CS 684 Algorithmic Game Theory	Scribe: Craig Bowles
Instructor: Eva Tardos	November 11, 2005

1. Learning and Rock-Paper-Scissors

To avoid repeating everything from last class, let's look at example game: Rock-Paper-Scissors (see Figure 1) where winner gets 1 and loser loses nothing (and thus not zero-sum).

	R	P	S
R	0,0	0,1	1,0
P	1,0	0,0	0,1
S	0,1	1,0	0,0

Figure 1: Traditional Rock-Paper-Scissors game with 0/1 payoff.

For every row or column, there is exactly 1 winning position. Let A, B be the payoff matrices for columns and rows, respectively. Assume both no regret learning for T steps. Row player has probabilities p_1, p_2, p_3 and column player has probabilities q_1, q_2, q_3 for rock, paper, scissor respectively which are empirical, i.e. p_1 is the fraction of time row played rock, etc. Let d_C be the value of best column with hindsight. Row played row i $p_i T$ times, so reward of different columns is $T \times pA$ and d_C is max entry in $T \times pA$. Completely analogously, d_R = max value of $T \times Bq$ is the value of the best row with hindsight.

From learning we showed that the column player gets at least $d_C - \frac{\epsilon}{2} d_C - \epsilon^{-1} \log n$ and the row player gets at least $d_R - \frac{\epsilon}{2} d_R - \epsilon^{-1} \log n$. As for the choice of ϵ, note that $d_R \leq T$ trivially. Thus column gets at least $d_C - \frac{\epsilon}{2} T - \epsilon^{-1} \log n$ and similarly for the row, so set $\epsilon = \sqrt{2 \frac{1}{T} \log n}$ and so the column and row players get, respectively, at least $d_C - 2\sqrt{2T \log n}$ and $d_R - 2\sqrt{2T \log n}$. If we further make the assumption that $T \geq \delta^2 \log n$, then we see that $2\sqrt{2T \log n} \leq O(\delta T)$.

Look at solutions with 0-regret (i.e. no error term). For example, consider $p_i = q_j = 1/3$ for all i, j, which happens to be the unique Nash Equilibrium for the game. No regret for the column means that the column's income is at least $T/3$ and similarly for the row player. What the players should do to make more is to plan moves so that they never, or rarely, play the same thing because no one wins then, and this is just what learning does! Learning goes in phases and generates roughly $T/2$ income each. Each time once a strategy gets dominant for one player, the losing player wakes up and picks a better strategy, but the previous winner was making so much on that spot that he put a lot of weight on that spot, so the other player makes up money for a while until the other player decided to put its weight somewhere else. This is the learning strategy and will produce a cycle on all the non-diagonal states.

Now consider there being a "giving state" as shown in Figure 2. Now there is still only one Nash, but it is a deterministic Nash, the deterministic choice $[G, G]$. Unfortunately, that Nash is not going to happen with learning. With learning the players will assign less and less probability to G and the previous learning strategy is learned with roughly $T/2$ income each, which is worse than the Nash.

	R	P	S	G
R	0,0	0,1	1,0	2,0.34
P	1,0	0,0	0,1	2,0.34
S	0,1	1,0	0,0	2,0.34
G	0.34,2	0.34,2	0.34,2	2.1,2.1

Figure 2: Modified Rock-Paper-Scissors that includes Giving state in which both players have positive payoff.

2. Correlated Equilibrium

A regular Nash is a probability p on rows and q on columns where (i, j) is played with probability $p_i \cdot q_j$. A *Correlated Equilibrium*, or *Correlated Nash*, is a probability distribution on pairs of strategy (one strategy per player) in that we tell players what to play, but don't tell them what state they're playing. It may be that one player is happy and one player is not with that state, but you as the player do not know whether you are the loser or the winner. For example, if the row is told to play scissor, he has no idea if the column is told to play rock or paper. By doing this we can prevent users from every playing the same strategy and thus both winning nothing. It is a Correlated Nash if a player does not gain by deviating.

Consider the Dare-Chicken game in Figure 3. Imagine this is a game where two cars are at a 4-way intersection and to "dare" means to drive through the intersection, and to "chicken" is to stop at the intersection. The off-diagonal entries are each deterministic Nashes, and there is a randomized Nash given by assigning probability 1/2 to both of these entries. The 1/2-1/2 off-diagonal strategy is a Correlated Equilibrium as well, but the 1/3-1/3-1/3 non-$[Dare, Dare]$ strategy is as well, and it is clear to see that this is a better strategy for both players than any Nash.

	Dare	Chicken
Dare	0,0	6,1
Chicken	1,6	5,5

Figure 3: Chicken-Dare game which offers a Correlated Equilibrium that is better than any Nash.

This type of equilibrium depends on a trust of the players to use the given strategy, such as given by a stop light. They are computationally nice as well, unlike Nash Equilibria.