Examples

1. Blackjack game

The game is detailed explained on chapter 5.1 – Monte Carlo Policy Evaluation of [1]. Assuming you already know how to play blackjack, the game can be summarized as:

- Face cards (Jack, Queen, King) have point value 10
- Aces can either count as 11 or 1, and it's called 'usable' at 11.
- The game starts with each (player and dealer) having one face up and one face down card.
- The player can request additional cards (hit=1) until they decide to stop (stick=0) or exceed 21 (bust)
- After the player sticks, the dealer reveals their facedown card, and draws until their sum is 17 or greater. If the dealer goes bust the player wins.
- If neither player nor dealer busts, the outcome (win, lose, draw) is decided by whose sum is closer to 21. The reward for winning is +1, drawing is 0, and losing is -1

The observation of the agent at every state is a Python 3-tuple of:

- The player current sum
- The dealer's showing car (1, 2, …, 10)
- Whether or not the player holds a usable ace (0, 1)

Check the detailed code implementation here.