

An Intentional AI for Hanabi

Markus Eger

Principles of Expressive Machines Lab
Department of Computer Science
North Carolina State University
Raleigh, NC
Email: meger@ncsu.edu

Chris Martens

Principles of Expressive Machines Lab
Department of Computer Science
North Carolina State University
Raleigh, NC
Email: crmarten@ncsu.edu

Marcela Alfaro Córdoba

Department of Statistics
North Carolina State University
Raleigh, NC
Email: malfaro@ncsu.edu

Abstract—Cooperative games with partial observability are a challenging domain for AI research, especially when the AI should cooperate with a human player. In this paper we investigate one such game, the award-winning card game Hanabi, which has been studied by other researchers before. We present an agent designed to play better with a human cooperator than these previous results by basing it on communication theory and psychology research. To demonstrate that our agent performs better with a human cooperator we ran an experiment in which 224 participants played one or more games of Hanabi with different AIs, and will show that our AI scores higher than previously published work in such a setting.

I. INTRODUCTION

When explaining the behavior of a complex system, humans often ascribe intentionality to the system under observation [1], where intentionality means goal-directedness. This view is justified because its only assumption is rationality. In other words, if we have an agent or AI system that behaves (mostly) rationally, humans will try to view it as working towards some goal. This intentional view presents both a challenge and an opportunity for the development of cooperative game AIs. On one hand, human players expect the AI to behave intentionally, which imposes a restriction on its design, but on the other hand, intentionality can also be actively used to make the AI more believable or even guide the player. In this paper, we present an AI for the award-winning [2] cooperative card game Hanabi [3], that uses a simplified model of intentionality. Hanabi is a unique game because players can only see the cards in their cooperators' hands, but not their own, and there are strict limits on what kind of communication can occur between players. Our approach uses these restrictions to convey intentions to the human cooperator in a natural way, based on communication theory. In particular, we utilize the work of H.P. Grice [4]. In his seminal work on logic and conversation he described four maxims that humans follow in conversation. The main inspiration for our work comes from two of these maxims: The maxim of relation, that states that communication should be relevant to the topic under discussion, and the maxim of manner, that states that communication should not be ambiguous or obscure. We will show that using these maxims to convey intentions, our AI compares favorably to a previously published AI when playing with a human cooperator.



Fig. 1. A typical Hanabi board during play (Source: BoardAgain Games)

The main contribution of this paper is twofold: First, we present how intentionality and communication theory can be integrated in agent design to produce an agent capable of playing the game of Hanabi reasonably well. Second, we also demonstrate that this has a direct effect on agent performance when playing with a human contributor. Since our AI and the baseline AI share the same basic design, this provides further evidence that intentional behavior in agents is a desirable property if they are to interact with humans.

II. BACKGROUND: TWO-PLAYER HANABI GAME RULES

Hanabi is a cooperative card game, with cards in five colors, and the ranks 1 to 5. Each player is dealt five cards, which they hold facing the other player, i.e. players only see the cards in the other player's hand. The goal of the game is to build fireworks of each color, as represented by a stack of cards in ascending order. Figure 1 shows a typical game state, with three partially built fireworks in red, blue, and green. On a player's turn, they have to choose one of three options:

- Play a card from their hand. If the card is the next in order for the stack corresponding to its color on the board (or a 1, if the stack is empty), it is placed on top of that stack. Otherwise it is put in the discard pile and the number of mistakes is increased by one.
- Give a hint to the other player. Hints can consist of telling the other player all cards of either a particular rank or a particular color that they have. For example, a player may tell the other player where all their red cards are. Giving

hints comes at the expense of hint tokens, of which there are initially 8.

- Discard a card to the discard pile. This regenerates one hint token.

After a play or discard action, the player draws a card from the deck to bring their hand size back up to 5. Play proceeds until either 3 mistakes have been made, or all cards from the deck have been drawn, after which every player gets one more turn. The score the players achieved is equal to the number of cards they successfully played, for a maximum score of 25.

Note that there are three of the 1s in each color, two of the 2s, 3s and 4s of each color, but only one of the 5s per color. This means that discarding a 5 will decrease the maximum obtainable score by 1, since it can not be recovered. More severely, discarding both 2s of a color prevents any higher rank in that color from being played, since the cards have to be played in order, reducing the maximum possible score by 4. We want to note that while there is no actual research on the average score achieved in a typical Hanabi game, from the authors' experience humans should expect to score between 15 and 20 points in their first game with an unfamiliar play partner, with a score over 20 possible if both players have significant board game experience. Players more familiar with the game routinely score 20 or more points, though ¹.

For the remainder of this paper, we will use the following terms:

- A card can be of one of the five *colors* red, green, blue, yellow, and white, and can have a *rank* of 1 to 5
- The cards in the AI player's hand will be referred to as A with the individual cards indexed as A_1 to A_5
- The cards in the human player's hand will be referred to as B with the individual cards indexed as B_1 to B_5
- A card's *identity* refers to the pair (c, n) of the card's color c and rank n . Note that a player may consider several identities possible for a single card in their hand
- A card is called *playable* if it is the card that is to be played next on its color's stack
- A card is called *useless* if it is not and will never be playable
- A card is called *expendable* if there is still a duplicate in the deck or a player's hand, i.e. discarding the card does not necessarily decrease the maximum possible score
- A hint is said to *positively identify* a card when the card matches the information in the hint
- A hint is said to *negatively identify* a card when the card does not match the information in the hint

The meaning of *positive identification* is that the card was "pointed at" by the hint, i.e. when hinting a player about all green cards in their hand, the green cards are the ones that are positively identified, and the non-green cards are negatively identified by the hint.

¹The official rules state that ending the game with 3 mistakes scores no points and is to be interpreted as a loss for every player. The AIs presented in the literature so far ignored this rule and reported the score as it was after the 3 mistakes, and for the sake of comparison we will do the same.

III. RELATED WORK

Our work draws from previous research in AI and human perception of agents, and from previous results on Hanabi.

A. *Intentional Agents*

As has already been noted, a default-expectation of humans is that an agent behaves intentionally [1]. Additionally, humans also have a model of other agents' mental models and their desires and how these lead to intentions [5]. Previous work has utilized this in the design of AI agents. For example, Pynadath et al. [6] use beliefs about the beliefs, desires and intentions of agents, a theory of mind, to model social interactions in a multi-agent system. In narrative generation, Riedl and Young use a model of intentions to generate stories with more believable agents [7]. Consequently, it has been argued that AIs that behave in an intentional, goal-directed manner and do not cheat is also desirable for the use in video games [8]. Our model of intentionality is based on Cohen and Levesque's [9]. They describe intention as an agent adopting a goal and committing to it by forming a plan to achieve it.

As far as communication is concerned, Young has argued that Grice's maxims should be applied to the design of video games and other digital entertainment to make for a better experience for the players [10]. Although his argument is for the design of narratives, we would argue that it could similarly be made for communication embedded in game mechanics. Young has also operationalized one of Grice's maxims, the maxim of quantity, to generate short descriptions of plans [11].

B. *Hanabi*

Hanabi has several properties that make it an interesting research subject. It is cooperative, the game state is only partially observable by the agents with different parts visible to the different agents, and the means with which information can be communicated are strictly regulated by the rules. Furthermore, the score at the end of the game is a straightforward way to measure and compare performance between agents. Williams et al. [12] include it in their ongoing research that aims to catalog cooperative partially observable games. It is therefore unsurprising that it has been the target of previous work. Perhaps surprisingly, Baffier et al. showed that the (generalized) game is NP-hard even with perfect information [13]. Mark van der Bergh wrote his Bachelor thesis about Hanabi, showing how many possible initial configurations (shuffles) of the deck can be won for a reduced version of the game, and investigating several strategies utilizing a variety fixed rules that score up to 13.1 points on average, depending on which configuration of rules is used [14]. The highest-scoring AI strategy for Hanabi in the literature comes from Cox et al. who view it as a hat-guessing game [15]. Assuming a 5 player game, their AI works by assigning a numerical value between 0 and 7 to each visible hand that describes what the player holding that hand should do (e.g. play the leftmost card), and then summing up those values and taking the sum modulo 8. A player giving a hint then uses a similar encoding to "say" the modulus that they obtained in this way,

from which *every* other player can calculate the value the hint-giving player assigned to their hand (because they can see the other players’ hands). In other words, by using a clever encoding, giving one hint actually conveys information to all four other players. While their AI scores over 24 points on average, it heavily relies on the specific encoding which is not very friendly for the use by human players. Additionally, one of the main benefits is being able to convey information to multiple players at once, which no longer exists in a two-player game. The main inspiration for our AI comes from the more human-friendly AIs described by Osawa [16]. He describes a whole set of AIs ranging from a completely random one to various versions of an AI that carefully considers which hints to give and how to interpret hints that it is given. We will mainly use what Osawa called the Outer State strategy as a baseline to compare our AI to, and describe how our AI fits into his framework in the next section.

IV. INTENTIONAL AI

Our AI follows the same general outline as the Outer State strategy presented by Osawa [16]:

- 1) If the player has a card that they know is playable, play that card.
- 2) If the player has a card that they know is useless, discard that card.
- 3) Give a hint to the cooperater, if possible.
- 4) Discard a card.

For steps 1 and 2, the AI keeps track of which identities a particular card in their hand can possibly have, which is updated with hints received from the cooperater, but also by counting cards visible to them. For example, if the player is told that they have a 5, and the 5s of four colors are visible to them, because they have been played or discarded or are held by the other player, they know which color their 5 has. Steps 3 and 4 are where our AI deviates from the ones presented by Osawa to account for how humans expect and perceive the agent to behave. We will also discuss how to change steps 1 and 2 to interpret information that is received consistently with how hints are given by the AI.

A. Mental State Representation

Our AI represents the players’ mental states by keeping track of which identities they believe possible for each of their cards, and how many of each are still otherwise unaccounted for. Initially, every player believes that every one of their cards can have any identity, with every 1 existing thrice, every 2 existing twice, etc. Table I shows what the initial mental state for a single card in a player’s hand looks like. Each cell in the table contains the number of exemplars of the identity it represents that are currently unaccounted for. Receiving a hint removes possible identities from each card, corresponding to which hint was given. For example, if a player is told that a card that they previously had no knowledge about is red, all possibilities which correspond to non-red identities are removed from their mental state, i.e. all cells corresponding to non-red identities are set to 0. Additionally, the values in

Rank	red	green	blue	white	yellow
1	3	3	3	3	3
2	2	2	2	2	2
3	2	2	2	2	2
4	2	2	2	2	2
5	1	1	1	1	1

TABLE I
THE INITIAL MENTAL STATE REPRESENTATION FOR A SINGLE CARD IN A
PLAYER’S HAND

cells corresponding to the identities of cards an agent can see are decreased accordingly. For example, if the AI knows that a card is a 1, but doesn’t know the color, and sees that two red 1s have already been played or discarded, the entry in the cell corresponding to the red 1 would be decreased by two. If the other player then draws the third red 1, the AI would decrease the value in the cell corresponding to the red 1 by one, ruling out the possibility that the card is a red 1.

In the remainder of this paper we will write M_A for the mental state of the AI in this representation and M_B for the mental state of the human player, with M_{A_i} and M_{B_i} used for the mental states of the individual cards.

B. Giving a hint

To give a hint, our AI uses the fact that humans expect it to behave intentionally and follow Grice’s maxims of communication. In particular, the maxim of relation is followed by only giving hints about cards that have some immediate relevance for game play, i.e. our AI will not give hints about cards that they don’t expect the human player to do anything with, while the maxim of manner is followed to make sure that hints that are given are unambiguous. These two maxims are used to communicate the intentions the AI has for what the human player should do with their cards. In contrast, Osawa’s AI gives hints that have no immediate use except for increasing the knowledge of the cooperater. However, because human players try to infer intentions on part of the AI agent, such hints will often be misinterpreted. For example, if a player has four red cards, but none of them is playable, telling them about all their red cards will give them a lot of information. However, a human player is likely to take other information into account, like the order in which cards were drawn, and conclude that one of the cards is playable. Existing AIs do not model such inference mechanisms, and only play a card when they know for a fact that it is playable. For such AIs giving hints that provide a lot of information is helpful, because it allows the AI to eliminate many possible cases at once.

To determine which hint to give, if any, our AI first determines possible goals for each of the human cooperater’s cards. For every possible hint action it then simulates how that action would change the human player’s mental state, and predicts what they would be likely do with that information. This is then compared to the goals and converted to a score. The AI then adopts the goals corresponding to the highest scoring hint action as its intentions, and executes its part of the plan to fulfill them by performing the hint action. Figure

```

1:  $goals \leftarrow \text{CalculateGoals}(B)$ 
2:  $maxscore \leftarrow -1, action \leftarrow nil$ 
3: for all  $c \in Colors$  do
4:    $move \leftarrow \text{Predict}(M_B, \text{HintColor}(c))$ 
5:    $score \leftarrow \text{Compare}(move, goals)$ 
6:   if  $score > maxscore$  then
7:      $maxscore \leftarrow score$ 
8:      $action \leftarrow \text{HintColor}(c)$ 
9:   end if
10: end for
11: for all  $n \in Ranks$  do
12:    $move \leftarrow \text{Predict}(M_B, \text{HintRank}(n))$ 
13:    $score \leftarrow \text{Compare}(move, goals)$ 
14:   if  $score > maxscore$  then
15:      $maxscore \leftarrow score$ 
16:      $action \leftarrow \text{HintRank}(n)$ 
17:   end if
18: end for
19: if  $maxscore > 0$  then
20:   return  $action$ 
21: end if

```

Fig. 2. Determining which hint to give to the human player, if any

Input: B : Human player’s hand B

Output: $goals$: A mapping from cards in B to goals

```

1: for all  $c \in B$  do
2:   if  $\text{Playable}(c)$  then
3:      $goals(c) \leftarrow play$ 
4:   else if  $\text{Useless}(c)$  then
5:      $goals(c) \leftarrow discard$ 
6:   else if  $\text{Expendable}(c)$  then
7:      $goals(c) \leftarrow maydiscard$ 
8:   else
9:      $goals(c) \leftarrow keep$ 
10:  end if
11: end for
12: return  $goals$ 

```

Fig. 3. Assigning goals to the human player’s cards

2 shows the outline of this process. In the following sections, we will describe how `CalculateGoals` determines what goals the AI has for the cards in the human player’s hand, how `Predict` predicts what action the human player will perform given a particular hint and how `Compare` scores the outcome of the prediction compared to the goals.

1) *Determining Possible Goals*: The first step the AI does when deciding which hint to give, is to determine what it wants the human player to do with each card in their hand. Currently, our AI can choose one of four possible goals for each card: *Play*, *discard*, *may discard* and *keep*. These potential goals are determined by a static rule based system, as shown in figure 3. It simply says that we want players to play playable cards and discard useless ones, as well as allowing them cards of which there are still duplicates.

2) *Predicting the player’s action*: To predict what the human player does with a hint that they receive, the AI uses its representation of the human player’s mental model M_B by applying the hint to it, and then determines what the human player is likely to do with this information. In other words, the AI looks at the human’s current knowledge M_B and the knowledge after giving the hint M'_B and uses that to form a prediction. Note that the game rules prohibit hints that would refer to zero cards, so it would not be valid to tell a player about all their red cards when they don’t have any, even though that would convey information, i.e. that all of their cards are non-red. Our AI rejects such hints as a possibility at this stage by returning a prediction of *NIL*. And while it would be legal to give a hint that does not give any new information to a player, i.e. $M_B = M'_B$, for example by telling them which of their cards are red twice in a row when their hand did not change, our AI will also reject those hints at this stage.

Predicting what the player will likely do with the information they receive is based on the assumption that they expect the AI to follow Grice’s maxims, in particular the maxim of relevance and the maxim of manner. Assuming the maxim of relevance, we predict that the player expects a hint to be about something that they can do with their cards, either playing them or discarding them. The maxim of manner refers to the expectation of the player that the hint is unambiguous. In our AI we take this to mean that a typical player expects a hint about a particular set of cards to actually give them information about these cards. For example, if a player is told that two of their cards are red, we typically expect them to act on these cards, and not draw some conclusion about other cards. We also assume that the player has the goal of increasing the score of the game, which can only be done by playing cards.

All these assumptions together actually result in a rather simple prediction mechanism: A player will play a card they have been directly hinted about, if the hint is consistent with the card being playable, or discard it if the hint is consistent with the card being expendable. For example, if a player is told about all their 2s, without having any other knowledge about these cards, and there are any 2s that are potentially playable, we assume that the player will play one of their 2s². This assumption is consistent with the maxim of relevance in that a player would not expect a hint about a card if that card is not relevant, and with the maxim of manner, since telling a player about some card B_i and actually wanting them to play some other card B_j would lead to ambiguity. Figure 4 shows how our AI predicts what the human player might do when they are given a particular hint. Note that the result of this algorithm is an assignment of a possible action to each card in their hand, and we don’t commit to which of these cards they will act upon.

3) *Scoring the player’s action*: To determine how well a player’s predicted potential action matches with the intentions

²Expert players often use conventions like always playing the leftmost card in their hand when they have two or more equivalent choices. However, since there is no universally agreed upon convention we assume that players make an arbitrary choice in such situations.

Input: M_B : Human’s current knowledge about their hand
Input: *action*: Hint action the AI considers performing
Output: *predictions*: A mapping for every card in B to a predicted action

```

1:  $M'_B \leftarrow \text{Apply}(M_B, \text{action})$ 
2: for all  $c \in B$  do
3:   if PositivelyIdentified( $c, \text{action}$ ) then
4:     if  $\exists id \in M'_{B_c} : \text{Playable}(id)$  then
5:        $\text{predictions}(c) \leftarrow \text{play}$ 
6:     else if  $\exists id \in M'_{B_c} : \text{Expendable}(id)$  then
7:        $\text{predictions}(c) \leftarrow \text{discard}$ 
8:     else
9:        $\text{predictions}(c) \leftarrow \text{keep}$ 
10:    end if
11:  else
12:     $\text{predictions}(c) \leftarrow \text{keep}$ 
13:  end if
14: end for
15: return  $\text{predictions}$ 

```

Fig. 4. Predicting what the human player will do with a hint they received

the AI considers for them we employ a simple comparison: If the player would play a card they are not supposed to play, the hint action that would result in this behavior is discarded. Likewise, if the player would discard a card they are supposed to play or keep, the hint action is discarded. Therefore, only actions that align with the AI’s goals are considered. Among these, a potential play is scored with 3 points, discarding a useless card is scored with 2 point and discarding an expendable card is scored with 1 point. The total score of the hint action is the sum of these scores over all cards in the human player’s hand.

C. Discarding a card

Our AI also differs from Osawa’s in how it decides which card to discard, by following an intentional model. Unlike the process for giving a hint, the model for discarding a card is much simpler as there is only one goal: Discard the card that has the lowest expected value of lost points. The representation of the AI’s mental model lends itself to directly estimate the probability which color and rank a particular card has. Assigning a value of how many points might be lost by discarding a particular card is also challenging. Discarding a card for which a duplicate still exists in the deck might not, in theory, mean that points are necessarily lost, but since it is unknown when a copy of the card will be drawn, the effect on the actual game may still be a loss of points. For example, discarding a green 2 early in the game does not necessarily mean that the green firework may never be finished, but if the other green 2 is at the bottom of the deck, that will be the effect. Calculating the precise expected value of discarding a card would therefore require calculating probabilities and play traces for every possible permutation of cards left in the deck, which is computationally infeasible. Playing a 5 also generates

a hint token for the players, which this process would also have to take into account.

Because of these challenges, we use a heuristic to approximate how many points are expected to be lost by discarding a particular card that captures the most important aspects:

- Cards that are useful sooner are considered to be more important than cards that will be useful later (i.e. a green 2 will be considered more valuable than a green 4, if the green firework stack currently has a 1).
- Cards that are not expendable are considered more valuable than expendable cards.
- Hints are valued at half a point, which is included in the loss of discarding a 5, and the gain of discarding a useless card

D. Receiving hints

A significant contribution of Osawa’s work is that an AI for Hanabi not only needs to be able to give reasonable hints, but also needs to interpret hints it receives from the other player. In his work, this is done by enumerating all possible hands given the AI’s current information and determining which hint the AI would have given itself for each hand. Any hand for which the AI would have given itself a hint that differs from the one it received is then no longer considered possible. The intuition behind this behavior is derived from Grice’s maxims. If a particular hand of cards would allow the cooperator to give a better hint than the one they gave, then it is reasonable to assume that they would have done so. However, in a real time environment, such as when playing with a human player, enumerating all possible hands is not feasible.

In our case, though, when giving a hint, we already estimate what we expect the human player to do with a hint we give them. It is reasonable to use the same logic to interpret hints that we receive from the human player. Note that the algorithm in figure 4 does not need knowledge of the actual content of the player’s hand, as it operates purely on the mental state M_B . To determine what to do with a hint the AI was given by the human player, we can therefore use this algorithm, by using the AI’s mental state M_A after it received the hint instead of the updated mental state of the human player M'_B . The algorithm also needs to determine which cards were positively identified by the hint action, but this is given by game play information. As a result, we then have a list of possible actions for the AI to take, one for each card in their hand. The AI then simply prefers playing over discarding over keeping a card, and uses the leftmost such card in case multiple are applicable.

V. RESULTS

Osawa provides results of his AI playing with an AI cooperator, with an average score of 14.53 and a standard deviation of 2.21 for the Outer State strategy that we use as our baseline. While the goal of our work was to build an AI that does well when playing with a human cooperator, we also ran simulations for how the AI plays with itself as a cooperator, as well as with others. Our AI actually comes in two variants: One only uses the intentional component to decide which hint

	Outer	Intentional	Full
Outer	12.8 (2.0)	13 (2.1)	6.9 (4.3)
Intentional		12.6 (2.6)	14.6 (2.7)
Full			17.1 (2.5)

TABLE II

SIMULATION RESULTS FOR 10000 GAMES WHERE EACH AI PLAYS WITH A PARTNER OF EACH AI, REPORTED AS AVERAGE SCORE WITH THE SAMPLE STANDARD DEVIATION

to give and what to discard, while the other also includes the interpretation of the intention behinds hints that it receives by the cooperater. We call this variants *intentional AI* and *full AI*, respectively. Note that this step is analogous to how Osawa improves the Outer State strategy to get the Self Recognition strategy, but unlike the latter, our version avoids enumerating all possible hands and works in real time. To compare how the different AIs play with each other, we ran every one of the Outer State, Intentional and Full AIs with a partner of every AI for 10000 games, where every combination played with the same 10000 random shuffles. The results of this simulation can be seen in table II. Of note is how the intentional AI playing with another intentional AI actually scores slightly lower on average than the outer AI playing with itself, but it actually enables higher scores when playing with either the outer AI or the full AI. The reason for this is that the intentional AI only plays cards when it is certain that they are playable, but expects its hint to be interpreted according to Grice’s maxims. When giving a hint to another intentional AI, that AI will not necessarily pick up on the information, and the game will stall out with neither of the two players getting sufficient information to play their cards. In contrast, when playing with the outer AI, the intentional AI will receive enough hints to play its cards, even when the outer AI won’t interpret the hints it gets correctly. The full AI, on the other hand, interprets hints exactly the way the intentional AI gives them, thus resulting in the higher score. When the full AI plays with another full AI, both of them will use this logic, further increasing the average score. Finally, the low score when the outer AI and the full AI can also be explained by how their hint giving and receiving modules interact. The full AI expects hints to follow Grice’s maxims, whereas the outer AI actually has a fall-back case of giving random hints, which the full AI will misinterpret. Additionally, as in the intentional/intentional case, the hints that are given by the full AI are not always enough for the outer AI to have full information about its cards, so it won’t play or discard them appropriately. This demonstrates the importance of using conventions that are understood by both players. In the following sections we will show that the conventions used by our AIs more closely align with what human players expect than the ones used by Osawa.

A. Experiment setup

To evaluate how our AI performs when playing with a human cooperater, we implemented a browser-based interface for a human to play Hanabi with any of a number of AIs. Figure 5 shows a screen shot of our UI during a typical game.

Players are told what action the AI performed on its turn, and are then able to choose which action they want to perform: To play or discard a card, they click on the appropriate link on that card in their hand. To give a hint about a particular rank or color, they need to select a card in the AI player’s hand of that rank or color and then click the “Hint Rank” or “Hint Color” links.

To test our hypothesis that adding the intentional behavior leads to a higher score when playing with a human player, as well as what effect of using the full AI has on the score, we assigned a random one of the three AIs to each test subject, without disclosing which one or how the AI they are playing with would behave and had them play one game with that AI. To account for the variance in difficulty of different initial configurations of the deck, each participant played with a deck order chosen randomly from only five possible configurations. After playing one game the participants were asked several questions about board game experience in general and experience with Hanabi in particular, as well as how recently they played. We also asked them to rate the AI in terms of enjoyment, how good at the game and how intentional they perceived it to be. At the conclusion of the study we allowed players to play more games without filling out any additional surveys, but with the games still being recorded for analysis. For each of these subsequent games they were assigned a randomly shuffled deck and a random one of the three AIs. Subjects were given the option to have their game logs and survey answers included in a public release of the data set.

B. Experiment results

To run the experiment, we recruited participants via social media, the website www.boardgamegeek.com and the Reddit forum [r/boardgames](https://www.reddit.com/r/boardgames). 224 participants finished at least one game, and played a total of 1211 games. These participants were aged between 18 and 64, and due to our recruitment method the population skewed towards participants familiar with board games. 152 of the participants self-identified as a gamer, while only 14 did not, with 52 opting not to answer the question. Additionally, 148 participants stated that they play board or card games very often, which was the highest available value on a 4-point Likert scale. Since our goal was not to teach players how to play the game, but rather how it plays with a player that is already familiar with the rules, we do not see this skewed sample as a limitation.

Among the 224 participants, 79 played with the baseline (outer state) AI, 73 played with the intentional AI and 72 played with the full AI for their first game. Figure 6 shows the distribution of scores for the three different AIs. As can be seen, players that played with the intentional AI scored higher on average (mean: 14.99, stddev: 4.17) than players that played with the outer state AI (mean: 11.09, stddev: 4.59). On the other hand, the peak of scores for participants that played with the full AI is slightly higher than for those that played with the intentional AI, but many players also scored significantly lower (mean: 12.88, stddev: 5.98). An ANOVA showed that the AI the participant played with was the determining factor for their

Hint tokens left: 1
 Mistakes made so far: 2
 Cards left in deck: 26

Discarded

green	yellow	white	blue	red
3	4	3	2	
		1	4	
		4	4	
		5	5	

Cards written in red have been discarded or misplayed since your last turn.

Other player



Board



You



Actions

AI just discarded white 3
 You just played blue 1
 AI hinted you about all your blue cards
 You discarded white 5
 AI hinted you about all your white cards
 You discarded yellow 4
 AI hinted you about all your 5s
 You discarded blue 4
 AI hinted you about all your blue cards
 You discarded green 3
 AI hinted you about all your 4s
 You discarded white 4
 AI hinted you about all your 4s
 You hinted the AI about all their 5s
 AI hinted you about all your green cards

Fig. 5. A screenshot from our browser-based implementation of Hanabi. At the top the cards in the AI player’s hand are shown, below that is the current state of the board and at the bottom there is a representation of the player’s unknown hand. Underneath each card in the AI and human players’ hands we see which hints the respective player got about that card.

score, even when controlling for age, board game experience, experience with Hanabi, recency of play or which of the 5 decks they played with ($p < 0.0001$). Using a Tukey test accounting for multiple testing, we found that the difference between the intentional AI and the outer state AI is statistically significant ($p < 0.0001$), the difference between the intentional AI and the full AI is also statistically significant ($p = 0.0295$), but the difference between the full AI and the outer state AI was only weakly statistically significant ($p = 0.0715$). We believe that these results indicate that the conventions used by our AI follow more closely what human players expect naturally of a cooperater than Osawa’s agents. Interestingly, as players got more familiar with the AIs and the setup, their performance with the full AI improved slightly (mean: 13.16, stddev: 5.76), while their performance with the outer state AI actually decreased (mean: 8.12, stddev: 4.87). Figure 7 shows the distribution of scores over all 1211 games played by the participants. As above, we ran a Tukey test accounting for multiple testing on the whole data set, and over all games played by the participants the difference between the outer state AI and the full AI is statistically significant ($p < 0.0001$), while the difference between the full AI and the intentional AI is no longer statistically significant. This is likely due to the fact that the human players learn the conventions used by the AI and expect them to be followed, but the outer state AI violates this assumption.

One minor difference between the intentional AI and the full AI is how participants rated them in the survey. When asked

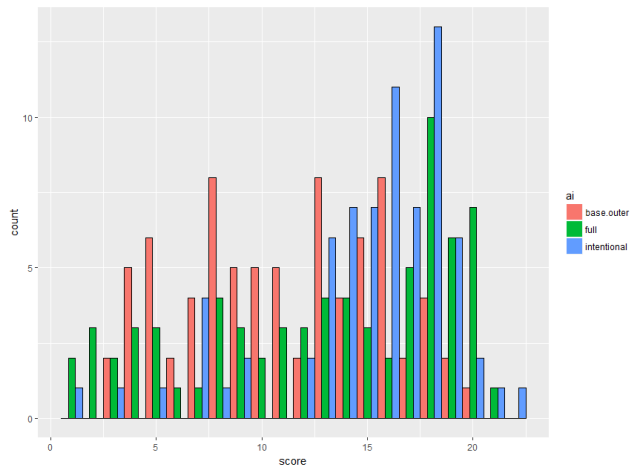


Fig. 6. The distribution of scores for the three different AIs for the first game of each participant

how much they enjoyed playing with the AI, participants that scored between 13 and 18 points rated the full AI higher than the intentional AI. We tested this by grouping the participants into categories depending on what 5-point score range, starting at 3 points, they fell into, and noting how many participants in each group rated their enjoyment as 3 or higher on a 5 point Likert scale. For the range 13 to 18 points, 10 out of 14 participants, or 71%, that played with the full AI did so,

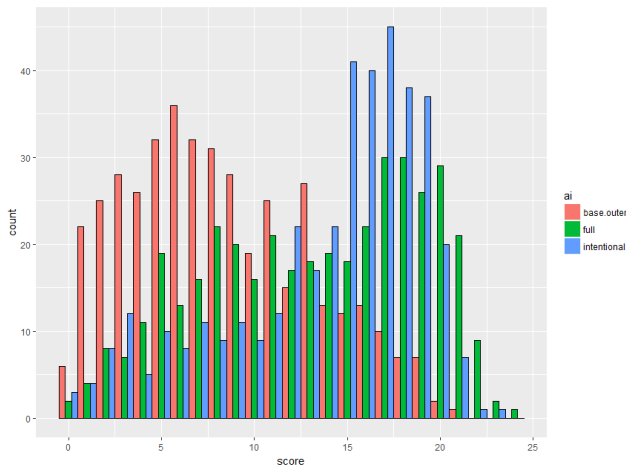


Fig. 7. The distribution of scores for the three different AIs for all games played by the participants

whereas only 7 out of 33, or 21%, gave a rating of 3 or higher to the intentional AI. We ran a Tukey test accounting for interactions to find which pairs of AI and score range were statistically significantly different, and for the 13 to 18 range we found that the difference we described is weakly statistically significant ($p = 0.0827$). No other statistically significant difference between results in the same score range were found, but there is a statistically significant difference between players that scored more than 18 points with the full AI, where 15 out of 17 participants, 88%, rated it with 3 or higher, and players that scored between 13 and 18 points when playing with the intentional AI. A possible explanation for this result is that players enjoy playing successful games more readily with the more complex AI, while being more easily frustrated by the simpler AI when they fall just short of a high score, but more data would need to be gathered to make this analysis fully conclusive.

Finally, the responses on the survey also indicated that there is a positive correlation between how intentional the behavior of the AI was rated and how much players liked playing with it (Kendall’s rank correlation $\tau = 0.45, p < 0.0001$), as well as between how intentional an AI was rated as having played and how high players rated its skill at playing the game (Kendall’s rank correlation $\tau = 0.52, p < 0.0001$). This provides additional evidence that intentionally acting agents are preferred by players, and also perceived to play better.

VI. CONCLUSION AND FUTURE WORK

We have presented an AI agent for the two player version of the cooperative card game Hanabi that is based on intentionality and communication theory. Our agent is based off of a previously published agent by Osawa, but rather than performing actions to only serve its own logic our agent strives to use the communicative actions present in the game to convey its intentions to the other player. We also described how the same logic that is used to predict what the other player will do with the information they receive can be used

to determine what to do with hints received by the other player. This led to two different agents: One that acts intentionally in the actions it performs, and one that additionally also interprets the intentions behind other player’s actions. We then showed that when playing with human cooperators our two agents performed significantly better than the baseline agent, and players also perceive the full AI to be better in some cases. 190 of the participants gave us permission to make their survey answers and game logs publicly available, and this data set is available on github³, along with the complete source code of our AIs, including the browser-based UI⁴, to be used for future work. For example, extending the AI to more than two players would be an interesting challenge since it adds the decision of whom to give hints to. The game logs could also be used with machine learning techniques to learn human responses to hint actions in particular situations, and use that as the prediction mechanism in our AI framework. Finally, we believe that the techniques we used for our AI and the results we gathered from the experiment can also be used to develop AIs for other games involving human/AI-interaction or -communication.

REFERENCES

- [1] D. C. Dennett, “Intentional systems,” *The Journal of Philosophy*, vol. 68, no. 4, pp. 87–106, 1971.
- [2] Spiel des Jahres award 2013,” 2013. [Online]. Available: <http://www.spieldesjahres.de/en/hanabi>
- [3] A. Bauza, “Hanabi,” 2010. [Online]. Available: <https://boardgamegeek.com/boardgame/98778/hanabi>
- [4] H. P. Grice, “Logic and conversation,” 1975, pp. 41–58, 1975.
- [5] A. Whiten, *Natural theories of mind: Evolution, development and simulation of everyday mindreading*. Basil Blackwell Oxford, 1991.
- [6] D. V. Pynadath, M. Si, and S. C. Marsella, “Modeling theory of mind and cognitive appraisal with decision-theoretic agents,” *Social emotions in nature and artifact: emotions in human and human-computer interaction*, pp. 70–87, 2011.
- [7] M. O. Riedl and R. M. Young, “Narrative planning: balancing plot and character,” *Journal of Artificial Intelligence Research*, vol. 39, no. 1, pp. 217–268, 2010.
- [8] A. Nareyek, “Review: Intelligent agents for computer games,” in *International Conference on Computers and Games*. Springer, 2000, pp. 414–422.
- [9] P. R. Cohen and H. J. Levesque, “Intention is choice with commitment,” *Artificial intelligence*, vol. 42, no. 2-3, pp. 213–261, 1990.
- [10] R. M. Young, “The cooperative contract in interactive entertainment,” in *Socially Intelligent Agents*. Springer, 2002, pp. 229–234.
- [11] —, “Using grice’s maxim of quantity to select the content of plan descriptions,” *Artificial Intelligence*, vol. 115, no. 2, pp. 215–256, 1999.
- [12] P. R. Williams, D. Perez-Liebana, and S. M. Lucas, “Cooperative games with partial observability,” 2016.
- [13] J.-F. Baffier, M.-K. Chiu, Y. Diez, M. Korman, V. Mitsou, A. van Renssen, M. Roeloffzen, and Y. Uno, “Hanabi is NP-complete, even for cheaters who look at their cards,” in *8th International Conference on Fun with Algorithms*, 2016.
- [14] M. van den Bergh, F. Spieksma, and W. Kusters, “Hanabi, a co-operative game of fireworks,” Bachelor’s thesis, 2015.
- [15] C. Cox, J. De Silva, P. Deorsey, F. H. Kenter, T. Retter, and J. Tobin, “How to make the perfect fireworks display: Two strategies for hanabi,” *Mathematics Magazine*, vol. 88, no. 5, pp. 323–336, 2015.
- [16] H. Osawa, “Solving Hanabi: Estimating hands by opponent’s actions in cooperative game with incomplete information,” in *Workshops at the Twenty-Ninth AAAI Conference on Artificial Intelligence*, 2015.

³<https://github.com/yawgmoth/HanabiData>

⁴<https://github.com/yawgmoth/pyhanabi>