An Epistemic Approach to Analyzing and Modeling Knowledge in Blufpoker

Ivar Mak (S2506580) Oscar de Vries (S2713748) Joost Warmerdam (S2591367)

1 Introduction

Blufpoker is a dice game which contains several interesting aspects with regards to knowledge, beliefs and bluffing. In this project, we implement Blufpoker and simulate agents playing with each other, using different different strategies. This game is implemented as a Multi-agent System where agents interact and knowledge about other agents is relevant for decision making. In order to reason about what an agent knows, a formalization of the knowledge is done using epistemic logic. Furthermore, a model is proposed that allows for reasoning about knowledge for an agent in Blufpoker. Using this model the knowledge is formalized using a Kripke model which is necessary to allow for reasoning in the simulation. An agent that uses this model to perform knowledge-reasoning using the obtained Kripke model. This model will contain the possible worlds which can be reached from the in-game situation and will be used for the decision making process of the knowledge-based agent. The aspect of bluffing (false public announcements) comes into play when a player rolls a bad turn. But before we address this, we will describe the rules of the game.

1.1 Game Description

Every round the players take turns rolling three dice, which are concealed under a cup (i.e. other players do not see the player's dice). The aim of the game is to roll a number that is higher that the number that was passed on by the previous player in order to not lose the round. The dice together depict a number in the hundreds. The highest ranking die translates to the hundred, the second to the ten, and the lowest ranking die translates to the single digit. For example, in a situation where a 5, 3, and 2 are rolled the corresponding number is 532. The exception are the "pokers" which are triplets (i.e. 111, 222, 333, etc). These are the highest possible rolls, 111 is higher than 665, and 666 is poker 6 and the highest possible score. Also, whenever a poker is rolled, this is handled differently than the other rolls, which will be described later in this section. Table 1 describes all possible combinations, ranked in descending order from highest to lowest.

666	664	652	632	552	532	441	332
555	663	651	631	551	531	433	331
444	662	644	622	544	522	432	322
333	661	643	621	543	521	431	321
222	655	642	611	542	511	422	311
111	654	641	554	541	443	421	221
665	653	633	553	533	442	411	211

Table 1: Set of possible dice combinations in the game of *Blufpoker*, ranked from top left to bottom right, descending per column.

When the round starts, an agent rolls the dice and announces a claim of what number he has rolled.

Apart from the first roll, this announcement has to be higher then the previous announcement. Now the next players' turn starts. It is up to this player to determine whether his adversary is telling the truth. If he believes the truth was told, the player rolls the dice to reach a higher number yet again. In the case he does not believe the previous player, he can call the bluff and lift the cup. If the number that was announced is actually on the table player 2 loses the round, if the number is not there; player 1 loses the round. Apart from the first roll, the agent can roll the dice one at a time and decide to not roll remaining dice. The dice that are not rolled, are shown and become common knowledge. Obviously, there is a possibility that he did not reach a number that is high enough. In this case the player needs to bluff, by announcing a number that he did not roll. Since some of the dice are concealed, he can say whatever number he deems suitable, and pass on the dice to the next player. When a player is caught bluffing, that player loses the round and receives one penalty point.

The first player to receive five penalty points loses the game. Therefore there will always be one loser and multiple winners. The number of penalty points chosen to finish can resemble a word such as "HORSE", receiving a letter for each loss.

1.1.1 Poker phase

The poker phase is slightly different than the usual rolling phase. Here we will describe how that phase is executed. Whenever some player announces a poker and the next player believes this (regardless of whether there truly is a poker), the game gets into the poker phase. In this phase, the player (that believes the poker) has three chances to roll an equal or higher poker, which happens openly (not under the cup). In these three rolls, the player can leave aside any dice he wants, or roll them all again, as long as it are three rolls in total.

If the player then does not manage to roll an equal or higher poker, he gets a penalty point. If he manages to roll an equal poker, no player gets a penalty point. Finally if he rolls a higher poker, the player who initially called the poker gets a penalty point.

1.2 Framework

First the implementation of the game is addressed, along with the strategies that are implemented. These strategies will include superficial strategies and one more complex strategy based on the epistemic knowledge model. Then the implementation of the epistemic model is shown as well as a description of the strategy that is modelled based on this epistemic model. The knowledge-based strategy is then used to play games in order to give results about its behavior versus more naive, superficial strategies. The model will give insight in how well the proposed knowledge-based strategy performs based on ratios of how often this agent does not lose and how many bluffs are called of the previous player. Furthermore, we report an example walk-through of one game to identify the situations where the epistemic model helps, and what limitations are present.

2 Methods

The Blufpoker simulation and the epistemic model are made in Python[1]. The game runs in the terminal and is visualized using the User Interface library *Tkinter*. To give insight in the knowledge-

[1] Python Software Foundation. Python Language Reference, version 2.7. Available at http://www.python.org

based strategy, the program contains an implementation of the probabilities of the possible worlds based on the public announcements and the shown dice.

2.1 Program

The Blufpoker game is implemented using the aforementioned rules (Section 1.1), along with several simplifications. In the real game, a player can bid with any number. In the simulations, if the bid is not truthful, it can only be higher than the previously announced bid. Also, in the real game the combinations $3, 2, 1$ is the highest roll. This is not the case in the simulation as it makes the knowledge-based strategy more complex in an unnecessary way. Furthermore, in a real game, a player can decide to not roll the dice at all and announce a higher bid. The is not possible in the simulation.
The program simulates three agents playing against each other. Different strategies can be executed by the agents, which are applicable to each of the three phases in a turn. One of these agents will use a knowledge-base for its decision making. This reasoning is described in Section 2.2. The other strategies are less complex and regarded as *naive strategies*. The simulation performs the game in a step-wise manner and assigns penalty letters to the losing agent after each round. The simulation will record, which player has lost and how often the knowledge-based agent wins against the naive strategies and how many bluffs the knowledge agent will identify correctly. This simulation is run for **Hoeveel runs?** to create data indicating the performance of the knowledge-based agent.

2.1.1 Strategies

Before the strategies are defined, a single turn is divided in three phases in which the strategies will differ:

1. *Calling phase*: The agent decides whether the previous agent is truthful about its announcement of the rolled dice.
2. *Rolling phase*: The agent rolls the dice.
3. *Bidding phase*: The agent forms a bid and announces this for others to decide to believe it or not.

Strategies influence an agents decision making in the Calling phase and the bidding phase. The strategies of the other 'naive' agents will differ in three manners: (1) deciding whether to believe the previous player's bid or not, (2) deciding which dice they roll and (3) deciding how they will bid (i.e. bid truthfully or not, and how to bluff). Eventually, these strategies will be used to measure how well the knowledge-based strategy performs in comparison to these different strategies.

An overview of the strategies of the naive agents that will be used in the games against the knowledge-based agent is the following:

1. Calling Phase:
    - Believe a bid with a certain probability
    - Believe any bid

- Believe no bid at all

2. Rolling Phase
   - Roll one of the dice, randomly chosen
   - Roll the n lowest dice, in which n is randomly determined
   - Roll only the lowest die
   - Greedy: Roll any die that is lower than 6
3. Bidding Phase:
   - Bid truthfully, if possible. Otherwise bid a number x above the previous value, where x is randomly determined in a small range
   - Always overbid previous bid: Bid a number x above the previous value, where x is randomly determined in a small range

A number of combinations of these strategies will be used to define naive agents. Following, testing different configurations of such naive agents in combination with a knowledge-based agent will be used for measuring the performance of the knowledge strategy.

2.1.2 Exceptions: First turn and poker phase

Since both the first turn and the poker phase are different than the aforementioned phases, yet do not allow for much variation, it was decided to keep these phases equally for each of the agents (including the knowledge agent):

- In the first turn, an agent will simply roll the dice and truthfully announce what he has rolled.
- In the poker phase, since the dice are rolled openly, it seemed like any player should apply the same strategy, which is quite straightforward. In this phase, a player will first roll the dice, and when this does not lead to a equal or higher poker in the first roll, the player will execute the following steps until he has no rolls remaining, or does reach a poker:
  1. First check if there are two dice with the same value that can form a higher or equal poker. If so, keep these and roll the other dice, otherwise execute step 2.
  2. Keep the highest die, given that this can form a higher or equal poker, and roll the remaining two in order to get that poker.
  3. If none of the dice can lead to a higher or equal poker, roll all the dice.

2.2 Epistemic Model

The epistemic model that is proposed for modelling the knowledge for the different agents consists of a Kripke Model. A Kripke Model[1] is defined as a tuple $\langle S, \pi, R_1, ..., R_m \rangle$ where:

1. S is a non-empty set of states and m is an agent in the set of agents $\boldsymbol{A}$.
2. $\pi : S \rightarrow (\boldsymbol{P} \rightarrow \{\boldsymbol{t}, \boldsymbol{f}\})$ is the truth assignment to the propositional atoms per state.

3. $R_i \subseteq S \times S(i = 1, ..., m)$ are the accessibility relations.

Every world is defined as a possible dice roll, so Table 1 shows all possible worlds, each containing a specific dice combination. To represent this, we define a formula $d_{x,y,z}$ with $0 < x \leq y \leq z \leq 6$ which show which dice are thrown. Here, x represents the hundreds, y represents the tens, and z represents the single values. So a world w_1 can contain a formula $d_{4,2,1}$ where the current dice roll is 421.
In this model, the set of agents consists of the players, i.e. $\boldsymbol{A} = 1, 2, 3$ when there are three players. The set of states S is defined by all the possible dice rolls, as mentioned in 1. The truth assignment of the propositional atoms is not used in the model and is therefore omitted in the definition. The accessibility relations are defined as follows: a world w_i can be accessed by world w_j by agent x if that agent considers it possible that when rolling a dice given the dice in w_i makes possible that the combination in w_j is rolled. For example if agents x knows that the current world contains $d_{3,2,1}$, then a world with $d_{4,3,2}$ is possible as it can be rolled by rerolling the singular die.

In the simulation, the number of possible worlds change per state. The reasoning that can be done on behalf of the formalization is modelled in the knowledge-based strategy.

2.3 Knowledge-based agent

One of the agents in the simulation will play the phases in its turn according to a more sophisticated strategy, involving knowledge and probabilities in the game. In contrast to the naive strategy agents, this knowledge-based agent will make decisions on the basis of available knowledge. In the game, any unrolled dice during the rolling phase should be put on top of the cup for all players to see. In the implementation, this means the unrolled dice are made available as common knowledge. Consequently, the knowledge-based agent can use this common knowledge to cross-out some of the possible worlds. The remaining possible world will be part of its knowledge and can be used for reasoning about its own decisions.

In the following part, the knowledge-based strategy in each of the three phases will be discussed. For each phase, we will explain how the agent uses knowledge for making a decision in that phase. This will be based on a regular turn (i.e. not a first turn of the game).

2.3.1 Phase 1: Believing or calling a bluff

First of all, the agent has to make the decision whether it believes the previous players bid or calls it a bluff. Now a description will be given of how the knowledge agent determines to call a bluff.

The agent will first evaluate whether the previous bid is possible with respect to its knowledge, by cross-checking whether the bid is within its set of possible worlds. If this is not the case, the agent will regard the previous bid as a bluff. Otherwise, the assessment will be done on the basis of probabilities. In that case, the agent will make a *joint probability distribution* (referred to JPD after this) of the possible worlds, where each possible worlds gets the probability for it to be rolled assigned to it. This probability depends both on the available knowledge, as well as on the number of combinations that make up for the roll, since there is only one possibility to roll (1,1,1), but multiple possibilities to roll (6,5,1). Hence, the probability P(Possible World | Knowledge $\wedge$ Number of Combinations) is determined as follows:

- **Common knowledge has 2 open dice:** Only one dice is unknown, meaning that there exist six possible worlds (in conjunction with the common knowledge). Every world that is impossible given the common knowledge can be disregarded and will not be used in the JPD, since their probability is equal to 0. Since there is only one possible way for each value to be rolled, each of the possible worlds will have a $\frac{1}{6}$ probability assigned to them in the JPD.

- **Common knowledge has 1 open dice:** There are two unknown dice. Again, every possible world in which the open die is not present can be disregarded and will not be used in the JPD, since these are impossible. Furthermore, there are two possibilities for the unknown dice, which each have a different probability. These are the following:

  (1) If both dice have an equal value, there is only one combination possible for worlds in which this is the case. This means these probabilities will be equal to $\frac{1}{6}^2 = \frac{1}{36}$.

  (2) If the unknown dice have an unequal value, there are two possible combinations of dice rolls that make up for the same possible world (E.G. when 5 is common knowledge, both (5,1,4) and (5,4,1) will be (5,4,1)). Consequently, the probabilities assigned to these worlds in the JPD will be equal to $2 * \frac{1}{6}^2 = \frac{2}{36}$.

- **Common knowledge has no open dice:** In this case all three dice are unknown, meaning that the possibilities are determined by three unknown factors. Now all the possible worlds have a probability assigned to them in the JPD. Here we have the following three possibilities:

  (1) Each of the three unknown dice have an equal value. In these cases the assigned probability in the JPD equals $\frac{1}{6}^3 = \frac{1}{216}$.

  (2) Two of the three unknown dice have the same value, but one does not. In these cases there are 3 possible combinations of dice rolls that make for the same possible world (E.G. (2,1,1), (1,2,1) and (1,1,2) will all be (2,1,1)). Hence the assigned probability in the JPD equals $3 * \frac{1}{6}^3 = \frac{3}{216}$.

  (3) All dice have a different value. In these cases there are 6 possible combinations of dice rolls that make up for the same possible world (E.G. (6,5,4), (6,4,5), (5,6,4), (5,4,6), (4,6,5) and (5,4,6) will all be (6,5,4)). Consequently the assigned probability in the JPD equals $6 * \frac{1}{6}^3 = \frac{6}{216}$.

Now that the JPD is constructed, the agent can determine the probability that the previous players bid is actually rolled. This is equal to the cumulative probability of the worlds in the JPD that are equal or higher to this bid (since higher dice rolls than the bid are also not a bluff).

Following, a believe threshold has to be set, which determines whether the agent will call a bluff or not. It was found that this threshold should not be static, since realistically it happens that the same bid in the same situation is sometimes believed and sometimes not. Also the common knowledge is a factor in determining roll probability, hence that should also be included in determining the believe threshold. Otherwise the knowledge agent will either believe too many bets or call too many bluffs. Therefore the believe threshold is determined as follows:

- **Common knowledge has 2 open dice:** The believe threshold is equal to a randomly picked value from a normal distribution with $\mu = 1/4$ and $\sigma = 1/12$.

- **Common knowledge has 1 open dice:** The believe threshold is equal to 0.75 times a randomly picked value from a normal distribution with $\mu = 1/4$ and $\sigma = 1/12$.

- **Common knowledge has no open dice:** The believe threshold is equal to 0.5 times a randomly picked value from a normal distribution with $\mu = 1/4$ and $\sigma = 1/12$.

Given that these thresholds are determined with some randomness, the knowledge agent will have a probability to both believe and not believe a variety of bids. This seemed like the most elegant solution, since when a static threshold was used, the agent's decision would be deterministic.[2]

Finally, when the probability retrieved from the JPD, given the bid and possible worlds, is lower than the believe threshold, the agent will call a bluff. In all other cases, if the knowledge agent believes the previous player not to be bluffing, it believes the bid and goes into the rolling dice phase.

2.3.2 Phase 2: Rolling dice

The knowledge agent has to decide which dice to roll (and which not to roll) in order to get a higher value than the previous bid. The following procedure is executed:

1. The agent will roll the lowest die first, since always at least one die has to be rolled. If this leads to a higher roll value than the previous bid, the knowledge agent will stop rolling and go to the bidding phase. Otherwise step 2 is executed.

2. If the unrolled dice are equal, there is a small chance of $\frac{1}{6}$ that the agent decides to go to the bidding phase and bluff on a poker, with the value of those open dice. This is because the agent reasons such that other players will see both open dice of that value as common knowledge when the agent stops rolling, hence making the bluff believable. If the dice are unequal, bluffing on the poker is not a viable option and step 3 is executed.

3. If there is at least one 6 in the current bid, move to step 4. Otherwise, if there is no 6 in the current bid (and it is not a poker), then the agent will try to roll one of the remaining dice to get a 6, since that has a reasonable chance to get the agent to a truthful higher bid. If still no 6 is rolled after this, the agent will roll the final die and try to get a 6. Also, if the agent had rolled a 6 already with the first try (and with that reached a higher possible bid), there is still a $\frac{1}{2}$ chance to roll the other die anyway, in order to try and reach a higher bid with it.

4. In this case there is exactly one 6 in the current bid (for two 6's go to step 5), and the agent has not reached a higher roll value than this bid with its first roll. Now the agent will determine whether it wants to roll another die, on the basis of what it has to roll in order to reach a higher bid, in accordance with the probability of rolling that die higher. This probability is equal to $\frac{1}{6}$ times the number of possibilities that are higher for that die. To clarify: when the die has value 1, there are 5 possibilities to roll the die higher than that 1, hence the probability becomes $\frac{5}{6}$. However, when the die is a 5, there is only one possibility to roll the die higher than that 6, causing the probability to become $\frac{1}{6}$. If the probability causes the knowledge agent not to roll the die, it will argue that bluffing with the common knowledge for the next agent is less risky than trying to roll higher.

5. Finally, in this case there are two 6s in the current bid, and the agent has not reached a higher roll value than this bid with its first roll. The agent will go to the bidding phase and bluff on poker 6.

[2]Note that this believe threshold will always involve some kind of arbitrariness, which is due to uncertainty. Hence, finding the correct normal distribution is quite hard. However, this is also the case for human player.

After executing (a number of) these steps, the agent will go into the bidding phase.

2.3.3 Phase 3: Bidding

In this step the knowledge agent decides how to bid smartly. Smartly in this sense means that the agent will use the common knowledge available for the next agent to determine what is a believable bluff. This is done with the following procedure:

1. Firstly, the knowledge agent will always bid truthfully, if it has a higher valued cup than the previous bet after the rolling phase. Otherwise, go to step 2.
2. In this step, a bluff is made on a poker, given that this was decided during the rolling phase. This possibility was the case in steps 2 and 5 of the rolling phase. If this was not decided, a bluff must be made in the following step.
3. A number of possible bluffs are determined on the basis of the common knowledge for the next player. This is done by excluding bets that are impossible given the common knowledge. Given the number of possible bluffs, one that is a small random number higher than the current bid is chosen. For example, a (6,6,2) will have the knowledge agent choose randomly between (6,6,3), (6,6,4) and (6,6,5), given that the common knowledge allows this.

The performance of and agent executing this knowledge-based strategy will be compared to agents using the naive strategies.

3 Results

The results we will gather are focused whether we will be able to efficiently construct a model and its accuracy. We will evaluate the decision making process based on this model by measuring the performance of our logic based agent compared to the other agents. Firstly we will measure this by looking at the ratio between won and lost games. Secondly, by looking at the amount of bluffs that are successfully called by our logical agent, drawing conclusions about in what way the knowledge in this model aids to making the correct decision.

4 Possible Extensions

If time allows for it, we have thought of several directions to take in order to extend our project and/or it's implementation.

- We want to evaluate different settings with regards to the strategies against which the logical player has to compete, and look at the way its performance changes.
- When possible, we would like to add more players using a logical model for their decision making process.
- We would like to add feedback prompts describing the decision making process of the logical agent, much like the thought process of a human player.
- More aesthetically, we would like to add visualizations of the dice and or the cup.

References

[1] John-Jules Ch Meyer and Wiebe Van Der Hoek. *Epistemic Logic for AI and Computer Science.* Cambridge University Press, 1995.