CS224M, Fall 2009-10
Homework #3 : Multi-Agent Learning & Protocols for Agents
Handout #4

Note: Homework #3 is due on **November 2, 2008** in class

1. **[25 points] (Fictitious Play)** Consider the normal form game given in Table 1. Player 1 is the row player, and player 2 is the column player. The two parts of this question are separate, but both concern fictitious play of this game, with ties broken arbitrarily.

	L	R
U	$(2, 1)$	$(3, 1)$
D	$(4, 0)$	$(1, 3)$

Table 1: A two-player, normal form game.

Note that this game does not have generic payoffs, and thus for part (b) you cannot apply Theorem 7.2.5 on page 199. Also, the intent was not for you to solve for all Nash equilibria (you did enough of that in the previous problem set), and it's not necessary to do so for this problem. If it helps, the set of all Nash equilibria for this game are the strategy profiles in which player 1 plays U with probability 1 and player 2 plays L with probability p and R with probability $(1 - p)$, for $p \leq 0.5$.

(a) [13 points] In this part of the problem, player 1 starts with initial beliefs over (L, R) of (2.5,1) – which means that he initially gives a weight of 2.5 to player 2's action L and a weight of 1 to player 2's action R – and player 2 starts with initial beliefs over (U, D) of (1,1).
Show the actions and updated beliefs for the first 6 rounds of fictitious play (starting at round 1) by filling in the following Table 2. The beliefs of an agent are updated at the end of a round. Thus, the beliefs in round i include the actions that were observed in the i-th round, while the actions in round i are based on the beliefs of round $i - 1$. You do not need to show any work for full credit (but, of course, it can lead to partial credit if you just make a math error).

Round	Player 1's action	Player 2's action	Player 1's beliefs	Player 2's beliefs
0(initial)	xxx	xxx	(2.5,1)	(1,1)
1				
2				
3				
4				
5				
6				

Table 2: Fictitious play.

(b) [12 points] In this part, we do not specify the initial beliefs of the players– we only restrict them. Specifically, in a player's initial beliefs, the weight he assigns to an action of the other player can be any value in the range $[1, 10]$. Of the four action profiles for this game, list all that could possibly be played in the millionth round of fictitious play. Briefly justify your answer.

2. **[25 points] (Rational Learning)** Suppose two agents are each using rational learning to guide their play in the infinitely repeated Prisoner's Dilemma game with limit average rewards. Each considers the set of possible strategies to be all stationary strategies, defined to be those strategies that are based only on the latest play of the game (that is, the set of 32 mappings from outcomes in the previous turn to actions plus an initial move; for instance $\{CC \to D, CD \to D, DC \to D, DD \to D$, initial move $D\}$ would be the "always defect" strategy).

   (a) [5 points] Using this representation define the grim strategy (cooperate until your opponent defects then defect forever) and the tit-for-tat strategy (cooperate first, then reply with your opponent's previous move) as stationary strategies. (Under the notation above, set the first action to be the agent's own action last turn, so DC indicates they defected and their opponent cooperated).

   (b) [10 points] Is there a pair of stationary strategies s from the set we are considering, one for each agent, and a belief P_i for each agent about the distribution of her opponent's strategies such that at each time the strategies played are best response for each agent to their current beliefs about the opponent's possible strategies, and μ_s is absolutely continuous with respect to μ_{P_i}? (Recall that μ_x is the distribution over histories induced by x).

   (c) [10 points] Assuming the agent starts with an initial uniform distribution over the 32 "pure" strategies, use Bayesian updating to calculate the resulting beliefs of the agent after observing the following play: CC, CD, DD, DC. Calculate the best repeated game strategy for the agent to adopt given these new beliefs.

3. **[25 points] (Evolutionary Learning)** Consider the two-player symmetric game in Table 3.

	A	B
A	$(3, 3)$	$(3, 0)$
B	$(0, 3)$	$(3, 3)$

Table 3: A two-player game

For a population of agents using the replicator dynamic, indicate whether each of the following states is a steady state, a stable steady state, an asymptotically stable steady state, or none of the above. The states are listed as (a, b) where a is percentage of the agents playing the pure strategy A and b is the percentage of the agents playing the pure strategy B.

Justify your answers.

   (a) [5 points] $(a = 1, b = 0)$?

   (b) [5 points] $(a = 0, b = 1)$?

   (c) [5 points] $(a = 0.5, b = 0.5)$?

   (d) [10 points] Suppose that now you treat the above states as *mixed strategies* (e.g. (c) means play A half the time, and B half the time). Which of the states above are ESS?

4. **[25 points] (Social Choice Functions)**

Consider the following social choice function, which we'll call benign dictatorship (BD). Under this social choice function, each voter submits a strict total ordering over the candidates, and the winning candidate is determined as follows. For each candidate, count the number of times that it appears at the top of a voter's ordering. If there is a single candidate with the maximum count, then it is declared the winner. Otherwise, there must be a tie. The tie is broken according to the preferences of voter #1 (who is, thus, the benign dictator). Note that in case of a tie, the winner need not be #1's top choice. Thus, BD is really a *plurality* rule with dictatorial tie breaking.

(a) Specify whether BD satisfies each of the following properties. Informally justify your answer if it does, or provide a counterexample otherwise.

i. [5 points] *unanimity*
ii. [5 points] *dictatorship*
iii. [5 points] *monotonicity*

(b) [10 points] Recall from the course reader the *Condorcet condition*, which requires that if a candidate x is chosen, then for any candidate $y \neq x$ it must be the case that at least half the voters prefer x to y. Does BD satisfy this condition? If so, informally justify why it does. Otherwise, present a counterexample with explanation.