Interference Detection amongst Interdependent Human-Machine Teams [Extended Abstract]

Eugene Santos Jr.[1], Hien Nguyen[2], Keum Joo Kim[1], Gregory Hyde[1], Clement Nyanhongo[1], Luke Veenhuis[2]

[1] Thayer Engineering School, Dartmouth College, Hanover, NH
[2] Dept. of Computer Science, UW-Whitewater, Whitewater, WI

Abstract

The mutual relationship between humans and machines has significantly changed due to the remarkable strides in the development of machine intelligence. Machines are being designed to solve challenging tasks that require problem comprehension, logical reasoning, and in some cases, domain expertise. In many scenarios, machines now act as human coworkers, rather than static, stagnant tools or objects (Johnson et al. 2011). To understand the evolving nature of human-machine cooperation, several studies have been conducted which cover aspects such as machine autonomy, degrees of freedom of machines, and cooperative planning (Barbosa et al. 2017; Bellet et al. 2011; Johnson et al. 2014). Most of these studies, however, are conceptual and qualitative, and there is a need to develop empirical and computational systems for capturing interdependence in complex human-machine teams.

Interdependence, one of the fundamental units of cooperation, is defined as the process by which interacting agents influence each other's experiences (Van Lange and Balliet 2014). This influence can either reinforce (positive interference) or threaten (negative interference) the intent or goals of the other agents in a team (Hoc 2001). For our work, we present a computational framework capable of detecting both positive and negative interference among cooperative agents. A computational understanding of interference allows us to empirically assess the fundamental requirements for building human-machine teams. Additionally, we can determine how machines should be designed and also evaluate how humans could be trained to understand and co-adapt with autonomous machines. This formalizes the overall effectiveness of human-machine teams.

Related to our work is an experiment by Johnson et al. (2012) in which human participants interacted with machine agents with varying degrees of autonomy (high to low). Interdependence was qualitatively assessed by each of the human participants through subjective rankings collected from surveys designed to determine interdependence factors such as team burden and performance. In a similar experiment conducted by Pearsall et al. (2010), teams of humans operated in an environment incentivized by either team-based rewards or individual-based rewards. Again, interdependence was evaluated by measuring human rankings and other game metrics. The benefit of our work is that our models are entirely computational and needs no reflection from the human to analyze. We believe this advance to be a strong characteristic of a system capable of detecting interference. With it, scenarios of negative interference can be caught and addressed in stride, rather than by self-reflection.

We use the Starcraft II Learning Environment (SC2LE) (Vinyals et al. 2017) as our testbed. The SC2LE environment is a real-time strategy game that encourages players to make complex decisions as they compete against their opponents (human or computer-generated). It provides players with the flexibility to tailor custom scenarios for collaborative team play under different environmental constraints. The testbed is well-suited for studying cooperation under high pressure, dynamic, multi-agent, and time-sensitive environments that simulate real-world battlefield scenarios.

To capture inference in a team of agents, we must first understand the intent of each individual agents. For a given cooperative task, we can then determine if the agents' intent aligned or deviated while they were engaging in the task. To model an agent's intent, we employ the Double Transition Model (DTM) (Yu 2013). The DTM is a graph structure that captures the decision-making processes of agents in a given domain. Its nodes represent the agent's cognitive states, and its edges represent the actions or decisions that caused transitions to other cognitive states. Each node is an instance of the cross product of two subgraphs, a Query Transition Graph (QTG) and a Memory Transition Graph (MTG), at a particular point in time. Each node in the QTG represents an agent's thought processes and queries at a given instant. The MTG represents an agent's knowledge base, which highlights its environmental perceptions, observations, and contextual representations from past cognitive states. The edges of the DTM are embedded with reward values, which reflect an agent's intent when transitioning from one state to another. These reward values are computed through Inverse Reinforcement Learning (IRL) on a dynamic Markov Decision Process (MDP), which is derived from the DTM. An MDP is defined as a 5-tuple (S, A, R, P, γ) where S represents the states in the environment, A represents the actions that an agent can take, R represents the reward function, P represents the probability transition function between states and γ represents the discount factor for prioritizing on future rewards over immediate rewards. A sequence of transitions from one state to another forms a trajectory, ϑ, of the form:

$$\vartheta = \{s_i, a_i, s_{i+1}, a_{i+1}, \dots, s_{i+k}\},$$

where $s \in S$, and $a \in A$. The derived MDP is dynamic since it is constantly updated as the agent navigates its environment. IRL was first proposed by Ng and Russell (2000) as a means to infer an agent's preferences or intent in the form of a reward function, which is computed from the agent's past behavior. For our experiments, we examine these rewards to evaluate decisions that the cooperating agents make as they interact to solve the given tasks.

For the SC2LE environment, we create scenarios where the human and the machine agents (same team) act as commanders who try to build and control units (and structures) to strategically defeat a computer-generated opponent (i.e. AI enemy). Using DTMs, we investigate interference during task allocation and execution. We construct two levels of DTMs, one for the team, and the other for the individual agents. The team-level DTM is based on the combined human-machine team, whilst the individual-level DTM covers each of the agents individually. This representation is very much akin to the given situation versus the effective situation discussed by Kelley and Thibaut (1978). All agents operate on a common frame of reference. Cooperation is dependent on the agents' ability to elaborate this frame of reference by concerning themselves with common goals, common plans, role allocations, action monitoring and evaluation, and common representations of the environment (Pacaux et al. 2011). To ensure these tasks are lined up, we present the single common goal of defeating the adversary AI to the human and their cooperative machine agents.

Let a team, T, be composed of $\{H, M\}$, where H represents a human, and M represents a set of machines. For simplicity, we only focus on a team composed of one human and one machine, but our analysis can be extended to teams with more than two human-machine agents. Based on the attribute hierarchy of SC2LE, we identify features relevant to each human. With each selected group of data, we can form DTMs representing the team, DTM_T; the human, DTM_H; and the machine, DTM_M. DTMs allow us to compare agents by looking at the graph structures of their QTGs and MTGs, and through transition rewards on the DTM's edges computed after IRL. From these rewards, we can compute the expected rewards for a series of sequential decisions (policies) that were made by the agents. This computation is integral for analyzing interdependence, which can be assumed to have a temporal element, as suggested by Van Lange and Balliet (2014). Specifically, these authors argue that temporally extended scenarios afford the expression of self-control and the resistance of instant gratification pertaining to the accomplishment of goals. Therefore, any piece of analysis that focuses only on single, time-step comparisons fails to gain a comprehensive picture of interdependence captured in sequential decision-making. Expected rewards (in nature) are a look into the horizon of an agent. It is an estimate of what is to come, and yields details about an agent's behavior well into the future. Reward expectation over an entire trajectory (or subsection of a trajectory) is evaluated using the following equation:

$$E(\vartheta) = \sum_{j=i}^{i+k-1} P_{a_j}(s_j, s_{j+1})\, R_{a_j}(s_j, s_{j+1})\gamma^{j-i}$$

This analysis on an agent's horizon allows us to make comparisons over the varying DTMs to assess interference. For example, if the expected long-term pattern (its policies) of DTM_H are not in line with DTM_M and/or DTM_T (or vise versa), we classify this as an example of negative interference. If they are inline, however, we can consider this as a form of positive interference.

We conduct two experiments for our analysis of human-machine cooperation. In the first one, the human agent has no control over how its machine teammate behaves but adapts to its behavior over time through multiple gameplays. In the second one, the human has control of the machine's behavior and can communicate with it to perform in a particular fashion (aggressive, defensive, etc.) For the first experiment, machine agents (rule-based) are built to suit different styles, and games are collected such that we have the following sets:

(individual collected games)
h = {human player}
d = {defense oriented machine}
a = {attack oriented machine}
m = {a machine that equally prioritizes attacking and defending}
i = {random, irrational machine}

(team collected games)
dh = {defensive machine and human team}
ah = {attack oriented machine and human team}
mh = {mixed strategy machine and human team}
ih = {irrational machine and human team}

For this experiment, we analyze how interference occurs when a human works with a machine but has no control over it. We are interested in seeing how the human adapts its behavior to satisfy its (and the machine's) goals. We are also interested in examining how cooperation occurs when the human cannot verbally (or through text) communicate with the machine but must use its intuition to determine how it should work with the machine.

For the second experiment, the human has partial control over the machine. The human can choose between various styles of play for the cooperative machine in the form of textual cues (i.e., "attack" makes the cooperative machine agent more aggressive, whereas "defend" defaults to more passive play). To ensure negative interference is present within the experiment, we introduce a forced deviation (at random instances) of the machine to the human's requests (i.e., the machine may act aggressively after the human tells it to play passive). This partial control affords us a meaningful ground truth to determine the negative interference for the given scenario.

To conclude, we present a quantifiable way to measure interference amongst agents in human-machine teams. This work is necessary to infer the complex relationships that are present in human-machine teams, which enables us to identify the strengths and shortfalls of these teams. These insights are invaluable for developing human-machine systems that are resilient to possible miscommunications and to thwarting unpredicted behavior amongst agents.

Acknowledgements

This work supported in part by Office of Naval Research Grant No. N00014-19-1-2211 and Air Force Office of Scientific Research Grant No. FA9550-20-1-0032.

References

Barbosa, I., Oliveira, M., Reis, P., Gomes, T., & da Silva, F. Towards Understanding the Relationships between Interdependence and Trust in Software Development: A Qualitative Research. *IEEE/ACM 10th International Workshop on Cooperative and Human Aspects of Software Engineering*, pages 66-69, 2017.

Bellet, T., Hoc, J., Boverie, S., & Boy, G. From Human-Machine Interaction to Cooperation: Towards the Integrated Copilot. In From *Human-Computer Interactions in Transport*, pages 129-156, 2011.

Johnson, M., Bradshaw, J., Feltovich, P. J., Hoffman, R., Jonker, C., Riemsdijk, B., & Sierhuis, M. Beyond Cooperative Robotics: The Central Role of Interdependence in Coadive Design. In *IEEE Intelligent Systems*, pages 81-88, 2011.

Johnson, M., Bradshaw, J., Feltovich, P. J., Jonker, C., van Riemsdijk, B., & Sierhuis, M. Autonomy and Interdependence in Human-agent-robot Teams. In *IEEE Intelligent Systems*, pages 43-51, 2012.

Johnson, M., Bradshaw, J., Hoffman, R., Feltovich, P. J., & Woods, D. Seven Cardinal Virtues of Human-Machine Teamwork: Examples from the DARPA Robotic Challenge. *IEEE Intelligent Systems*, 29, pages 74-80, 2014.

Hoc, J. Towards a Cognitive Approach to Human-machine Cooperation in Dynamic Situations. In *Human-Computer Studies*, 54, pages 509-540, 2001.

Kelley, H. H., & Thibaut, J. W. Interpersonal Relations: A Theory of Interdependence. *New York, NY: Wiley*, 1978.

Ng, A. Y., & Russell, S. Algorithms for Inverse Reinforcement Learning. In *ICML*, 2000.

Pacaux, M. P., Debernard, S., Godin, A., Rajaonah, B. Anceaux, F. & Vanderhaegen, F. Levels of Automation and Human-machine Cooperation: Application to Human-robot Interaction. In *IFAC*, 2011.

Pearsall, M. J., Christian, M. S. & Ellis, A. P. J., Motivating Interdependent Teams: Individual Rewards, Shared Rewards, or Something in Between? In *Journal of Applied Psychology*, 95, pages 183-191, 2010.

Van Lange, P. A. M., & Balliet, D. Interdependence Theory. In *APA Handbook of Personality and Social Psychology*, 3, 2014.

Vinyals, O., Ewalds, T., Bartunov, S., Georgiev, P., Sasha, A., Yeo, M, Tsing, R. Starcraft II: A New Challenge for Reinforcement Learning. arXiv preprint arXiv:1708.04782, 2017.

Yu, F. A Framework of Computational Opinions. Dartmouth College, 2013.