NETS 412: Algorithmic Game Theory	February 9, 2017
Lecture 8	
Lecturer: Aaron Roth	*Scribe: Aaron Roth*

Correlated Equilibria

Consider the following two player traffic light game that will be familiar to those of you who can drive:

	STOP	GO
STOP	(0,0)	(0,1)
GO	(1,0)	(-100,-100)

This game has two pure strategy Nash equilibria: (GO,STOP), and (STOP,GO) – but these are clearly not ideal because there is one player who never gets any utility.

There is also a mixed strategy Nash equilibrium: Suppose player 1 plays $(p, 1-p)$. If the equilibrium is to be fully mixed, player 2 must be indifferent between his two actions – i.e.:

$$0 = p - 100(1-p) \Leftrightarrow 101p = 100 \Leftrightarrow p = 100/101$$

So in the mixed strategy Nash equilibrium, both players play STOP with probability $p = 100/101$, and play GO with probability $(1-p) = 1/101$. This is even worse! Now both players get payoff 0 in expectation (rather than just one of them), and risk a horrific negative utility. The four possible action profiles have roughly the following probabilities under this equilibrium:

	STOP	GO
STOP	98%	<1%
GO	<1%	≈ 0.01%

A far better outcome would be the following, which is fair, has social welfare 1, and doesn't risk death:

	STOP	GO
STOP	0%	50%
GO	50%	0%

But there is a problem: *there is no set of mixed strategies that creates this distribution over action profiles.* Therefore, fundamentally, this can never result from Nash equilibrium play.

The reason however is not that this play is not rational – it is! The issue is that we have defined Nash equilibria as profiles of mixed strategies, that require that players randomize independently, without any communication. In contrast, the above outcome requires that players somehow correlate their actions.

Drivers of course do this all the time – the correlating device is a traffic light. The traffic light suggests to each player whether to STOP or GO, and (at least when roads are busy), conditioned on the advice it gives you following its advice is a best response for everyone involved.

This idea can be generalized:

Definition 1 *A* correlated equilibrium *is a distribution $\mathcal{D}$ over action profiles A such that for every player i, and every action a_i^*:*

$$\mathrm{E}_{a\sim\mathcal{D}}[u_i(a)] \geq \mathrm{E}_{a\sim\mathcal{D}}[u_i(a_i^*, a_{-i})|a_i]$$

In words, a correlated equilibrium is a distribution over action profiles a such that after a profile a is drawn, playing a_i is a best response for player i conditioned on seeing a_i, given that everyone else will play according to a. For example, in the traffic light game, conditioned on seeing STOP, a player knows that his opponents see GO, and hence STOP is indeed a best response. Similarly, conditioned on seeing GO, he knows that his opponents see STOP, and so GO is a best response.

Nash equilibria are also correlated equilibria – they are just the special case in which each player's actions are drawn from an independent distribution, and hence conditioning on a_i provides no additional information about a_{-i}. But as we saw above, the set of correlated equilibria is strictly richer than the set of Nash equilibria.

We can define an even larger set still:

Definition 2 *A* coarse correlated equilibrium *is a distribution $\mathcal{D}$ over action profiles A such that for every player i, and every action a_i^*:*

$$\mathrm{E}_{a \sim \mathcal{D}}[u_i(a)] \geq \mathrm{E}_{a \sim \mathcal{D}}[u_i(a_i^*, a_{-i})]$$

The difference is that a coarse correlated equilibrium only requires that following your suggested action a_i when a is drawn from $\mathcal{D}$ is only a best response in expectation *before* you see a_i. This makes sense if you have to commit to following your suggested action or not up front, and don't have the opportunity to deviate after seeing it. A coarse correlated equilibrium can for example occasionally suggest that players play obviously stupid actions. Consider the following game, and distribution over action profiles:

	A	B	C
A	(1,1)	(-1,-1)	(0,0)
B	(-1,-1)	(1,1)	(0,0)
C	(0,0)	(0,0)	(-1.1,-1.1)

	A	B	C
A	1/3		
B		1/3	
C			1/3

The payoff for each player for playing according to this distribution is:

$$(1/3) \cdot 1 + (1/3) \cdot 1 - (1/3) \cdot 1.1 = 0.3$$

In contrast the payoff a player would get by playing the fixed action A or B while his opponent randomized would be:

$$(1/3) \cdot 1 - (1/3) \cdot 1 + (1/3) \cdot 0 = 0$$

and the payoff for playing C would be strictly less than zero. Hence, the given distribution is a coarse correlated equilibrium *even though* conditioned on being told to play C, it is not a best response. This means that the given distribution is a coarse correlated equilibrium, *but not* a correlated equilibrium, proving that coarse correlated equilibria are a strictly larger set of distributions.

To recap, we have so far considered several solution concepts: Dominant strategy equilibria (DSE), Pure strategy Nash equilibria (PSNE), mixed strategy Nash equilibria (MSNE), correlated equilibria (CE), and coarse correlated equilibria (CCE), and we know the following strict containments:

$$DSE \subset PSNE \subset MSNE \subset CE \subset CCE$$

where starting at Mixed Nash equilibria, the solution concept is guaranteed to exist (but may still be hard to find). We want to show that starting at Correlated equilibria, not only is the solution concept guaranteed to exist, but we can always efficiently compute one.

Lets now characterize these new equilibrium concepts using the notion of regret that we saw last lecture.

Definition 3 *For a* strategy modification rule $F_i : A_i \to A_i$ *and an action profile* $a \in A$*:*

$$Regret_i(a, F_i) = u_i(F_i(a_i), a_{-i}) - u_i(a)$$

i.e. it is how much player i regrets not applying F_i to change his action.

We say that F_i is a constant strategy modification rule if $F_i(a_i) = F_i(a_i')$ for all $a_i, a_i' \in A_i$.

We can give an equivalent definition of coarse correlated equilibrium using this notion of regret:

Definition 4 *A distribution $\mathcal{D}$ is a coarse correlated equilibrium if for every player i and for every* constant *strategy modification rule F_i:*

$$\mathrm{E}_{a\sim\mathcal{D}}[Regret_i(a, F_i)] \leq 0$$

Note that one immediate consequence of this definition is that if $a^1, \ldots, a^T$ are a sequence of actions with $\Delta(T)$ regret, then $\bar{a} = \frac{1}{T}\sum_{t=1}^{T} a^t$ forms a $\Delta(T)$-approximate coarse correlated equilibrium. This means in particular, that if everyone plays an (arbitrary) game with the polynomial weights algorithm, after T steps they will have generated a sequence of plays that corresponds to a $\Delta(T) = 2\sqrt{\log k/T}$-approximate Coarse correlated equilibrium. In other words, after $T = 4\log(k)/\epsilon^2$ rounds (independent of the number of players!) they will have converged to an ϵ-approximate coarse-correlated equilibrium. This in particular means that unlike Nash equilibria, *coarse correlated equilibria* don't just exist, but are easy to compute in arbitrary games.

Can we say the same thing for correlated equilibria? A natural approach is to first characterize them in terms of regret:

Definition 5 *A distribution $\mathcal{D}$ is a correlated equilibrium if for all players i and for* all *strategy modification rules F_i:*

$$\mathrm{E}_{a\sim\mathcal{D}}[Regret_i(a, F_i)] \leq 0$$

To see that this corresponds to our first definition, note that a strategy modification rule F_i lets player i consider different deviations for each suggested action a_i, and so if there are no beneficial deviations of this sort, player i must be playing a best response even conditioned on seeing his suggestion.

Are there learning algorithms that efficiently converge to correlated equilibrium? A natural strategy (by analogy to how we can find coarse correlated equilibria) is to try and find an experts algorithm that has the following guarantee:

Given any k experts and an arbitrary sequence of losses $\ell^1, \ldots, \ell^T$, the algorithm chooses a sequence of experts $a_1, \ldots, a_t$ such that:

$$\frac{1}{T}\sum_{t=1}^{T} \ell_{a^t} \leq \frac{1}{T}\sum_{t=1}^{T} \ell_{F(a^t)} + \Delta(T)$$

for all strategy modification rules F and for $\Delta(T) = o(1)$.

This guarantee is known as having no *swap-regret.* Rather than guaranteeing that the algorithm merely does as well as the best fixed action in hindsight, it guarantees that the algorithm could not have done better even by applying any swap function – i.e. "Every time I bought Microsoft, I should have bought Apple. But every time I bought Google I should have bought Comcast."

Next lecture, we will see that (remarkably!) there is an efficient algorithm that gives this guarantee!