

## Lie Game Final Project Report

### Problem Statement

The game of 'Lie' involves players placing cards and sometimes bluffing about what they placed. The AI agent must be able to not only bluff, but also "guess" when another player is also bluffing. Cards are dealt equally to each player out of a deck of 52 cards. Each player takes turns placing cards in the middle in ascending order. They may place 1-4 of the next numbered card. They may also lie about which numbered cards they placed down. If the next player believes they are lying, they can call "Lie". The cards are revealed. If the player who placed the cards was telling the truth, then the player who called "Lie" takes all the cards in the middle. If the player placing the card lied, they take the stack in the middle. The game ends when one player has gotten rid of all their cards.

The agent is designed to detect opponents bluffing, decide when to bluff, maintain beliefs about the opponents, and adapt its strategy as the game proceeds. The game is partially observable because the player can only see what cards they have in their hand. The agent has to decide when and how to lie and whether or not other players lie.

### Related to Similar Problems

A similar method can be used to model poker using Bayesian inference and Monte Carlo search trees. Bayesian inference is used to model what hand each opponent is believed to have based on information revealed in the game, similar to our game. A Monte Carlo search tree is used to model probability distributions that can be used to predict opponent behavior.

### State Space Description

State Space:

- Cards currently in the agent's hand
- Number of cards each player has
- Current number in ascending order of placing cards
- Previous plays by players including what cards they claimed to have played
- Probability of what cards a player current holds
- $S = (H_i, N, R, C, B)$ 
  - $H_i$  is the agent's hand,  $H_i \subseteq C$  (the deck)
  - $N = (n_1, n_2, \dots, n_k)$  how many cards a player has, up to  $k$  players
  - $R \in \{1, 2, \dots, 13\}$  the current number being played

- $C = (p, r, c)$  the most recent claim of player  $p$  who claimed to play  $c$  cards of rank  $r$
- $B(i, x) = P(\text{player } i \text{ holds card } x)$  a belief distribution over opponents' hands.  $x \in C$

Action Space:

Step 1:

- Play 1-4 cards of the correct number (tell the truth)
- Play 1-4 cards of the incorrect number (lie)
  - Including playing cards of different numbers

Step 2

- Call BS on the player prior to them
- Decline to call BS
- $A = A_{\text{play}} \cup A_{\text{challenge}}$ 
  - $A_{\text{play}} = \text{all truth/bluff actions}$
  - $A_{\text{challenge}} = \text{Call BS or Decline}$

State Transitions:

- Player plays a card
  - After a player plays 1-4 cards of the current number, they can tell the truth or lie. The game records the information about what they said they played, what they actually played (truth or lie), moves to the next player, and moves to the next number in the sequence.
  - For  $a = (c_{\text{actual}}, r_{\text{claim}}, q)$  by player  $p$ :  $S' = T(S, a) = (H_i', N', R', C', B')$ 
    - $H_i' = \begin{cases} H_i - c_{\text{actual}} & \text{if } p=i \\ H_i & \text{else} \end{cases}$
    - $n_p' = n_p - c_{\text{actual}}$
    - $N' = (n_1, \dots, n_p', \dots, n_k)$
    - $R' = (R \bmod 13) + 1$
    - $C' = C + (p, r_{\text{claim}}, q, c_{\text{actual}})$
    - $B' = \text{updateBelief}(B, p, c_{\text{actual}}, r_{\text{claim}})$
- Calling BS
  - If a player calls BS, then the most recent cards are revealed. If the player was telling the truth, then the player who called BS has to take the card, else the player who played the cards has to take the pile. Then the game moves to the next player

- For  $a = BS$ 
  - $S' = T(S, BS) = (H_i', N', R', C', B')$ 
    - If the lie is caught ( $r_{claim} \neq r_{actual}$ )
      - $H_i' = \int H_i \cup D$  if  $c = i$   
 $\setminus H_i$  else
      - where  $D =$  discard pile
      - $n_p' = n_c + D$
    - If truth ( $r_{claim} = r_{actual}$ )
      - $H_i' = \int H_i \cup D$  if  $p = i$   
 $\setminus H_i$  else
      - where  $D =$  discard pile
      - $n_p' = n_p + D$
    - $H_i' = H + \text{challenge}(c, p, \text{outcome})$
    - $B' = \text{updateBelief}(B, c_{actual})$
  - No challenge (the player does not call BS)
    - $T(S, \text{no challenge}) = S$

Observations:

- What other players claimed to have played
- How many cards players currently hold
- Whether someone called “BS” and whether the player was telling the truth in that case or not
  - $Z(s', a, o)$  probability that the agent observes  $o$  after moving to state  $s'$

## Solution Method

Steps:

- At the beginning of the game:
  - Initialize probability tables
  - Set up opponent models with default assumptions
- Update all beliefs:
  - Card probabilities based on my hand
  - Opponent trust scores based on recent outcomes
  - Game phase based on cards remaining

- If we decide to challenge a claim:
  - Get opponent's historical bluff rate
  - Run Monte Carlo simulation of claim feasibility
    - Count known cards of the rank in our hand and previously revealed
    - Calculate the remaining possible cards
    - If it is unlikely to be true then return that it is impossible
    - For 200 simulations:
      - Randomly distribute remaining possible cards among players
      - Weight distribution by current probability estimates
      - Check if the claimer has at least the claimed count copies in this sample
      - Calculate the probability the claim is physically possible
  - If the probability the are lying is high enough then challenge the claim
- Else play:
  - Check if we have the cards in the current rank
  - If we don't have the right cards then bluff
  - Else:
    - Calculate the bluff probability considering the game phase, pile risk, opponent tendencies, and recent success
    - If the bluff probability is high enough, choose the bluff size and select which cards to play
    - Else play truthfully