Evaluating a Plan Recognition Agent for the Game Pandemic with Human Players

Pablo Sauma-Chacón
Escuela de Computación e Informática
Universidad de Costa Rica
San Jose, Costa Rica
pablo.saumachacon@ucr.ac.cr

Markus Eger
Department of Computer Science
CalPoly Pomona
Pomona, CA, USA
meger@cpp.edu

***Abstract*—Cooperation between AI agents and humans is of ever greater importance. In this paper we present an AI agent for the game Pandemic that was specifically designed to play cooperatively with a human player. Our agent utilizes planning to determine which actions to perform, and plan recognition to determine the current goal of its cooperator in order to assist them. We also present an experiment we performed with human participants, and how our agent performs at a level that is comparable to other AI agents playing with themselves, when playing with a human player, as well as the impact of plan recognition on how the participants perceive the AI agent.**

***Index Terms*—cooperative games, planning, plan recognition, human-AI cooperation**

I. Introduction

AI agents that cooperate with humans have become of increasing interest in recent years, from virtual assistants to agents that cooperate in games with human players. Perhaps the most thoroughly researched cooperative game is Hanabi, which combines partial information with limited communication to provide a challenge for AI agents and humans alike, and even more so for a combination of the two [1]. However, Hanabi is far from the only game in which collaboration between players is required. Another classical cooperative game is Pandemic, which - at its core - does not revolve around communication, but instead only defines actions the players can take, and which goals they have to work together to achieve.

In a social setting, players will, of course, communicate freely, which may often result in one player dominating the decision making process (also called "quarterbacking"). However, by simply observing the game actions themselves, a sophisticated player can discern what their team mates are likely up to, and how to aid them with their goals. While free-form communication is an interesting problem to consider, our work instead focuses on this latter approach, by developing an AI agent that plays the game in a way that is understandable for a human collaborator, and also deduces that collaborator's intermediate goals and reacts to them appropriately.

In this paper, we present our planning and plan-recognition based agent for the game Pandemic, which was designed to play with a human player. In order to determine how successful our agent is in its collaboration, and which effect recognizing the human player's likely course of action has on the agent's performance, we performed an experiment with human participants, which is also described in this paper. Our agent is based on work we have previously published, but the use of plan recognition, as well as the evaluation with human subjects are novel contributions in this paper. Before we describe our own agent design, though, we will first provide a short summary of the rules of Pandemic, followed by a brief review of the relevant literature.

978-1-6654-3886-5/21/$31.00 ©2021 IEEE

A. Pandemic

Pandemic [2] is a cooperative board game for two to four players. In this game, players are members of a task-force in charge of discovering the cures for four deadly diseases. The players win the game once the cure for each of the four diseases has been discovered, and lose if they either run out of cards in one of the decks when they need to draw (representing "running out of time"), the diseases spread too much, or more than seven *outbreaks* occur (representing "a worldwide panic").

The game is played on a board with a world map which highlights forty-eight major cities. Each city has one of four colors (red, yellow, black or blue), which specifies which of the four diseases is endemic to it. The players, represented by meeples on the board, must travel to the cities to treat the diseases in order to prevent them from spreading. At the beginning of the game, each player is assigned a specific role, which grants them special actions and effects, which, in turn, affect their play style. The game also contains two decks: The *Player deck* and the *Infection deck*, both of which contain one card corresponding to each city, with the Player deck additionally containing a number (4-6, depending on the desired difficulty of the game) of special *Epidemic* cards.

Each turn a player can perform up to four actions with different options to choose from: travel to another city, treat a disease in the current city, build a research station, transfer cards to/from another player in the same city or discover a cure for a disease. Afterwards, the player draws two player cards from the Player deck. The player cards are kept by the player (up to a maximum of seven) and are used to discover the cure for diseases and required to perform certain actions. However if a player draws an Epidemic card as one of their player cards, that card is discarded and an epidemic occurs in a city: The bottom card from the Infection deck is taken, and infected with its disease. Then, the cards in the discard pile of the Infection deck are shuffled and put on top of that deck. Regardless of whether an epidemic occurred or not, a player then proceeds to draw a given number of cards from the Infection deck, which are immediately put into the discard pile and each of the cities named on the cards gets infected.

Whenever a city gets infected, a *disease cube* of the corresponding color is added to it, to represent the level of infection in that city. If a city were to get infected by a disease when it already has three cubes of that disease an outbreak occurs instead. During an outbreak in a city, all neighboring cities get infected by the disease causing the outbreak and an outbreak counter goes up by one. Adding disease cubes to neighboring cities, may lead to chain reactions, if one or more of these cities already have three disease cubes, leading to additional outbreaks. However, once a city was affected by an outbreak during such a chain reaction it cannot be affected again.

In order to win, the players need to collectively find a cure for each of the four diseases. To find a cure for a disease, a single player must possess five player cards of the same color, and be located in a city with a research station (initially only Atlanta), and discard those five cards. Cards can be exchanged between players, but only if both players are in the city corresponding to the card they want to exchange. The main challenge posed to the players is therefore that while the disease cubes are distributed across the entire map, and players must control the spread of the diseases, actually finding cures requires coordination, as it is highly unlikely that a single player draws five cards of the same color. Additionally, the mechanics of the Epidemic card result in the same cities being infected over and over again, as every time an Epidemic card is drawn, the discarded cards from the Infection deck, which are exactly the cards corresponding to already infected cities, are put back on top of that deck.

In our work, we actually look at a slightly reduced version of Pandemic, by applying two changes to the original rules. First, event cards were removed from the game. These cards are special player cards which can be used by the players to perform one-time special actions. The reason for the removal of these cards was that, while powerful, they can be used at any given moment, even between turn phases, and have a wide variety of possible options to choose from which greatly increases the search space of the game. The second change was the removal of three of the seven player roles that can be assigned to the players. The Dispatcher and the Operations Expert roles were removed due to having special movement actions which add multiple options to each action, which increases the search space significantly. The Contingency Planner role was removed because its special actions had to do with event cards, which left the role purposeless without these cards. The removal of the roles has no impact on core mechanics of the game, but the removal of the event cards makes the game somewhat harder for the players (since they can no longer use the powerful cards and have fewer turns to win the game). However, together these changes significantly reduce the search space, while leaving the core idea of the game intact.

II. Related Work

Prior work that we based our agent on generally falls into two categories, with some overlap between them: First, we will present some prior research into agents for cooperative games, including agents that play with human players, and agents that play Pandemic. Then we will briefly discuss the relevant literature about planning, which forms the basis for the decision making process of our agent.

As mentioned above, cooperation has been of increasing interest recently, with the game AI community focusing strongly on Hanabi [3]. The main challenge with Hanabi is that the cooperators have disparate knowledge bases, and a limited communication channel which they can use to coordinate their moves. Most research has thus focused on how to optimize the use of this communication channel, with approaches ranging from encoding logical puzzles [4] to machine learning approaches that determine good communication protocols, either among themselves [5] or with different cooperators [6]. There has even been work that explores using high-resolution timing information to improve communication [7]. However, all of these approaches do not capture the nuances of playing with human players, and, indeed, end up performing actions that are incomprehensible to most humans. A different branch of research, which is more relevant to our work, has therefore been working towards building AI agents that are specifically tailored to playing with human players. These approaches often use concepts from communication theory, such as conversational implicature [8] to interpret communicated action in the context of the game. Another, related approach uses the concept of intentionality [9], [10] to connect communicative actions with (short-term) goals the players ought to collaboratively work towards [11]. While our work does not feature explicit communication, we posit that a similar approach can be used to interpret entirely implicit communication, which is conveyed through the actions performed by the cooperating player.

While Hanabi has been the main focus point of game AI research in recent years, there has already been some interest in Pandemic as well. In contrast to Hanabi, no explicit communication is present, but the game itself has a much larger state- and action-space, features "more" randomness, and has therefore already been proposed as a new domain of interest for AI research [12]. There have already been AI agents that successfully play the game with themselves as collaborators using a Rolling Horizon Evolutionary Algorithm (RHEA) [13], which win about 22% of the games on easy difficulty, on a similarly restricted version of Pandemic as ours. Our own implementation, based on planning, achieves a win rate of about 34% when playing with itself as a collaborator [14]. While these results represent significant advances in addressing the challenges of the domain, neither of these agents was evaluated with human players. We can use these win rates as a comparison baseline, though, and will compare our agent's performance when playing with human players against them below.

Our agent uses planning to determine which actions it should perform on its own turn. Planning is a process, in which an agent is given a formal description of the current state of the game, actions it may perform, and a goal it should work towards. In classical formulations, such as STRIPS, actions are described using preconditions, i.e. which states they are applicable in, and effects, i.e. how they change a state to produce a successor state [15]. More generally, planning can be seen as a search process on an implicitly defined graph, where the nodes are states and the (directed) edges are actions the agent can take [16]. While there are standards on how to describe planning problems, such as PDDL [17], our agent instead uses this more generic view to implement a custom planner. We do this, because encoding the entire game domain in a logical language would be prohibitively cumbersome, and to address the randomness inherent in the game by integrating a Monte-Carlo sampling procedure [18], [19] in the planning process as described below.

Finally, in order to cooperate with a human player, our agent performs a plan recognition process on the actions performed by the other player. Plan recognition is the problem of determining which plan an observed agent is attempting to perform given some observed actions performed by that agent [20]. The way our agent performs this process is by transforming it into a planning problem [21], where it compares plans it would perform for each of the candidate goals with the observed actions to determine the most likely plan followed by the cooperator. We will now describe how our agent works in more detail.

III. Our Agent

In this section we will describe our agent, which uses AI planning to determine which actions it should perform on each of its turns. In order to apply planning to Pandemic, we need to define the game in terms of states and actions, provide the agent with a suitable goal or set of goals, and guide the search process with a heuristic. In the following sections we will briefly describe these parts, as we implemented them in our existing agent. The main novelty of this article is the addition of a plan recognition module in order to determine which goals the cooperator is likely pursuing, as well as a Unity interface to allow human players to play with the AI agent, which we will present in more detail.

A. State and Action Representation

Pandemic, as a game, readily affords a state representation of the public game state: The location of the players, research stations and disease cubes on the board, the roles of each player, as well as which cards they hold are all known to all players. The hidden information, on the other hand, requires some finesse. While we could model the infection deck to be in random order, it actually consists of two parts that are independently randomized: The bottom cards, which have never been seen by the players, and the top cards, which were put there from the infection card discard pile as the result of an epidemic card. As this information is essential for strategic play (players have some idea which cities are likely to be infected next), we track these two parts separately.

Actions the players perform then take one such game state and transform it into a new game state: On each of their turns, a player performs four actions which are unaffected by the order of cards in either deck, and then have to draw cards from each of the decks, which involves these random orders. Our implementation decouples the *actual* state of the game from what players *know* about it: A player may perform a forward simulation step and when the game requires cards from either of the two decks the game state will instead be generated as one in which the decks are reshuffled while taking the player's observed knowledge, i.e. which cards are in the top and bottom part of the infection deck, into account.

B. Possible Goals

The ultimate goal in Pandemic is to discover the cure to the four diseases ravaging the world. However, when playing the game the players face one additional challenge: they have to avoid losing the game. This challenge requires players to balance their priorities between a duality of goals when deciding their actions: they can either focus on discovering the cures or on fighting the diseases to stop them from spreading.

To handle this challenge, our agent performs an evaluation of the current game state whenever it needs to find a new plan: If it is possible to discover a new cure with the cards currently in play it will try to find a plan to do so. Otherwise it chooses the "survive" goal, which is defined as "not losing the game". The agent will then try to find a plan for the next n turns, our *planning horizon*, that reaches the chosen goal as closely as possible.

C. Planning Heuristic

When performing the planning process, our agent uses a heuristic as part of its evaluation of the "quality" of the states it visits, by measuring how "far" from a solution each state might be, i.e. states with lower heuristic values are considered to be closer to an overall positive outcome of the game. This heuristic is used to choose the best actions to explore during planning, and as a tie breaker between multiple possible final states/plans. Our heuristic is composed of a number of equations, each of which evaluates a different aspect of the game state. Before presenting the equations used, some of their expressions will be explained. The expression $\mathrm{active}\,(k)$ evaluates whether the k colored disease is currently active. The expression $\mathrm{distance}\,(p, c)$ represents the minimum distance (as number of movement actions required) between the player p and the city c. In a similar fashion, $\mathrm{infection}\,(c)$ represents the number of disease cubes present in the city c. The expression $\mathrm{cards}\,(p, k)$ stands for the number of k colored cards currently in player p's hand, and the expression

$\text{discard}(k)$ counts the number of cards of k color currently in the discard pile. Finally, the constant R_p represents the number of cards required by a player to discover a cure (five for every role except for the Scientist which only needs four cards for this action).

Equation 1 evaluates the number of active diseases which is directly related to the overall goal of the game, i.e. if h_{cures} starts out at 4, and if it reaches 0, the players have found a cure for all four diseases and won the game.

$$h_{cures} = \sum_{k \in Color} \text{active}(k) \quad (1)$$

Equations 2 and 3 evaluate the distance of the players to places of interest. Equation 2 measures the distance of the players to each city weighted by the number of disease cubes in that city. This measurement is then divided by the total number of disease cubes on the board. This equation has the property of increasing in value as the players move further away from the nexuses of infection on the board. Equation 3, on the other hand, measures the distance of the players to the closest city with a research station on it, which are the only places where they can discover cures.

$$h_{dsurv} = \sum_{p \in Player} \frac{\sum_{c \in City} \text{distance}(p, c) \cdot \text{infection}(c)}{\sum_{c \in City} \text{infection}(c)} \quad (2)$$

$$h_{dcure} = \sum_{p \in Player} \min_{c \in C_{rs}} \text{distance}(p, c) \quad (3)$$

Equations 4 and 5 evaluate the cards available to the players. Equation 4 measures the minimum number of missing cards for each of the active diseases in the players' hands required to discover a cure. This equation has the property of decreasing in value as player manages to concentrate more cards of the same color in their hand. The equation 5 evaluates the number of discarded cards for each active disease, as this value increases it becomes more difficult for the players to win the game (there are fewer cards remaining of a given colored disease).

$$h_{cards} = \sum_{k \in Color} \text{active}(k) \cdot \min_{p \in Player} R_p - \text{cards}(p, k) \quad (4)$$

$$h_{disc} = \sum_{k \in Color} \text{active}(k) \cdot \text{discard}(k) \quad (5)$$

Equation 6 evaluates the number of disease cubes currently active across the board. This equation is directly related to the goal of preventing the spread of the infections by the players. Furthermore, the reduction of disease cubes also motivates agents to eradicate diseases, since that will lead to future states with fewer infections. Finally, equation 7 evaluates the average distance between all the cities in the game multiplied by the ratio of turns remaining. The resulting value is not constant because as the players build research stations in cities, the cities become connected between them (because players can "fly" between research stations), reducing the average distance among cities. The impact of this distance reduction however, decreases in importance as the game approaches its end, with fewer actions remaining.

$$h_{inf} = \sum_{c \in City} \text{infection}(c) \quad (6)$$

$$h_{dist} = \sum_{c_1 \in City} \sum_{c_2 \in City} \frac{\text{distance}(c_1, c_2)}{48 \cdot 47} \cdot \frac{turns_{remaining}}{turns_{max}} \quad (7)$$

All of these equations are then added into a single equation using weights. The weights used were selected after performing a grid search to test the performance of the planning agent playing with itself with different values. Equation 8 is the one used by our planning agent as an heuristic to evaluate the states of the game.

$$\begin{aligned} h_{state} = \; & 0.5 \cdot h_{dsurv} + 0.5 \cdot h_{dcure} + 1 \cdot h_{cards} + \\ & 0.5 \cdot h_{disc} + 0.6 \cdot h_{inf} + 0.6 \cdot h_{dist} + 24 \cdot h_{cures} \end{aligned} \quad (8)$$

D. Plan Recognition

Our agent incorporates the use of a plan recognition module. The objective of this module is to identify, through the observation of the actions taken by the other player, their current goal. The intention behind it is to allow the planning agent to incorporate the other player's goal into its own planning process, by assuming the other player will continue with their current plan. This allows the agent to formulate plans that enhance its cooperation with the other player.

At the beginning of its planning process, the agent performs a plan recognition pass. There are two possible goals the agent assumes the player may pursue: Eliminating diseases (i.e. avoiding losing the game), or trying to discover a cure (i.e. working towards winning the game). The agent takes the state as it was at the beginning of the other player's turn, and performs its own planning process with each of these two possible goals as a goal, resulting in two possible plans. It then compares the actions it actually observed the other player perform against the actions it would have performed itself for each possible goal, to determine which goal the other player is more likely to pursue. Assuming the other player uses a similar strategy to decide which actions to perform, the plan that matches more of the observed actions is the one they are more likely to be executing.

With this information, our agent then adapts its plan assuming the player will continue forth with their plan as well. For example, when the other player performs actions that the agent would have performed to find a cure for a disease, the agent detects that. As our agent accounts for the other player's likely future actions when calculating its own plan (i.e. it assumes the player will continue working towards the goal that was just recognized), this results in actions such as passing a necessary card to the other player, or meeting up with the other player so that they can give the agent their cards.

It should be noted that this process requires the execution of the planning process multiple times during the agent's turn. In order to keep the response time of the agent within an acceptable range for play with a human player (around 20-30s), the planning process used for plan recognition only plans actions for a single turn, and uses the planning heuristic described above to evaluate each state that was reached.

E. User Interface

To evaluate the agent's performance when playing with human players, we implemented a graphical user interface[1] in Unity, shown in figure 1. It is designed to run in a web browser, and connects to the actual implementation of the game using HTTP calls. The server it connects to contains the aforementioned representation of the game state and actions, as well as the planning agent, with or without the plan recognition module. This user interface allows human participants to play with our AI agents, while allowing us to record game play information. We also use the same Unity application to present survey questions to participants, as described below.

IV. Evaluation

The evaluation of our agent was two-fold: First, we were interested how well the agent plays with human players in general. Second, we were investigating if plan recognition would add significant capabilities to the agent, with regard to the outcome of the game as well as how the participants perceive it. For this evaluation, we performed an experiment with volunteer participants playing our implementation of Pandemic with our agents. We will first describe how we set up the experiment before reporting the results we obtained.

A. Experiment Setup

We recruited participants via snowball sampling on social media, including Facebook, Twitter, and the boardgame community on reddit, to play Pandemic with our AI agents. Participants were randomly and secretly assigned to one of two agent types: The basic planning agent, or the planning agent with plan recognition, which we will call "plan recognition agent" from now. After being shown a consent form, participants were asked to read a brief explanation of the user interface, and optionally the game rules, and then played one game with the agent type assigned to them. After the conclusion of the game, participants were prompted to answer survey questions about their experience with board games in general and Pandemic specifically. They were also asked to rate the AI agent they just played with in terms of how well it played, how helpful it appeared to their actions, and how well the participant understood what the agent was trying to do, each on a 5-point Likert scale. 51 participants finished their game when playing with the basic planning agent, while 65 participants did so with the plan recognition agent. Participants were assigned to the agents randomly but evenly, and we could only speculate about what caused the disparity in completion between the two agents.

[1]Available in the GitHub repository: https://github.com/BlopaSc/PAIndemic

Our participants represented a relatively balanced cross-section of players, with around 40% having never played the game before, 21% having played between 11 and 50 times, and 10% having played more than 50 times. While there is no published data on an average win rate of Pandemic, most of our participants reported never winning the game (48%), or winning it about one in four times (27%). Only about 25% of our participants report winning the game every one in two-three times, or more often than one in two games. This underscores that Pandemic is a challenging game, even for pure human player teams, with access to the special event cards.

B. Results

The first metric we looked at was the percentage of games participants won when playing with our agents. Overall, 26.7% of games were won by participants with either agent, with no statistically significant difference between the two agent types (planning agent: 25.49%, plan recognition agent: 27.69%). This falls between the results obtained by the agent-agent teams of the RHEA agents (22%) and our planning agent playing with itself (34%), and compares favorably to the participants self-reported win-rate.

Of the 116 participants, only 77 answered the post-game survey (32 for the planning agent, 45 for the plan recognition agent). We compared the responses of these participants to our three evaluation metrics, play skill, helpfulness and understandability, using a χ^2-test and applied a Holm-Bonferroni correction to account for multiple testing. Only the results for how well participants understood the agent showed a statistically significant difference ($p < 0.05$), with the participants evaluation of the two agents shown in figure 2. We also performed a Mann-Whitney-U test on these results to determine whether there was a difference in the mean response, but were unable to detect one. As can be seen, the plan recognition agent was rated as being understood "sometimes" or "often" more often than the planning agent, while the planning agent was rated as being understood "Rarely" more often. However, at the extremes, of being understood "Never", or "Always" the planning agent outperforms the plan recognition agent. We interpret this as meaning that while the plan recognition agent does make an effort at performing actions a human player would expect and understand, there are some cases in which it completely fails to determine what the human player is trying to do, or how to support them in their endeavors.

Another result we want to highlight is how players rated the helpfulness of the two agents, shown in figure 3. While we were unable to show a statistically significant difference in the distribution of the two ratings, as they are very similar for most ratings, there is a significant difference in how many participants rated the two agent types as "Always" being helpful, with 13% of participants giving this rating to the plan recognition agent, but only 3% doing so for the planning agent. This indicates that our plan recognition agent, when it manages to pick up on the human player's plans, is perceived as being actively helpful.

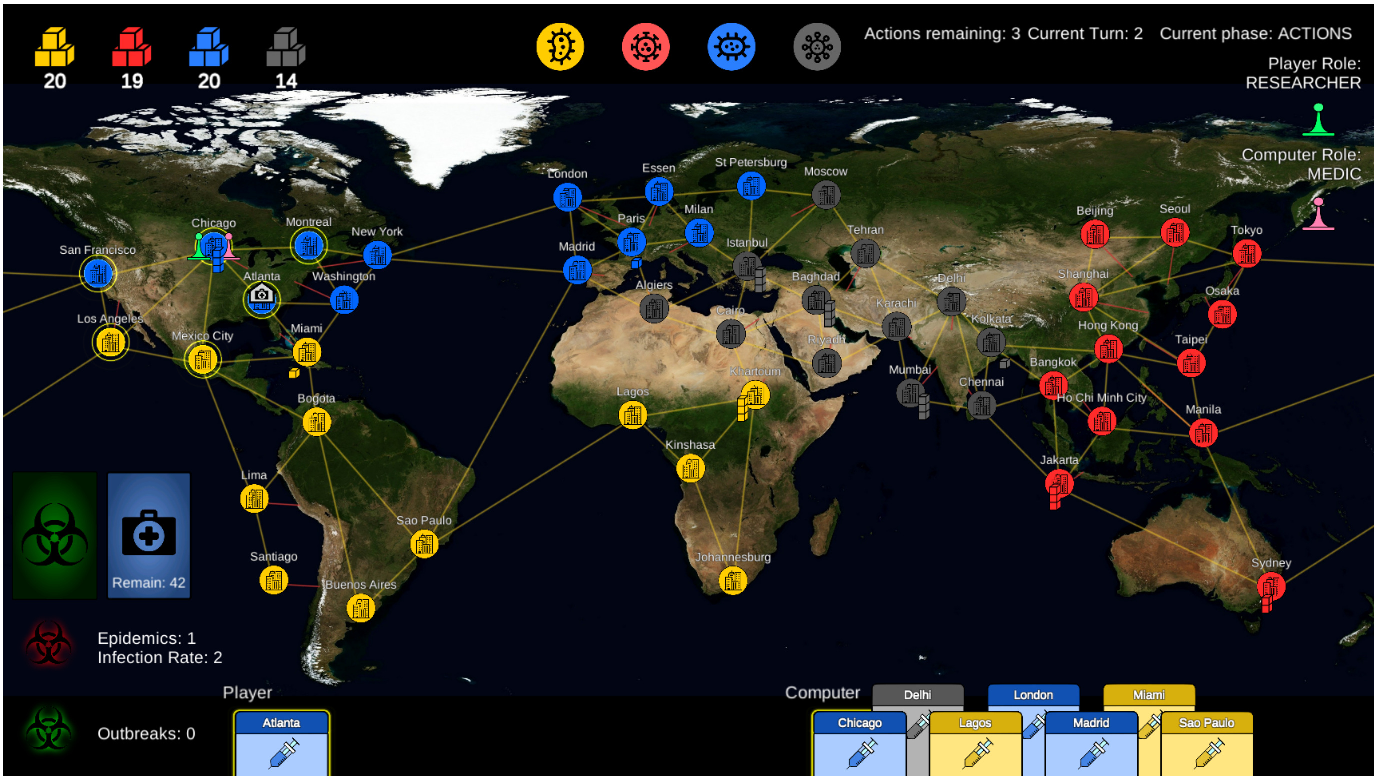

Fig. 1. Screenshot of the game client for Pandemic depicting the game board, players' meeples and cards, and other game information.

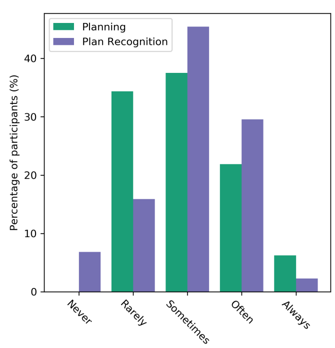

Fig. 2. How often the participant understood the agent.

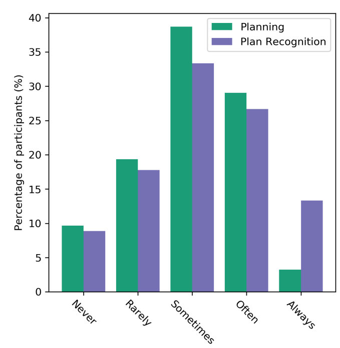

Fig. 3. How often the agent was helpful towards the participant.

Additionally, we investigated if there were any interesting correlations between the ratings given for different metrics. For both the planning and the plan recognition agents there is a strong correlation (0.69 and 0.76 respectively) between the perceived skill level of the agent and the perceived helpfulness, which indicates that human players perceive a good player as one that is willing to help others, further supporting the necessity for better cooperation.

Finally, while all participants were asked to play one game, they were given the option to continue playing additional games with a random agent. In order to avoid the influence of learning effects, aforementioned results only take the first game played by each participant into account. However, subsequent games by the participants were also recorded, and we also calculated their win rate. Of these 71 additional games, participants won 40.8%, with no statistically significant difference between the two agent types (planning agent: 41.02%, plan recognition agent: 40.63%). This indicates that our agents are actually capable of assisting a human player on a competent level. It also suggests that players need some time to adjust to, either, our user interface or the agents.

Qualitative analysis of the game information showed that, while there was no statistical difference between the kinds of actions actions used between the two agents, the plan recognition agent seemed to "time" its actions in a way that better integrated with the other player's actions. For example, the planning agent transfers cards whenever it is possible and immediately increases the value of the players' hands (as evaluated by the heuristic), while the plan recognition agent behaved in a more deliberate way, delaying the transfer of cards until the discovery of a cure was actually at hand.

V. Conclusion

In this paper we presented our implementation of a plan recognition agent for the game Pandemic, designed to play the game with a human cooperator. We explained the different challenges presented by the game and how they are handled by our agent. We also presented the results obtained by the agent when playing with human participants using a web-based implementation of the game.

The results of our experiment were two-fold. First, we demonstrated that our agent is able to successfully play the game with human cooperators, at a level that is comparable to teams consisting only of AI-agents, as well as the self-reported average performance of our participants. Second, the results indicate that there are some advantages of performing plan recognition in order to better assist human players. On one hand, participants understood our agent better when it performed plan recognition in most cases, with some error cases where the recognized plan apparently did not match what the player was actually doing. On the other hand, we also found a general correlation between the perceived "skill" and "helpfulness" of our agent by human players.

These results indicate that AI agents can play a cooperative game like Pandemic successfully with human players, and that assisting them is an important aspect of this cooperation. However, in future work we want to focus more on the nuances of different plans in order to better determine what the human player is actually trying to do. Our agent currently very coarsely tries to peg the human player as either trying to "win" or "not lose", when there would be many different variants of each, as well as shades in between. For example, a player may have a choice between multiple regions of the map to control diseases in ("not lose"), which may be influenced by the availability of another player and their cards, making a potential cure be found more easily ("win").

Our implementation also has the limitation of requiring significant expert knowledge to construct and tweak the heuristic used to evaluate game states and guide the planning process. In lieu of having to rely purely on expert knowledge, we plan on using actual game logs to train a machine learning model to improve this heuristic.

Finally, while our work focuses on Pandemic, we believe that the general structure of our agent is applicable to more general domains. While our agent uses a custom planning approach, we are currently investigating the use of a more general planning implementation in order to support multiple different domains more easily in the future. Another topic for research could focus on the exploration of the space of different heuristics weights, studying their impact in more detail, as well as different methods for choosing them.

References

[1] N. Bard, J. N. Foerster, S. Chandar, N. Burch, M. Lanctot, H. F. Song, E. Parisotto, V. Dumoulin, S. Moitra, E. Hughes *et al.*, "The hanabi challenge: A new frontier for ai research," *Artificial Intelligence*, vol. 280, p. 103216, 2020.

[2] M. Leacock, "Pandemic," https://boardgamegeek.com/boardgame/30549/pandemic, 2008, online; accessed 25 February 2019.

[3] A. Bauza, "Hanabi," https://boardgamegeek.com/boardgame/98778/hanabi, 2010.

[4] B. Bouzy, "Playing hanabi near-optimally," in *15th International Conference on Advances in Computer Game (ACG15)*. ICGA, 2017.

[5] H. Hu and J. N. Foerster, "Simplified action decoder for deep multi-agent reinforcement learning," *arXiv preprint arXiv:1912.02288*, 2019.

[6] R. Canaan, J. Togelius, A. Nealen, and S. Menzel, "Diverse agents for ad-hoc cooperation in hanabi," in *2019 IEEE Conference on Games (CoG)*. IEEE, 2019, pp. 1–8.

[7] M. Eger and D. Gruss, "Wait a second: playing hanabi without giving hints," in *Proceedings of the 14th International Conference on the Foundations of Digital Games*, 2019, pp. 1–7.

[8] C. Liang, J. Proft, E. Andersen, and R. A. Knepper, "Implicit communication of actionable information in human-ai teams," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–13.

[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] M. E. Bratman, "What is intention," *Intentions in communication*, pp. 15–32, 1990.

[11] M. Eger, C. Martens, P. S. Chacón, M. A. Córdoba, and J. H. Cespedes, "Operationalizing intentionality to play hanabi with human players," *IEEE Transactions on Games*, 2020.

[12] P. Sauma-Chacón and M. Eger, "Pandemic as a challenge for human-AI cooperation," in *Proceedings of the AIIDE workshop on Experimental AI in Games*, 2019.

[13] K. Sfikas and A. Liapis, "Collaborative agent gameplay in the pandemic board game," in *Proceedings of the 15th International Conference on the Foundations of Digital Games*, 2020.

[14] P. Sauma-Chacón and M. Eger, "Paindemic: A planning agent for pandemic," in *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, vol. 16, no. 1, 2020, pp. 287–293.

[15] R. E. Fikes and N. J. Nilsson, "STRIPS: A new approach to the application of theorem proving to problem solving," in *Proceedings of the 2Nd International Joint Conference on Artificial Intelligence*, ser. IJCAI'71. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1971, pp. 608–620. [Online]. Available: http://dl.acm.org/citation.cfm?id=1622876.1622939

[16] B. Bonet and H. Geffner, "Planning as heuristic search," *Artificial Intelligence*, vol. 129, pp. 5–33, 2001.

[17] D. McDermott, M. Ghallab, A. Howe, C. Knoblock, A. Ram, M. Veloso, D. Weld, and D. Wilkins, "Pddl-the planning domain definition language," 1998.

[18] N. Metropolis and S. Ulam, "The monte carlo method," *Journal of the American Statistical Association*, vol. 44, no. 247, pp. 335–341, 1949, pMID: 18139350.

[19] D. P. Kroese, T. Brereton, T. Taimre, and Z. Botev, "Why the monte carlo method is so important today," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 6, pp. 386–392, 2014.

[20] H. Kautz, "A formal theory of plan recognition," 1987.

[21] M. Ramırez and H. Geffner, "Plan recognition as planning," in *Proceedings of the 21st international joint conference on Artifical intelligence. Morgan Kaufmann Publishers Inc*, 2009, pp. 1778–1783.