CS747: Assignment 2

Yashvardhan

October 2022

1 Task1

1.1 Planner.py

- I implemented all algorithms i.e. value interaction, howard policy improvement and linear programming
- In my testing phase VI performed the fastest among the algorithm and hence also default algorithm for task 2.
- HPI was the slowest among all
- policy evaluation and find V* given a policy was implemented as a sub-routine for the HPI

2 Task2

2.1 Encoder.py

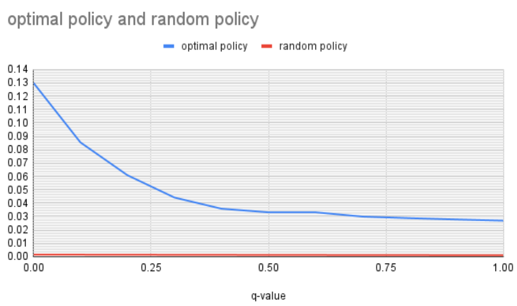
- We were given as n-states of the form bbrr. Here encoded it as 2*n+2 state mdp corresponding to balls and runs left in the game and the player at stricker end.
- 2 extra states corresponding to winning or losing (could have done with fewer states :)
- Transition probability table was completed using p1-parameter.txt
- Reward of 1 is given when the winning state is achieved

2.2 Decoder.py

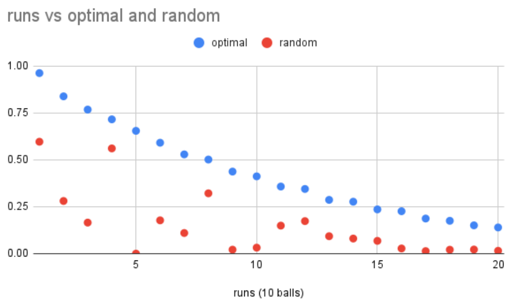
- There is a one-to-one translation of states. Just reverse-engineered the mdp states to cricket states.

2.3 Graphs

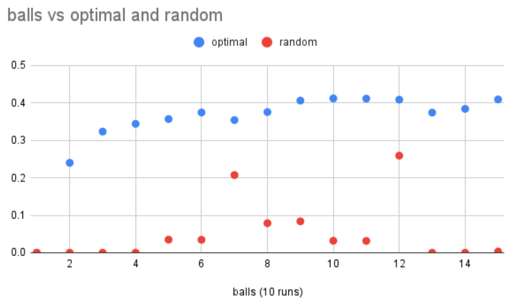
2.3.1 15 balls, 30 runs, Optimal vs random win prob



2.3.2 fix balls, vary runs to be scored



2.3.3 fix runs, vary balls left



Inference:

- In every graph we can see that random policy can not guarantee better performance than the optimal policy
- Optimal policy follow our intuition of fewer runs to be scored and more balls left, the higher the win probability
- In graph 1 changing increasing q, reflected in the weakness of the tail-ender, hence the probability monotonically decreased for both optimal and sub-optimal policy
- In graph 2, with increasing, runs to score, win probability decreased
- In graph 2, random policy player always performed worse than the optimal one, but saw irregularity in win probability. This can be due to him taking the best/better or most suitable action at that state.
- In graph 3, for optimal we regular increase in win probability with an increasing number of balls and some randomness for sub-optimal policy player
- In all, all these graphs show that optimal policy outperforms sub-optimal policy for every state. This empirically demonstrates existence of **"the optimal policy"** such that V* ¿+ V for all states and policy