Coordinating Human and Agent Behavior in Collective-Risk Scenarios

Elias Fernández Domingos,[1,2,3] **Juan Carlos Burguillo,**[3] **Ann Nowé,**[1] and **Tom Lenaerts**[1,2]

[1] AI lab, Computer Science Department, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium
[2] MLG, Département d'Informatique, Université Libre de Bruxelles,
Boulevard du Triomphe CP212, 1050 Brussels, Belgium
[3] Department of Telematic Engineering. University of Vigo. 36310-Vigo, Spain
Corresponding: eliferna@vub.ac.be and tlenaert@ulb.ac.be

Abstract

Various social situations entail a collective risk. A well-known example is climate change, wherein the risk of a future environmental disaster clashes with the immediate economic interest of developed and developing countries. The collective-risk game operationalizes this kind of situations. The decision process of the participants is determined by how good they are in evaluating the probability of future risk as well as their ability to anticipate the actions of the opponents. Anticipatory behavior contrasts with the reactive theories often used to analyze social dilemmas. Our initial work can already show that anticipative agents are a better model to human behavior than reactive ones. All the agents we studied used a recurrent neural network, however, only the ones that used it to predict future outcomes (anticipative agents) were able to account for changes in the context of games, a behavior also observed in experiments with humans. This extended abstract aims to explain how we wish to investigate anticipation within the context of the collective-risk game and the relevance these results may have for the field of hybrid socio-technical systems.

Many real situations, wherein humans interact among themselves or through technologies in hybrid socio-technical systems (Hamann et al. 2016), resemble social dilemmas. In particular, scenarios in which individuals have to decide between short-term personal profits and long-term social benefits are determined by the risk that this future collective gain entails. Therefore, the behavioral outcome in those dilemmas is very much dependent on how successful the participants are in calculating the risk associated to the uncertainty of future rewards and on anticipating the opponents' choices (Pezzulo et al. 2008).

Take for instance the question of climate change. Alleviating (or even reverting) this severe phenomenon requires the cooperation of several countries with different ideologies, customs, and economical perspectives for their industries, which are in many cases still very dependent on fossil energy. Measures that need to be taken will have a high impact on industrialized countries as well as the so-called new economies. However, if the transition to renewable sources of energy keeps being postponed, the consequences are most certainly terrible. As long as the risk of a climate or environmental disaster is perceived to be low, individuals, or countries in this case, will be more likely to act to maximize their welfare over that of the collective. Only in high-risk situations will they be persuaded to make sufficient investments to ensure that the disaster is avoided. This situation has been operationalized in game theory as the collective risk game (Milinski et al. 2008). In this game participants are each given an endowment and they need to decide whether to contribute or not, up to a predefined amount, to the common good for a fixed number of rounds. If the joint contributions of all the participants over those rounds is above a certain threshold, which is achieved when everyone gives half or more of the predefined amount, then the disaster is adverted and they receive as a reward the remainder of the endowment (hence the dilemma). Yet, when the target is not reached the disaster can occur (meaning that they lose the remainder of the endowment) with a certain probability, which is defined by a risk parameter. The experiments show that people only tend to contribute to avoid the disaster if they perceive the risk to be high (Milinski et al. 2008).

Peer-to-peer energy markets or cloud-computing architectures are technology related examples of such collective-risk scenarios, with less disastrous outcomes. In those situations, it would be a tragedy when the energy-market fails to achieve the benefits of cooperation, leading to its demise, or the loss/over-consumption of computing-resources leads to a period of unavailability for the many users of the system. In addition, in complex societies defined by autonomous multi-agent systems, where non-human agents need to deal with uncertain situations, similar problems can arise, as was discussed before (Artikis and Pitt 2001).

In order to guide both humans and agents in the previous examples towards an outcome beneficial for the collective, we need to have a clear understanding of the mechanisms that define individual behavior at the micro-level, and how these behaviors may aggregate into the dynamics one can observe at the macroscopic level. To achieve this goal, we can either focus on performing data analysis, when sufficient quantities of information are publicly available, or we can construct theoretical models that can be verified and improved through behavioral experiments. Data analysis in itself is limited in providing causal dynamic insights without having a model that can show that the knowledge extracted from the data can indeed lead to the observed macroscopic

behavior. Hence, we focus on the latter approach, linking it to the mathematical and computational modeling of strategic behaviors, i.e. linking game theory and artificial intelligence research.

Anticipative agents are able to account for changes in the context of games

In (Lalev and Grinberg 2006), the authors compared the behavior of backwards- and forward-looking agents in the iterated prisoner's dilemma (PD). Both types of agents used a recurrent neural network (RNN) (Elman 1990; Hausknecht and Stone 2015) to generate an online model of the opponent, the environment and the agent itself. Their results showed that, even when the aggregated behavior of these models were similar for individual games, only the forward-looking agents were able to account for changes in the payoff matrix of the game, a behavior also observed in experiments with human participants. Likewise, we performed simulations with similar forward- and backward-looking agents on the Anticipation Game (AG) (Zisis et al. 2015), a gift-giving game where players are divided into two groups: dictators and receivers. During the game, players of both groups are matched in a pairwise manner and the receivers are asked to decide whether or not to play the game with the dictator, based on her reputation. If the receiver does not accept, both players receive zero payoff. Otherwise, the dictator has to decide how much of a given endowment she wishes to share with the receiver. The amount shared becomes part of the reputation of the dictator, which will be used by the next matched receiver to make the decision of playing the game. Clearly, the dictator is aware of this and will adjust her behavior to maximize the chances of being accepted while keeping as much of the endowment as possible. This game describes a situation in which being able to anticipate the effects of ones actions on the future acceptance is key for the dictators' success. The difference with the previous PD work is that the strategy-space is larger, as the agents have more actions from which they can choose, and their effects don't reflect on their payoff immediately. Nevertheless, again, only the anticipative neural network-based agents were able to adapt their behavior to changes in the context of the game, something that was also observed in the behavioral experiments performed in (Zisis et al. 2015). These results are available in (Fernández Domingos, Burguillo, and Lenaerts 2017).

Discussion and future work

Notwithstanding the importance of anticipatory behavior for human decision-making and the potential it has for the development of hybrid socio-technological systems, little attention has been given to the development of its theories, systems and applications recently, an issue that we wish to address with the progress of our research.

Our main hypothesis resides within the assumption that human behavior in collective risk scenarios is highly determined by their capacity to anticipate future outcomes and the effect of their actions, which we can replicate and analyze through the development of forward-looking behavioral models. Humans behave differently depending on the level of risk they face (Milinski et al. 2008), which requires us to understand also how humans perceive and forecast risk, and the effect of their actions on their opponents and partners. Up to now, we have ascertained that anticipative agents are a better model to human behavior in some specific scenarios, however we want to extend this conclusion to the collective risk game.

In order to further advance our research, we plan to perform new behavioral experiments that will be focused on understanding how humans anticipate and behave under risk scenarios. Moreover, we plan to test the application of the state of the art in machine learning to identify, extract and classify different behaviors that may be present in the data we obtain. Additionally, we will progressively adapt and study our models in more complex scenarios and, eventually, perform large-scale multi-agent simulations.

References

Artikis, A., and Pitt, J. 2001. A formal model of open agent societies. In *Proceedings of the fifth international conference on Autonomous agents*, 192–193. ACM.

Elman, J. L. 1990. Finding structure in time. *Cognitive Science* 14(2):179–211.

Fernández Domingos, E.; Burguillo, J. C.; and Lenaerts, T. 2017. Reactive versus anticipative decision making in a novel gift-giving game.' In *Proceedings of the thirty-first AAAI conference*, (to appear). AAAI.

Hamann, H.; Khaluf, Y.; Botev, J.; Divband Soorati, M.; Ferrante, E.; Kosak, O.; Montanier, J.-M.; Mostaghim, S.; Redpath, R.; Timmis, J.; Veenstra, F.; Wahby, M.; and Zamuda, A. 2016. Hybrid Societies: Challenges and Perspectives in the Design of Collective Behavior in Self-organizing Systems. *Frontiers in Robotics and AI* 3(April).

Hausknecht, M., and Stone, P. 2015. Deep Recurrent Q-Learning for Partially Observable MDPs. *arXiv preprint arXiv:1507.06527*.

Lalev, E., and Grinberg, M. 2006. Backward vs. forward-oriented decision making in the iterated prisoner?s dilemma: A comparison between two connectionist models. In *Workshop on Anticipatory Behavior in Adaptive Learning Systems*, 345–364. Springer.

Milinski, M.; Sommerfeld, R. D.; Krambeck, H.-J.; Reed, F. a.; and Marotzke, J. 2008. The collective-risk social dilemma and the prevention of simulated dangerous climate change. *Proceedings of the National Academy of Sciences of the United States of America* 105(7):2291–2294.

Pezzulo, G.; Butz, M.; Castelfranchi, C.; and Falcone, R., eds. 2008. *The Challenge of Anticipation*. Springer.

Zisis, I.; Di Guida, S.; Han, T.; Kirchsteiger, G.; and Lenaerts, T. 2015. Generosity motivated by acceptance-evolutionary analysis of an anticipation game. *Scientific reports* 5:18076.