Lecture 4: Further solution concepts

Motivations

- We reason about multiplayer games using solution concepts, principles according to which we identify interesting subsets of the outcomes of a game

- Nash equilibrium is the most important solution concept

- There are also a large number of others:
  - Maximin and minmax strategies
  - Minimax regret
  - Removal of dominated strategies
  - Rationalizability
  - Correlated equilibrium
  - Trembling-hand perfect equilibrium

Maxmin and minmax strategies

Definition (Maxmin)

The maxmin strategy for player i is $s_i^* = \arg\max_{s_i} \min_{s_{-i}} u_i(s_i, s_{-i})$ and the maxmin value for player i is $\max_{s_i} \min_{s_{-i}} u_i(s_i, s_{-i})$

- The ***maxmin strategy*** of player i in an n-players game is a strategy that maximizes $i's$ worst – case payoff, in the situation where all the others players happen to play the strategies which cause the greatest harm to i
- The ***maxmin strategy*** is a sensible choice for a conservative agent who wants to maximize his expected utility without having to make any assumptions about the other agents
- The ***maxmin value*** (or security level) of the game for player i is that minimum amount of payoff guaranteed by a maxmin strategy
- It is strategy that **defends against** other agents (defensive strategy)
- Player i set the mixed strategy ⇒ player $-i$ observe this strategy (not an action) and choose their own strategies to minimize i's expected payoff

(temporal interpretation)

Maxmin and minmax strategies

Definition (Minmax, two-player**)**

In an two-player game, the *minmax strategy* for player i against player $-i$ is $s_i^* = \arg\min_{s_i} \max_{s_{-i}} u_{-i}(s_i, s_{-i})$ and the minmax value is $\min_{s_i} \max_{s_{-i}} u_{-i}(s_i, s_{-i})$

- The *minnmax strategy* of player i in an two-players game is a strategy that keeps the maximum payoff of $-i$ at a minimum
- The *minmax value* of player $-i$ is that minimum
- It is strategy that attack against other agents (offensive strategy)

Maxmin and minmax strategies

In agent $i's$ perspective

$$\boxed{\max_{s_i}} \min_{s_{-i}} u_i(s_i, s_{-i})$$

- Agent always maximizes its payoff
- Defensive strategy (if max is first)

$$\boxed{\min_{s_i}} \max_{s_{-i}} u_{-i}(s_i, s_{-i})$$

- Agent always maximizes its payoff
- offensive strategy (if min is first)

Maxmin and minmax strategies

Definition (Minmax, n-player)

In an n-player game, the minmax strategy for player i against player $j \neq i$ is i-th component of the mixed-strategy profile s_{-j} in the expression $\arg\min_{s_{-j}} \max_{s_j} u_j(s_j, s_{-j})$. As before, the minmax value for player j is $\min_{s_{-j}} \max_{s_j} u_j(s_j, s_{-j})$

- Here, we assume that all the players other than j choose to "gang up" on j
  - They are able to coordinate appropriately when there is more than one strategy profile that would yield the same minimal payoff for j

Maxmin and minmax strategies

- An agent's maximum strategy nor his minmax strategy depends on the strategies that the other agents actually choose
- the maximin and minmax strategies give rise to solution concept
- Call $s = (s_1, \ldots, s_n)$ a maxmin strategy profile of a given game if s_1 is a maxmin strategy for player 1, s_2 is a maxmin strategy for player 2 and so on.
  - ➢ Similar to minmax strategy profile
- In two-player games, a player's minmax value is always equal to his maxmin value

$$\min_{s_{-i}} \max_{s_i} u_i(s_i, s_{-i}) = \max_{s_i} \min_{s_{-i}} u_i(s_i, s_{-i})$$

- For games with more than two players, a weaker condition holds:

$$\max_{s_i} \min_{s_{-i}} u_i(s_i, s_{-i}) \leq \min_{s_{-i}} \max_{s_i} u_i(s_i, s_{-i})$$

- See that player $-i$ chooses first, allowing player i to best respond to it.

Maxmin and minmax strategies

- An agent’s maximum strategy nor his minmax strategy depends on the strategies that the other agents actually choose
- the maximin and minmax strategies give rise to solution concept
- Call $s = (s_1, \dots, s_n)$ a maxmin strategy profile of a given game if s_1 is a maxmin strategy for player 1, s_2 is a maxmin strategy for player 2 and so on.
  - Similar to minmax strategy profile

- In two-player games, a player’s minmax value is always equal to his maxmin value

$$\min_{s_{-i}} \max_{s_i} u_i(s_i, s_{-i}) = \max_{s_i} \min_{s_{-i}} u_i(s_i, s_{-i})$$

- For games with more than two players, a weaker condition holds:

$$\max_{s_i} \min_{s_{-i}} u_i(s_i, s_{-i}) \leq \min_{s_{-i}} \max_{s_i} u_i(s_i, s_{-i})$$

$$\max_{s_i} \min_{s_{-i}} u_i(s_i, s_{-i}) \leq \min_{s_{-i}} \max_{s_i} u_i(s_i, s_{-i})$$

$$\min_{s_{-i}} u_i(s_i^{\max}, s_{-i}) \leq u_i\left(s_i^{\max}, s_{-i}^{\min}\right) \leq \max_{s_i} u_i(s_i, s_{-i}^{\min})$$

  - See that player $-i$ chooses first, allowing player i to best respond to it.

Minimax theorem (von Neumann, 1928)

Theorem (Minmax theorem by von Neumann, 1928)

In any finite, two-player, zero-sum game, in any Nash equilibrium each player receives a payoff that is equal to both his maxmin value and his minmax value.

- Minmax theorem states that in a two-player zero-sum game:

maximin value = minmax value = Nash equilibrium value

- Any *maximin* strategy profile is a Nash equilibrium. Furthermore, these are all the Nash equilibria
  - Consequently, all Nash equilibria have the same payoff vector

Minimax theorem (von Neumann, 1928)

Proof:

- Let's assume (s_i', s_{-i}') be an arbitrary Nash equilibrium and denote v_i to be the i's equilibrium payoff
- Denote i's maxmin value as $\overline{v_i} = \max_{s_i} \min_{s_{-i}} u_i(s_i, s_{-i})$
- Denote i's minmax value as $\underline{v} = \min_{s_{-i}} \max_{s_i} u_i(s_i, s_{-i})$
- First, we show that $\overline{v_i} = v_i$

✓ $\overline{v_i} \leq v_i$

$$\overline{v_i} = \max_{s_i} \min_{s_{-i}} u_i(s_i, s_{-i}) \leq \max_{s_i} u_i(s_i, s_{-i}') < v_i$$

✓ $\overline{v_i} \geq v_i$

$$v_{-i} = \max_{s_{-i}} u_{-i}(s_i', s_{-i})$$

$$-v_{-i} = \min_{s_{-i}} -u_{-i}(s_i', s_{-i}) \qquad \because \max f(x) = -\min\{-f(x)\}$$

since the game is zero sum, $-v_{-i} = v_i$ and $u_i = -u_{-i}$, thus

$$v_i = \min_{s_{-i}} u_i(s_i', s_{-i})$$

$$\overline{v_i} = \max_{s_i} \min_{s_{-i}} u_i(s_i, s_{-i}) \geq \min_{s_{-i}} u_i(s_i', s_{-i}) = v_i$$

➢ As a result, $v_i = \overline{v_i}$

Minimax theorem example

		Player 2	
		q Heads	$1-q$ Tails
Player 1	p Heads	$1,-1$	$-1,1$
	$1-p$ Tails	$-1,1$	$1,-1$

- Player 1's maxmin value :

$$\bar{u}_1 = \max_{s_1} \min_{s_2} u_1(s_1, s_2)$$
$$= \max_{p} \min_{q} \{pq - p(1-p) - (1-p)q + (1-p)(1-q)\}$$

- Player 1's minmax value :

$$\underline{u}_1 = \min_{s_2} \max_{s_1} u_1(s_1, s_2)$$
$$= \min_{q} \max_{p} \{pq - p(1-p) - (1-p)q + (1-p)(1-q)\}$$

Minimax theorem example

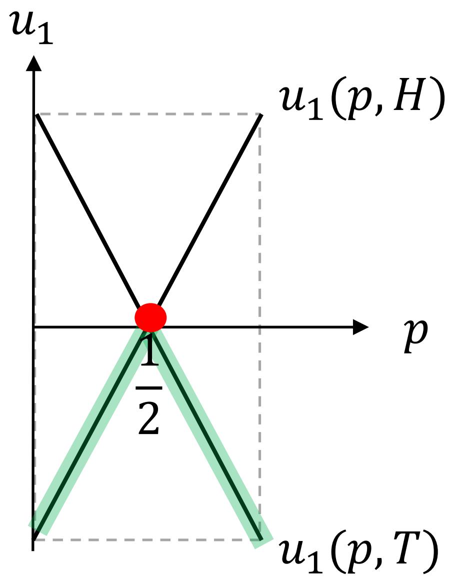
- Player 1's maxmin value : $\bar{u}_1 = \max_{s_1} \min_{s_2} u_1(s_1, s_2)$
  $= \max_{p} \min_{q} \{pq - p(1-p) - (1-p)q + (1-p)(1-q)\}$

- For any p set by player 1, player 2 tries to chooses q **deterministically** to minimize u_1
- $\min_{q}\{pq - p(1-p) - (1-p)q + (1-p)(1-q)\} \Rightarrow$

$$\min_{q\in\{0,1\}} \{pq - p(1-p) - (1-p)q + (1-p)(1-q)\} = \min\{2p-1, 1-2p\}$$

  - When player 2 plays Heads $(q = 1)$: $u_1(p, H) = 2p - 1$
  - When player 2 plays Tails $(q = 0)$: $u_1(p, T) = 1 - 2p$

- Thus, $\bar{u}_1 = \max_{p} \min\{2p-1, 1-2p\} = 0$



- Player 1's maxmin strategy:

$$\bar{s}_1 = \operatorname{argmax}_{s_1} \min_{s_2} u_1(s_1, s_2) = \left(\frac{1}{2}, \frac{1}{2}\right)$$

- Player 1's maxmin value:

$$\bar{u}_1 = \max_{s_1} \min_{s_2} u_1(s_1, s_2) = 0$$

Minimax theorem example

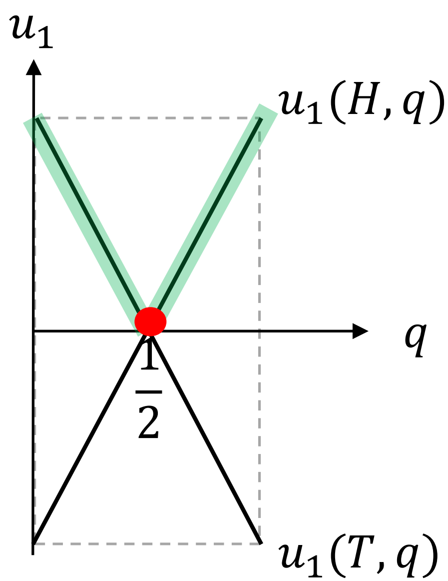
- Player 1's minmax value : $\underline{u}_1 = \min_{s_2} \max_{s_1} u_1(s_1, s_2)$
$$= \min_{q} \max_{p} \{pq - p(1-p) - (1-p)q + (1-p)(1-q)\}$$

- For any q set by player 2, player 1 tries to chooses p **deterministically** to maximize u_1
- $\max_{p} \{pq - p(1-p) - (1-p)q + (1-p)(1-q)\} \Rightarrow$

$$\max_{pq \in \{0,1\}} \{pq - p(1-p) - (1-p)q + (1-p)(1-q)\} = \min\{2q-1, 1-2q\}$$

- When player 1 plays Heads $(p = 1)$: $u_1(H, q) = 2q - 1$
- When player 1 plays Tails $(p = 0)$: $u_1(T, q) = 1 - 2q$

- Thus, $\underline{u}_1 = \min_{q} \max\{2q-1, 1-2q\} = 0$



- Player 2's minmax strategy:

$$\underline{s}_2 = \operatorname{argmin}_{s_2} \max_{s_1} u_1(s_1, s_2) = \left(\frac{1}{2}, \frac{1}{2}\right)$$

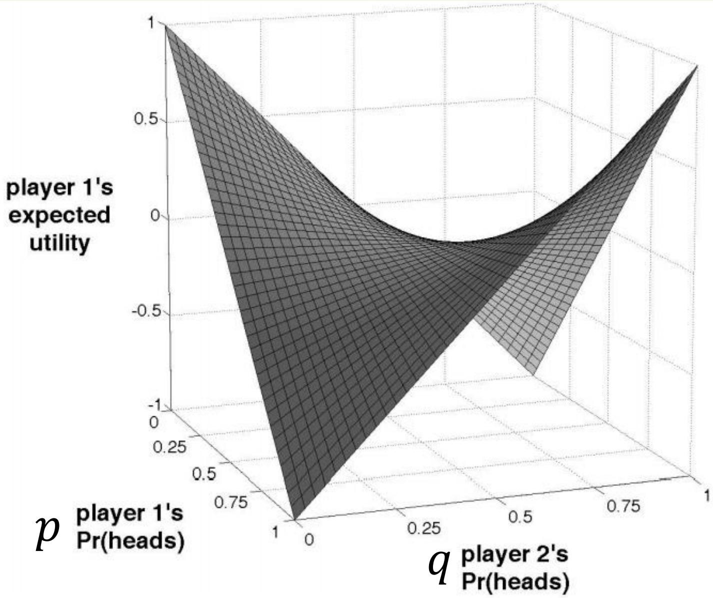
- Player 1's minmax value:

$$\underline{u}_1 = \min_{s_2} \max_{s_1} u_1(s_1, s_2) = 0$$

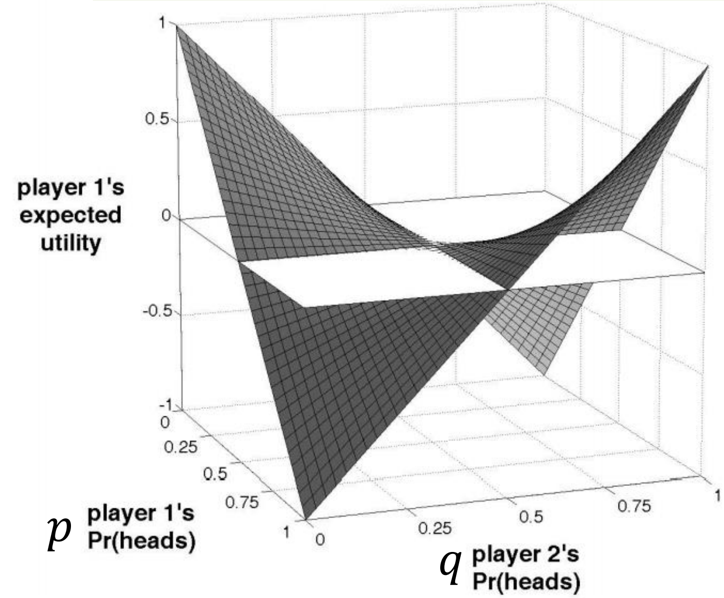
Minimax theorem graphical representation

		Player 2: q Heads	Player 2: $1-q$ Tails
Player 1	p Heads	$1, -1$	$-1, 1$
	$1-p$ Tails	$-1, 1$	$1, -1$

$u_1(p,q) = pq - p(1-p) - (1-p)q + (1-p)(1-q)$



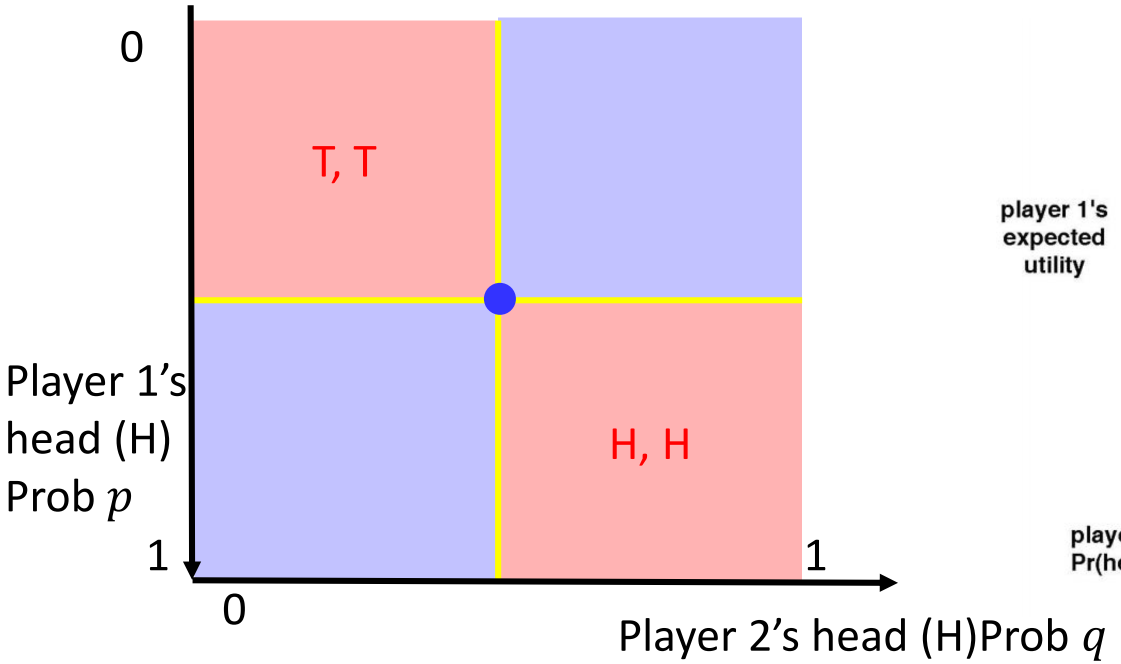
$u_1(p,q) = pq - p(1-p) - (1-p)q + (1-p)(1-q)$

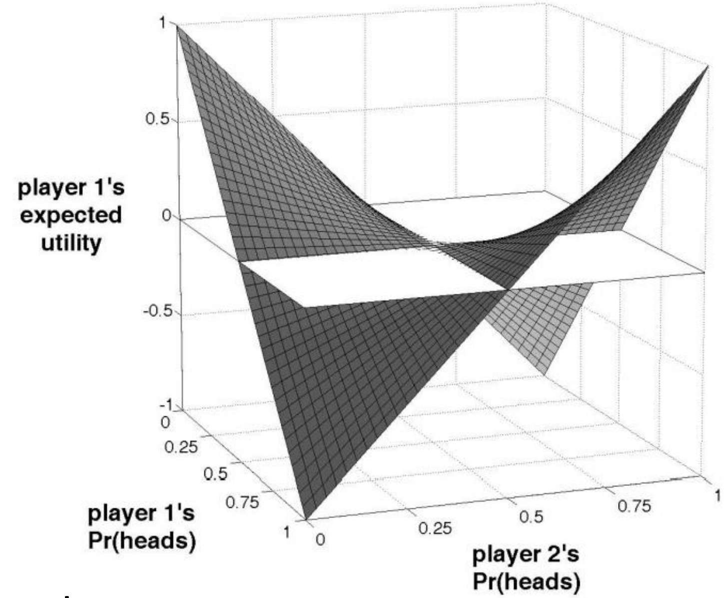


- Nash equilibria in zero-sum games can be viewed graphically as a “saddle” in a high-dimensional space
- At a saddle point, any deviation of the agent lowers his utility and increases the utility of the other agent.

Minimax theorem graphical representation

		Player 2	
		q Heads	$1-q$ Tails
Player 1	p Heads	$1, -1$	$-1, 1$
	$1-p$ Tails	$-1, 1$	$1, -1$





- Nash equilibria in zero-sum games can be viewed graphically as a "saddle" in a high-dimensional space
- At a saddle point, any deviation of the agent lowers his utility and increases the utility of the other agent.

Minimax regret

		Player 2	
		L	R
Player 1	T	$100, a$	$1-\epsilon, b$
	B	$2, c$	$1, d$

- We argued agents might play maxmin strategies to achieve good payoffs in the worst case

- Player 1's maximin strategy is to play B (to receive 1 rather than $1-\epsilon$):
  - If player 1 play T, then player 2 will chose R to minimize player 1's payoff: $u_1 = 1-\epsilon$
  - If player 1 play B, then player 2 will chose R to minimize player 1's payoff: $u_1 = 1$
  - Thus, maximin strategy for player 1 is to play B, giving him a payoff of 1

Minimax regret

		Player 2	
		L	R
Player 1	T	$100, a$	$1-\epsilon, b$
	B	$2, c$	$1, d$

- However, the other agent is not believed to be malicious, but is instead unpredictable
- In this case, agents might care about minimizing their worst-case losses, rather than maximizing their worst case payoffs

- Player 1's Minmax regret strategy is to play T:
  - If player 2 were to play R, then it would not matter very much how player 1 plays
    - ✓ The most he could lose by playing the wrong way would be ϵ
  - If player 2 were to play L, then player 1's action would be very significant
    - ✓ If player makes wrong choice, his utility would be decreased by 98
  - Thus, given that player can maximize your regret, player 1 might choose to play T in order to minimize his worst-case loss

Minimax regret

Definition (Regret)

An agent i's regret for playing an action a_i if the other agents adopt action profile a_{-i} is defined as

$$\left[\max_{a_i' \in A_i} u_i(a_i', a_{-i})\right] - u_i(a_i, a_{-i})$$

- In words, this is the amount that i loses by playing a_i, rather than playing his best response to a_{-i}. Of course, i does not know that actions the other players will take.
- But, we can consider those actions that would give him the highest regret for playing a_i

Minimax regret

Definition (Regret)

An agent i's regret for playing an action a_i if the other agents adopt action profile a_{-i} is defined as

$$\left[\max_{a_i' \in A_i} u_i(a_i', a_{-i})\right] - u_i(a_i, a_{-i})$$

- In words, this is the amount that i loses by playing a_i, rather than playing his best response to a_{-i}. Of course, i does not know that actions the other players will take.
- But, we can consider those actions that would give him the highest regret for playing a_i

Definition (Max Regret)

An agent i's regret for playing an action a_i if the other agents adopt action profile a_{-i} is defined as

$$\max_{a_{-i} \in A_{-i}} \left(\left[\max_{a_i' \in A_i} u_i(a_i', a_{-i})\right] - u_i(a_i, a_{-i}) \right)$$

- This is the amount that i loses by playing a_i rather than playing his best response to a_{-i}, if the other agents chose the a_{-i} that makes this loss as large as possible

Minimax regret

Definition (Regret)

An agent i's regret for playing an action a_i if the other agents adopt action profile a_{-i} is defined as

$$\left[\max_{a'_i \in A_i} u_i(a'_i, a_{-i})\right] - u_i(a_i, a_{-i})$$

- In words, this is the amount that i loses by playing a_i, rather than playing his best response to a_{-i}. Of course, i does not know that actions the other players will take.
- But, we can consider those actions that would give him the highest regret for playing a_i

Definition (Max Regret)

An agent i's regret for playing an action a_i if the other agents adopt action profile a_{-i} is defined as

$$\max_{a_{-i} \in A_{-i}} \left(\left[\max_{a'_i \in A_i} u_i(a'_i, a_{-i})\right] - u_i(a_i, a_{-i}) \right)$$

- This is the amount that i loses by playing a_i rather than playing his best response to a_{-i}, if the other agents chose the a_{-i} that makes this loss as large as possible

Definition (Minmax Regret)

Minmax regret actions for agent i are defined as

$$\arg\min_{a_i \in A_i} \left[\max_{a_{-i} \in A_{-i}} \left(\left[\max_{a'_i \in A_i} u_i(a'_i, a_{-i})\right] - u_i(a_i, a_{-i}) \right) \right]$$

- Minmax regret actions are one that yields the smallest maximum regret

Removal of dominated strategies

Definition (Domination)

Let s_i and s_i' be two strategies of player i, and S_{-i} the set of all strategy profiles of the remaining players. Then,

1. s_i strictly dominates s_i' if for all $s_{-i} \in S_{-i}$, it is the case that $u_i(s_i, s_{-i}) > u_i(s_i', s_{-i})$
2. s_i weekly dominates s_i' if for all $s_{-i} \in S_{-i}$, it is the case that $u_i(s_i, s_{-i}) \geq u_i(s_i', s_{-i})$, and for at least one $s_{-i} \in S_{-i}$, it is the case that $u_i(s_i, s_{-i}) > u_i(s_i', s_{-i})$
3. s_i very weekly dominates s_i' if for all $s_{-i} \in S_{-i}$, it is the case that $u_i(s_i, s_{-i}) \geq u_i(s_i', s_{-i})$

- Domination is comparison between two strategies s_i and s_i' given others $s_{-i} \in S_{-i}$

Definition (Pareto domination)

Strategy profile s Pareto dominates strategy profile s' if for all $i \in N$, $u_i(s) \geq u_i(s')$, and there exists some $j \in N$ for which $u_j(s) > u_j(s')$.

Removal of dominated strategies

Definition (Dominant strategy)

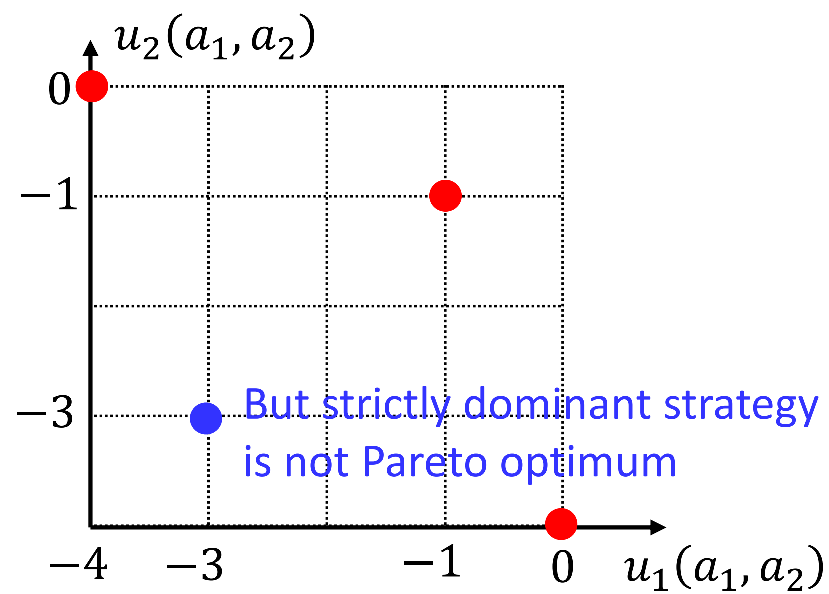
A strategy is strictly (resp., weekly; very weakly) dominant for an agent if it strictly (weakly; very weakly) dominates any other strategy for that agent.

- A strategy profile $(s_1, \ldots, s_n)$ in which every s_i is dominant for player i (whether strictly, weakly, or very weakly) is a Nash equilibrium.
- A strategy profile consisting of dominant strategies for every player must be a Nash equilibrium
  - ➢ An equilibrium in strictly dominant strategies must be unique.

	C	D
C	-1, -1	-4, 0
D	0, -4	$\boxed{\text{-3, -3}}$

Strictly dominant strategy profile

Prisoner's Dilemma game



Removal of dominated strategies

Definition (Dominated strategy)

A strategy s_i is strictly (weakly; very weakly) dominated for an agent i if some other strategy s_i' strictly (weakly; very weakly) dominates s_i

- Note that it is easy to check if a strategy is dominated because we need to find any strategy that dominate it

Removal of dominated strategies

	L	C	R
U	$3, 1$	$0, 1$	$0, 0$
M	$1, 1$	$1, 1$	$5, 0$
D	$0, 1$	$4, 1$	$0, 0$

- R is dominated by L

Removal of dominated strategies

	L	C	R
U	$3, 1$	$0, 1$	$\cancel{0, 0}$
M	$1, 1$	$1, 1$	$\cancel{5, 0}$
D	$0, 1$	$4, 1$	$\cancel{0, 0}$

- R is dominated by L

Removal of dominated strategies

	L	C
U	$3, 1$	$0, 1$
M	$1, 1$	$1, 1$
D	$0, 1$	$4, 1$

Removal of dominated strategies

	L	C
U	$3, 1$	$0, 1$
M	$\cancel{1, 1}$	$\cancel{1, 1}$
D	$0, 1$	$4, 1$

- M is dominated by the mixed strategy that selects U and D with equal probability

Removal of dominated strategies

	L	C
U	$3, 1$	$0, 1$
D	$0, 1$	$4, 1$

- No other strategies are dominated.

Removal of dominated strategies

	L	C	R
U	$4, 3$	$5, 1$	$6, 2$
M	$2, 1$	$8, 4$	$3, 6$
D	$3, 0$	$9, 6$	$2, 8$

- Find an equilibrium by yourself

Removal of dominated strategies

- This process preserves Nash equilibria.
  - strict dominance: all equilibria preserved.
  - weak or very weak dominance: at least one equilibrium preserved.

- Thus, it can be used as a preprocessing step before computing an equilibrium
  - Some games are solvable using this technique.
  - Example: Prisoner's Dilemma!

- What about the order of removal when there are multiple dominated strategies?
  - strict dominance: doesn't matter.
  - weak or very weak dominance: can affect which equilibria are preserved.

	L	C
U	$1, 1$	$2, 1$
D	$1, 2$	$3, 1$

- Remove the action of the column player first
- Remove the action of the row player first

What is the result?

Removal of dominated strategies

Cournot duopoly

- Two identical firms, players 1 and 2, produce some good
- Firm i produce quantity q_i
- Cost for production is $c_i(q_i) = 10q_i$
- Price is given by $d = 100 - (q_1 + q_2)$
- The profit of company 1 is $u_1(q_1, q_2) = (100 - q_1 - q_2)q_1 - 10q_1 = 90q_1 - q_1^2 - q_1q_2$

What should firm 1 do in order to maximize their profit?

Removal of dominated strategies

Cournot duopoly

- Two identical firms, players 1 and 2, produce some good
- Firm i produce quantity q_i
- Cost for production is $c_i(q_i) = 10q_i$
- Price is given by $d = 100 - (q_1 + q_2)$
- The profit of company 1 is $u_1(q_1, q_2) = (100 - q_1 - q_2)q_1 - 10q_1 = 90q_1 - q_1^2 - q_1q_2$

What should firm 1 do in order to maximize their profit?

- As the payoff is concave in q_1, the maximum is obtained by imposing the derivative of the payoff with respect q_1 for any given value of q_2

$$q_1 = \frac{90 - q_2}{2}$$

  - That is, for any given q_2 chosen by company 2, company maximize its payoff

- The same applied to company 2

$$q_2 = \frac{90 - q_1}{2}$$

Removal of dominated strategies

Cournot duopoly

- The profit of company 1 is $u_1(q_1, q_2) = (100 - q_1 - q_2)q_1 - 10q_1 = 90q_1 - q_1^2 - q_1q_2$
- As the payoff is concave in q_1, the maximum is obtained by imposing the derivative of the payoff with respect q_1 for any given value of q_2

$$q_1 = \frac{90 - q_2}{2}$$

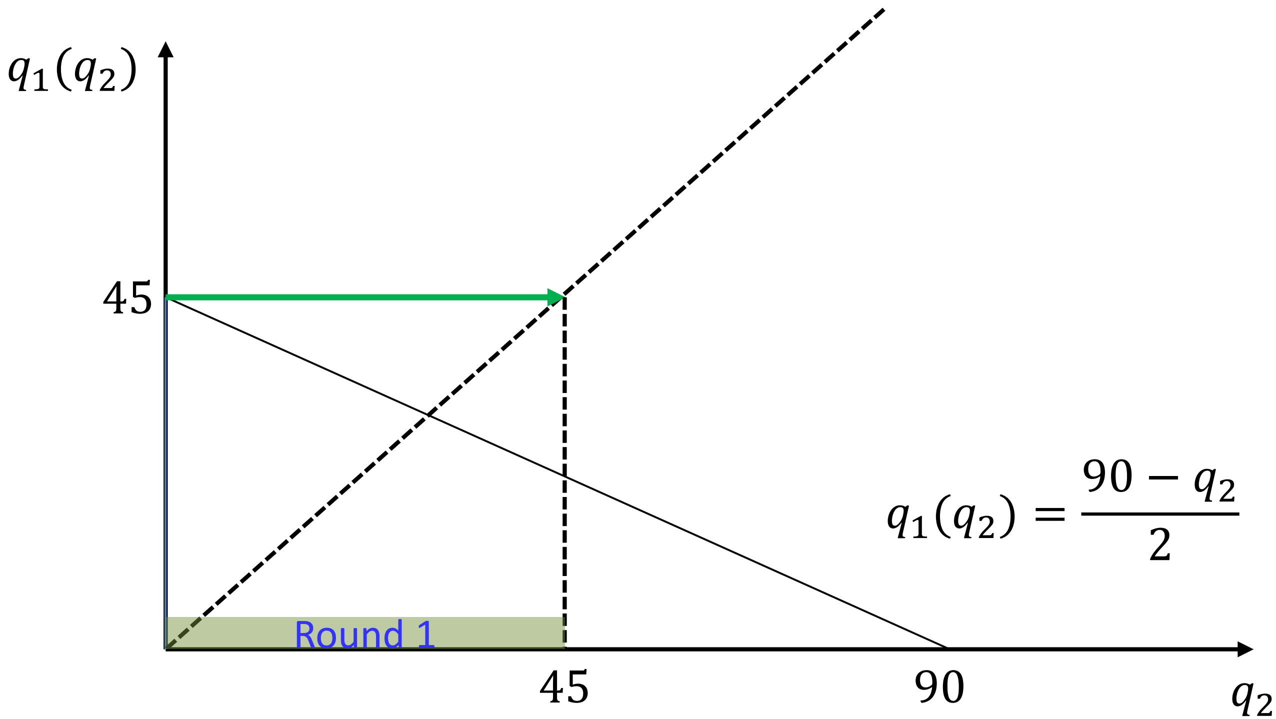
- Company 1 will never choose to produce more than $q_1 > 45$ because any quantity $q_1 > 45$ is strictly dominated by $q_1 = 45$ as follows:
  - $u_1(q_1 = 45, q_2) = (100 - 45 - q_2)45 - 450 = 2025 - 45q_2$
  - $u_1(q_1, q_2) = (100 - q_1 - q_2)q_1 - 10q_1 = 90q_1 - q_1^2 - q_1q_2$
  - $u_1(45, q_2) - u_1(q_1, q_2) = 2025 - q_1(90 - q_1) - q_2(45 - q_1) > 0$ for any $q_1 > 45$ regardless of q_2
- Due to symmetry, any $q_2 > 45$ is strictly dominated by $q_2 = 45$

- **The first round of iterated elimination**:
  - A rational produces no more than 45 units, implying that the effective strategy space that survives one round of elimination is $q_i \in [0,45]$ for $i \in \{1,2\}$

Removal of dominated strategies

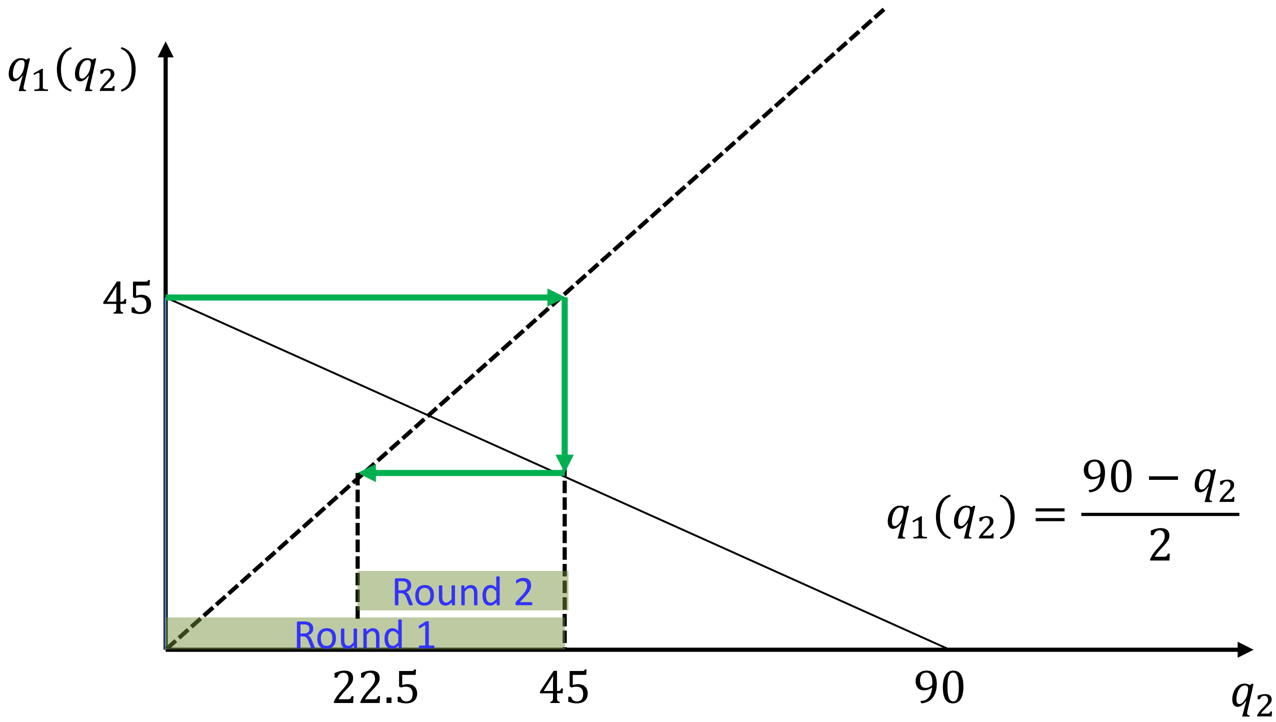
Cournot duopoly

- **The first round of iterated elimination**:
  - $q_2 > 45$ is strictly dominated by $q_2 \leq 45$



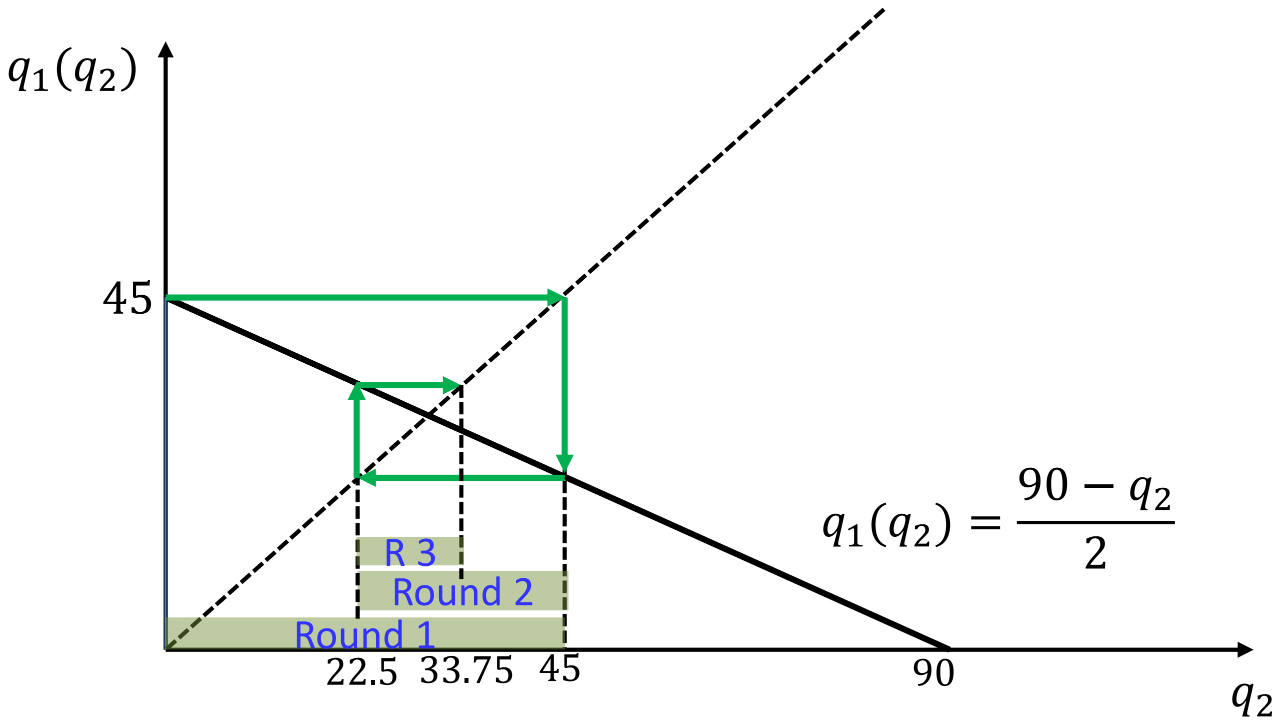
Cournot duopoly

- **The second round of iterated elimination**:
  - Because $q_2 \leq 45$, the equation $q_1 = \frac{90-q_2}{2}$ implies that company 1 will chose $q_1 \geq 22.5$
  - Symmetric argument applies to $q_2 \geq 22.5$
  - Therefore the second round of elimination implies that the surviving strategy sets are $q_i \in [22.5, 45]$ for $i \in \{1,2\}$



Cournot duopoly

- **The third round of iterated elimination**:
  - Because $q_2 \geq 22.5$, the equation $q_1 = \frac{90-q_2}{2}$ implies that company 1 will chose $q_1 \leq 33.75$
  - Symmetric argument applies to $q_2 \leq 33.75$
  - Therefore the second round of elimination implies that the surviving strategy sets are $q_i \in [22.5, 33.75]$ for $i \in \{1,2\}$



Rationalizability

- Rather than ask what is irrational, ask what is a best response to some beliefs about the opponent
  - assumes opponent is rational
  - assumes opponent knows that you and the others are rational
  - ...
- Examples
  - is heads rational in matching pennies?
  - is cooperate rational in prisoner's dilemma?
- Will there always exist a rationalizable strategy?
  - Yes, equilibrium strategies are always rationalizable.
- Furthermore, in two-player games, rationalizable $\Leftrightarrow$ survives iterated removal of strictly dominated strategies.

Correlated equilibrium

If there is intelligent life on other planets, in a majority of them, they would have discovered correlated equilibrium before Nash equilibrium.

– Roger Myerson

Correlated equilibrium

- Consider again Battle of the Sexes.
  - The values of each player under mixed Nash equilibrium is 2/3

		Player 2	
		p TF	$1-p$ LA
Player 1	q TF	2, 1	0, 0
	$1-q$ LA	0, 0	1, 2

$$u_1(\mathrm{TF}) = u_1(\mathrm{LA})$$
$$2 \times p + 0 \times (1-p) = 0 \times p + 1 \times (1-p)$$
$$p = \frac{1}{3}$$

$$u_2(\mathrm{TF}) = u_2(\mathrm{LAL})$$
$$1 \times q + 0 \times (1-q) = 0 \times q + 2 \times (1-q)$$
$$q = \frac{2}{3}$$

- The mixed Nash equilibrium is $s^* = (s_1^*, s_2^*) = \left\{\left(\frac{2}{3}, \frac{1}{3}\right), \left(\frac{1}{3}, \frac{2}{3}\right)\right\}$
- The expected payoff under s^* are $u_1^* = \frac{2}{3} = u_2^*$

Can we do better?

Correlated equilibrium

- Consider again Battle of the Sexes.
  - The values of each player under mixed Nash equilibrium is 2/3

			Player 2	
			p TF	$1-p$ LA
Player 1	q	TF	2, 1	0, 0
	$1-q$	LA	0, 0	1, 2

- We could use the same idea to achieve the fair outcome in battle of the sexes.
  - Intuitively, the best outcome seems a 50-50 split between (TF, TF) and (LA, LA).

$$u_1^{CE} = \frac{1}{2}(2+1) = \frac{3}{2} > u_1^{NE} = \frac{2}{3}$$

$$u_2^{CE} = \frac{1}{2}(2+1) = \frac{3}{2} > u_2^{NE} = \frac{2}{3}$$

We show that no player has an incentive to deviate from the "recommendation" of the coin.

Correlated equilibrium

- Another classic example: traffic game

	Go	Wait
Go	-100, -100	10, 0
Wait	0, 10	-10, -10

Traffic game

- What is the natural solution here?
  - A traffic light: a fair randomizing device that tells one of the agents to go and the other to wait.
- Benefits:
  - the negative payoff outcomes are completely avoided
  - fairness is achieved
  - the sum of social welfare exceeds that of mixed Nash equilibrium

Correlated equilibrium

- More complex example

	L	R
U	5, 1	0, 0
D	4, 4	1, 5

- There are two pure strategy Nash equilibria: (U, L) and (D, R).
- This implies that there is a unique mixed strategy equilibrium with expected payoff (5/2,5/2).
- Suppose the players find a mediator who chooses $x \in \{1, 2, 3\}$ with equal probability 1/3. She then sends the following messages:
  - If $x = 1$, player 1 plays U, player 2 plays L.
  - If $x = 2$, player 1 plays D, player 2 plays L.
  - If $x = 3$, player 1 plays D, player 2 plays R.

Actions are correlated

- Our first example presumed that everyone perfectly observes the random event; not required.
- More generally, some random variable with a commonly known distribution, and a private signal to each player about the outcome.
  - signal doesn't determine the outcome or others' signals; however, correlated:
    - ✓ Actions for agents are jointly determined by a drawn random variable

Correlated equilibrium

	L	R
U	5, 1	0, 0
D	4, 4	1, 5

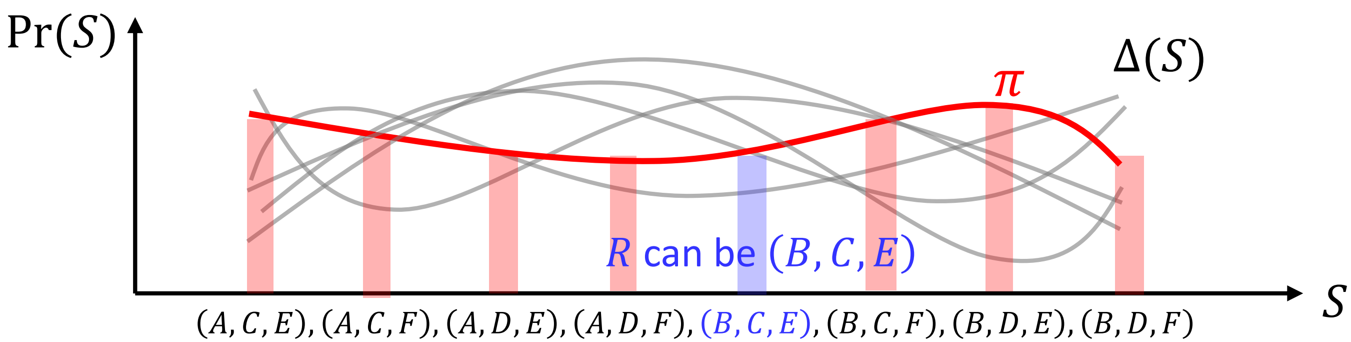
- If $x = 1$, player 1 plays U, player 2 plays L.
- If $x = 2$, player 1 plays D, player 2 plays L.
- If $x = 3$, player 1 plays D, player 2 plays R.

- We show that no player has an incentive to deviate from the "recommendation" of the mediator:
  - If player 1 gets the recommendation U, he believes player 2 will play L, so his best response is to play U.
  - If player 1 gets the recommendation D, he believes player 2 will play L, R with equal probability, so playing D is a best response.
  - If player 2 gets the recommendation L, he believes player 1 will play U, D with equal probability, so playing L is a best response.
  - If player 2 gets the recommendation R, he believes player 1 will play D, so his best response is to play R.
- Thus the players will follow the mediator's recommendations.
- With the mediator, the expected payoffs are (10/3, 10/3), strictly higher than what the players could get by randomizing between Nash equilibria.

Correlated equilibrium

- The preceding examples lead us to the notions of correlated strategies and "correlated equilibrium".
- Let $\Delta(S)$ denote the set of probability measures over the set S. Let R be a random variable taking values in $S = \Pi_{i=1}^{n} S_i$ distributed according to $\pi \in \Delta(S)$.
  - An instantiation of R is a pure strategy profile and the i th component of the instantiation will be called the recommendation to player i.
  - Given such a recommendation, player i can use conditional probability to form a posteriori beliefs about the recommendations given to the other players.

- $S_1 = \{A, B\}, S_2 = \{C, D\}, S_3 = \{E, F\}$
- $S = \{(A, C, E), (A, C, F), (A, D, E), (A, D, F), (B, C, E), (B, C, F), (B, D, E), (B, D, F)\}$
- $\Delta(S)$ is a set of probability mass function (PMF) over S
- $\pi \in \Delta(S)$ is a PMF over S
- $R \sim \pi(\mathrm{S})$ is a random variable distributed according to π and represents the joint action



Correlated equilibrium

Definition (Correlated equilibrium)

A correlated equilibrium of finite game is a joint probability distribution $\pi \in \Delta(S)$ such that if R is random variable distributed according to π then

$$\sum_{s_{-i} \in S_{-i}} \mathrm{Prob}(R = s | R_i = s_i) u_i(s_i, s_{-i}) \geq \sum_{s_{-i} \in S_{-i}} \mathrm{Prob}(R = s | R_i = s_i) u_i(t_i, s_{-i})$$

For all players i, all $s_i \in S_i$ such that $\mathrm{Prob}(R_i = s_i) > 0$, and all $t_i \in S_i$

- A distribution π is defined to be a correlated equilibrium if no player can ever expect to unilaterally gain by deviating from his recommendation, assuming the other players play according to their recommendations.
  - s_i is a recommendation by R drawn from $\pi \in \Delta(S)$
  - t_i is a deviation from this recommendation

Correlated equilibrium

Proposition

A joint probability distribution $\pi \in \Delta(S)$ is a correlated equilibrium of a finite game if and only if

$$\sum_{s_{-i} \in S_{-i}} \pi(s) u_i(s_i, s_{-i}) \geq \sum_{s_{-i} \in S_{-i}} \pi(s) u_i(t_i, s_{-i})$$

For all players i, all $s_i \in S_i$, $t_i \in S_i$ such that $s_i \neq t_i$

Proof:

$$\mathrm{Prob}(R = s | R_i = s_i) = \frac{\pi(s_i, s_{-i})}{\pi(s_i)} = \frac{\pi(s)}{\sum_{t_{-i} \in S_{-i}} \pi(s_i, t_{-i})}$$

$$\sum_{s_{-i} \in S_{-i}} \mathrm{Prob}(R = s | R_i = s_i) u_i(s_i, s_{-i}) \geq \sum_{s_{-i} \in S_{-i}} \mathrm{Prob}(R = s | R_i = s_i) u_i(t_i, s_{-i})$$

$$\sum_{s_{-i} \in S_{-i}} \frac{\pi(s)}{\sum_{t_{-i} \in S_{-i}} \pi(s_i, t_{-i})} u_i(s_i, s_{-i}) \geq \sum_{s_{-i} \in S_{-i}} \frac{\pi(s)}{\sum_{t_{-i} \in S_{-i}} \pi(s_i, t_{-i})} u_i(t_i, s_{-i})$$

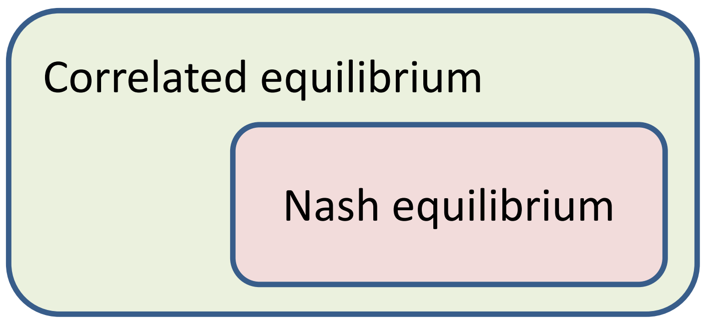
- The denominator does not depend on the variable of summation so it can be factored out of the sum and cancelled

Correlated equilibrium

Theorem (Correlated equilibrium)

For every Nash equilibrium σ^* there exists a corresponding correlated equilibrium σ

- Not every correlated equilibrium is equivalent to a Nash equilibrium, (e.g., Battle of Sex game)
  - Correlated equilibrium is a strictly weaker notion than Nash



- Any convex combination of the payoffs achievable under correlated equilibria is itself realizable under a correlated equilibrium

Trembling-hand perfect equilibrium

Definition (Trembling-hand perfect equilibrium)

A mixed strategy profile is a (trembling-hand) perfect equilibrium of a normal-form game G if there exists a sequence $s^0, s^1, \ldots$ of fully mixed strategy profiles such that $\lim_{n\to\infty} s^n = s$, and such that for each s^k in the sequence and each player i, the strategy s_i is a best response to the strategies s^k_{-i}.

- Rationalizability is a weaker concept than Nash equilibrium, but perfection is a stronger one
- Perfect equilibria are relevant to one aspect of multiagent learning
- It requires that the equilibrium be robust against slight errors
- One's action out to be the best response not only against the opponent's equilibrium strategies, but also against small perturbation of those.

ϵ − Nash equilibrium

- Players might not care about changing their strategies to a best response when the amount of utility that they could gain by doing so is very small.

Definition (ϵ − Nash equilibrium)

Fix $\epsilon > 0$. A strategy profile $s^* = (s_1^*, \ldots, s_n^*)$ is an ϵ-Nash equilibrium if, for all agents i and for all strategies $s_i \neq s_i^*$, $u_i(s_i^*, s_{-i}^*) \geq u_i(s_i, s_{-i}^*) - \epsilon$

	L	R
U	$1, 1$	$0, 0$
D	$1 + \frac{\epsilon}{2}, 1$	$500, 500$

A game with interesting ϵ − Nash equilibrium