Implementation of the *“Cheat!”* card game based on public announcements Group 10

Rares Oancea (s3974537)
Andra Minculescu (s3993507)
Adrian Serbanescu (s3944735)

June 2023

1 Introduction

This report conducts an analysis of the card game known as *“Cheat!”*. Traditionally, the game is usually played with three or more players, and a full deck of 52 cards is used. The deck is shuffled and then the cards are dealt one by one to each player, one card to one player at a time. The number of players in the game is not relevant, as the player that is dealt the first card will be the one to start the game. To this extent, it does not matter whether the total number of cards is divisible by the player count or not. This means that if there is an even number of players, then every player will have the same amount of cards. On the other hand, if there is an odd number of players taking part in the game, then the player that is dealt the first card (i.e. the one that will also begin the game) will have one extra card compared to the other players. Once the cards have been dealt, an order is selected, starting from the player who received the first card, either by going from left to right or counter-clockwise within the group. The first player starts by setting down one or more cards onto a common pile and states what is the rank of the cards he played. In this game, the ranks that are stated are 2, 3, 4, 5, 6, 7, 8, 9, 10, Jack (J), Queen (Q), King (K), and Ace (A). The important aspect that has to be kept in mind is that the player who begins the round has to claim the same rank for the played card(s). This means that, if one player puts three cards down for instance (e.g. three Queens or one 2, one Ace, and one 10), he can only claim a single rank for all cards. The claim made by the player can either be truthful, partially true (only one or some of the cards correspond to the claimed cards), or completely false (the claimed rank does not correspond to any of the cards played). Whether the card represents an Ace of Hearts or a 10 or Spades for instance is not relevant. The next player has to match the previous claim and place at least one card of the same rank onto the common pile or play one or more cards that have an incorrect rank while claiming it/they have the correct rank.

The main trick of the game is that a player can either play a card of the correct, claimed rank, play a wrong card while lying about its rank, or call the previous player’s bluff. If the player calling the bluff proves to be correct (i.e. the previous player lied about the played rank), then the player who lied has to take the entire common pile. On the other hand, if the player who was called out played truthfully (i.e. played card(s) that match the rank stated at the beginning of the round), then the player who called a bluff has to take the pile. The game goes on in this described manner, with players placing cards on the common pile and having to state each time the (alleged) rank of the played cards, the winner being the first player to get rid of all the cards in his/her hand.

The reason why we chose this specific card game is that it displays the concept of public announcements, re-occurring in each round. More specifically, it tackles the topic of *lying in* *public announcements*, and since this is a very commonly played game, we have decided to implement it along with this mechanic and observe the behavior and performance of logical agents playing it. The public announcements in question are the announcements made by the agents with regards to the rank of the card(s) they claim to have played, which can be seen as announcements of the type "Player 1 holds an Ace." before playing the card.

The mechanics of the card game that we chose to investigate require a proper definition of what exactly is a lie when considering logical agents making truthful or false claims in public announcements. Such an overview can be found in the study conducted by Van Ditmarsch (2014), which provides a thorough definition of the modal dynamics of lying. The authors consider an atomic proposition p and the process of lying involves some agent X telling an agent Y that p when X believes p to be false, with the intention of making Y believe p is true. If the intention of X is considered to be a success, then Y believes that p and also that X's announcement of p was truthful. There are furthermore some preconditions with regards to lies made by X to Y that p. Those preconditions can be formulated with the help of dynamic epistemic logic (Van Ditmarsch et al., 2007), and for instance, as stated by Van Ditmarsch (2014), the precondition for "X is lying to Y that p" is the following: $B_X \neg p$. This can be contrasted to the announcement itself, as when X believes $\neg p$ but still announces p, Y believes p when X announces p, To this extent, the same precondition $B_X \neg p$ applies to the announcement as well. Intuitively, the precondition for a truthful announcement made by X will be $B_X p$. According to Van Ditmarsch (2014), there is also a third scenario, as "in modal logic there are always three instead of two possibilities", and therefore one can be uncertain regarding the truth value of p. For the scenario described in this paragraph, this would mean that Y is uncertain about p, and this corresponds to the following precondition $\neg(B_X p \vee B_X \neg p)$. The authors of the paper refer to this precondition as a "bluffing" announcement.

This project will investigate the behavior and performance of three logical aspects when tasked with playing this game. Each agent will make use of a different strategy when playing, namely an agent who will adopt a "Trusting" strategy, another agent who will make use of a "Distrusting" strategy, and finally an agent with a "Hybrid" strategy, which is a combination of the two previous strategies. These strategies that the agents will use mainly revolve around the way they will play during the game regarding whether they will believe the public announcements made by the other agents or not. They will be explained in detail in Section 2.3. To this extent, we came up with a research question for which we aim to provide an answer by running simulations of the game and observing the overall performance of each type of agent. The research question we are proposing is "How will each of the playing strategies, trusting, distrusting, and hybrid, impact the overall performance of logical agents playing the *"Cheat!"* card game?".

2 Methods

2.1 Simplified Version of the Game

The game itself and the logical representation with Kripke models will be overly complex if we used the traditional version of the game. To this extent, in order to achieve a manageable implementation, design the logical model, and focus on analyzing the difference in performance between the three strategies, we are going to implement a simplified version of the *"Cheat!"* card game.

In this version of the game, there will only be three (AI) agents that play the game. There will also be only two ranks of cards, namely Aces (A) and Queens (Q) and there will only be three cards of each rank in play. For every round, therefore, every agent will start with two cards. The three types of agents will be: "Trusting" agent, "Distrusting" agent, and "Hybrid" agent. The types of agents will be explained in Section 2.3. When the game starts, a random agent is assigned to play the first hand. The game ends once an agent does not have any more cards in their hand.

Another simplification that will be made to the game revolves around how many cards an agent can play during its turn. For the traditional *"Cheat!"* card game players can play more than one card on the same turn. For the sake of simplicity, however, the agents in our simulation will only be able to play one card at a time. This will also prevent the game from ending too suddenly since there are only six cards in play.

An agent can only play a card of the correct rank (i.e. the rank declared by the agent starting the round) if it has a card with that specific rank. If this is not the case, the agent must then lie about the card it plays or call a bluff on the previous player. When an agent is caught lying about the card it played, it must then take all the cards from the common pile and the agent who called the bluff begins the next round. On the other hand, if an agent is truthful about the card it played, but another agent calls it a bluff, the agent that called the bluff must take the cards from the common pile and the agent that played the correct card begins the next turn. Another important rule is that one can only call a bluff on the previous agent that played a card. This also means that an agent can choose to call a bluff only during its turn.

2.2 Decision Points

Throughout the simulation, the three agents will have to tackle three main decision points. All scenarios stem from the situation in which an agent has to play a card when its turn comes around. This can either happen in the middle of a round, with some cards already having been placed on the common pile, or at the start of the round when the agent places the first card for the round. The decision points can be described as follows:

1. What to do when you have the card matching the true rank?
   (a) Be truthful: agent plays the correct card.
   (b) Lie: agent plays a card with the wrong rank while claiming he has played a card with the correct rank.
   (c) *"Cheat!"*: agent calls a bluff on the agent who has played a card previous to him.
2. What to do when you don't have a card matching the true rank?
   (a) Lie: agent plays a card with the wrong rank while claiming he has played a card with the correct rank.
   (b) *"Cheat!"*: agent calls a bluff on the agent who has played a card previous to him.
3. What to do when you start?
   (a) Be truthful: agent plays a card and declares the true rank of the card.
   (b) Lie: agent plays a card and declares a different rank.

2.3 Agent Types

Our simulation of the *Cheat* card game will make use of the following three agent types:

1. **Trusting Agent:**
   The "Trusting" agent believes all the public announcements to be true until there is a contradiction in the gathered knowledge and the current world/state. Thus, as it considers all announcements (of the type "X holds an Ace") to be truthful, it will only call *"Cheat!"* when it encounters a contradiction. For example, if a pile of four cards was announced

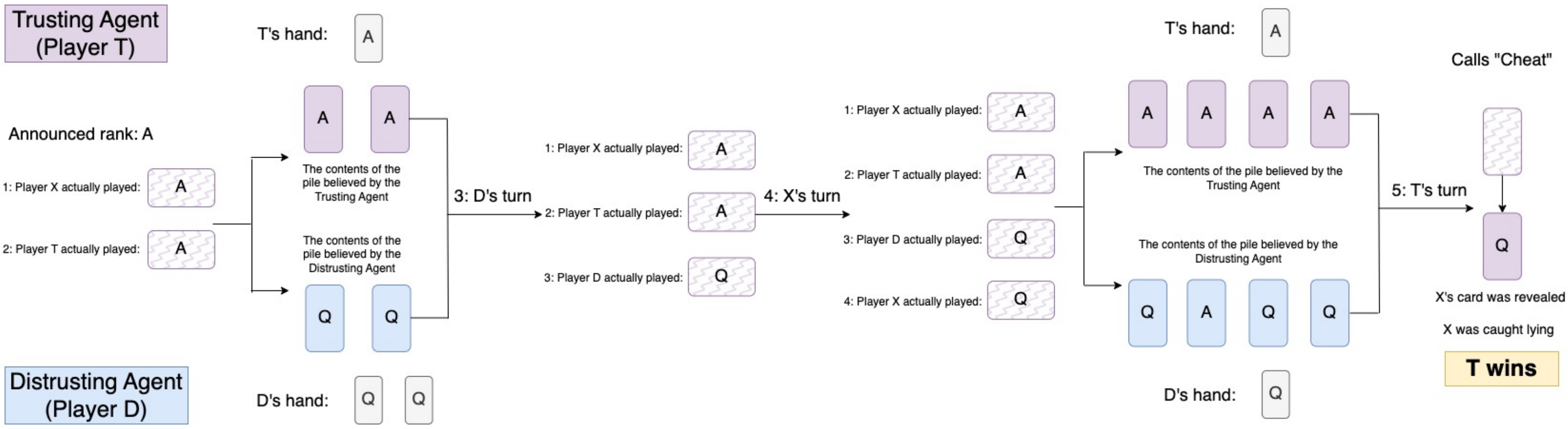

Figure 1: An example of a game of "Cheat" with three agents. The initial deck of cards was of length six, containing three Aces and three Queens. The example focuses on two agents following the Trusting (T) and Distrusting (D) strategies. It displays the belief regarding the pile contents before either of their turns occurs. The remaining player (X) is unknown and serves the role of a random opponent.

by the other players only to contain *Aces*, but the trusting agent knows that there are only 3 *Aces* in the game, it will call *"Cheat!"* based on the encountered contradiction.

Thus, as Figure 1 depicts, the Trusting player (T) was the second to place a card. As it did not encounter any contradiction based on its beliefs regarding the pile contents, it did not call bluff on the previous player and placed a card truthfully. However, when its turn came again (in the fifth round), as it believed all announcements, in its representation of the pile there were four Aces. Moreover, it also still had an Ace in its hand. Thus, as the believed number of Aces is above the total number of cards of this rank in the game, the "Trusting" agent has to remove the current state where it is considered the announcement truthful. Thus, it will now assume that at least two players lied and will call "*Cheat!*" on the previous one. In the given example the game ends after this round as the X player was caught lying resulting in T's turn again. It only has one remaining card and can choose the rank of the next round. Thus, it will place the remaining Ace and announce it truthfully. It does not matter if the next player calls *"Cheat!"* or not as the announcement was truthful and T remained without cards, winning the game.

2. **Distrusting Agent:**
   This agent believes that all public announcements are lies. Thus, if a player announced that it plays an *Ace* (which translates into "The player X has an Ace"), the distrusting agent will believe that a *Queen* was actually played. Thus, this agent is more prone to calling *"Cheat!"* unless a contradiction is encountered. For example, if a player announces that it plays an *Ace* and the distrusting agent has in its hand two *Queens* and believes that the remaining player has the other *Queen*, then it will believe that an *Ace* was actually played and will not call *"Cheat"*.

   In Figure 1, this behavior is depicted. The Distrusting agent (D) was the third one to have its turn. It would normally call *"Cheat!"* on the previous opponent as it believes all announcements to be false. However, its belief regarding the pile's contents contains two Queens while it also has two Queens in its hand. Thus, as it knows that there are only three Queens in the game, it removes the current state where it considered the previous announcement false. Therefore, it will not call *"Cheat!"* and instead update its belief regarding the pile (as it now believes that T played truthfully an Ace).

3. **Hybrid Agent:**
   The "Hybrid" Agent will alternate between the trusting and distrusting policies in a

weighted random manner. To implement the weighted alternation, a success factor is assigned to each strategy based on its relative performance. Specifically, if the "Trusting" strategy outperforms the "Distrusting" strategy, the weight factor for the "Trusting" strategy ($\frac{\text{wins\_trusting}}{\text{wins\_total}}$) is increased, resulting in a higher probability of selecting the "Trusting" strategy in the next game. Conversely, if the "Distrusting" strategy performs better, the weight factor for the "Distrusting" strategy ($\frac{\text{wins\_distrusting}}{\text{wins\_total}}$) is increased, leading to a higher likelihood of using the "Distrusting" strategy in the following games.

The weights are updated dynamically after each game based on the performance comparison. The update process ensures that the weight factors reflect the recent performance history, allowing the "Hybrid" strategy to adapt to changing conditions and exploit the more successful strategy. The alternating nature of the "Hybrid" strategy enables it to leverage the strengths of both the "Trusting" and "Distrusting" strategies.

As it was previously mentioned in Section 1, "in modal logic, there are always three instead of two possibilities". For the game we are investigating, this would relate to truthful, lying, and bluffing announcements, similar to what was stated in Van Ditmarsch (2014). However, the types of agents that we are modeling can either believe the public announcements as true or lies. To this extent, the project will only consider two of the precondition, the first one being for a lying announcement, considering an agent X who believes $\neg p$ (i.e. $B_X \neg p$). The second precondition relates to true announcements, in which we have an agent X that believes p (i.e. $B_X p$). The precondition concerned with bluffing announcements (i.e. being uncertain whether p; $\neg(B_X p \vee B_X \neg p)$) will not be considered.

The manner in which the agents have been implemented enables them to figure out when there is a possible contradiction between their knowledge, beliefs, and the current state of the world. This feature will alter the manner in which the agents will play the game, ponder on the truth value of public announcements, and whether they choose to play a card or call bluff. The way in which agents figure out when contradictions occur is similar to the examples provided in the book published by Van Ditmarsch et al. (2015). In chapter three of the book, there are several puzzle examples portraying a scenario with two children, Alice and Bob, that have been playing outside, and when they come back home, their father notices they have been playing in the mud because Bob has mud on his face. The father proceeds to ask each one in turns whether they know they are muddy or not. The two children are only able to see if the other is muddy. The main underlying mechanism here is that each child figures out whether he or she is muddy based on the public announcement made by the other.

In the scenario of the card game of *"Cheat!"*, the agents do not realize facts about themselves when other agents make public announcements but rather figure out the validity of the public announcements based on the current state of the world, their own knowledge, and the announcements themselves. In this sense, the "Trusting", "Distrusting" and "Hybrid" agents use the public announcements to figure out whether a contradiction took place or not. More specifically, the "Trusting" agent tries to identify contradictions in order to know when to call *"Cheat!"* rather than play a card, the "Distrusting" agent aims to identify contradictions to figure out when to not call a bluff instead. The "Hybrid" agent alternates between the two strategies. It does not change strategy mid-game however.

2.4 Kripke Model

In this section, we are going to translate the card game *"Cheat!"* to Kripke models. The initial states of the Kripke model depends on the total number of card ranks in the game and the total number of cards for each rank. Thus, keeping track of all the possible distribution of cards each player can have in hand is impossible. In terms of Kripke models, the game is impossible to track if all the cards are used. Firstly, we are going to introduce the overall formal model using all the ranks and all the cards.

Let N be the set of players with $|N| = n$.
Let P be the set of propositions.
Let C be the set of card ranks of length $|C| = 13$ with $C = \{2, 3, 4, ..., J, Q, K, A\}$.
Let public announcements be: "player j plays the card c" with $c \in C$. This type of public announcement can be seen as an announcement of "player j holds card c" before the card is removed from play. The public announcement will then be denoted as p_{jc}, with p standing for player, j is the number of the player and c is the card that the player has in hand.

We can now define the following Kripke model:

- $M = \langle S, R, \pi \rangle$
- S represents the set of all possible states, which in our case translates as the set of all possible distributions of cards among the players and the pile on the table. S can therefore be defined as $S \in \{s_1, s_2, ..., s_m\}$
- R is the set of accessibility relations. $R = \{R_1, R_2, ..., R_m\}$.
- $\pi : P \rightarrow (S \rightarrow \{T, F\})$ assigns a truth value to all propositions P. For example, $\pi(s_1)(p_{1A}) = T$ or $\pi(s_2)(p_{2A}) = F$. In this example, T stands for true and F stands for false.

One aspect to take into account about the accessibility relations in this model is that they are reflexive, but not symmetrical. Each agent considers their current state, based on the knowledge of their own cards. However, in the case of symmetry, this would mean that if state **s1** is accessible from state **s2**, then state **s2** should also be accessible from state **s1**. However, because in the card game *"Cheat!"* a player could be lying, lying in public announcements, the accessibility relations cannot be symmetrical. Each agent has knowledge about their own cards and different knowledge about the other agents. However, this different knowledge is based on what each agent claims to have played, which can vary based on the type of agent.

2.4.1 Simplified Kripke model

A simplified version of the game is used in order to visualize a Kripke model of a *"Cheat!"* game. The simplified version of the card game *"Cheat!"* consists of only three ranks (Aces, Kings and Queens) with one card for each of these ranks. The Kripke model focuses on the public announcements made by each agent, in order to see how these announcements affect the model. We used this version, in order to be able to take into consideration all the possible distribution of cards among the agents. Firstly, we are going to introduce all the possible states and the accessibility relations between these states, then we are going to look through an example gameplay to see how the Kripke model changes per turn.

- $M = \langle S, \pi, R_1, R_2, R_3 \rangle$ where we have $S = \{s_1, s_2, s_3, s_4, s_5, s_6\}$.
- $|N| = n$, where $n = 3$. $N \in \{p_1, p_2, p_3\}$.
- $C \in \{K, A, Q\}$; $|C| = 3$.
- P is a set of propositions; "player j holds card c" where $c \in C$ and $j \in N$.
- $\pi : P \rightarrow (S \rightarrow \{T, F\})$; π assigns a truth value to all propositions in P.
- $R = \{R_1, R_2, R_3\}$.
- $R_1 = \{\langle s_1, s_2 \rangle, \langle s_3, s_4 \rangle, \langle s_5, s_6 \rangle\}$

- $R_2 = \{\langle s_1, s_6\rangle, \langle s_2, s_4\rangle, \langle s_3, s_5\rangle\}$
- $R_3 = \{\langle s_1, s_3\rangle, \langle s_2, s_5\rangle, \langle s_4, s_5\rangle\}$
- The relations are not symmetrical, but instead are reflexive.
- $s_1 = \{\langle p_1, 1A\rangle, \langle p_2, 1K\rangle, \langle p_3, 1Q\rangle\}$
- $s_2 = \{\langle p_1, 1A\rangle, \langle p_2, 1Q\rangle, \langle p_3, 1K\rangle\}$
- $s_3 = \{\langle p_1, 1K\rangle, \langle p_2, 1A\rangle, \langle p_3, 1Q\rangle\}$
- $s_4 = \{\langle p_1, 1K\rangle, \langle p_2, 1Q\rangle, \langle p_3, 1A\rangle\}$
- $s_5 = \{\langle p_1, 1Q\rangle, \langle p_2, 1A\rangle, \langle p_3, 1K\rangle\}$
- $s_6 = \{\langle p_1, 1Q\rangle, \langle p_2, 1K\rangle, \langle p_3, 1A\rangle\}$

The initial Kripke model can be seen in Figure 2. It contains all the possible states and the accessibility relations between those states. Each node contains the distribution of cards in that state, the order matters. For example, in state **s5**, agent 1 has one Q, agent 2 has one A and agent 3 has one K, this results in **QAK**. Essentially, we use the index (starting from 1) in order to see which card each agents has (e.g. Q corresponds to index 1, which means agent 1 has one Q).

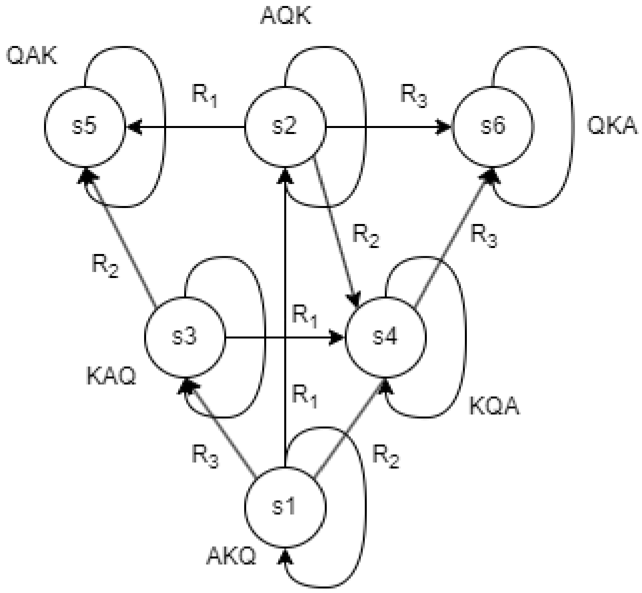

Figure 2: Graphical representation of the initial Kripke model with 3 agents, 3 ranks and one card per rank.

Gameplay In this gameplay, there will be all three types of agents: one "Trusting", one "Distrusting", and one "Hybrid". In this scenario, the first player is a distrusting agent, the second player is a trusting agent and the third player is a hybrid agent. The game starts and the cards are dealt to each player. Player 1 receives one Ace, player 2 receives a King and player 3 receives a Queen. This means that we have p_{1A}, p_{2K} and p_{3Q} in the real world. From these relations and the fact that each player knows its own cards, we have the following relations that hold:

- $K_1 p_{1A}$

- $K_2 p_{2K}$
- $K_3 p_{3Q}$
- $\neg K_1(p_{2K} \wedge p_{3Q})$
- $\neg K_2(p_{1A} \wedge p_{3Q})$
- $\neg K_3(p_{1A} \wedge p_{2K})$

Turn 1: It's player 1's turn and it plays an Ace, claiming to be an Ace. This could be seen as a public announcement "Player 1 holds an Ace", which would remove the following states: **s3**, **s4**, **s5** and **s6**. In the removed states, player 1 has a different rank in hand. The revised Kripke model after this announcement can be seen in Figure 3.

Turn 2: Now it's player 2's turn. Because player 2 is a trusting agent, it believes the announcement "Player 1 holds an Ace". However, because player 2 does not have any Aces in hand, it will play a King and claim it to be an Ace. So, we can see this as a public announcement ¬"Player 2 holds an Ace." Because this a lie, this announcement will not be used to eliminate states, so the Kripke model remains unchanged.

Turn 3: Now it's player 3's turn. Player 3 is a hybrid agent, which means that based on the weighted average, it will decide whether to use a trusting or distrustful strategy. Let's say it will choose a trusting strategy, which means that based on the previous announcements (in the game there is only one Ace, but there were two Aces announced, so it is a contradiction), it will call a bluff on player 2. Now, the cards from the pile are turned and player 2 was caught lying and it takes the cards. Now, each agent knows what the other agents have in hand, so the final Kripke model, see Figure 4, would consist of only state **s1**, which is the real world.

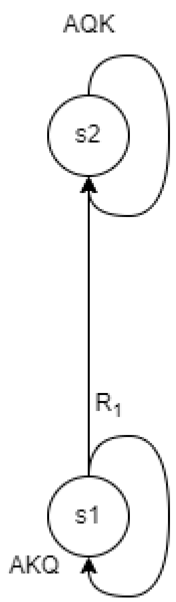

Figure 3: Graphical representation of the Kripke model after Turn 1, with only two remaining states.

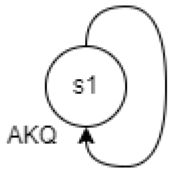

Figure 4: Graphical representation of the final Kripke model with only one state remaining.

2.5 Experiments

The aim of this study is to gauge the performance of the three agent types we are modeling in order to see which of the three strategies is most appropriate for the portrayed scenario playing the card game of *"Cheat!"* with two cards per agent and accounting for a total of two ranks (Aces and Queens). To achieve this, we will be conducting a total of 5 runs in which the agents will play 100.000 games per run. The metric that we will use to measure the performance of each strategy will consist in the number of games won. As it was previously mentioned in section 2.3, the "Hybrid" agent will make use of a weighted random variable which helps it alternate between the "Trusting" and "Distrusting" strategies in order to benefit from both policies over the course of multiple games. This weighted random variable will be reset at the beginning of each of the 5 runs such that every time a run of 100.000 games commences, the "Hybrid" agent starts from scratch. This will prevent the "Hybrid" strategy from gaining an unfair advantage over the other two strategies as the runs unfold. A barplot will be made portraying the performance of the agent for each run.

3 Project Delivery

A website was developed for this project in order to showcase an overall summary, the report, as well as the GitHub repository on which the implementation can be found. The website was designed with the help of GitHub Pages, using JavaScript, CSS, HTML, Markdown and YAML components. The page was tested for the Chrome, Internet Explorer, and Firefox browsers and contains three tabs containing a short summary of the project, the pdf with the report, and a link which takes the user to the GitHub repository which contains the code for the game implementation. The final version of the web page can be found here.

4 Results

After running experiments on the simplified model of the *"Cheat!"* card game, we were able to extract the numerical data that shows us the overall performance of the three agent types.

This numerical information can be seen in Table 1, which shows the exact number of games won by each of the three agents across 5 runs of 100000 games each. Having a look at those numbers, we can state that the "Trusting" and "Hybrid" agents outperform the "Distrusting" one by a significant margin. Out of all three strategies, the "Hybrid" agent seems to win most games across the five runs, getting rid of all cards in about 44500-44600 games. Not far behind, the "Trusting" agent displays a good overall performance, consistently winning around 44000 games. The "Distrusting" strategy however showed the worst performance in the long run, only managing to win around 11000 games during a run.

The discrepancy between the three strategies can also be seen in Figure 5, which shows the winning percentage for the three agents during the course of 100000 games. Just as the figure shows, the "Trusting" and "Hybrid" strategies are really close performance-wise since both of them won more than 40% of the 100000 games. The "Distrusting" agent on the other hand

Agent Type	Number of Games Won				
	Run 1	Run 2	Run 3	Run 4	Run 5
Trusting	44103	44105	44187	44214	44211
Distrusting	11261	11427	11280	11124	11330
Hybrid	44636	44468	44533	44662	44459

Table 1: Results from the 5 runs for every Agent Type

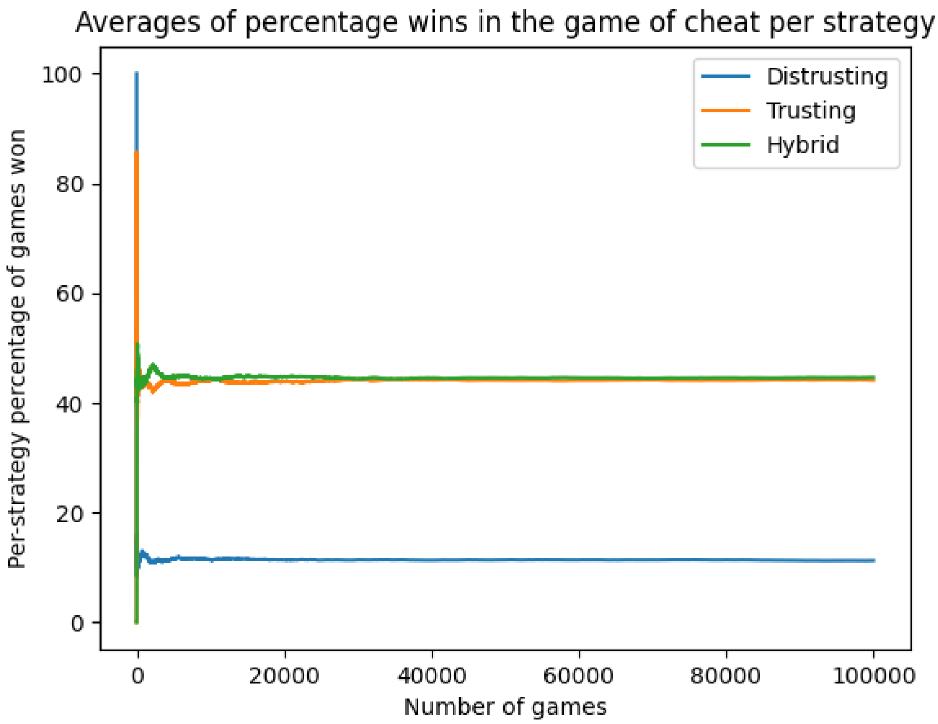

Figure 5: Plot displaying the performance of the three agents across 100000 games

performed significantly worse compared to the other two strategies having won less than 20% of the games.

Figure 6 shows the same overall outcome, this time around considering 5 runs of 100000 games each. In a consistent manner throughout each of those runs, the “Trusting” and “Hybrid” agents win most of the games, the latter outperforming the former by a very small margin each run. In a similar manner, the “Distrusting” strategy under-performs, showing the worst performance out of all three strategies.

Furthermore, performing a Chi-Square test on the number of winnings on each of the five runs of 100000 games each result in a p-value smaller than .00001 (X^2(8, N=500000) = 29719.87). Thus, there is a significant difference between the three strategies in terms of the number of won games (in the five experiments of 100000 games each). An analysis of the contingency table indicates that there is a stronger correlation between the Trusting and Hybrid strategies (justified by the Hybrid's architecture) and a significant difference between these two and the Distrusting agent.

5 Discussion

After analyzing the results obtained from the experiments, we can clearly observe that for this simplified context of the *“Cheat!”* card game, played with three agents, two ranks, and two cards per player, considering all public announcements to be true unless a contradiction is encountered

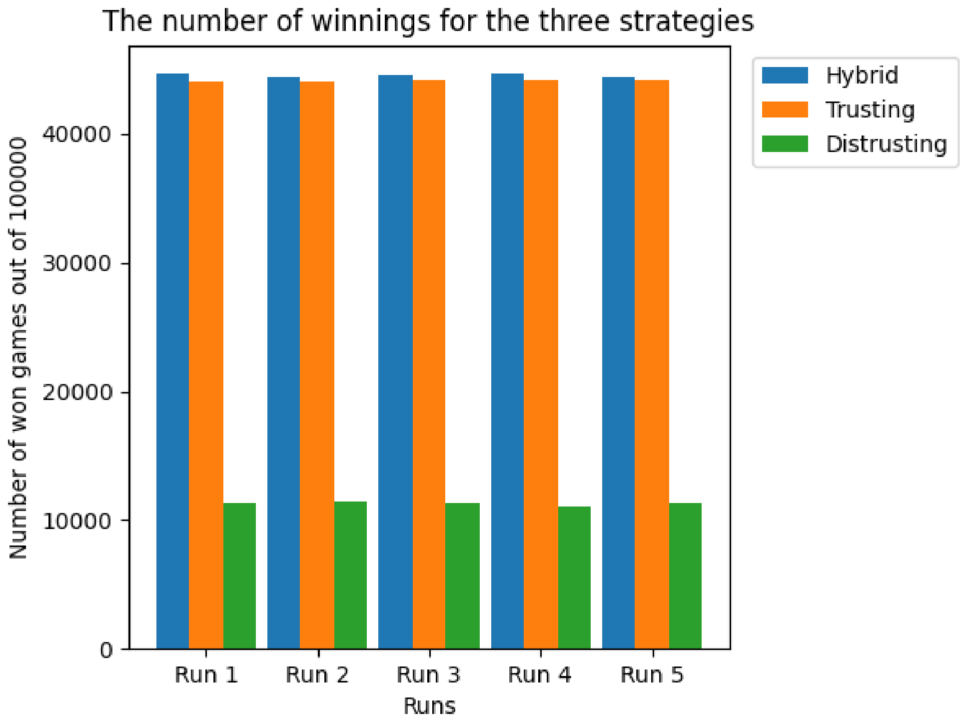

Figure 6: Barplot displaying the number of games won by the three agents across 5 runs of 100000 games each

seems to pay the most dividends in the long run. We can consider this conclusion to be solid since the outcome is consistent across all of the 5 runs, with the "Trusting" and "Hybrid" agents outperforming the "Distrusting" by a significant margin. Moreover, the results from the statistical analysis suggest that there is a significant difference between the "Distrusting" agent and the other two (that are close in performance).

Taking into account both Figures 6 and 5, we can observe that after around 10000 games, the "Hybrid" agent begins to solely adopt the strategy of the "Trusting" agent. This goes on to prove that the strategy of the "Trusting" agent is the best within this context, as it shows consistent performance during each run. We can therefore conclude that trusting all public announcements unless a contradiction is identified results in better outcomes overall, compared to regarding all announcements as false beside the cases in which it detects contradictions. To this extent, taking into account the research question we proposed in Section 1 ("How will each of the playing strategies, trusting, distrusting, and hybrid, impact the overall performance of logical agents playing the *"Cheat!"* card game?"), we can state that the "Trusting" strategy is the one which impacts performance positively. This is also backed up by how the "Hybrid" agent behaves since it prefers to adopt the strategy of the "Trusting" agent in the long run.

5.1 Future work

For future research, it would be essential to look firstly at games that contain all the ranks and all four types of cards. It would be interesting to see if the results will change significantly because of the addition of multiple ranks and all four types of cards for each rank.

Additionally, we saw that the order in which the agents are playing matters with respect to the percentage of games won. This happens due to the fact that an agent can only call *"Cheat!"* to the previous player. Figure 7 depicts the difference between the agents' performance based on their order that remains constant in each of the 1000 games per run. As it can be seen in Figure 7 the agent that is followed by the "Distrusting" player has an advantage. This can be explained by the fact that the "Distrusting" agent considers all announcements false (calling

"Cheat!") until a contradiction is encountered. Thus, as the "Trusting" and "Hybrid" agents do not lie unless they do not have the correct card, they will be falsely accused often. After a false accusation, the turn of the accused agent will follow again. Therefore, they will have the opportunity to discard cards more often, resulting in an increased number of won games. This is the reason why, in our implementation, after each game, the order of the players is randomly shuffled. An interesting future research direction could be an in-depth investigation regarding the order in which the agents start each game and how can this order affect the overall knowledge and performance of the other agents.

Furthermore, an integrated web implementation of the card game *"Cheat!"* could prove useful for testing purposes. This web implementation would also show based on the game turns, the respective Kripke model, from which we can get additional knowledge with respect to what each agent thinks the other players have in hand. This could come in handy if we would like to look into the uncertainty of the public announcements, and how this uncertainty can be tackled by the agents.

Moreover, the addition of multiple strategies or agents varying these strategies through a run of multiple games can be seen as a good future research direction. Some new strategies may look into the uncertainty and decide based on that what strategy to use for future games.

References

Van Ditmarsch, H. (2014). Dynamics of lying. *Synthese*, *191*(5), 745–777.

Van Ditmarsch, H., Kooi, B., van Ditmarsch, H., & Kooi, B. (2015). *One hundred prisoners and a light bulb*. Springer.

Van Ditmarsch, H., van Der Hoek, W., & Kooi, B. (2007). *Dynamic epistemic logic* (Vol. 337). Springer Science & Business Media.

Appendix

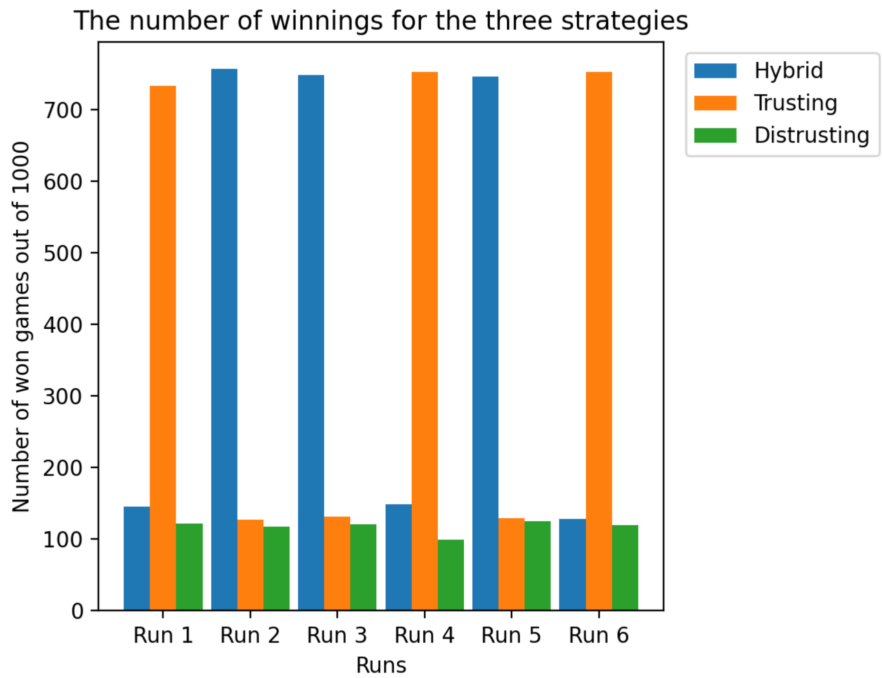

Figure 7: This bar plot displays the number of won games per strategy in 6 different runs of 1000 games each. The runs differ due to the player's order within the game. Each run is constant regarding this aspect. Thus, in all games from Run 1, the "Hybrid" agent begins, followed by the "Trusting" agent, and lastly by the "Distrusting" player. In Run 2, the order is: "Hybrid", "Distrusting" and "Trusting". In Run 3, the order is: "Trusting", "Hybrid", and "Distrusting". In Run 4, the order is: "Distrusting", "Hybrid", and "Trusting". In Run 5, the order is: "Distrusting", "Trusting", and "Hybrid". Lastly, in Run 6, the order is: "Trusting", "Distrusting", and "Hybrid".