Quantifying Cooperation Between Rule-Based Hanabi Agents Using Information Theory

Patricia WOLLSTADTa,1, Matti KRÜGER a, and Christiane B. WIEBEL-HERBOTH a

aHonda Research Institute Europe, Offenbach am Main, Germany

Abstract. With increasing abilities of artificial systems and agents, it has been proposed to design those systems to behave cooperatively towards human users. However, it is often unclear how to evaluate whether a cooperation in human-machine interaction takes place or how to even quantitatively describe such a cooperation. We have previously proposed a framework to quantify the degree of cooperation in an interaction between two agents, using novel measures from information theory. We here extend the initial evaluation of this framework by applying the proposed measure to rule-based artificial agents playing Hanabi. We show that the measure correlates with the number of points scored by the agents and that the framework allows to describe uni-directional interactions between stronger and weaker players. The proposed framework may be used to evaluate and guide HMI design towards more cooperative interactions, which is believed to lead to a more pleasant user experience.

Keywords. human-machine interaction, cooperation, information theory, partial information decomposition, Hanabi

Introduction Today, many intelligent systems are developed towards operating in interaction with human users instead of acting autonomously. It has been proposed to enable these systems to act cooperatively [17,14,2,3], as cooperation is thought to be beneficial with respect to task success, trust, user satisfaction, and counters negative effects of pure automation [7,17,3,8,6]. However, cooperation in this setting is not dichotomous—rather, cooperation is a continuous phenomenon such that different degrees of cooperative behavior have to be distinguished [5]. Today, it is still difficult to benchmark and evaluate cooperative behavior in human-machine interaction (HMI) [2,16,15,12]. To successfully design cooperative HMI, it is therefore desirable to develop metrics that reliably identify different qualities of interactions and in particular, identify different degrees of cooperation. Ultimately, such a measure may also be used to guide agent behavior, as it has been shown that optimizing e.g., for task success, may not always lead to the most satisfying user experience [18]. We have previously proposed a measure for cooperation and have provided a first, successful evaluation of the measure in a toy system [21,22]. Here, we extend previous results by applying the proposed measure in the popular cooperation benchmark Hanabi and demonstrate its ability to differentiate between degrees of cooperative behavior in different agent strategies.

1Corresponding Author: Patricia Wollstadt, Honda Research Institute Europe, Carl-Legien-Str. 30, 63073 Offenbach am Main, Germany; E-mail: patricia.wollstadt@honda-ri.de.
Shared and unique contributions

Synergistic contribution

Total contribution of joint actions

I(A0, A1 ; T)

Goals

Cooperation

Sensors

Effectors

Internal model

A1

A2

Ishd(A0, A1 : T)

Iunq(A1 : T)Iunq(A0 : T)

Isyn(A0, A1 : T)

BC

0.0

0.1

0.2

VDB Int

0.15

0.20

0.25

0.30

10 15 20

score

Self-play

10 15 20

score

Other-play

D

r=0.9045* r=0.9869**

Figure 1. A) Illustration of PID framework. B-C) Synergy versus score for agent self-play and the Piers agent playing other strategies. D) Synergy towards individual task success for VDB versus Internal agent.

Methods  In prior work [21,22], we have proposed an information-theoretic measure of cooperative behavior using the partial information decomposition (PID) framework [20]. PID allows to decompose the contribution of two or more input variables towards a target variable into unique, redundantly shared and synergistic contributions. Here the synergy describes a contribution that is exclusively provided by both inputs together and that can not be obtained from one input alone (Figure 1A). We proposed the synergy as the operationalization for a common notion of cooperation in HMI, namely the “facilitation” of agents’ actions towards a goal [13,9]. We successfully demonstrated our approach in a toy system. We here extend this evaluation in the popular benchmark for cooperative interaction, Hanabi [1]. We used an existing implementation of the Hanabi learning environment and five rule-based agents presented in [4] (originally proposed in [19]). To estimate the synergy between agents’ actions and the score in each round, we use a measure of the synergy proposed by Ince [10] which is implemented in the dit Python toolbox [11]. We estimated the synergy between the first agent’s action, A0, the subsequent action by the second action, A1, and whether this second action resulted in an error, no point, or a point. We distinguished between agents’ actions Reveal (either color or number), Discard, and Play. The investigated agents were the Internal, Piers, IGGI, Outer, and VDB (see [19] and [4] for a detailed description).

Experiments and Results  We simulated 1000 episodes of Hanabi in different experimental settings. First, we evaluated the synergy as a marker of cooperative behavior in agent self-play, where we found a strong correlation between the synergy and the average score per episode (Fig. 1B). Second, we estimated the synergy for games in which the Piers agent (the most successful agent) played all other agents, and again found a strong correlation with the average score (Fig. 1C). Last we investigated whether the synergy was able to reflect also asymmetric interactions between a well-performing and a poorly performing agent. Here, we estimated the synergy between the VDB and Internal agent (Fig. 1D), where the former achieved higher scores in all games compared to the latter. We estimated the synergy once between VDB’s and Internal’s actions with respect to scores assigned after VDB’s action and once with respect to scores assigned after Internal’s actions. We found that the synergy towards Internal’s actions was higher than towards VDB’s actions, indicating a stronger one-sided or asymmetric cooperation towards the weaker Internal agent.

Conclusion  We have recently introduced a novel framework for quantifying cooperative behavior using methods from information theory. Here we extend previous findings by providing evidence for its ability to differentiate between different cooperative strategies in a famous benchmark of cooperative behavior for artificial intelligence [1].
We show that the synergy measure correlates closely with points scored during a game, which is a direct outcome of cooperative game plays. Additionally, we show that the synergy allows insights into uni-directional interactions.

While it could be argued that in this particular case there are little additional insights from estimating the synergy, in other contexts, the relationship between cooperative behavior and task success may be less clear; i.e., cooperation may not necessarily lead to the highest task success or may not be necessary for solving a task. Nevertheless, we may be interested in evaluating whether cooperative behavior takes place for other reasons than optimal performance, e.g., user satisfaction [18]. Here, our proposed measure fills a gap as it allows to make a statement about the quality of the interaction, independent of overall task success, and thus may be an alternative guidance for designing HMI systems towards a more pleasant user-experience. An evaluation of our approach in human-machine interaction will be subject to future work.

References


