Proceedings of the Sixteenth AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE-20)

PAIndemic: A Planning Agent for Pandemic

Pablo Sauma-Chacón, Markus Eger
University of Costa Rica
San José, Costa Rica
{pablo.saumachacon, markus.eger}@ucr.ac.cr

Abstract

Cooperative games are an important challenge for AI research. One example of this genre of games is the board game Pandemic, characterized by the challenges it presents, including its considerable search space and hidden information, especially when the cooperators are not able to communicate explicitly. The game pushes players to prioritize between short term and long term objectives, which forces them to not only plan their own individual actions but also to cooperate with the other players in order to win the game. In this paper we present a planning agent which uses state-space planning with a domain-specific heuristic, combined with a Monte Carlo sampling approach to predict possible outcomes in the face of hidden information. We performed several experiments with our agent, including a comparison with a baseline version that does not use planning. Our experiments showed that our agent is able to win about a third of the games played in a realistic game setup.

Introduction

Games provide an ideal environment for AI research, because they grant researchers a controlled environment with a specific set of rules and, usually, a scoring system that eases the evaluation of the agents (Bowling et al. 2006; Bellemare et al. 2013).

Previous applications of AI research in games were mainly focused on competitive game play, be it as part of the game itself (controlling the non-playable characters) or as a way to compete against experienced players in various games, as is the case of Deep Blue for Chess (Campbell, Hoane, and Hsu 2002), Alpha Go for Go (Silver et al. 2016), Alpha Star for Starcraft II (Vinyals et al. 2019), or OpenAI Five for DotA 2 (Berner et al. 2019). However, when dealing with cooperative environments, in which players must work together to achieve a common goal, research has been more limited and often focused on explicit communication.

In this paper we will describe agents that use planning to play the cooperative game Pandemic, which we previously proposed as a new domain for AI research because provides new challenges regarding cooperation (Sauma Chacón and Eger 2019). The game presents the players with the problem of competing goals, hidden information and a large action space which make the problem challenging. We treat the game as a closed system, explicitly excluding communication between the players about the game state, in order to focus on cooperation by observation. We will first explain the game, and the restrictions we placed on it, then describe our agent design and how it approaches the game, before we present the results of our experiments.

Pandemic

Pandemic is a cooperative board game developed by Matt Leacock (Leacock 2008) in which a group of two to four players must work together to discover the cures to four diseases spreading through the world. The game is played on a board with the image of a map of the world with forty-eight important cities connected to one another through paths. Each city has a specific color which defines which of the four diseases will appear in the city when an infection takes place. Each player is represented by a meeple, which shows the current location (=city) of the player on the board. There are seven roles that give the players bonuses. Each players is assigned a different role at the beginning of the game.

The game makes use of two main card decks: the infection deck and the player card deck. The *infection cards* make up the infection deck and are used to randomize the process in which the next city to get an infection is chosen. There is one infection card for each city. The *player cards*, together with the *event cards* and *epidemic cards* are grouped into a single deck to make up the player card deck. The *player cards*, of which there is also one per city, are used by the players to discover the cures to the diseases, build research stations, can be traded between players and can be used to move quickly from one place to another. The *event cards* allow a player to perform a special action particular to each such card. The *epidemic cards*, when revealed, cause an epidemic to happen in a city, spreading one of the diseases to that city. The city which is being infected this way is selected by getting the bottom card of the infection deck.

The game keeps track of the infections in each city by using disease cubes of four different colors to represent infections of the four illnesses. When a city is getting infected a disease cube of the disease's respective color is added to the city. Whenever a city is going to get infected and would get a total of more than three cubes of the same color an *outbreak* occurs. During an outbreak each of the neighboring cities is infected and also gets another disease cube. This may cause chains of outbreaks to occur, if one of the neighboring cities already had three cubes on it, which must be resolved in order to continue. An epidemic causes a city to be infected by three cubes of its own color and then the infection deck discard pile is shuffled and added on top of the infection deck. This results in the same cities that were already affected by the diseases to be affected again when the next cards are drawn.

The players have to work together to discover the cure to the four diseases. The challenge of the game Pandemic comes from keeping the diseases at bay while the main objective of finding the cures is being completed. The players can lose if any one of three conditions is met: The diseases spread too much (there are more than twenty-four disease cubes of the same color on the board), a "world panic occurs" (eight or more outbreaks happen throughout the game) or the players run out of time (there are no more cards in the player card deck and a player needs to draw).

Each turn, one of the players (taking turns in order) will perform four actions, will draw two player cards and draw a given number of cards from the infection deck, infecting the corresponding cities. For their actions, the players may choose among four different types of movement actions (some of them requiring the use of a player card or having cities with research stations), treating a disease (removing an infection cube from the city they are currently in), building a research station in the current city (by discarding a player card with its name), transferring player cards to or from other players (if both players are on the same city and the transferred card is the one of the current city) and discovering a cure (which requires five cards of the same color).

The players are assigned one of the seven different roles at the beginning of the game which grant them special bonuses:

- The *Scientist* can discover cures using only four cards of the same color.
- The *Medic* removes all infection cubes of a color when treating a disease and does this for free after the cure for the disease is discovered.
- The *Quarantine Specialist* prevents infections from happening in their current city and all neighboring cities.
- The *Researcher* can transfer their cards to other players regardless of the city name.
- The *Dispatcher* can move other players' meeples as their own and rally a player to another's position.
- The *Operations Expert* can build research stations without discarding cards and move from a city with a research station to any other.
- The *Contingency Planner* can keep an additional event card in their hand and reuse discarded event cards a second time.

To be successful, the players must move around the cities on the board, trying to prevent the spread of the diseases, transfer cards among them (which requires them to be in the same city as that card), and manage to discover cures before any of the loss conditions are triggered. This forces the players to make decisions about whether to spread out across the board to contain the disease or stay together to help each other and exchange cards. The players must reevaluate this decision every turn to adapt to the new situation, which makes the game an interesting challenge for players and as a research problem.

Restricted Pandemic

Pandemic, as a game, presents a major challenge for AI research. In order to reduce the complexity of the game, while still keeping the core gameplay intact, we placed some limitations on the game. First of all, the event cards, which a player can play at any moment (even between steps of a phase) were removed. The removal of these cards does not impact the core mechanics of the game since the main objectives and player actions remain the same. However, their removal has a slight impact on the game, since having fewer cards in the player deck results in fewer turns for the players to win the game (aside from not having the event cards which are an additional tool for the players). Many of the event cards, aside from being playable at any given moment, have effects that can target the whole map (setting a research station at any city or moving a player anywhere on the board) which greatly increased the complexity of the game by the amount of options available to the player and the possible times in which it could be played.

We limited our work to four out of seven roles, which will still provide significant variation in game play. The roles removed from the original game were the Dispatcher and the Operations Expert because of how they greatly increase the number of available movement actions. The Contingency Planner was removed too, since it has abilities related to the use of event cards.

The focus of our research was game play and the cooperation itself, and as such we view every player as its own, independent entity. In a tabletop setting, players may talk freely, but our version of the game does not allow for any explicit communication between the players, and instead collaboration must rely on observing the other players' actions and reacting to them appropriately.

Related Work

There has been an increasing interest in the study of cooperative games in recent years. Perhaps one of the most studied games in this area is the game Hanabi (Bauza 2010), which faces the players with the challenge of working together in a game of restricted communication and unknown information. Hanabi has been proposed as an ongoing challenge for AI research (Bard et al. 2019). Research on Hanabi focuses mainly in its communicative aspects. Therefore agents have mainly relied on using protocols to convey information (Bouzy 2017) or communication theory to address the problem (Eger, Martens, and Cordoba 2017).

There have also been efforts to utilize planning for AI agents in games. Planning is the process by which, given an initial state, a set of actions that can be performed and a goal condition, a sequence of actions is found that transform the initial state into a state that satisfies the goal (Fikes and Nilsson 1971; Weld 1994). Planning has been used in games with great success as is the case of the game F.E.A.R were agents are able to plan ahead of the player to devise a plan to defeat them (Orkin 2006). This approach proved successful especially because of the versatility it provided towards different types of agents: the planning process can be performed regardless of the particular characteristics of the various enemies in the game, and a single planner can therefore be used by all the different kinds of enemies designed for the game.

Similar approaches to planning have been taken in future games, like in Shadow of Mordor (Higley 2015) and in Tomb Rider (Conway 2015). While these planning agents were competing with the human player, they implemented some degree of cooperation between the different non-player characters by communicating their plans.

However, in the previous mentioned examples, the planning agents engage in systems with little to no hidden information, which is not the case in Pandemic. For games with hidden information, which make the future unpredictable, other approaches have proven more useful, such as Monte Carlo Tree Search (Coulom 2006) which performs simulations of the game, building up information to decide the best course of action. This kind of approach has been used successfully in card games like in Poker (Van den Broeck, Driessens, and Ramon 2009) and Magic: The Gathering (Cowling, Ward, and Powley 2012).

While there has been some prior treatment of Pandemic in a research contest, the game has been approached in a rather different way than what we propose. Berland et al. used the game to study computational thinking among groups of human players (Berland and Lee 2011), while Wallace et al. evaluated the impact of automation in digital tabletop board games in various degrees (Wallace et al. 2012). The game has previously been proposed as a domain of interest for artificial intelligence research (Sauma Chacón and Eger 2019). More recently, the development of an agent for the game was performed using the Rolling Horizon Evolutionary Algorithm (RHEA) obtaining a winrate of up to 22% in games with randomized setups (Sfikas and Liapis 2020). This approach used a macro-action encoding to reduce the complexity of the plans and used an evolutionary algorithm to develop a better plan to be taken by the agent. Similar to our work, the authors also opted to remove event cards, and simplify the game. However, our approach does not need a macro-action encoding, and outperforms their agents in almost all scenarios.

Planning Agent

This section describes the agent we developed to play Pandemic, which uses a planning-based approach that we augmented with a sampling strategy to account for the unknown information. We will first describe how game states are encoded in our game and how game actions can be use to expand these states resulting in a state-space encoding of the game. Afterwards, we will discuss which goals our agent may choose to pursue, and how it determines which one to use in a particular situation. To handle the unknown information present in the game, we will then elaborate on the sampling procedure used by the agent, before describing the state evaluation heuristic we use to aid the search process.

State Space Search

In order to perform a planning procedure, our agent needs a representation of the game state and the actions it can use, and which would be the resulting state of applying an action. We can think of a game state as a node in a graph and all the possible actions as the edges which connect a game state to all "neighboring" game states which would result from taking each specific action in the current state. This representation allows us to view the problem as that of traversing the graph to a desired destination. To this end, we perform planning using the A* algorithm which requires a heuristic, to assign a value to each game state, and goal function to evaluate if a given state satisfies the desired objectives. When a state that satisfies the goal is found, the sequence of actions (edges) taken by the agent to arrive at it is the resulting plan to be executed.

However, there are some challenges to planning which are imposed by the game itself that must be taken into account. Not all of the information is known to the player at any given moment. The two card decks of the game are a source of unknown information for the player. This increases the difficulty planning by making the outcome of a given plan unpredictable.

Our game state representation allows the agent to know which cards are currently in each players' hand, their position, the different infections among cities, the cures discovered, the current turn and actions remaining. For the unknown information we only keep track of the cards still remaining in the decks, and any known order should there be one (e.g: after shuffling the discard into the top of the deck in an epidemic). This representation also allows the agent to know which actions are currently available at that point in the game and the states that would be obtained for performing each of those actions (e.g: treating a disease would reduce the number of infections in that city).

In many cases, trying to plan until the end of the game becomes unfeasible given how quickly the search space grows with each action to be taken by the players (about six-fold for each action). Furthermore, the game presents the player with the problem of deciding which objective to focus on: should they focus on discovering the cures or should they focus on fighting the diseases. This objective is quickly changing during the game, as the players change and adapt their strategies each time new information is revealed.

Competing Goals

We defined two different goals for our agent to decide between: the *discover cure* goal and the *survive* goal. The *discover cure* goal is directly related to the main objective of the game, discovering all four cures. The goal is satisfied when a state is found that has one more discovered cure than the current game state. The *survive* goal is directly related to the maintenance goals required to not lose the game: the prevention of the spreading of the diseases. This goal is considered to be satisfied when the agent managed to reach a state that is two turns ahead of the current state (it managed to *survive*). However, as there will usually be several states satisfying this condition, the agent uses its heuristic evaluation of state to choose the most favorable one.

With these two goals available, our agents uses a rule based approach to choose its current goal. Whenever the agent must perform an action, and new information has become available since the last planning pass, an evaluation of the game state will be performed to choose the agent's goal. The agent evaluates if it is possible to discover a cure with the current cards. If that is the case, it chooses the *discover cure* goal, otherwise it chooses the *surviving* goal.

Unknown Information

As previously mentioned, the two decks used in Pandemic (player and infection decks) make the game have a degree of uncertainty because their order is (mostly) unknown. This hidden information makes the results of turns unpredictable since players rarely know for certain which cards are going to show up during their draws and as such are forced to "hope for the best" for their player cards and "prepare for the worst" for the infection cards.

To handle the problem of the uncertainty related to the card draws, a Monte Carlo sampling approach is used by our agent to simulate the card draws performed at the end of the turn. During the planning process, the drawing and infection phases are skipped, only coming up with a plan with the information that is currently known. When a possible goal state is encountered it is evaluated performing a hundred rollouts simulating possible card draws and evaluating the resulting states. The expected value is then used as the value of the goal state and the planning continues, allowing for other plans which may have a better expected result to be found before concluding the planning.

State Evaluation Heuristic

Our agent depends on a way to evaluate the states it can visit to guide the search procedure and when comparing different potential goal states it reaches. This is necessary as its goals do not represent the end of the game and therefore a way to find more desirable states (closer to winning the game) is necessary. For this purpose our agent uses a heuristic function which evaluates a game state using different terms which it tries to minimize.

The agent uses two main in-game distance values which are used depending on the current goal of the agent. The first one, seen in equation 1 measures the distance to all the cities times the number of disease cubes (*infection*) present in them, these are then averaged over the total number of disease cubes in the game. The purpose of this equation is to motivate the agent to remain closer to groups of cities with high infection count when pursuing its "survive" goal. The second distance term, seen in equation 2, calculates the distance to the closest city with a research station in it. This rewards the player for staying close to research stations when trying to discover a cure ("discover cure" goal).

$$h_{dsurv} = \sum_{p}^{Player} \frac{\sum_{c}^{City} \text{distance}(p,c) \cdot \text{infection}(c)}{\sum_{c}^{City} \text{infection}(c)} \tag{1}$$

$$h_{dcure} = \sum_{p}^{Player} \min_{c \in City \wedge c_{RS}} \text{distance}(p,c) \tag{2}$$

To measure the value of the player cards in the game, the agent uses an equation for the cards in the players' hands and another equation for the cards in the discard pile. The value of the players' hands is calculated, as seen in equation 3, as the minimum number of cards missing to discover a cure for each disease color among the players' hands (R being 4 for the Scientist and 5 for every other player). This has the effect of motivating the grouping of cards of the same color. The value of the cards in the discard is calculated, as shown in equation 4, as the sum of the number of discarded cards for each of the active diseases still missing a cure.

$$h_{cards} = \sum_{k}^{Color} active_k \cdot \min_{p \in Player} R - \text{cards}(p,k) \tag{3}$$

$$h_{disc} = \sum_{k}^{Color} active_k \cdot \text{discard}(k) \tag{4}$$

The total number of infections, as seen in equation 5 is used as a means to motivate the agent to control the propagation of the diseases. As to give a value to the construction of research station, whose value is more related to its ability to reduce the required number of future actions, the agent uses the value present in equation 6, which calculates the average distance required to move from each city to another city with a linearly decreasing value associated with the number of turns remaining (taken from the amount of cards still remaining in the player deck). Lastly, the final measure, presented in equation 7, counts the number of active diseases which are still lacking a cure. This value is strictly related to the main objective of the game which requires the players to discover all four cures (minimize the value of the term to 0) to win.

$$h_{inf} = \sum_{c}^{City} \text{infection}(c) \tag{5}$$

$$h_{dist} = \sum_{c_1}^{City} \sum_{c_2}^{City} \frac{\text{distance}(c_1,c_2)}{48 \cdot 47} \cdot \frac{turns_{remaining}}{turns_{max}} \tag{6}$$

$$h_{cures} = \sum_{k}^{Color} active_k \tag{7}$$

When evaluating a state, the agent takes the terms defined by these equations and multiplies each of them by a weight to give different importance to the different terms, and sums them to calculate the heuristic value of a state. These weights

Table 1: Winrate for the heuristic agent over 3000 games with varying amount of players and difficulty ($\pm 0.54\%$ with a confidence of 95%).

Epidemics	4	5	6
2-players	2.27%	0.83%	0.37%
3-players	0.60%	0.10%	0.03%
4-players	0.03%	0.00%	0.00%

can be seen as hyper-parameters of our heuristic. These values are used in the search process when executing the A* algorithm to perform the planning and when performing the rollouts of the possible card draws to obtain the expected values.

In order to determine the weights we performed a simple grid search procedure resulting in the overall definition of our heuristic as shown in equation 8. The values used in the grid search were chosen from preliminary tests.

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

Results

In order to evaluate the performance of our agent, we performed an experiment in which it played 3000 games with itself for each of several different setups, varying the number of players and the number of epidemic cards in the player deck. Concretely, we performed experiments with each combination of 2, 3 and 4 players, using random possible role assignments of the four roles Scientist, Researcher, Quarantine Specialist and Medic to the players (which means, that for games with 2 players there were $\binom{4}{2} = 6$ different combinations, while a 4 player game only affords on possible combination), as well as a number of pandemics equal to 4, 5 and 6.

We implemented a simple heuristic agent as baseline, based on the state evaluation heuristic presented above. The heuristic agent will greedily choose its actions based on the immediate best option, as measured by the evaluation function (equation 8), minimizing the cost of the next state.

We measure the success of the agent using two metrics: the winrate and the average number of cures that were discovered in a game. The winrate is calculated as the percentage of games won by the agent, while the average discovered cures per game allow us to predict how close the players were to winning the game. The winrate for the heuristic agent is shown in table 1 and the winrate of the planning agent over 3000 games is shown in table 2, while the results for the average cures per game can be seen in the table 3.

When comparing the winrate results of the heuristic and the planning agent it becomes evident that a naive, greedy approach does not perform very well, but that planning greatly increases the amount of games won. When playing with two players it is still possible for the heuristic agent to win the game, though it is unlikely, only winning around 1 out of 50 games. Planning seems to be a crucial part of this game since it is required for cooperation (what will my teammate do? how can I assist with it?), while also playing an important role in preparing for the possible outcome of events.

Table 2: Winrate for the planning agent among 3000 games with varying amount of players and difficulty ($\pm 1.70\%$ with a confidence of 95%).

Epidemics	4	5	6
2-players	34.27%	17.63%	5.67%
3-players	21.73%	8.73%	2.53%
4-players	15.60%	5.13%	1.70%

Table 3: Average number of cures discovered by the planning agent among 3000 games with varying amount of players and difficulty (± 0.04 with a confidence of 95%).

Epidemics	4	5	6
2-players	2.80	2.27	1.67
3-players	2.41	1.83	1.30
4-players	2.22	1.54	0.92

As for the results of both of the tables of the planning agent, two different patterns can be seen: as the number of players increases both the winrate and the average number of discovered cures decrease; and the same happens when increasing the number of epidemics in the game. The reduction of the winrate and the average discovered cures caused by the increment of the epidemics was to be expected, simply by the fact that increasing a game's difficulty tends to lower the players' winrate.

The agent's performance in games with more players is caused by each individual player performing fewer actions, which would require more direct cooperation, and perhaps even communication, than our agents are currently attempting. It can also be explained by the fact that the planning agent applies limits to how far ahead it plans (when "surviving"). We limit the search to 2 turns ahead, which means that with a larger amount of players, our agent doesn't take into account the possible plays performed by the other agents.

A final observation regarding these results is how the average number of discovered cures impacts the winrate. For example, in the case of the games with 4 epidemics, the difference between the averages of the games with 2 players and the ones with 4 players is 0.58, however this small reduction represents a reduction of 18.67% in winrate.

The agent's performance varies depending on the assigned roles. In figure 1 the different winrates obtained by the agent in the games with 2 to 4 players and 4 to 6 epidemics with the different combinations are presented.

The Scientist (S) role has an impact in the 2 player games, combinations containing the Scientist have a greater winrate than the others. This overall increase in winrate can be explained by the fact that the Scientist can discover cures with only four cards of the same color, rather than 5 compared to the rest of the players, reducing the number of cards required to win the game up to 20%.

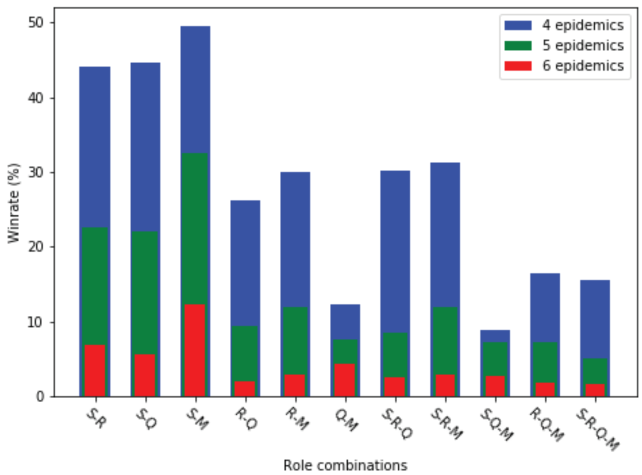

Figure 1: Winrate for the planning agent when using different role combinations. ‘S’ for Scientist, ‘R’ for Researcher, ‘Q’ for Quarantine Specialist and ‘M’ for Medic.

The Researcher (R) role, which can hand their cards to other players more easily than other roles, also seems to have an important effect on the winrate on the combinations without Scientist. The difference can be explained by the fact that while the Scientist and Researcher are cure-focused roles, they lack abilities to help contain the diseases, so while their abilities might help them to win the game more easily, they are failing at preventing the spread of the diseases and losing the game before they can discover the cures. In contrast, the Scientist-Medic combination allows the Scientist to focus solely on discovering the cures while the Medic (M) focuses on preventing the spread of the diseases.This explains why the S-M combination gets a larger winrate than S-R, even though the Researcher plays the biggest role in non-Scientist 2 player combinations. The worst result was obtained by the Medic-Quarantine Specialist (M-Q) combination, which lacks any role specialized in discovering cures which might explain why the combination makes it difficult to win the game.

In the 3 player game however, the Scientist role seems to become second in importance to the Researcher role. This is supported by the fact that when there are more players in the game, each player has fewer turns which also affects the number of cards they draw. The Researcher role can give other players their cards which helps the other players to acquire the cards required to discover cures and win the game, while the Scientist suffers from the reduction in drawn cards. The two combinations of 3 player games in which the Scientist and Researcher participate are the ones with the highest winrates because they can focus on discovering cures while the third player focuses on treating the infections. However, when the Scientist plays with the Medic and Quarantine Specialist it gets the worst result of the four combinations, likely due to the reduction of card draws by the Scientist and the treatment-focus of the team.

Finally, for the 4 player game with a single combination the obtained results are still better than some combinations of the 2 player and 3 player game. This once again can be explained by the fact that, even if there’s a reduction in the number of turns and cards drawn by each player, the synergy of the Scientist and Researcher roles can work in favour of the team while the other players focus on fighting the spread of the disease.

Conclusions

In this paper we presented our implementation of a planning agent for the game Pandemic. We explained how we decided to simplify the mechanics of the game to reduce its complexity, as well as the different challenges an agent faces in the game, while also detailing how we approach each of them.

We presented the results for our agent with varying numbers of players and epidemics, obtaining a winrate of over a third of the played games for the two-player four-epidemic scenario. This agent demonstrates a new way of facing cooperative challenges with unknown information through the use planning and sampling. We demonstrate how planning provides a significant improvement over greedy, heuristic-based approaches when dealing with complex scenarios with competing goals.

There is still ground for improvement in our agent, new heuristic functions can be tested which take into account other factors of the game or use different ways of measuring values. A different set of goals can be defined which might impact the results, as well as refining the goal choosing heuristic. The agent could be upgraded to be able to interact with a higher number of players, since currently it just takes the next player’s turn into account when planning. To this end, plan recognition could be considered as a way to increase the effectiveness of the cooperation by the agent, by allowing it perform teammate modelling.

Cooperative game play with humans can be considered as a future project. This would be done by measuring how the interaction between a human player and a planning agent change the results obtained. To this end, plan recognition could be a useful tool in improving cooperation with human players.

Finally, other approaches without planning can be used for developing an agent for Pandemic. The planning agent can be used as part of a training stage for a neural network approach, be it through reinforced learning or supervised learning. For this reason, the source codes of the agents and experimental setup have been made available publicly for free in the repository: https://github.com/BlopaSc/PAIndemic.

References

Bard, N.; Foerster, J. N.; Chandar, S.; Burch, N.; Lanctot, M.; Song, F.; Parisotto, E.; Dumoulin, V.; Moitra, S.; Hughes, E.; Dunning, I.; Mourad, S.; Larochelle, H.; Bellemare, M. G.; and Bowling, M. 2019. The Hanabi challenge: A new frontier for AI research. *CoRR* abs/1902.00506.

Bauza, A. 2010. Hanabi. https://boardgamegeek.com/boardgame/98778/hanabi. Online; accessed 25 May 2020.

Bellemare, M. G.; Naddaf, Y.; Veness, J.; and Bowling, M. 2013. The arcade learning environment: An evaluation platform for general agents. *Journal of Artificial Intelligence Research* 47:253–279.

Berland, M., and Lee, V. R. 2011. Collaborative strategic board games as a site for distributed computational thinking. *Int. J. Game Based Learn.* 1:65–81.

Berner, C.; Brockman, G.; Chan, B.; Cheung, V.; Debiak, P.; Dennison, C.; Farhi, D.; Fischer, Q.; Hashme, S.; Hesse, C.; Józefowicz, R.; Gray, S.; Olsson, C.; Pachocki, J.; Petrov, M.; Pinto, H.; Raiman, J.; Salimans, T.; Schlatter, J.; and Zhang, S. 2019. Dota 2 with large scale deep reinforcement learning.

Bouzy, B. 2017. Playing hanabi near-optimally. In *ACG*.

Bowling, M.; Fürnkranz, J.; Graepel, T.; and Musick, R. 2006. Machine learning and games. *Machine Learning* 63(3):211–215.

Campbell, M.; Hoane, A.; and Hsu, F.-H. 2002. Deep blue. *Artificial Intelligence* 134(1):57 – 83.

Conway, C. 2015. GOAP in Tomb Rider. In *Game Developers Conference*.

Coulom, R. 2006. Efficient selectivity and backup operators in Monte-Carlo tree search. In *In: Proceedings Computers and Games 2006*. Springer-Verlag.

Cowling, P. I.; Ward, C. D.; and Powley, E. J. 2012. Ensemble determinization in monte carlo tree search for the imperfect information card game magic: The gathering. *IEEE Transactions on Computational Intelligence and AI in Games* 4(4):241–257.

Eger, M.; Martens, C.; and Cordoba, M. A. 2017. An intentional AI for hanabi. In *2017 IEEE Conference on Computational Intelligence and Games (CIG)*, 68–75.

Fikes, R. E., and Nilsson, N. J. 1971. STRIPS: A new approach to the application of theorem proving to problem solving. In *Proceedings of the 2Nd International Joint Conference on Artificial Intelligence*, IJCAI'71, 608–620. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Higley, P. 2015. GOAP from FEAR to Shadow of Mordor. In *Game Developers Conference*.

Leacock, M. 2008. Pandemic. https://boardgamegeek.com/boardgame/30549/pandemic. Online; accessed 25 May 2020.

Orkin, J. 2006. Three states and a plan: The A.I. of F.E.A.R.

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

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

Silver, D.; Huang, A.; Maddison, C. J.; Guez, A.; Sifre, L.; van den Driessche, G.; Schrittwieser, J.; Antonoglou, I.; Panneershelvam, V.; Lanctot, M.; Dieleman, S.; Grewe, D.; Nham, J.; Kalchbrenner, N.; Sutskever, I.; Lillicrap, T.; Leach, M.; Kavukcuoglu, K.; Graepel, T.; and Hassabis, D. 2016. Mastering the game of go with deep neural networks and tree search. *Nature* 529(7587):484–489.

Van den Broeck, G.; Driessens, K.; and Ramon, J. 2009. Monte-carlo tree search in poker using expected reward distributions. In Zhou, Z.-H., and Washio, T., eds., *Advances in Machine Learning*, 367–381. Berlin, Heidelberg: Springer Berlin Heidelberg.

Vinyals, O.; Babuschkin, I.; Chung, J.; Mathieu, M.; Jaderberg, M.; Czarnecki, W.; Dudzik, A.; Huang, A.; Georgiev, P.; Powell, R.; Ewalds, T.; Horgan, D.; Kroiss, M.; Danihelka, I.; Agapiou, J.; Oh, J.; Dalibard, V.; Choi, D.; Sifre, L.; Sulsky, Y.; Vezhnevets, S.; Molloy, J.; Cai, T.; Budden, D.; Paine, T.; Gulcehre, C.; Wang, Z.; Pfaff, T.; Pohlen, T.; Wu, Y.; Yogatama, D.; Cohen, J.; McKinney, K.; Smith, O.; Schaul, T.; Lillicrap, T.; Apps, C.; Kavukcuoglu, K.; Hassabis, D.; and Silver, D. 2019. AlphaStar: Mastering the real-time strategy game StarCraft II. https://deepmind.com/blog/alphastar-mastering-real-time-strategy-game-starcraft-ii/. Online; accessed 5 September 2019.

Wallace, J. R.; Pape, J.; Chang, Y.-L. B.; McClelland, P. J.; Graham, T. N.; Scott, S. D.; and Hancock, M. 2012. Exploring automation in digital tabletop board game. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work Companion*, CSCW '12, 231–234. New York, NY, USA: ACM.

Weld, D. S. 1994. An introduction to least commitment planning. *AI magazine* 15(4):27.