Bayesian Affect Control Theory in the Iterated Networked Prisoner's Dilemma

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

BayesACT, a generalisation of Affect Control Theory (ACT), combines affective reasoning with expected utility maximization (rationality) [4]. BayesACT allows for the creation of agents that are both emotionally guided and goal-directed. We have simulated BayesACT agents in the Iterated Networked Prisoner's Dilemma (INPD), and shown that four out of five known properties of human play in the INPD [3] are replicated by these socio-affective agents. In contrast, previously used imitation-based agents are only able to replicate one of the five properties.

The five properties described by Grujić et al. [3] are as follows. 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 [7] 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 [3]. 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 [4] and one measured in a human study by [6]. We also applied several different timeouts (0, 1, and 10 seconds) to BayesACT's Monte Carlo search. 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.

Ultimately, it was found that, compared to imitation-based agents, BayesACT agents displayed as emergent properties more of the human qualities identified by [3] in the INPD. In particular, we observed the 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), while imitation-based agents displayed only MCC. Full results may be found in [5]. 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.

References

[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] 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.

[4] 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.

[5] Joshua D A Jung and Jesse Hoey. Socio-affective agents as models of human behaviour in the networked prisoner's dilemma, 2017. arXiv:1701.09112.

[6] 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.

[7] 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.