# Final Project Proposal: Hit or Stand?

**Naixiang Gao** ngao4@stanford.edu
**Jack Ren** zr11@stanford.edu
**Jiaqi Shao** jiaqis7@stanford.edu
*AA228/CS238, Stanford University*

## 1. Goal

Blackjack is a card game where players compete against the dealer to have a hand that's closer to 21 without going over. Players are always required to make decisions under uncertainty (Hewig, 2007). In our game, face cards such as Jack, Queen, and King have a value of 10. Aces have a value of 1 or 11, and other cards have their value. Usually, there is no upper limit for the number of participants in this game. For the simplicity of demonstrating a sequential decision-making process, we choose to have two players for our game. There are two types of players, which are agents and dealers. At the start of the game, two random cards will be distributed to each player. The dealers need to reveal the first received card of these two cards. Then, the agents draw the cards first and don't need to reveal the cards to the rest of the players. The dealers draw the card later. In our game, we will always simulate the role of the agent.

For this project, We will generate an algorithm that maximizes our chance of winning. The algorithm will base on the current value of our cards and the value of the agent's revealed card to decide the next step, which is to hit or stand. To hit means requesting another card and to stand means hold your total value and end your turn. (Grosvenor, 2021)

## 2. Decision Making

Our project aims to use Q-learning to instruct the player in making optimal decisions in the game of Blackjack. Blackjack is a well-known and classic sequential decision-making problem. In this game, the player makes continuous decisions between **hitting** or **standing**, each of which influences the subsequent game state and ultimately affects the outcome (win, tie, or loss). Specifically, the player's goal is to make a series of decisions that bring their hand as close to 21 as possible without exceeding it, while contending with the dealer's actions, which are randomly determined.
In our project, we will simplify the environment to consist of a simulated blackjack table with only two players (the agent plays against the dealer). Additionally, there are only two actions: hit and stand. The model(as a MDP) of this problem is shown below:

1. **States:** The total hand value of the agent, whether the agent has an ace in hand, the value of the deal's visible card, and whether the value of the deal's visible card is ace.
2. **Actions:** Hit and stand.
3. **Transitions:** When the agent chooses to "hit," it will randomly draw a card from the deck and update its total hand value. When the agent chooses to "stand," their turn ends, and the dealer begins to draw cards according to the Stand on Hard 17

rules. Furthermore, if either the player or the dealer busts (exceeds a total of 21), the other automatically wins the game.

4. **Rewards:**
   (a) Win: +1 (general case) / +2 (when the dealer's visible card is an Ace).
   (b) Loss: -1 (general case) / -2 (when the dealer's visible card is an Ace).
   (c) Tie: 0 (general case) / +0.5 (when the dealer's visible card is an Ace).

## 3. Sources of Uncertainty

There are several sources of uncertainty which might influence the action of player. The first one is the dealer's hidden card: the player would only see one of the dealer's cards and another one is unknown. The player's action to hit or stand is strongly influenced by guesses about the dealer's hidden card. The second one is the uncertainty about what value will come from the next draw if they hit. The player's hand value influences decisions like whether to hit or stand. For example, if the player has a total of 16, they might face a higher risk of busting by hitting but might feel they need to hit if they believe the dealer has a strong hand. The third one is the remaining cards in the deck: if the player knows that many high-value cards (like 10s and face-cards) have already been drawn, they may be more inclined to hit since there's a lower chance of busting. On the other hand, they might choose to stand if the deck still contains many high-value cards.

We could model all those uncertainties in a Bayesian Network. The nodes could be: player's current hand value, the dealer's visible card, the dealer's hidden card, the remaining cards in the deck, the player's action, the next card for the player, the dealer's action, the dealer's final hand, player busts, dealer busts, and outcome of the game. Some edges that we could build are: player's action is influenced by the player's current hand value, dealer's visible card, estimate of dealer's hidden card, and remaining cards in the deck; remaining cards influence the next card for player in the deck and player's action.

## References


Grosvenor. Blackjack strategy: Hit or stand?, 2021. URL `https://blog.grosvenorcasinos.com/blackjack-strategy-hit-or-stand#:~:text=When%20playing%20blackjack%20there%20are,total%20and%20end%20your%20turn.`

Johannes Hewig. Decision-making in blackjack: An electrophysiological analysis. *Cerebral Cortex*, 17(4):865–77, 2007.