Midterm Quiz II

EL-GY 9233 Reinforcement Learning for Complex Networks

Spring 2015, New York University

1:00pm – 2:00pm, May 15, 2015

- The exam consists of 2 problems, and there are 20 pts in total.
- You do not need calculators for the exam.

Problem 1 (10pts)**.** Consider the following Rock-Paper-Scissor game. Both players are maximizers with the following payoff matrix

	R	P	S
R	(0,0)	(0,1)	(1, 0)
P	(1,0)	(0,0)	(0,1)
S	(0,1)	(1,0)	(0, 0)

Let $x = (x_R, x_P, x_S)$ and $y = (y_R, y_P, y_S)$ be the mixed strategies of P1 and P2, respectively. The regularized expected payoff to P1 and P2 can be written as $U_1(x, y) = x'Ay + H(x)$ and $U_2(x, y) = x'By + H(y)$, where H is the entropy term defined by $H(z) := \sum_{i=1}^{3} z_i \log z_i$.

- Find the best response functions of the two players.
- Write down the fictitious play learning algorithm for estimating x and y, i.e., specify the update functions F and G in the following updates:

$$\hat{x}_i(k+1) = \hat{x}_i(k) + F(k, a_k, \hat{x}, \lambda_k, A), \;\; i \in \{R, P, S\},$$

$$\hat{y}_j(k+1) = \hat{y}_j(k) + G(k, b_k, \hat{y}, \mu_k, B), \;\; j \in \{R, P, S\},$$

  where λ_k, μ_k are learning rates, and a_k, $b_k \in \{R, P, S\}$ are the actions taken by P1 and P2 at time k, respectively. Explain the mechanism of the learning algorithm.
- Write down the continuous-time ODE approximation of the above stochastic processes.

Problem 2 (10pts)**.** Consider the following stochastic game with two players P1 and P2. P1 is the maximizer, and P2 is the minimizer. The game has two states $S = \{s_1, s_2\}$. At state s_1 and s_2, the payoff matrices are

	L	R
U	3	6
D	2	1

,

	M
M	0

, respectively,

with transition probability

$$P_{12} = \begin{array}{c|c|c} & \text{L} & \text{R} \\ \hline \text{U} & 0 & \frac{2}{3} \\ \hline \text{D} & 0 & 0 \end{array}, \; P_{11} = \begin{array}{c|c|c} & \text{L} & \text{R} \\ \hline \text{U} & 1 & \frac{1}{3} \\ \hline \text{D} & 1 & 1 \end{array}, \; P_{21} = 0 \text{ and } P_{22} = 1.$$

- Write down the dynamic programming equation for this zero-sum stochastic game.
- Find the value function and the stationary mixed strategies for P1 and P2.
- Write down the multi-agent Q-learning algorithm associated with this game.