Lecture 7.2

Theorem: Let $\varepsilon > 0$. Let T be such that the time-averaged external regret of every player is at most ε. Define $\sigma_t = \prod_{i \in N} p_i^t$ and $\sigma = \frac{1}{T} \sum_{t=1}^{T} \sigma_t$. Then σ is an ε-CCE.

Proof:

$$\sigma = \frac{1}{T} \sum_{t=1}^{T} \sigma_t$$

$$\mathbb{E}_{s \sim \sigma}[u_i(s)] = \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}_{s \sim \sigma_t}[u_i(s)]$$

For every $i \in N$, every $s_i' \in S_i$

$$\frac{1}{T}\left[\sum_{t=1}^{T} \mathop{\mathbb{E}}_{(s_i,s_{-i})\sim\sigma_t}\left[u_i(s_i',s_{-i})\right] - \sum_{t=1}^{T} \mathop{\mathbb{E}}_{s\sim\sigma_t}\left[u_i(s)\right]\right] \leqslant \varepsilon$$

$$\mathop{\mathbb{E}}_{(s_i,s_{-i})\sim\sigma}\left[u_i(s_i',s_{-i})\right] - \mathop{\mathbb{E}}_{s\sim\sigma}\left[u_i(s)\right] \leqslant \varepsilon$$

∎

Hence, σ is an ε-CCE.

Q. Does there exist similar natural learning dynamic which enables player to converge to a CE?

Swap-Regret:

"Modification rule" $(\pi_1, \ldots, \pi_t, p_1, \ldots, p_t, a_1, \ldots, a_t)$ history.

takes the history of the play as input and outputs a sequence of action for t iterations.

External regret tries to perform as good as any fixed action; i.e. the modification rules $\{ F_a = (\underbrace{a, \ldots, a}_{t}) \mid a \in A \}$

Swap regret tries to perform as good as any modification rule $\{ F_{a,b} : a, b \in A, a \neq b \}$

whenever the player played a, it outputs b and vice versa.

No Swap-Regret: A learning algorithm is said to have no swap regret if the time-averaged swap regret

$$\frac{1}{T}\left(\max_{\delta \in \mathcal{F}^{swap}} \sum_{t=1}^{T} \sum_{a \in A} p_t(a) \cdot \pi_t(\delta(a)) - \sum_{t=1}^{T} \sum_{a \in A} p_t(a)\, \pi_t(a)\right)$$

goes to 0 as T goes to ∞.

- The connection of no swap-regret algorithm and CE:

Theorem: Let $\varepsilon > 0$. Each player runs a no-swap-regret algorithm.

$$A_i = S_i, \quad \pi_i(s_i) = u_i(s_i, p_{-i}) = \sum_{(s_j)_{j \neq i} \in S_{-i}} u_i(s_i, s_{-i}) \cdot p_j(s_j)$$

Let T be such that the time-averaged swap-regret of every player is at most ε. Define $\sigma_t = \prod_{t=1}^{T} p_t$ and $\sigma = \frac{1}{T} \sum_{t=1}^{T} \sigma_t$. Then σ is an ε-CE of Γ.

Proof: By definition of σ,

$$\mathbb{E}_{s\sim\sigma}[u_i(s)] = \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}_{s\sim\sigma_t}[u_i(s)]$$

For any $i \in N$ and any $\delta : S_i \longrightarrow S_i$

$$\frac{1}{T}\left[\sum_{t=1}^{T}\sum_{a\in A} p_i^t(a)\,\pi_t(\delta(a)) - \sum_{t=1}^{T}\sum_{a\in A} p_i^t(a)\,\pi_t(a)\right] \leqslant \varepsilon$$

$$\Rightarrow \frac{1}{T}\left[\sum_{t=1}^{T}\mathbb{E}_{s\sim\sigma_t}\left[u_i(\delta(s_i), s_{-i})\right] - \sum_{t=1}^{T}\mathbb{E}_{s\sim\sigma_t}[u_i(s)]\right] \leqslant \varepsilon$$

$$\Rightarrow \mathbb{E}_{s\sim\sigma}\left[u_i(\delta(s_i), s_{-i})\right] - \mathbb{E}_{s\sim\sigma}[u_i(s)] \leqslant \varepsilon$$

$\blacksquare$