Black Jack

Corrado Possieri

Machine and Reinforcement Learning in Control Applications

Problem

Learn to play black jack.

Problem formulation

- The goal is to obtain cards whose sum is the closest to 21.
- All face cards count as 10.
- An ace counts as 1 or as 11.
- Each player competes independently against the dealer.
- Begins with two cards dealt to both dealer and player.
- One of the dealer's cards is face up and the other is face down.

Winning and losing the game

- The player can request additional cards (hit), one by one.
- If the player exceeds 21, he goes burst an loses the game.
- If the player stops (sticks), it becomes the dealer's turn.
- The dealer plays according to a fixed and known strategy
  - stick on any sum of 17 or greater, and hit otherwise
- If the dealer goes burst, the player wins; otherwise, the outcome is determined by whose final sum is closer to 21.

Modeling black jack

- Reward
  - $+1$ for winning.
  - -1 for losing.
  - 0 for drawing.
- All rewards within a game are zero.
- Do not discount (γ=1).
- Player actions are to hit or to stick.
- Assume that cards are drawn from an infinite deck.

State in black jack

- An ace is *usable* if it can count as 11 without going burst.
- If the sum 11 or less, there is no decision to be made because the player should always hit.
- The player takes decision on the basis of 3 variables
  - the current sum of its hand (12-21);
  - the card shown by the dealer (1-10);
  - whether there is an usable ace (0-1).

Assignment #3

- Consider driving a race car around a turn like those shown in Figure 5.5 of the textbook.
- Go as fast as possible without going out the track.
- The car is at one of a discrete set of grid positions.
- Velocity is discrete: number of grid cells moved horizontally and vertically per time step.
- Actions are increments to the velocity components.
- Velocities restricted to be less than 5.
- Episodes start at random on started line.
- Episodes terminates if you cross the finish line.
- The rewards are -1 for each step.
- If you hit the boundary, reset randomly on the starting line.