university of groningen / faculty of science and engineering / artificial intelligence

The game of skull testing strategies against random play

Bachelor's Project Thesis

Max Oortman Gerlings, s3159187, m.pa.oortman.gerlings@student.rug.nl,
Supervisors: Dr. H.A. de Weerd

Abstract: Games have been a subject in various studies, and one of these well-researched subjects is Theory of Mind. Theory of Mind allows a person to understand the thoughts and mental states of another person. In games, a player will want to predict and know as much about the opponents' thoughts and their possible strategies. For example, theory of mind is needed to bluff effectively, by reasoning about how the beliefs of others can be incorrect. Research has found that bluffing can be a vital part of a player's strategy in games like Poker and Liar's Dice. A similar type of game with little research is the auction game of Skull. In this three to six player game, the players bid on the number of cards they believe they can turn over without revealing a skull card. The study aims to determine whether there is a good strategy against the bluff of another player to win the game of Skull. To measure this, an agent model was constructed to play these games. These agents can play the game with bluffing or counter strategies. The model generated data by playing 1000 games using different agent setups. The results showed that bluffing was a good strategy for a player and improved the chances of winning. However, the counter strategy was not considered to be a better approach. Future research could look into a more complete strategy to counter bluffing players.

1 Introduction

Bluffing can be found in many areas of life, even though we tend to associate it with games more often. One of the games bluffing is always associated with is poker. This has been a topic for many studies. Such as the research of Friedman (1971), which focuses on the optimal bluffing strategies in poker. He showed that calling a potential bluff half of the time can be a good strategy. In another study by Guazzini & Vilone (2009), it was shown that agents in multiple adaptive models would quickly learn to bluff. It even showed that the agent who learned bluffing the best would win more often.

Another type of game where bluffing plays a part is auction games. Bluffing becomes a part of auction games in the bidding stage. During the bidding stage, having the skill of bluffing can be an advantage to a player. One of these auction games is Liar's Dice, also known under the name 'Bluff'. Sum & Chan (2003) looked at this version and how the players should be modelled to include lying and telling the truth. In the model, they went for an approach using probabilities. Another study by Freeman (1989) looked at the tactics and strategies of Liar's Dice. The study did find tactics that allowed the player to catch a liar more often. However, it mentioned that bluffing strategies need to be researched more in the future.

As shown from this previous research, bluffing can significantly benefit a player. When a player is bluffing, it can be hard for an opponent to read this player. Reading other people and knowing their thoughts is called Theory of Mind. Various research has been done on Theory of Mind in a game environment. A study by Goodie et al. (2012) focused on the levels of Theory of Mind players would use in a game. Lower levels imply a player focusing more on their own strategy and less on the opponents. However, in a game where bluffing plays a significant role, it is important that a player focuses on their opponent. In this study, they found that in simpler games, players were able to make realistic assumptions about their opponents by using Theory of Mind reasoning. To prevent other people from reading you, people will use bluffing strategies to deceive their opponents. Bluffing can make it difficult for the opponents to know what a player is doing. Therefore, a counter-strategy that can read a player's bluffing strategy can be beneficial.

In this thesis, we will investigate the game of Skull (Marly, 2011), an auction game in which bluffing can take a significant role. Therefore, the question of this research is, "Does a bluff predicting strategy improve a player's chances of winning the game of Skull?".

1.1 The game of Skull

A game that has excellent potential for bluffing is Skull (Marly, 2011). This card game is played with three to six players and feels similar to Perudo or Liar's Dice. The game Skull is played with four cards per player, consisting of three roses and one skull. The first stage begins with the players taking turns laying down one of their cards face down on their own stack. After laying down the first card, players can lay down a second card or decide to make a bid. This bid is the number of roses a player thinks can be turned over on the table without picking a skull card. When a player makes a bid, the other players can no longer lay down a new card, and the game moves on to the next stage.

In the second stage, players take turns either raising the current highest bid or passing their turn. A player must consider their bid well. When a bid is too high, the player may not be able to satisfy this bid. However, when a bid is too low, another player may win the game instead. In this stage, bluffing can be crucial in tricking others or securing the bid for yourself. The different types of bluffing will be explained in the methods section. Once all but one player has passed, the game continues to the last stage with the highest-bidding player.

In the third stage, the player with the highest bid can start turning over cards and try to satisfy the bid. First, the player will turn over their own cards. If a player has played a skull card themselves in the first stage, they will have to turn over that skull and thus lose their bid. However, if the player did only play roses and turns them over, they can continue to the cards of the opposing players. The player can choose which top card they want to turn over from any of the opponents' stack of cards. The cards on the table may only be turned over from top to bottom. When a player finds a skull, the player loses one randomly selected card from their hand. If the correct number of roses is found, the player earns one point, and a new round begins. The game continues until one player has won two points or when all but one player lost all their cards.

2 Methods

To determine whether detecting bluffing players can be beneficial in the game of Skull, we implement an agent-based model capable of playing this game. The agent model was implemented in Python. The model will provide data on how many wins a player has, the number of points a player has won and how many times a player may have lost a bid.

2.1 Model

The game of Skull has two main parts, which each have their own function in the game. These parts are the players and the game itself. The game part will handle the different stages, keep track of the players, and other important information. The players will perform different actions per stage and will keep track of their own information. How the player makes a decision will be handled differently per strategy.

2.1.1 Game Class

For the game, a class has been made to handle the stages and hold the information. The game keeps track of the following variables:

- `winner`: A Boolean that keeps track if a player has won the game
- `players`: A list of all the active players
- `deadPlayers`: A list of players who have lost all their cards
- `totalCardsPlaced`: Total amount of cards placed on the table
- `winningPlayer`: The winning player
- `passedPlayers`: The number of players who have passed
- `bluffMemory`: A dictionary of the number of times a player lost due to bluffing

- `totalRoundsHighestBidder`: A dictionary of the number of times a player was the highest bidder

These variables are needed to play a round of the game, and some are used for the strategy, which will be explained later.

The game class has the following functions:

- `setupGame()`
- `playRound()`
- `reorderPlayers()`
- `resetGame()`

The `setupGame()` function is used to set up the players and add them to the list of players. After setup, the game will start playing by using the `playRound()` function. The `playRound()` function will play one round of the game by going through all three stages of the game.

2.1.2 Player Class

The player class represents the players and the information they need to keep track of. This class keeps track of the following variables:

- `Name`: The player's name
- `Strategy`: A Boolean that implies if the user will use strategy or not
- `Deck`: A list of the cards that remain in the deck (hand)
- `Placed`: A list of the cards that have been placed on the table
- `Points`: The number of points a player has won that game
- `Passed`: A Boolean that is true when a player has passed
- `Bid`: The last bid that a player has made on the cards on the table
- `RemovedCards`: A list of the cards that the player has lost
- `LostBid`: A Boolean that is true when a player has lost/grabbed a skull
- `PassProbability`: A variable that is 1 or 0 depending on whether a skull has been played
- `OBSLevel`: A variable used to implement the overbid bluffing strategy for the players
- `SBSlevel`: A variable used to implement skull bluffing strategy for the players

These variables are paired with the following functions to make the game work:

- `SetBluffLevels`: Initializes the different bluff levels a player can use
- `ResetPlayer`: This function resets players after each round
- `MakePlay`: Chooses a card to put on the table
- `MakeBid`: Generates a bid for the current number of cards
- `PlayStageOne`: Plays the first stage of the game
- `PlayStageTwo`: Plays the second stage of the game
- `FindBidProbability`: Standard function used to find probability to bid
- `StrategyBidProbability`: Finds probability to bid depending on the strategy
- `PlayStageThree`: Plays the third stage of the game

2.2 Game Stages

The game consists of three stages:

- The first stage: the card placing stage
- The second stage: the bidding stage
- The third stage: the card flipping stage

In the following section, we will discuss each of these stages in more detail.

2.2.1 Stage One

The first stage of the game starts with players laying down cards on their own stack in turns and ends when one player makes the first bid. This bid is the number of rose cards a player thinks can be turned over on the table. If a player decides to make the first bid on the number of roses, the Boolean variable `challenge` is set to true. The variable `challenge` receives a new state after the player has played a turn in the first stage. This turn is done by calling the `playStageOne()` function of the player class. This function has two possible actions: laying down a card or making a bid. No particular strategy has been implemented for this stage. Therefore, when placing a card, this card is chosen at random from the cards left to the player. The first bid can only be made when all players have at least placed one card on their stack. If a bid has been made, the first stage is over, and the second stage begins. The first is made randomly in the range of cards currently placed on the table.

2.2.2 Stage Two

The second stage of the game centres around the bidding stage. The players will start bidding on the number of roses they think can be turned over on the table. This is done in turns, and the players will have to think about whether bidding higher is a smart thing to do. To implement rational thinking in this decision, there are two important things to keep in mind. First, the probability of the cards being a rose. As three out of the four cards a player starts with are a rose, the likelihood that any given card is a rose is high. Adding to this rationality, a player should not want to win the bidding if this player has placed a skull on their stack. One of the reasons is that a player would rather have an opponent turn over this skull and lose a card. The second reason is that when a player wins the bidding, all their own cards must be turned over first. This would mean that playing a skull and winning the bid ensures that you turn over your own skull card. For this scenario, the `passProbability` variable is used to decrease the probability that a player will bid higher than the current bid. The decision to raise the bid or pass is made through a threshold. A random number is generated between 0 and 100. When this random number is lower than a given threshold, a player will raise the bid. If this number is higher than the threshold the player will pass. These thresholds are calculated through the use of multiple formulas. The function will check what type of strategy the player uses and will choose the correct probability function for the threshold.

2.2.2.1 Strategies

The model has five different strategies a player can use during the bidding stage. First, there is the Probability-Based Strategy (PBS). This strategy is used for the base model and has players make rational decisions based only on the cards on the table and what the player knows about their own cards. This strategy was also described in the previous section. There are three different bluffing strategies. The first bluffing strategy is called the Overbid Bluffing Strategy (OBS). In this strategy, a player using it is more likely to raise the bid even if the probability of succeeding is not high. The player will take more risks in overbidding. The second bluffing strategy focuses on the skull card. As mentioned previously, when a player has placed a skull card on their own stack, winning the bid will mean that this player must turn over their own skull card. Therefore, a rational player will not be as likely to overbid in this scenario. However, overbidding in this scenario can be considered an excellent bluffing move if it causes another player to bid even higher. A player using the Skull Bluffing Strategy (SBS) is more likely to raise the bid in the case of having placed a skull card and using this bluffing strategy to try and trick other players into potentially picking a Skull. The third bluffing strategy combines the first two types of bluffing. This strategy is called the Dual Bluffing Strategy (DBS). The last strategy a player can use is the Bluff Tracking Strategy (BTS). This is the strategy that will be researched. The Bluff Tracking Strategy will keep track of the bluffing habits of all the players. The section of stage three explains how these bluffing habits are tracked. The player using the BTS strategy will use the gathered information to predict whether an opponent is bluffing. This prediction is taken into account in the decision to overbid or not.

2.2.2.2 Probabilities

There are two different functions used for the probability of bidding higher than the current bid. Both functions are extended from the following base formulas for the rational part of the decision. The probability will be calculated for a bid that is one higher than the current higher bid. For this, formula 2.1 is used. Then the probability that the new bid can be satisfied is calculated first. This probability is found using formula 2.2, which looks at the number of cards a player can already turn over and how many cards still need to be found in the cards of the other players. In formula 2.2, it is assumed that the cards are only roses. For the chance that a player has played a skull themselves, the `passProbability` (ρ_{pass}) variable is used in the player class. When a player plays a skull in the first stage of the game, this variable is set to 1. If not, it stays at 0. In formulas 2.3 and 2.5, it is shown how this `passProbability` influences the threshold. For the two functions, this base threshold will be increased or decreased depending on what a player's strategy is. The PBS strategy only uses these base formulas to decide the threshold.

$$Bid_{new} = Bid_{highest} + 1 \tag{2.1}$$

$$\rho_{rational} = \frac{Bid_{new} - self_{placed}}{cardsOtherUsers} \tag{2.2}$$

Increasing this threshold hold will increase the chance that a player will raise the bid, and this is the direction used for bluffing players. The players that bluff will take more risk and will not take all rationality into account. The first probability function is `findBidProbability()`. This function is used by the players using one of the bluffing strategies of OBS, SBS, or DBS. In this function, the following formulas are used to calculate a probability.

$$\rho_{bluff} = \rho_{rational} - (0.5 - SBS_{level}) * \rho_{pass} + OBS_{level} \tag{2.3}$$

In this formula, bluffing is implemented by increasing the probability. When used, the `OBSLevel` and `SBSlevel` variables are set to 0.25 at the beginning of the game. In the formula, these probabilities can then increase the threshold. This number was decided through testing different settings. These different settings showed that 0.25 was the better option (see Appendix A). In the function, the `passProbablity` will reduce the threshold when a skull card has been played. The two types of bluffing strategies are Overbid Bluffing Strategy and the Skull Bluffing Strategy (the third type Dual Bluffing Strategy combines the two). The first type uses the `OBSLevel` variable and is a constant increase of 0.25 on the threshold. The Skull Bluffing Strategy is used in the scenario a player has placed a skull card in the first stage. This setting uses the `SBSLevel` variable to reduce the `passProbability`. In the formula, the middle part shows how to decrease the chance of overbidding if a player has played a skull card. In SBS, this `passProbability` is decreased. This makes the player more likely to overbid in the case they had played a skull, which is a part of the Skull Bluffing Strategy.

Decreasing the threshold will be used for the players using BTS. The probability of making a new bid will be influenced by the likelihood that the current highest bidder is bluffing. If the current highest bidder is bluffing, the player should be less likely to overbid the current bid. The second function is `StrategyBidProbability()` which uses the formulas 2.4 and 2.5. This function is used for the player using the Bluff Tracking Strategy (BTS). In this formula, the player uses the two dictionaries that the game class keeps track of. These dictionaries include the `bluffMemory` and `totalRoundsHighestBidder`. The first dictionary provides the number of times a player is believed to have been bluffing. The players are classified to have been bluffing during the third stage. This classification of bluffing will be explained more in the section on stage three. And the second dictionary provides how many times a player has been the highest bidder and made a mistake. With these two numbers, the function calculates the percentage of times a player has ended up bluffing. This percentage is then used in the formula to lower the probability that the player will bid higher than the current bid. In BTS, the player only uses the rational type of the `passProbability`. In this version, the threshold is lowered by 0.5 when a player has played a skull in the first round. Continuing, in this case, would be a risky move.

$$\rho_{bluffs} = \frac{bluffs_{highestBidder}}{bidsLost_{highestBidder}} \tag{2.4}$$

$$\rho_{Strategy} = \rho_{rational} - (0.5 * \rho_{pass}) - \rho_{bluffs} \quad (2.5)$$

2.2.3 Stage Three

In the third stage, the highest bidder will try to turn over the number of roses corresponding with the bid they have made. This stage will end when the highest bidder is able to find all the roses needed to satisfy the bid or when the highest bidder has picked a skull card. The `playRound()` function keeps track of the bid the player has made during this stage. When the player picks a rose, this bid is reduced by one. If the bid is reduced to zero, the bid is satisfied, and the player receives a point. However, if a player does pick a skull card before reducing the bid to zero, this player will lose the bid, and the `lostBid` Boolean is set to true. Setting the `lostBid` Boolean to true will end the while loop and, with it, the third stage. If a player wins a second time and thus earns a second point, this player will win the game, and the winner Boolean is set to true, ending the game.

The player uses the `playStageThree()` function to perform the actions for this stage. This function will first pop all cards from their own stack. A player must always take the top card from a pile. Therefore, the stack data type is used. When a player has been able to pop all their own cards without turning over a skull, the player can move on to the piles of other players. The player will pick a card randomly from one of the piles left. If a rose is picked, the bid is reduced by one, and the player will pick another card. However, if a skull is picked, there are two scenarios. The first scenario is when a player chooses the skull from their own pile; this is classified as a bluff. As playing a skull and continuing to bid is a very risky move, the player will likely try to trick other players into raising the bid. When a player does end up as the highest bidder, the bluff did not succeed, and the player ends up picking their own skull card. The other scenario is when a player picks the skull from another player. This is not always classified as bluffing, as a player can make a rational bid but reveal a skull very quickly. Therefore, the model looks at how many cards are left to satisfy the bid, when 90 percent of the bid has been met, a player has overplayed themselves. A player has overplayed themselves, as they bluffed or played with high risk believing that almost all the cards on the table would be a rose. When a loss is classified as a bluff, this is saved in the `bluffMemory` dictionary of the game class and available to all players using it in their strategy. This will be used for the strategy in stage two.

When a round is won, the `playRound()` function will reorder the players so that the highest bidder starts the next round. The original order is used starting from the highest bidder. After reordering the players, the `resetPlayer()` function is called to reset the variables for the next round. In each round, the player will get their cards back. However, if a player has revealed a skull in stage three, they must hand in a card. This is checked in the `resetPlayer()` function through the `lostBid` Boolean; if true, a random card is taken from the deck and added to the `removedCards` list. This list will ensure that the cards will remain removed in future rounds, as the decks are reset to the original four cards every round in the function. Other variables that need to be reset are the `placedCards` list, the `passed` Boolean, the `bid`, and the `passProbability`. After this, the round is finished, and a new round can be played.

2.2.4 Game End

If the game has ended due to a player attaining two points and winning the game, the whole game must be reset. Resetting the game is done in the game class. As explained earlier, the function `resetGame()` will reset all the players, the `winner` Boolean, and the `winningPlayer` variable. However, it does not reset the `bluffMemory` and `totalRoundsHighestBidder`. These two dictionaries are kept for future games to be played. Adding more game results to these dictionaries improves the knowledge that the player has over the other players.

Once the game has been reset, the next game can be started. The model plays as many games as requested in the terminal.

2.3 Data Collection

The research will be done based on the results of the first player on the list. This player is called Sophie. In all the games Sophie will be using the Bluff

Tracking Strategy (BTS), except for the base game. From each game, certain values about Sophie will be stored. These values are generated through playing multiple games of Skull. After each game, the following variables are saved to a panda data frame:

- `Wins`: if Sophie has won or not
- `Points`: the number of Points Sophie has won that game
- `lostBids`: the number of times Sophie's bid was wrong
- `roundsPlayedUntilwinner`: the number of rounds played until a player won
- `nameOfWinner`: name of the winning player

The setup per game can be adjusted at the start of the model. When running the model, these settings will be asked for in the terminal. The settings are the following:

- Number of players: between 3 and 6
- Type of bluffing: No bluffing (PBS), Skull bluffing (SBS), Overbid bluffing (OBS), and Dual Bluffing (DBS)
- Number of players using BTS strategy
- Number of games to play

For the results, the following standard settings were used. The number of players was set at 6. The number of games played was set at 1000. For the baseline results, the model was set up with all players using PBS. The Bluff Tracking Strategy is tested against the baseline model and all bluffing strategies. These setups are run in two different ways, the first being where only Sophie uses BTS and the second where all but one player uses BTS. In table 2.1, the different setups that were used are shown. The second column shows if the opponents used PBS, OBS, SBS, or DBS. The last column shows the number of players using BTS. This means that with one player using BTS, the other players use a different strategy.

The names are set up as follows: the name of the strategy that the opponents use - the number of opponents using this strategy. The remaining players will use the Bluff Tracking Strategy (BTS). This group can be only Sophie or Sophie and four other players that will use BTS. For example: in Overbid-5, five players will use the Overbid Bluffing Strategy and the one remaining player (Sophie) will use the Bluff Tracking Strategy.

Table 2.1: Settings used to run different games

Name	Opponent Strategy	BTS Users
PBS-6 (Base)	PBS (6 Players)	0 Players
PBS-1	PBS (1 Players)	5 Player
PBS-5	PBS (5 Players)	1 Player
SBS-1	SBS (1 Player)	5 Players
SBS-5	SBS (5 Players)	1 Player
OBS-1	OBS (1 Player)	5 Players
OBS-5	OBS (5 Players)	1 Player
DBS-1	DBS (1 Player)	5 Players
DBS-5	DBS (5 Players)	1 Player

3 Results

In the methods section, it is explained for which different settings the model has generated data. This data will focus on the results that Sophie has produced. By comparing these different settings, results can show how effective the Bluff Tracking Strategy (BTS) was for Sophie. As mentioned in the methods, the model plays 1000 games per setting. The data for all of the settings are saved in a CSV file and analysed in R.

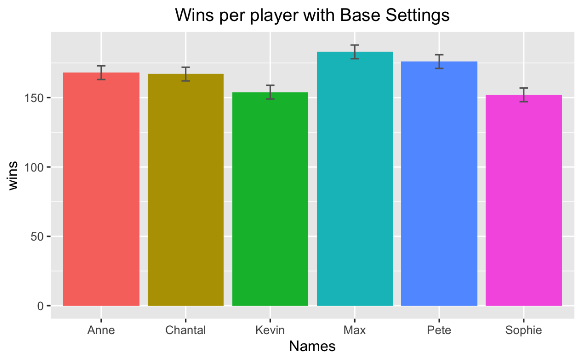

Figure 3.1: Model win results with only PBS is used and no tactics or bluffing

In figure 3.1, the wins are shown per player. These results were from the model run where all players used the Probability Based Strategy. These results show how the players would perform if only rational actions were taken. In the graph, it is visible that there is not one player that outperformed the other players. The dispersion of wins is very similar. The graph shows that Sophie was able to win 152 out of the 1000 games played.

3.1 Game Wins

First, the results of the wins will be inspected to see whether there has been a performance increase. An increase in wins would imply that the performance has increased.

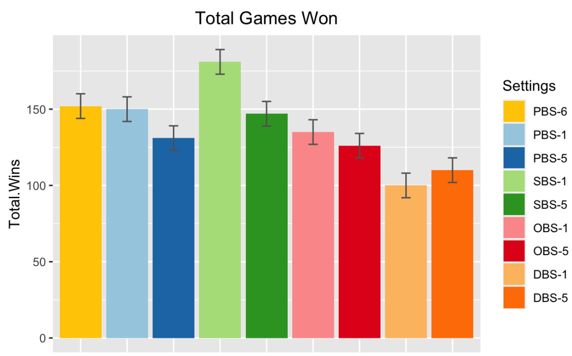

Figure 3.2: Number of games Sophie won in each different setting out of a 1000 games. Each bar represents a different setting.

In Figure 3.2, the number of games that Sophie has won is displayed. In all of these settings, Sophie is using the BTS strategy. The difference between settings is the types of bluffing strategies and the number of agents using them. For each bluffing strategy, there are two different settings, the first is where five players use the bluffing strategy, and only Sophie uses the BTS. Second is the case where five players (including Sophie) use the BTS, and only one player uses the bluffing strategy.

The yellow bar represents the base setting where all players use PBS as their strategy.

First, the blue bars show the wins when the opponents use the PBS strategy. The light blue bar represents the setting where five players use the BTS strategy, and in the dark blue bar only Sophie uses BTS. These bars show that the number of wins did not drop as much as for the last bars/settings where DBS was used. This could mean that the BTS strategy does react more when other players are bluffing, which it was designed for.

Second, the green bars represent the wins for the setting where the opponent used the Skull Bluffing Strategy. By continuing the bid after playing a skull, the players would increase the chance of having to pick their own skull and lose. In this setting, Sophie did win more. However, this was only the case for the setting when five of the six players were using the BTS strategy.

Looking at the settings where the opponents use the Overbid Bluffing Strategy represented by the red bars. The results also did not drop as much as the orange bars for DBS. Indicating that the strategy did not decrease the performance much for the setting with the Overbid Bluffing Strategy.

Lastly, in the figure, there are two orange bars for the settings where the Dual Bluffing strategy was used. Here it shows that Sophie wins less than in the base setting, where PBS was used (the yellow bar). The number of wins drops by almost a third. It does seem that when the opponent bluffs, Sophie does not have an edge by using the BTS strategy. This does not imply that she does not learn more from the bluffing players. It can also mean that she is more careful in bidding, and thus the other players are able to play more often for the points.

From these results, it still needs to be shown why the strategy decreases and increases the wins for certain settings. Therefore, the results that were found on the number of points and the number of bids lost will be looked at next.

3.2 Number of Points Won

Looking at the points that Sophie has won during the games will show whether Sophie did make more good calls on the bid. Sophie wins a point when she is able to satisfy the bid she made during the second stage and won the bidding. Making more points shows that the performance increased overall.

For the blue bars in the first setting, Sophie does show a decrease in points won. Indicating that Sophie did not do as well as in the base setting.

In the second setting (green bars), where the players only used SBS, Sophie did not increase her points as well. However, it does show that in this setting, Sophie did obtain more points when five players were using BTS as opposed to only Sophie using BTS. For the other settings, when all play-

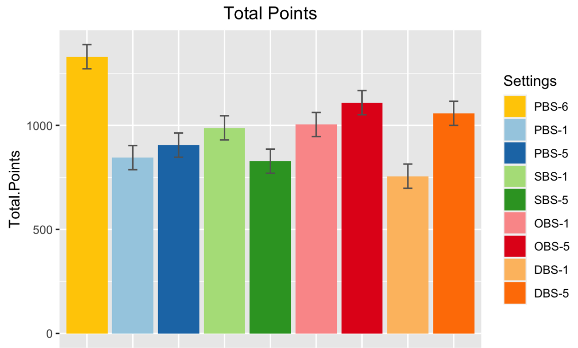

Figure 3.3: Number of points Sophie won in total for each setting. The total of points over a 1000 games. Each bar represents a different setting

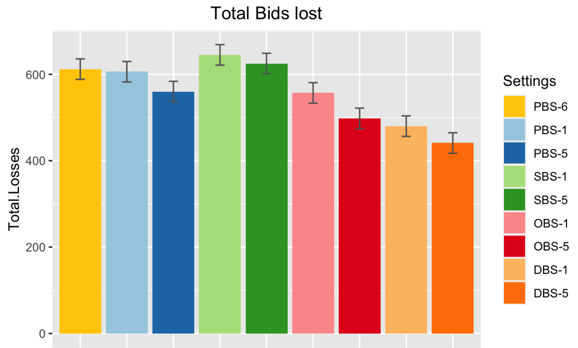

Figure 3.4: Total number of times Sophie's bid was wrong over a 1000 games. Each bar represents a different setting.

ers but Sophie were bluffing, Sophie would obtain more points. The result of the last settings aligns with the wins Sophie obtained as Sophie was able to win more games in the last setting.

In the third setting (red bars), where opponents used OBS, Sophie also had a decrease in performance. The points obtained were more than in the prior settings where the opponents used DBS and PBS. However, it does show that Sophie did not make more points than in the base settings either.

In the last settings (orange bars) where DBS was used, Sophie obtained fewer points than in the base setting (yellow bar). This would mean that either Sophie has passed more on bids and thus has made fewer points, or Sophie revealed more a skull more often in stage three.

3.3 Bids Lost

When a bid is lost, this means that the player revealed a skull card and thus lost the bid. If Sophie could decrease the number of lost bids, it would increase the performance.

Figure 3.4 shows the number of bids Sophie lost in stage three for the different settings.

In the settings where the PBS strategy was used (blue bars), Sophie does make around the same mistakes as she did in the base model. This would imply that the BTS strategy was not able to prevent Sophie from making fewer mistakes in this setting. As the players were not bluffing, the BTS strategy could be making the wrong calls by assuming that the players are bluffing.

The next setting (green bars) shows that Sophie has lost more bids than before. Losing more would imply that the BTS strategy needs to learn better when the opponent is using the SBS strategy. However, as players may only sometimes play a skull card, more games could be played with a less risk-involved strategy by most players. Therefore, Sophie would have the chance to win the auction in stage two more often. In this setting Sophie loses more bids and wins more points, and this also increased her wins overall. This shows that being the highest bidder more often, increases the chance of winning. It does show that Sophie has also made the most mistakes in this setting.

For the third setting (red bars), Sophie did make fewer mistakes as well. However, the mistakes did not decrease as much as in the last setting (orange bars). This is due to how Sophie classifies bluffing in the strategy. If a player loses a bid with ten percent of the bid left or less, this loss is classified as Overbid Bluffing by Sophie. This would imply that a player has made a bid close to the total amount of cards on the table, which is a high-risk play. However, as the chance that a player loses early on in the third stage is also a possibility, Sophie may have yet to pick up the bluffing tendencies of the opponents as well as with the other setting.

Looking at the last two settings (orange bars), the BTS strategy does help Sophie to make fewer mistakes than in the base game. Making fewer mistakes would happen because Sophie is more careful and passes earlier in the second stage.

Overall looking at the figure, Sophie does make fewer mistakes when she is the only player using the BTS strategy. Sophie can classify other players as bluffing when this could not be the case. The strategy will become less reliable when more players anticipate other players to bluff by using the BTS strategy.

3.4 Overall Auctions Won

Combining the last two graphs will show what Sophie did overall with the auction. The previous results can also be explained through the percentage of the total auctions won. When a player wins more auctions, this gives a chance for more wins or mistakes.

Settings	Bids lost	Points Won	Total
PBS-6 (Base)	612	1330	1942
PBS-1	606	845	1451
PBS-5	560	905	1465
SBS-1	645	988	1633
SBS-5	625	828	1453
OBS-1	557	1004	1561
OBS-5	498	1109	1607
DBS-1	480	756	1236
DBS-5	441	1058	1499

Table 3.1: Number of bids correct and incorrect. Bids lost represent the times that Sophie was not able to satisfy a bid. The point won represents the times that Sophie was able to satisfy the bid. The total is the sum of these two.

First, the total number of auctions will be looked at. The table above shows the total number of points and bids lost. Together these numbers represent the number of times Sophie was the highest bidder.

The first settings where PBS was used do show a decrease in the points that Sophie was able to win. However, looking at the mistakes that Sophie made, there only is a big decrease for setting PBS-5. The overall times that Sophie was the highest bidder and able to convert this to a point decreased. This can also be seen in the percentages in table 3.2.

The table shows that Sophie lost more bids for the SBS setting than for other settings. However, the total number of auctions won remains are similar to the other settings (except for base). Therefore, Sophie will have made more mistakes with the auctions and should have passed more.

In the third setting, with Overbid Bluffing, Sophie won more auctions than in the first two settings. This is due to the classification of bluffing for BTS. Because a loss will be classified as Overbid Bluffing when the last 10 percent of the bid is not satisfied, this may not happen very often. This would mean Sophie will pick up on bluffing less, however, this will then also influence her probability of raising the bid less.

Sophie lost fewer bids in the last settings than in the base setting. However, the total shows that Sophie also has a decrease in the total auctions won. This would imply that Sophie has passed on more bids in the game. Due to the players using DBS, the BTS strategy will decrease the chance of Sophie overbidding. This will, in turn, make Sophie win fewer auctions.

The wins and loss percentages will show whether the BTS strategy made helpful decisions for Sophie.

Moving on to table 3.2, the percentages of the wins and losses are shown. These percentages tell whether Sophie made the correct decision to play or not play a bid.

Settings	Bids lost	Points Won	Total
PBS-6 (Base)	31.51%	68.49%	100%
PBS-1	41.76%	58.24%	100%
PBS-5	38.23%	61.77%	100%
SBS-1	39.50%	60.50%	100%
SBS-5	43.01%	56.99%	100%
OBS-1	35.68%	64.32%	100%
OBS-5	30.01%	69.01%	100%
DBS-1	38.83%	61.17%	100%
DBS-5	29.42%	70.58%	100%

Table 3.2: Percentage of bids correct and incorrect. Bids lost represent the times that Sophie was not able to satisfy a bid. The point won represents the times that Sophie was able to satisfy the bid. The total is the sum of these two.

Table 3.2 shows that the bids lost have increased in the overall percentages. The only time Sophie has had an increase in wins percentage-wise is in the last setting. This also being the setting where Sophie has won more games. Furthermore, table 3.1 showed that Sophie won fewer auctions in the second stage when using the BTS strategy. This, combined with the percentages in table 3.2, shows that Sophie also loses more bids.

Overall this table verifies the decrease in performance due to the BTS strategy. As can be seen in the table, the percentage of won auctions that were converted into points decreased. Even showing that the percentage of these won auctions lost was higher than in the base setting.

4 Conclusions

After looking at the results, the following conclusions can be made. First, the results showed that using the BTS strategy did not improve the performance of Sophie. It only increased the number of wins for Sophie in one of the settings. There are two possible explanations for these results. First, Sophie would try to anticipate the opponents' decisions and whether they are bluffing or not. By trying to anticipate this, Sophie is taking less risk when bidding in the second stage. By taking less risk, Sophie gives herself fewer chances to win the auction and make it to the third stage. Even when Sophie managed to win an auction and move on to the third stage, Sophie would reveal a skull more times than before and lose. However, this can also be approached from a different perspective.

Looking at the players that use the different bluffing strategies, they did succeed in letting others make more mistakes. Sophie's results show that she made more mistakes and should have passed more often. Making these mistakes does imply that the bluffing players may have tricked the other players correctly. What also helps the bluffing players is that BTS and the bluffing strategies work in opposite ways, as the BTS strategy will make Sophie more likely to pass. The bluffing players will take more risks in these plays and, therefore, have more chances to win a point as these players win more auctions.

Overall it shows that bluffing in the game of Skull will improve the chances of winning. As players will only need to get two points to win the game, playing with more risk can, most of the time, pay off. The players start a game with four cards and only lose one card if they pick a skull. Therefore, playing with high risk and having 50 percent of the auctions correct would have you win the game. The high-risk strategy will thus give an excellent chance to win. It is essential to be able to play the third stage, as with a game where the winning condition only is two points, the game is over before you may have had the chance to go to the third stage.

5 Discussion

Overall this study showed that the BTS strategy does not help the player win more games. The study gave a positive insight into bluffing strategies and suggested that a counter-strategy needs more work. Looking at the BTS strategy, the results show that this strategy is incomplete. As Theory of Mind suggests, reading an opponent and predicting their moves are more complex than only tracking their previous moves. More than tracking their moves alone will be necessary to improve the results. The player must also focus on their own strategy and really play the game themselves. By using the BTS strategy, Sophie was more focused on what other players might be doing than deciding what she thought could be a reasonable bid. It showed that the players using bluffing strategies would win more games. This comes back to the fact that these players are focused on their own game and thoughts. The level of Theory of Mind these players use is very low. This would suggest that in a game like Skull, a low level of Theory of Mind is a better approach. The study was mentioned earlier by Goodie et al. (2012), and how this study looked at the levels of Theory of Minds players would use. This study found that most players tended to use a lower level in more complex games. Looking at the results from this study, there could be a trend where players find that focusing less on their opponents' moves is a good strategy. Focusing on what levels of Theory of Mind a player should use in a game is a fascinating subject for future research.

Looking at the future of the BTS strategy, some possible improvements can be made. Currently, the BTS strategy only looks at whether the player should overbid, depending on whether the opponent is bluffing. The BTS strategy could be extended to encourage a player to use bluffing by looking at the situation and whether it was successful in previous situations. Extending the BTS strategy with bluffing is a way of changing the level of Theory of Mind. The BTS strategy will also be able to focus more on its own decision about a reasonable bid. The model will have to keep track of more statistics, such as the number of cards placed, the highest bid and whether this bid was won or not. With this information, a player can make a more balanced decision on whether to overbid or not.

The model showed that the bluffing players did improve over the BTS strategy players and were able to win more games. However, for future research, it could be interesting to look at how much risk is still beneficial for players while bluffing. Furthermore, looking at how the players can give their bluff more thought in the games could show whether there are scenarios that decrease your chances. Looking at the formulas used, the levels of bluffing were static numbers. In a newer model, the players could learn from previous games, and the bluffing levels could become a variable. This variable number could be deducted from previous results by using a similar approach to the BTS strategy that Sophie used. This could be a new strategy based on the BTS strategy, a strategy where the bluffing player will keep track of the situations when bluffing was not favourable, for example.

In a future study, it could be looked into whether losing cards has a significant decrease in the chances of winning than it has in winning. As previously mentioned that playing with high risk and being willing to sacrifice some cards to get more opportunities to advance to the third stage would increase chances to win. Thus determining whether losing cards and having fewer cards to use in the first round is worth it overall.

References

Freeman, G. H. (1989). The tactics of liar dice. *APPLIED STATISTICS-JOURNAL OF THE ROYAL STATISTICAL SOCIETY SERIES C*, *38*(3), 507-516.

Friedman, L. (1971). Optimal bluffing strategies in poker. *Management Science*, *17*, 764-771.

Goodie, A. S., Doshi, P., & Young, D. L. (2012). Levels of theory-of-mind reasoning in competitive games. *Journal of Behavioral Decision Making*, *25*(1), 95-108.

Guazzini, A., & Vilone, D. (2009). The emergence of bluff in poker-like games. *arXiv: Physics and Society*.

Marly, H. (2011). Skull game.

Sum, J., & Chan, J. (2003). On a liar dice game - bluff. *Proceedings of the 2003 International Conference on Machine Learning and Cybernetics (IEEE Cat. No.03EX693)*, *4*, 2179-2184 Vol.4.

A Appendix

A.1 Bluffing Probability Settings

To decide on the settings for the `OBSLevel` and the `SBSLevel`, we looked at the games won by Sophie for the different settings. We took a range from 0.05 to 1. In the graph below, the different settings are shown for `OBSLevel` and `SBSLevel`, where five players use a bluffing strategy and only Sophie uses BTS. The level of 0.25 was chosen for the model, as we can see in the graph (first red bars) that Sophie was able to perform her best against bluffing with this setting. The graph does show that Sophie wins more games in `OBSLevel` when the level is 0.75 or 1 than in the setting of 0.25. Closer inspection reveals that for high values of `OBSLevel`, the bluffing players consistently win the bidding phase by overbidding and lose the subsequent turning phase. Since this puts bluffing players at an unrealistic disadvantage, the value of `OBSLevel` was set to 0.25 instead.

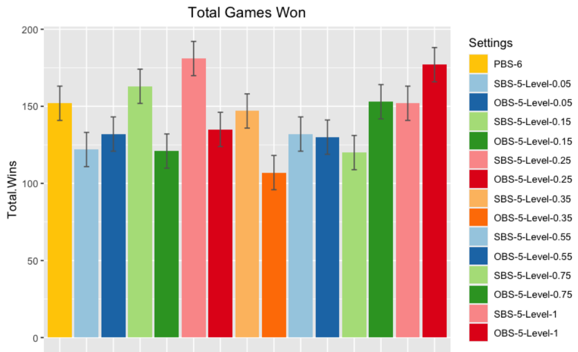

Figure A.1: The total games won by Sophie using BTS. The five opponents use either SBS or OBS. The different levels are used for SBSLevel and OBSLevel

A.2 Bids Won and Lost Graphs

The graphs show the same numbers as in the tables 3.1 and 3.2 in the results section. Here it is shown in a stacked bar graph.

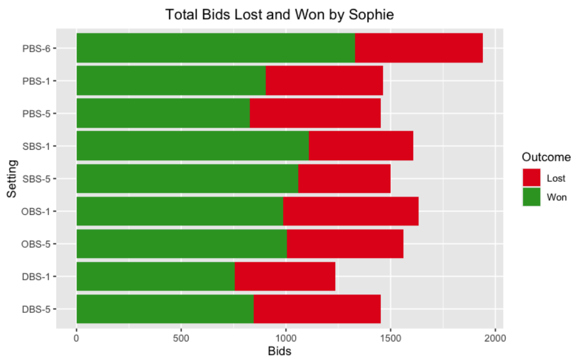

Figure A.2: The bars represent the number of bids Sophie has lost and won. The green bar is the bids won and the red bar the bids lost.

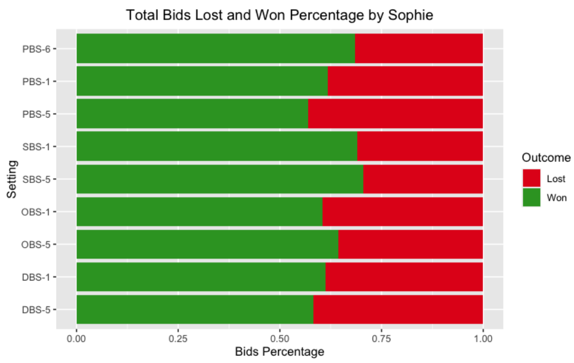

Figure A.3: The bars represent the number of bids Sophie had wrong and right. It is shown as a percentage of the total of bids Sophie got to play in the third stage. The green bar is the bids Sophie had correct and the red bar the bids Sophie had wrong.