Lecture 3.1

Iterative Elimination of Dominated Strategies

Definition (Strongly Dominated Strategy): Given $\Gamma = \langle N, (S_i)_{i \in N}, (u_i)_{i \in N} \rangle$ a strategy $s_i \in S_i$ for a player $i \in N$ is called a strongly/weakly dominated strategy if there exists a mixed strategy $\sigma_i \in \Delta(S_i)$ s.t.

$$u_i(s_i, s_{-i}) \;\substack{< \\ \leq}\; u_i(\sigma_i, s_{-i}) \quad \forall s_{-i} \in S_{-i}$$

Lemma: In any game $\Gamma = \langle N, (S_i)_{i\in N}, (u_i)_{i\in N} \rangle$ if a pure strategy $s_i \in S_i$ is strongly dominated, then, in every MSNE $(\sigma_i^*)_{i\in N}$ of Γ, we have $\sigma_i^*(s_i) = 0$.

Proof: by contradiction; Suppose there exists an MSNE $(\sigma_i^*)_{i\in N}$ of Γ such that $\sigma_i^*(s_i) \neq 0$ and s_i is strongly dominated by a mixed strategy $\sigma_i \in \Delta(S_i)$,

We have $\sigma_i \neq s_i$

Consider a mixed strategy $\underline{\pi \in \Delta(S_i)}$ as follows

$$\pi(s_i') = \sigma_i^*(s_i') + \sigma_i^*(s_i)\cdot\sigma_i(s_i')/(1-\sigma_i(s_i)) \quad \forall\, s_i' \in S_i \setminus \{s_i\}$$

$$\pi(s_i') \geqslant 0 \quad \forall\, s_i' \in S_i \quad , \quad \pi(s_i) = 0$$

$$\sum_{s_i' \in S_i} \pi(s_i') = \sum_{\substack{s_i' \in S_i \\ s_i' \neq s_i}} \left(\sigma_i^*(s_i') + \sigma_i^*(s_i) \frac{\sigma_i(s_i')}{1-\sigma_i(s_i)} \right)$$

$$= \sum_{\substack{s_i' \in S_i \\ s_i' \neq s_i}} \sigma_i^*(s_i') + \boxed{\frac{\sigma_i^*(s_i)}{1-\sigma_i(s_i)} \boxed{\sum_{\substack{s_i' \in S_i \\ s_i' \neq s_i}} \sigma_i(s_i')}}$$

$$= \sum_{s_i' \in S_i} \sigma_i^*(s_i') = 1$$

We have

$$u_i(\sigma_i, \sigma^*_{-i}) > u_i(s_i, \sigma^*_{-i}) \quad \text{[by our assumption]}$$

$$u_i(\pi, \sigma^*_{-i}) = \sum_{s_i' \in S_i} u_i(s_i', \sigma^*_{-i}) \cdot \pi(s_i')$$

$$= \sum_{\substack{s_i' \in S_i \\ s_i' \neq s_i}} \sigma^*_i(s_i') \cdot u_i(s_i', \sigma^*_{-i}) + \sum_{\substack{s_i' \in S_i \\ s_i' \neq s_i}} u_i(s_i', \sigma^*_{-i}) \cdot \boxed{\frac{\sigma^*_i(s_i)\,\sigma_i(s_i')}{1 - \sigma_i(s_i)}}$$

$$= \text{—"—} + \frac{\sigma^*_i(s_i)}{1 - \sigma_i(s_i)} \boxed{\sum_{\substack{s_i' \in S_i \\ s_i' \neq s_i}} \sigma_i(s_i')\, u_i(s_i', \sigma^*_{-i})}$$

$$= \text{—"—} + \frac{\sigma^*_i(s_i)}{1 - \sigma_i(s_i)} \cdot \Big[\underbrace{u_i(\sigma_i, \sigma^*_{-i}) - u_i(s_i, \sigma^*_{-i})}_{} \Big]$$

$$= u_i(\sigma_i^*, \sigma_{-i}^*) - \underline{\sigma_i^*(s_i')\, u_i(s_i, \sigma_{-i}^*)} + \frac{\sigma_i^*(s_i)}{1-\underline{\underline{\sigma_i(s_i)}}} \boxed{\left[u_i(\sigma_i, \sigma_{-i}^*) - u_i(s_i, \sigma_{-i}^*)\right]}$$

$$= u_i(\sigma_i^*, \sigma_{-i}^*) + \frac{\cancel{u_i(s_i, \sigma_{-i}^*)}\, \sigma_i^*(s_i) - \sigma_i^*(s_i)\, \cancel{u_i(s_i, \sigma_{-i}^*)} + \sigma_i(s_i) . \sigma_i^*(s_i') . u_i(s_i, \sigma_{-i}^*) + \sigma_i^*(s_i) .\, u_i(\sigma_i, \sigma_{-i}^*)}{1 - \sigma_i(s_i)}$$

$$= u_i(\sigma_i^*, \sigma_{-i}^*) + \frac{\sigma_i(s_i)\, \sigma_i^*(s_i')\, u_i(s_i, \sigma_{-i}^*) + \sigma_i^*(s_i)\, u_i(\sigma_i, \sigma_{-i}^*)}{1 - \sigma_i(s_i)}$$

$$> u_i(\sigma_i^*, \sigma_{-i}^*)$$

This contradicts our assumption that $(\sigma_i^*)_{i\in N}$ is an MSNE. $\blacksquare$