EE488C Special Topics in EE <Deep Learning and AlphaGo> November 9, 2016
EE, KAIST, Fall 2016 Due: 1:00pm, November 16

Problem Set #3

Warning: Homeworks will not be graded if submitted after the deadline. For all problems, show detailed reasoning.

1. (Two-armed bandit problem) Consider a two-armed bandit problem, where the reward for the first arm is Bernoulli(p) and the reward for the second arm is Bernoulli(q), where $0 < p, q < 1$. Assume $A_1 = 1$ and $A_2 = 2$, i.e., you choose the first arm at time 1 and choose the second arm at time 2. Knowing the outcomes of the two tries, which arm should you choose at time $t = 3$ to maximize the chance of getting the reward of one at $t = 3$? Hint) You can also use a random strategy. Since the values of p and q are assumed to be unknown, try to maximize the minimum of the two expected rewards, one assuming p and q are swapped and the other assuming they are not. Maximizing the minimum of the two rewards will make your strategy symmetric and not dependent on the values of p and q.

2. (Markov property) Find an example of three binary random variables X, Y, Z such that X is Bernoulli($\frac{1}{5}$), Y is Bernoulli($\frac{2}{5}$), Z is Bernoulli($\frac{7}{15}$), and $X - Y - Z$ form a Markov chain in that order. Hint) If A is Bernoulli(p), B is Bernoulli(q), and A and B are independent, then $A \oplus B$ is Bernoulli($p(1-q) + (1-p)q$), where $A \oplus B$ is the XOR of A and B.

3. (Value functions) Find the state-value function $v_\pi(s)$ for the continuing task given in the figure in the right side in page 8 of lecture notes #16. Assume π is an optimal policy and $0 < \gamma < 1$.

4. (Maze game) What is the minimum number of value iterations needed to learn to find the shortest path in the 20×20 maze game in page 19 in lecture notes #15? Assume $v_0(s)$'s are initialized to zero, the starting state is $(1, 9)$ (i.e., the 9th cell from left in the first row), and the terminal state is $(20, 20)$ (i.e., right bottom corner cell). The reward is 1 when the terminal state is reached from $(20, 19)$ and is 0 for all other transitions. Possible actions are up, down, left, and right. Assume the state is unchanged if an action is taken whose movement is blocked by an wall. If the action is not blocked by a wall, then you move by one cell. Assume $0 < \gamma < 1$.