Homework 1: Nov. 11

Lecturer: Yishay Mansour

Homework number 1.

Question 1

Consider a multiplicative model, where the return function of the reward sequence $r_1, \ldots, r_N$ is $\prod_{i=1}^{N} r_i$. Develop an algorithm that evaluates a given deterministic Markovian policy for such a return function.

Question 2

Suppose we have two actions. Action **D** has always reward 0, while action **R** has reward $+1$ with probability $1/2$ and -1 with probability $1/2$. Our aim is to maximize the probability that the sum of the rewards is at least 1. What is the optimal policy. (Hint: Show that there exists an optimal policy that first performs a sequence of R operations and only latter performs a sequence (possibly empty) of D operations.)

Question 3

Consider the following digit game. We have $N+1$ slots, marked by 0 to N. At each round (of $N+1$ rounds) a random digit is chosen. (I.e. we chose a random number from $\{0, 1, \ldots, 9\}$.) We can put the digit in one of the unoccupied slots, and that slot becomes occupied. Our aim is to maximize the expected value of the sequence of digits, when viewed as a number. (I.e. if we have d_i in slot i then the final value is $\sum_{i=0}^{N} d_i 10^i$.)

Example: Suppose $N = 2$. At the beginning we have three empty slots. Lets denote an empty slot by $\times$, then we have $\times\times\times$. Suppose the first digit is 5, and we decide to put it in the second slot, then we have $\times 5 \times$. Suppose the second digit is 4 and we decide to put it in the first slot, we have $\times 54$. Finally, the last digit is 4 again, and our final number is 454.

1. Build an MDP for this problem. (Hint: you may incorporate the random digit in the state.)
2. Show that the optimal policy depends only on the number of empty slots.
3. Compute the optimal policy for $N = 2$. (I.e the number has three digits.)

The homework is due in two weeks