] found five properties of human play in the Iterated Networked Prisoner's Dilemma. First, human play is invariant to network structure. Second, global cooperation rates decline over time, but remain non-zero. Third, cooperation is anti-correlated with reward. Fourth, most humans exhibit "moody conditional cooperative" behaviour, and fifth, human play is stratified into four major groups. We compared BayesACT agents (as defined in [1]) to standard imitative strategies [11] across a range of different network structures and payoff matrices.

For each test, 169 agents of one type (i.e. BayesACT or imitation) were arranged on a static network to play the Iterated Prisoner's Dilemma with their neighbours. These games each lasted for 60 individual rounds (or iterations), a number comparable to those of the largest human studies [5]. For each setting of our test parameters, 20 independent games were played, resulting in 3060 total simulations. Each round, agents chose between cooperation and defection and relayed that choice to each of their partners (network neighbours).

Testing was performed for three different network types (Grid, and Erdös-Rényi for two densities) and three different reward matrices. Additionally, each of the two agents tested had their own unique parameters. In the case of BayesACT, we chose to vary the initial EPA distribution between the original set as presented by [8] and one measured in a human study by [9]. We also applied several different timeouts (0, 1, and 10 seconds) to BayesACT's Monte Carlo search. We use the shorthand BACT[X][Y] to refer to a particular parameter setting, where X is one of D or S (default or study EPA settings) and Y is one of 0, 1, or 10 (timeout). For the imitation-based agents, we varied q, the probability of randomly selecting any neighbour instead of the highest scorer, from 0% to 100% in 10% intervals. A larger value of q therefore reduces the tendency of the network to settle, but introduces more erratic behaviour. We identify this via the shorthand IM[X], where $X \in [0, 100]$ is the value of q.

For all parameter settings of the BayesACT agents, we do not find evidence that network structure impacts agent behaviour. This is demonstrated by the consistently high p-values obtained when performing a G-test of cooperation rate per round across the 3 network types for a particular reward matrix. In particular, across all parameter settings, BACTD and BACTS agent cooperation rates were the same across network structures ($p > 0.05$) for 96.1% and 90.7% of rounds, respectively. For IM agents, only 5.6% of rounds show statistically similar rates, most gradually moving towards full defection (at different rates for different network settings).

In human studies, the global cooperation rate has been observed to drop from 55%-70% to less than 20%-40% after around 20 rounds of play, after which it remains approximately constant [5]. In general, we do not observe this behaviour in BayesACT agents, which produce relatively stable cooperation rates over time. The imitation-based agents tended to display one of two other non-human behaviours: either the cooperation rate decayed to zero, or it increased to some constant. Figure 1 shows an example of the cooperation rate over time for human players (taken from [3]) and for two simulations each of imitation-based and BayesACT agents. The slow decline but stabilisation of human players can be seen, while the IM simulations were selected to demonstrate the two behaviours described above. On the other hand, BayesACT simulations produce more stable cooperation rates, some of which are relatively close to those of humans (after stabilization).

We calculate the Pearson Correlation between the cooperation rates of individual agents and their scores to find the correlation between earnings and cooperation. We find that 100% of BACTD settings and 74.1% of BACTS settings display the desired anti-correlation ($p < 0.05$). The imitator agents showed cooperation-score anti-correlation in only 30.3% of settings, with the best setting, IM10, succeeding in 44.4% of matrix/network combinations.

Moody Conditional Cooperation has two requirements: hysteresis (i.e. an agent must be more likely to cooperate if it cooperated on the last turn) and conditionality (i.e. an agent must be more likely to cooperate if its neighbours were predominantly cooperators on the last turn) [4]. BACTD agents display a strong hysteresis (in 100% of test settings), while BACTS agents do not (only 4%). We believe that this is most likely a result of the larger difference between the EPAs of the cooperate and defect actions in the default set resulting in higher deflections and hence more severe reactions. BACTD0 displays conditionality in 44% of network/matrix combinations, while BACTS1 does so in 67% of them. All imitator agents have both strong hysteresis and conditionality.

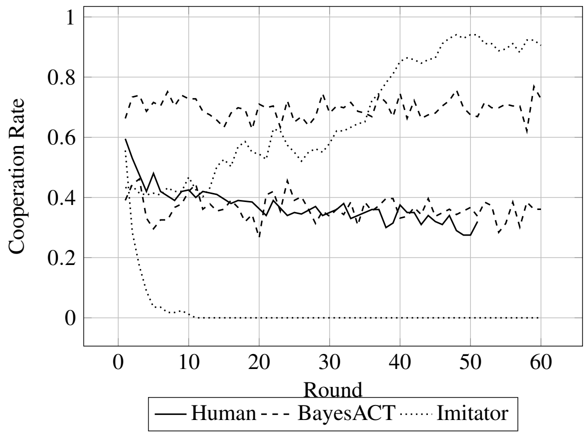

Figure 1: Cooperation rates for human players and for BayesACT (bottom: BACTS1/ER5, top: BACTS0/Grid) and imitator (bottom: IM50/Grid, top: IM100/Grid) agents.

Human players can be broadly classified into 5 groups [4]: those who only cooperate, those who mostly cooperate (at least two times in three), mixed cooperators, those who mostly defect (at least two times in three), and those who only defect. Averaged across all reward matrices and all network types, BACTD0 agents demonstrate this stratification in 100% of cases, while the best IM agent, IM100, does so in only 33% of network/matrix combinations.

We have shown that, compared to imitation-based agents, agents based on the social-psychological Affect Control Theory (BayesACT agents) display as emergent properties more of the human qualities identified by [5] in the Iterated Networked Prisoner's Dilemma (INPD). In particular, we show how the observed human behaviours of network structure invariance, anti-correlation of cooperation and reward, player type stratification, and (in $2/3$ of the cases we have considered) moody conditional cooperation (MCC), are all emergent properties of these agents, while imitation-based agents display only MCC. Our work moves a step closer to reproducing human behaviour in the INPD, and may find application both in domains that require human-like behaviour, and those that probe human reasoning. Our future work involves comparisons with additional agent models (e.g. [2]), and application to other networks.

[1] Nabiha Asghar and Jesse Hoey. Intelligent affect: Rational decision making for socially aligned agents. In *Proc. Conf. on Uncertainty in Artificial Intelligence*, pages 12–16, 2015.

[2] Ernst Fehr and Klaus M Schmidt. A theory of fairness, competition, and cooperation. *The quarterly journal of economics*, 114(3):817–868, 1999.

[3] Carlos Gracia-Lázaro, Alfredo Ferrer, Gonzalo Ruiz, Alfonso Tarancón, José A Cuesta, Angel Sánchez, and Yamir Moreno. Heterogeneous networks do not promote cooperation when humans play a prisoner's dilemma. *Proceedings of the National Academy of Sciences*, 109(32):12922–12926, 2012.

[4] Jelena Grujić, Constanza Fosco, Lourdes Araujo, José A Cuesta, and Angel Sánchez. Social experiments in the mesoscale: Humans playing a spatial prisoner's dilemma. *PloS one*, 5(11):e13749, 2010.

[5] Jelena Grujić, Carlos Gracia-Lázaro, Manfred Milinski, Dirk Semmann, Arne Traulsen, José A Cuesta, Yamir Moreno, and Angel Sánchez. A comparative analysis of spatial prisoner's dilemma experiments: Conditional cooperation and payoff irrelevance. *Scientific reports*, 4:4615, 2014.

[6] David R Heise. Social action as the control of affect. *Behavioral Science*, 22(3):163–177, 1977.

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

[8] 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.

[9] Joshua D A Jung, Jesse Hoey, Jonathan H Morgan, Tobias Schröder, and Ingo Wolf. Grounding social interaction with affective intelligence. In *Canadian Conf. on A.I.*, pages 52–57. Springer, 2016.

[10] Charles E Osgood. The nature and measurement of meaning. *Psychological bulletin*, 49(3):197, 1952.

[11] 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.