# Final Project Proposal: Optimal Poker Strategies

**Matthew Simpson** mssimpso@stanford.edu
**Michael Buchar** mbuchar@stanford.edu
*AA228/CS238, Stanford University*

## 1. Goal

Within this project, methodologies for policy generation will be explored in an attempt to optimize a single hand of poker to maximize profit. The number of moves required for each game is limited; thus, the computational time required to arrive at solutions will allow for multiple concepts discussed in this class to be implemented and contrasted to find the best policy that fits our expectations. The game is simple to execute and has a finite set of discrete states that remove complexity in attempting to fit continuous data sets, providing easy compatibility with methods that have been discussed. Success within this project can be defined by generating an MDP model (Kochenderfer et al. (2022)) for a single round of a poker game and extrapolating that framework to an entire game. This model shall achieve higher profits than a stochastic player and similar profits to known poker models that can be used to compare our model.

## 2. Decision Making

The initial state of the problem is two variables are assigned a discrete set of possible values of equal probability. The bet size and current cards on the table will also be captured in the state space. It may be possible to include a model for the value of a player's bets and what can be inferred by mapping those bets to probabilities, but this additional state variable depends on whether the computer players are stochastic in their implementation. These two variables dictate the decision to generate actions to "check," "raise," "call," or "fold" at each subsequent state based on the certainty of variables known as the "flop," "turn," "river," and the uncertainty outlined in the next section. The transition between states occurs when the "flop," "turn," and "river" become known to the player, for which new probability models can be computed based on the known card variable values and possible unknown card variable values. At each state, more information becomes available to the player, making the process sequential and narrowing down the possible ending states as the game progresses. The final state is a deterministic comparison of combinations to determine the winner based on each player's cards. The reward function will be based on the size of the pot or total bets on the table, subtracted from the amount that was bet by the player. The objective of the game is to make decisions that take into account the known and unknown cards to manage risk and ultimately generate a return on investments made into the game.

## 3. Sources of Uncertainty

There is a single source of uncertainty in the game that manifests in multiple ways. The uncertainty is bounded the number of suits and possible values that each card may possess. The uncertainty of the value of the cards held by each player dictates the possible combinations with cards on the board, while cards on the board dictate the possible winning states. The probability of each winning state determines which states are more favorable in the determination of a single winner. As each game progresses, there is less uncertainty in the final outcome state, but there is persistent uncertainty in the value of bets that players could issue, increasing risk as the game closes.

## References

Mykel J. Kochenderfer, Tim A. Wheeler, and Kyle H. Wray. *Algorithms for Decision Making*. MIT Press, 2022.