Introduction to Game Theory 3

Solution Concepts: Regret Minimisation, Minimax and Maximin, and Nash Equilibrium

Eric Pauwels (CWI & VU)

Version: November 14, 2023

Game Theory Part 1: Outline

1. Game Theory: Science of Strategic Thinking
2. Examples of interesting games
3. Formalising games
4. Solution concept 1: Weak optimality
5. **Solution concept 2: Strategies with (weak) guarantees**
    - Regret minimisation
    - Maximin
    - Minimax
6. **Solution concept 3: Nash equilibrium**

Table of Contents

Strategies with (weak) guarantees

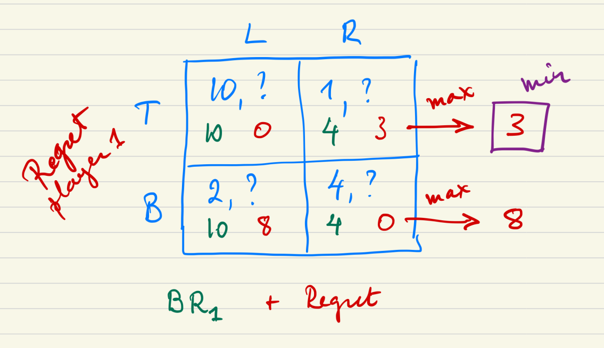
- Focus on outcomes that can be guaranteed by your own actions
- **Reasons:**
  - Opponent-agnostic: Opponent's utilities might not be known
  - Following a different/hidden agenda:
    e.g. threatening or punishing opponent;
  - ...others??
- **Strategies**
  1. Regret minimisation
  2. Safety (maximin) and Punishment (minimax) strategies

Regret minimisation

- What is my maximal possible regret if I take this action?
- Regret: diff. btw pay-offs best response and current action:

$$R_1(s_1, s_2) = u_1(BR_1(s_2)) - u_1(s_1, s_2)$$

- Take action to minimise max regret



Regret minimisation

- **Regret** (for agent i) is the difference between the actual and maximal pay-off for a given action profile (s_i, s_j)

$$R_i(s_i, s_j) = u_i(BR_i(s_j)) - u_i(s_i, s_j) = \max_{s_i'} u_i(s_i', s_j) - u_i(s_i, s_j)$$

- For each action s_i , there is maximum regret depending on s_j:

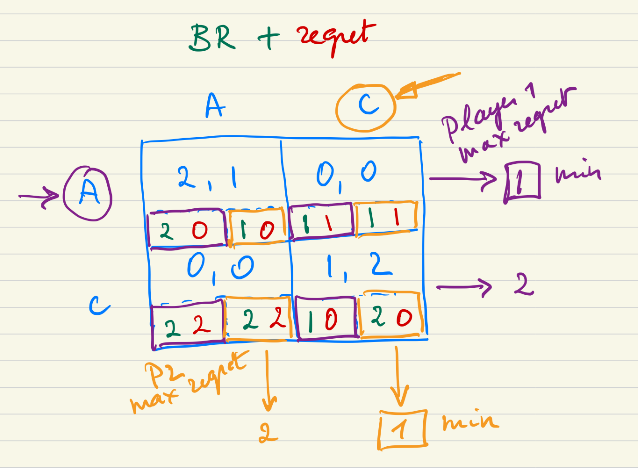
$$R_i^{max}(s_i) = \max_{s_j} R_i(s_i, s_j)$$

- **Regret minimisation** (minimax regret): agent i picks action s_i that mimimises max regret:

$$s_i^{rm} := \arg\min_{s_i} R_i^{max}(s_i) = \arg\min_{s_i} \max_{s_j} R_i(s_i, s_j).$$

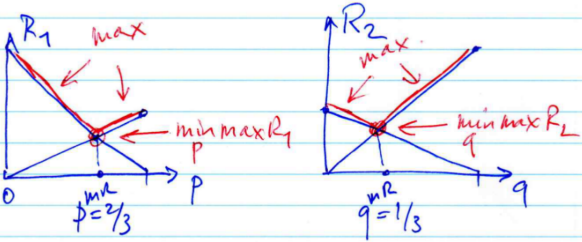
- Actual solution concepts: allows an agent to choose a strategy with specific guarantees/properties;

Example of regret minimisation for BoS



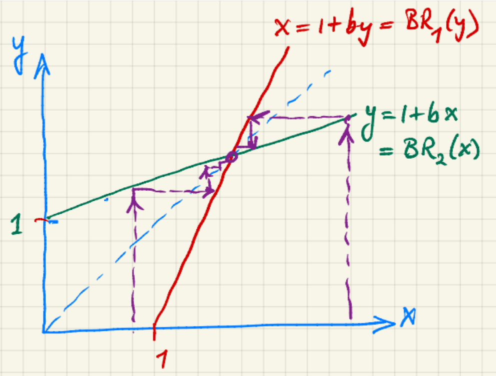
BoS: Regret minimisation for mixed strategies

	q	$1-q$	
p	$0,0$	$1,1$	$1-q$
$1-p$	$2,2$	$0,0$	$2q$
	$2(1-p)$	p	



Regret minimisation for partnership game

- Recall:



Regret minimisation for partnership game

Pay-off for player 1:

$$u(x,y) = 2(x+y+bxy) - x^2$$

$0 \leq b < 1$ $\qquad$ $0 \leq x,y \leq 4$

Recall: $\bar{x} = BR_1(y) = 1 + by$

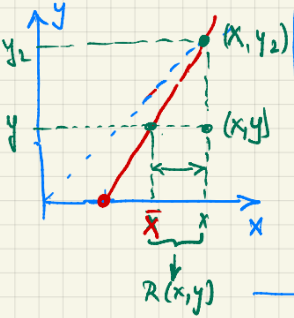
Regret for player 1 if he plays x, opp. y

$$\begin{aligned} R(x,y) &= u(\bar{x},y) - u(x,y) \\ &= 2(\bar{x}-x + b(\bar{x}-x)y) - (\bar{x}^2 - x^2) \\ &= (\bar{x}-x)\left[2(1+by) - (\bar{x}+x)\right] \end{aligned}$$

Regret minimisation for partnership game

$$R(x,y) = (\bar{x}-x)\left[2(1+by) - (\bar{x}+x)\right]$$

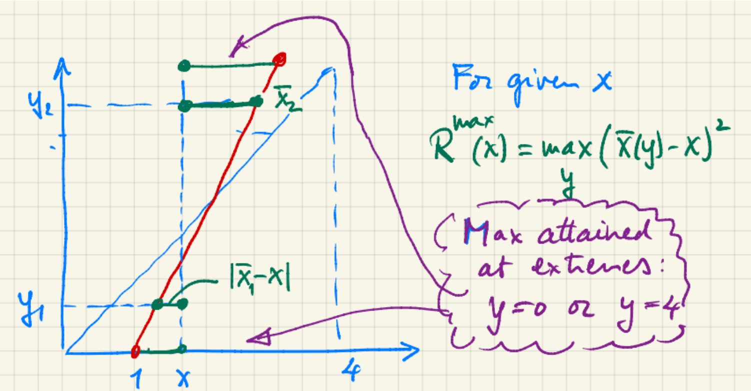
$$(\bar{x} = 1+by)$$



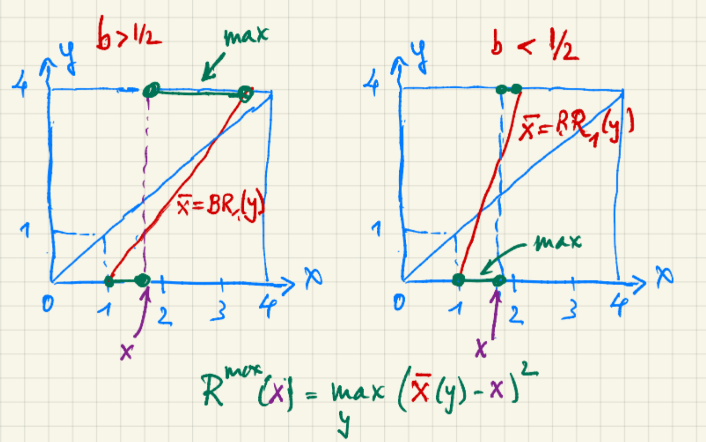
$$= (1+by-x)\left[2(1+by) - (1+by) - x\right]$$

$$= (\underbrace{1+by}_{\bar{x}}-x)^2 = (\bar{x}-x)^2$$

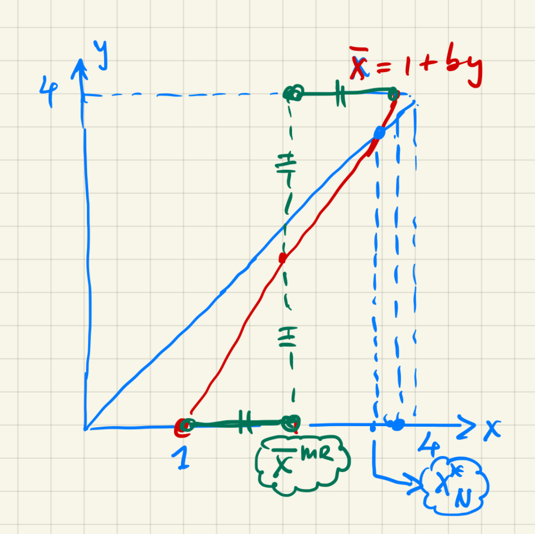
Regret minimisation for partnership game



Regret minimisation for partnership game

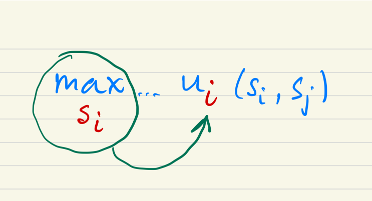


Regret minimisation for partnership game



Maximin Value and Safety Strategy

- Safety strategy: What is the best outcome i can secure for myself, no matter what the opponent does?
- No need to know the corresponding opponent's pay-offs!
- We consider player i's point of view: (i.e. maximise over s_i)



Maximin strategy (safety strategy): Algorithm

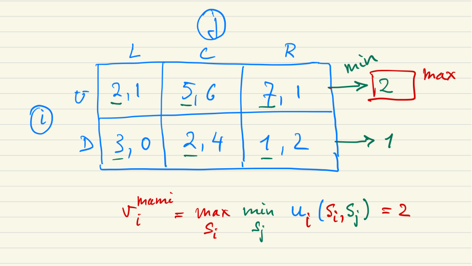
1. Ag_i computes for all his actions the worst possible outcome:

$$s_i \longrightarrow \min_{s_j} u_i(s_i, s_j)$$

2. Next, ag_i chooses action s_i to maximise his minimal pay-off:

$$v_i^{mami} := \max_{s_i} \min_{s_j} u_i(s_i, s_j)$$

Agent tries to maximise pay-off of worst possible outcome



Maximin Value and Strategy

Context: 2-player, general sum game

- Maximin **value** (or security level for WORST CASE) is the guaranteed minimal pay-off for agent i playing strategies in S_i:

$$v_i^{mami} := \max_{s_i \in S_i} \min_{s_j \in S_j} u_i(s_i, s_j)$$

- Maximin **strategy** (or safety strategy) for agent i maximizes his worst case pay-off (and therefore yields at least v^{mami} as utility):

$$s_i^{mami} = \arg \underbrace{\max_{s_i} \min_{s_j} u_i(s_i, s_j)}_{\text{security level}}$$

Maximin (safety) value and strategy

- **Why play maximin strategy?**
  - Pay-off guarantee! Highest pay-off agent i can guarantee for himself irrespective of the actions taken by other agent(s).
  - Worst case analysis: assume that opponent is malicious!

Example: both agents play maximin-strategy

Both players play maximin.

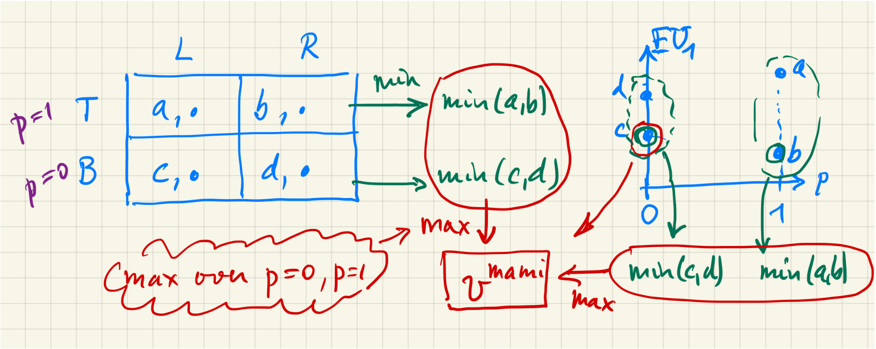
1 (rows), 2 (columns)

	L	CL	CR	R	min
U	3, 1	4, 6	8, 1	5, 2	$\boxed{3}$ max.
M	4, 2	7, 4	1, 5	4, 5	1
D	6, 2	1, 3	7, 0	0, 4	0
min	1	$\boxed{3}$ max	0	2	

$$\left.\begin{aligned} v_1^{\text{mami}} &= 3, \quad s_1^{\text{mami}} = U \\ v_2^{\text{mami}} &= 3, \quad s_2^{\text{mami}} = CL \end{aligned}\right\} \rightarrow \boxed{\begin{aligned} u_1(U,CL) &= 4 \\ u_2(U,CL) &= 6 \end{aligned}}$$

From Pure to Mixed safety strategies

- safety value for pure safety strategy

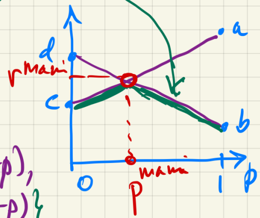


From Pure to Mixed safety strategies

		L	R
$p=1$	T	$a, \cdot$	$b, \cdot$
$p=0$	B	$c, \cdot$	$d, \cdot$
	$pT+(1-p)B$	$ap+c(1-p)$	$bp+d(1-p)$

$$\xrightarrow{\min} \min\{ap+c(1-p),\ bp+d(1-p)\}$$

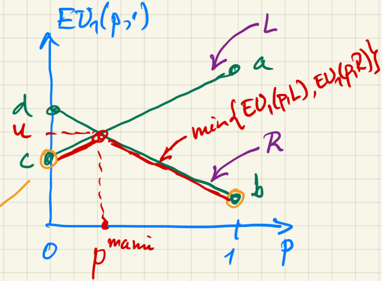
$$\boxed{v^{mami} = \max_p \min\{ap+c(1-p),\ bp+d(1-p)\}}$$



Mixed safety strategies

- If $EU_1(p,\cdot)$ graphs intersect, it might be better to look at mixed strategies

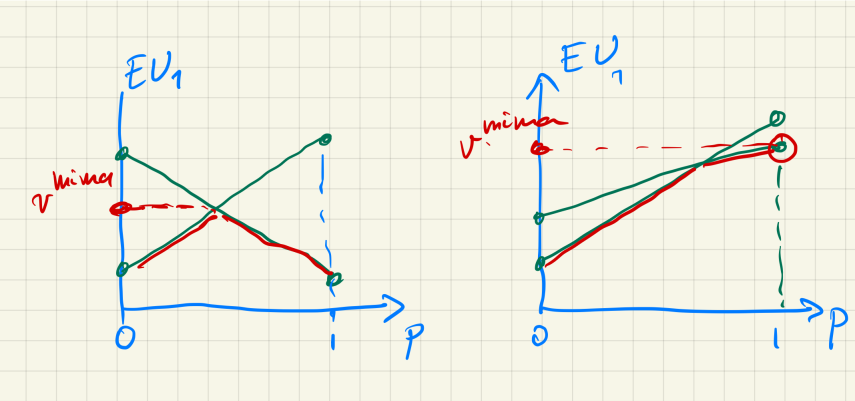
		L	R
T	p	$a, \cdot$	$b, \cdot$
B	$1-p$	$c, \cdot$	$d, \cdot$



pure: $v^{maxmin} = c$

mixed: $v^{maxmin} = u = \max_p \min\{EU_1(p,L), EU_1(p,R)\}$

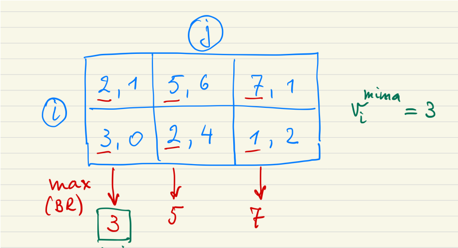
Sometimes pure safety strategy is best



Punishment (or minimax) Strategy (of player j) yields Minimax Value (for player i)

1. Player i computes best response utility for each of j's strategies s_j: $BR_i(s_j) = \max_{s_i} u_i(s_i, s_j)$;
2. Then player j (who wants to punish i) picks action to minimise player i best pay-off. This punishment strategy of **player j** results in the minimax value for **player i**:

$$v_i^{mima} := \min_{s_j} \max_{s_i} u_i(s_i, s_j) = \min_{s_j} u_i(BR_i(s_j))$$



Algo: Minimax Value and Punishment Strategy

- **Minimax value** for agent i:
  - Given the strategy s_j of his opponent, agent i will play its **best response**, resulting in a pay-off:

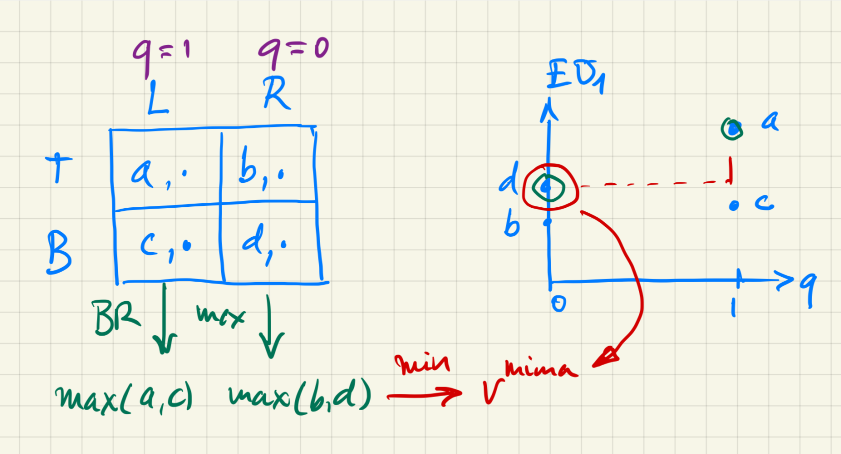
$$\max_{s_i} u_i(s_i, s_j)$$

  - The opponent is aware of this and wants to *"punish"* i by *minimizing* this pay-off, yielding the minimax value for player i:

$$v_i^{mima} := \min_{s_j} \max_{s_i} u_i(s_i, s_j)$$

- The corresponding minimising strategy s_j^{mima} is called the minimax strategy for player j.
- If j plays his minimax strategy s_j^{mima}, then i cannot do better than v_i^{mima} (even if i plays best response $BR_i(s_j^{mima})$.

Punishment strategy: Pure

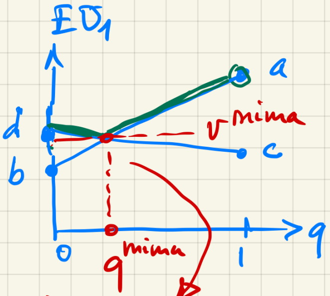


Punishment strategy: Mixed

	L	R	$qL+(1-q)R$
T	$a,\cdot$	$b,\cdot$	$aq+b(1-q)$
B	$c,\cdot$	$d,\cdot$	$cq+d(1-q)$

$$\downarrow$$

$$\max\{aq+b(1-q),\ cq+d(1-q)\} \xrightarrow{\min_q} v_M^{mima}$$



Comparing minimax and maximin value for player i

- $\forall s_i, s_j$:

$$\underbrace{\min_{s_j} u_i(s_i, s_j)}_{\phi(s_i)} \leq u_i(s_i, s_j) \leq \underbrace{\max_{s_i} u_i(s_i, s_j)}_{\psi(s_j)}$$

- Since the above inequality holds for all s_i and s_j, it follows:

$$\max_{s_i} \phi(s_i) \leq \min_{s_j} \psi(s_j)$$

- Hence:

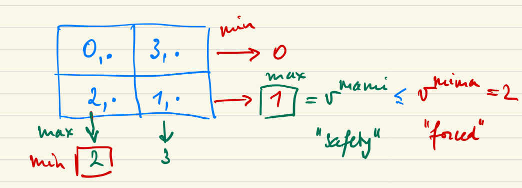
$$v_i^{mami} = \max_{s_i} \min_{s_j} u_i(s_i, s_j) \quad \leq \quad \min_{s_j} \max_{s_i} u_i(s_i, s_j) = v_i^{mima}$$

Minimax and Maximin Value for Player i

- In general:

$$v_i^{mami} = \max_{s_i} \min_{s_j} u_i(s_i, s_j) \quad \leq \quad \min_{s_j} \max_{s_i} u_i(s_i, s_j) = v_i^{mima}$$

- Your guaranteed pay-off is a is lower bound for the worst your opponent can force onto you!
- The worst your opponent can force onto you is an upper bound on your guaranteed pay-off.



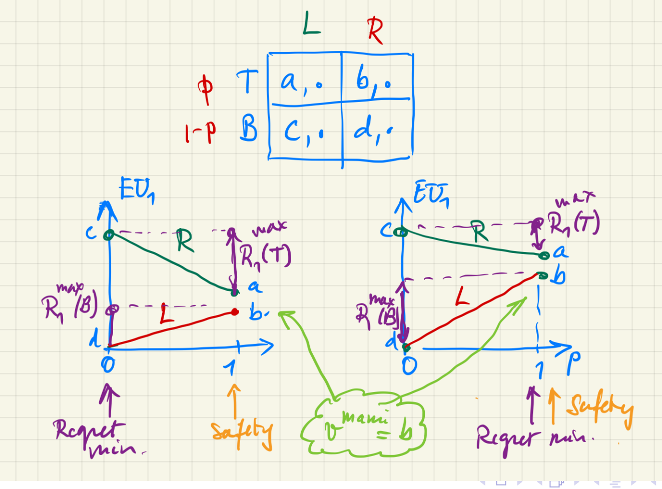
Recap Safety and Punishment Strategy

- **Player i's maximin strategy is safety strategy**
  - player i concerned about his own safety
    - Malicious or adversarial opponent
    - Multi-agent setting
  - strategy yields highest guaranteed outcome for player i
  - Viable solution algorithm.

- **Player i's minimax strategy is punishment strategy**
  - i's strategy is directed against player j
  - player i tries to minimize best (i.e. maximum) pay-off for j
  - Useful as threat (e.g. in repeated games);
  - i's maximin **strategy** gives rise to j's maximin **value**

Which is better? Regret Minimisation or Safety?

- Neither, not comparable!



Which is better? Regret Minimisation or Safety?

- Comparison for $(0 \leq T \leq 2, -1 \leq S \leq 1)$ parametrisation for social dilemmas.

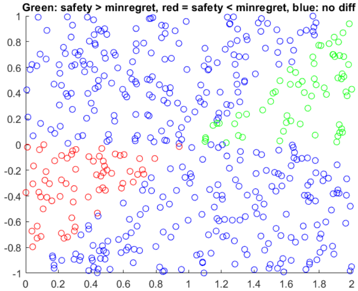


Figure: Utilities (assuming both agents play same strategy): : green: safety $>$ min regret, red $=$ min regret $>$ safety, blue: no difference

Table of Contents

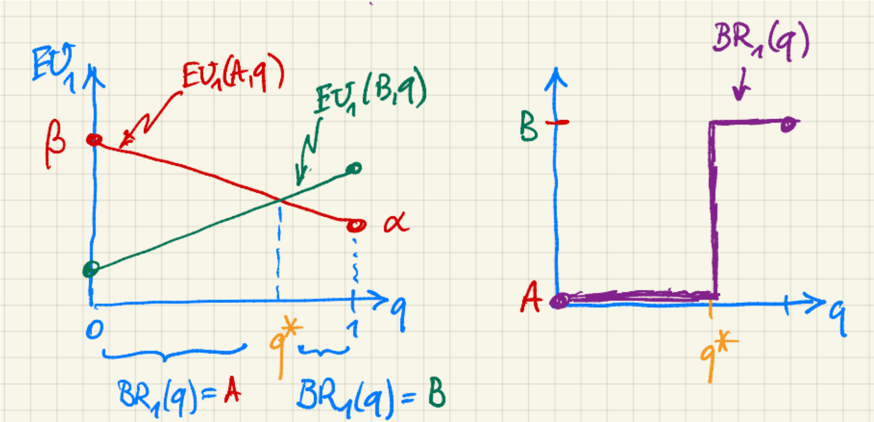
Nash equilibrium: Context

- Von Neumann's minimax theorem established game theory as a discipline;
- Nash equilibrium: Extension of von Neumann's minimax theorem
    1. From two person to n person game;
    2. from zero sum to general utilities
- For more info on the minimax theorem: See addendum;

Nash Equilibrium: Intro

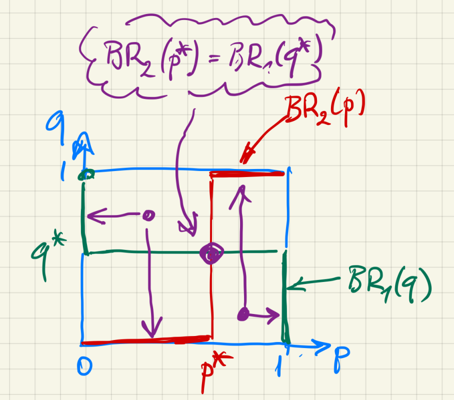
- Optimal choice no longer applies, as your outcome depends on opponent's actions.
  - Opposing "forces" at play;
- From optimum to equilibrium:
  - Opposing "forces" balance out;
- **Recall:** Best response (BR) dynamics . . .
  - sometimes converged to equilibrium position;
  - sometimes gave rise to cycles . . . Nash to the rescue!

Recap: Best response to mixed strategy



Nash equilibrium: intro

- Consider general mixed strategies in 2p-game, with two actions each;
- Imagine one player is allowed to change his mind!



Nash equilibrium (1)

- A **Nash equilibrium** (NE, 1950) is a solution concept based on *conditions* instead of an *algorithm*.
- **Mutual best response:** NE is joint strategy profile s^* such that **for each agent** i the strategy s_i^* is a **best response** to s_{-i}^*;
- **Formally:** A strategy profile $s^* = (s_1^*, s_2^*, \ldots, s_n^*)$ is a **strict NE** if:

$$\forall \text{agents}\, i,\ \forall s_i' \neq s_i^* : u(s_i^*, s_{-i}^*) > u(s_i', s_{-i}^*).$$

- **Strict** ($>$) versus **weak** ($\geq$) NE
- **No Regret/Self-enforcing:** a (strict) NE is a stable strategy profile for which no agent has an incentive to **unilaterally deviate**;

Nash equilibrium: Computation of NE (pure strategy)

Find mutual best responses:

- **Battle of the Sexes** (two pure strategy NEs)

	action	comedy
action	$\underline{2}, \underline{1}$	$0, 0$
comedy	$0, 0$	$\underline{1}, \underline{2}$

- **Prisoner's dilemma** (single NE, not Pareto-optimal!)

	hush	confess
hush	$-1, -1$	$-12, \underline{0}$
confess	$\underline{0}, -12$	$\underline{-8}, \underline{-8}$

Nash's Theorem

Existence of Nash Equilibrium (Nash, 1950)

A **finite strategic game** (i.e. finite number of players and actions) always has **at least one Nash equilibrium** (allowing mixed strategies).

- A **pure** Nash equilibrium can be strict or weak;
- A **mixed** Nash equilibrium is necessarily weak;

Nash equilibrium: Nash's Theorem

- A **finite strategic game** is a game with a finite number of agents and a finite number of actions;
- A game may have **zero**, **one**, or **more pure-strategy NE**.
- If there's a **single** NE: natural solution concept, but might be (Pareto) sub-optimal!
- If there are **multiple NEs**: there might be no compelling reasons to pick a particular one; but . . .
  - Utility dominant NE, Schelling's focal points
- Since humans are not always rational, Nash equilibria might not agree with experiments or observations..

Note on **multiple** Nash Equilibria

- A unique best alternative is exceptional.
- In general, non-cooperative game theory is plagued by a multiplicity of equilibria.
- Hence, prescriptions of how to act without any coordination or cooperation are in general impossible.
- Non-cooperative game theory tells us what to exclude from choice.
- Since in non-coop GT, there are no binding contracts, a player's announcements and promises in a pre-play phase are credible only, if they are totally in line with his best interests.

Computation of NE

- **Pure NE for each agent** i the strategy s_i^* is a **best response** to s_{-i}^* (mutual best response);
  - Matrix games (discrete state/action) space;
  - Continuous action space
- **Mixed NE**: make opponent indifferent (matrix games only);

Nash equilibrium: Computation of **mixed** NE

Matching pennies (zero-sum game)

- **No pure strategy NE . . .**

	heads	tails
heads	$\underline{1}, -1$	$-1, \underline{1}$
tails	$-1, \underline{1}$	$\underline{1}, -1$

- . . . hence, at least one **mixed strategy NE!**
- **Intuitively** this is obvious; play each action with prob $= 1/2$:

$$s_1 = s_2 = \{(H, 1/2), (T, 1/2)\}$$

- **Expected utility (pay-off)**:

$$u_1(s_1, s_2) = \frac{1}{4}u_1(H, H) + \frac{1}{4}u_1(T, H) + \frac{1}{4}u_1(H, T) + \frac{1}{4}u_1(T, T) = 0.$$

Matching pennies: Computation of mixed NE

$$
\begin{array}{cc|c|c|}
 & & q & 1-q \\
 & & H & T \\
\hline
p & H & 1,-1 & -1,1 \\
\hline
1-p & T & -1,1 & 1,-1 \\
\hline
\end{array}
$$

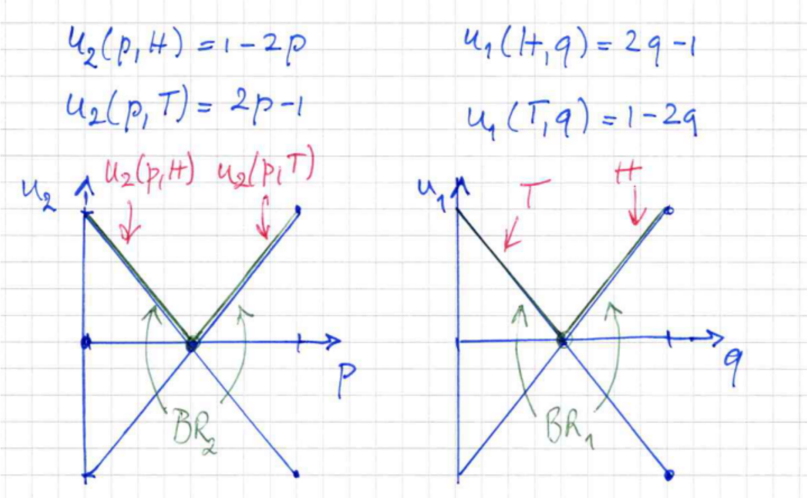
$$u_2(p,H) = (-1)p + 1.(1-p) = 1-2p.$$

$$u_2(p,T) = p + (-1)(1-p) = 2p-1$$

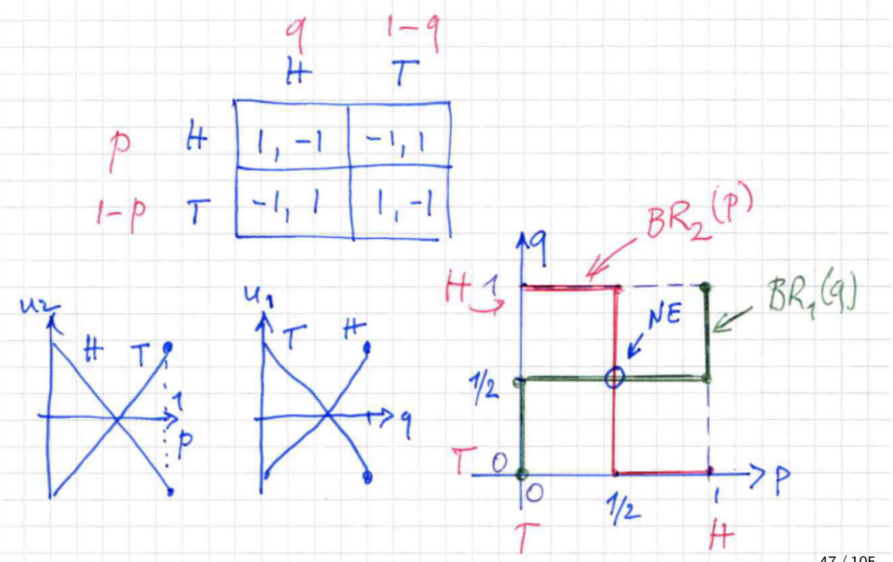
$$u_1(H,q) = q-(1-q) = 2q-1$$

$$u_1(T,q) = -1.q + (1-q) = 1-2q$$

Matching pennies: Computation of best response



Matching pennies: Best response graph



Matching pennies: Mixed Nash Equilibrium

Nash equilibrium characteristics:

- At the intersection point ($p = 1/2, q = 1/2$), players are simultaneously playing best response to each other;
- No player can do strictly better by unilaterally deviating:
  - If player 2 keeps playing $q = 1/2$ then player 1's utility $u_1(p, q = 1/2)$ can be computed as follows:

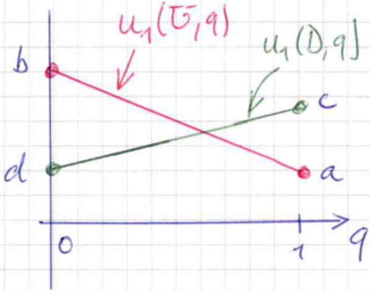
$$\begin{aligned} u_1(p, 1/2) &= (1 \cdot p + (-1) \cdot (1-p) + (-1) \cdot p + 1 \cdot (1-p)) \cdot \frac{1}{2} \\ &= 0 \qquad \textit{(independent of } p\textit{)} \end{aligned}$$

    Player 1:no incentive to change his strategy (i.e. change p)

  - Same consideration for player 2.

Computation of mixed Nash equilibrium

		q	$1-q$
		L	R
p	U	a, A	b, B
$1-p$	D	c, C	d, D



(Expected) utility for player 1:

$$u_1(p,q) = apq + bp(1-q) + c(1-p)q + d(1-p)(1-q)$$

Linear in p (for fixed q) and vice versa.

Computation of mixed Nash equilibrium

$$u_1(p,q) = apq + bp(1-q) + c(1-p)q + d(1-p)(1-q)$$

Player 1: no incentive to change

$$0 = \frac{\partial u_1}{\partial p} = \underbrace{\left(aq + b(1-q)\right)}_{=\,u_1(U,q)} - \underbrace{\left(cq + d(1-q)\right)}_{=\,u_1(D,q)}$$

$$u_1(U,q) - u_1(D,q) = 0$$

Player 2 needs to pick q such that player 1 is <u>indifferent</u> btw U and D.

Computing mixed Nash equilibrium

Player 1: T or B, player 2: L or R

- **Strategies:** Player 1: T or B, Player 2: L or R
- Player 1: fix $p = p^*$, such that player 2 is indifferent (between L and R):

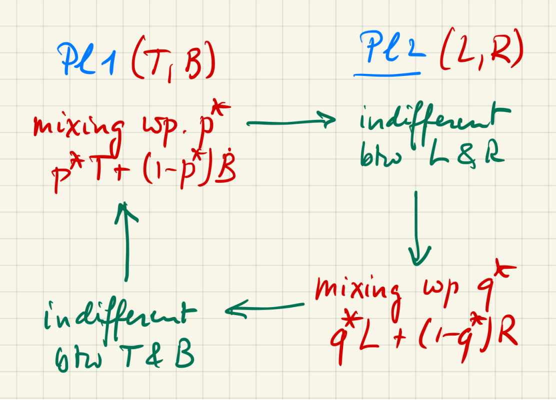
$$EU_2(p^*, L) = EU_2(p^*, R)$$

- Player 2: fix $q = q^*$, such that player 1 is indifferent (between T and B):

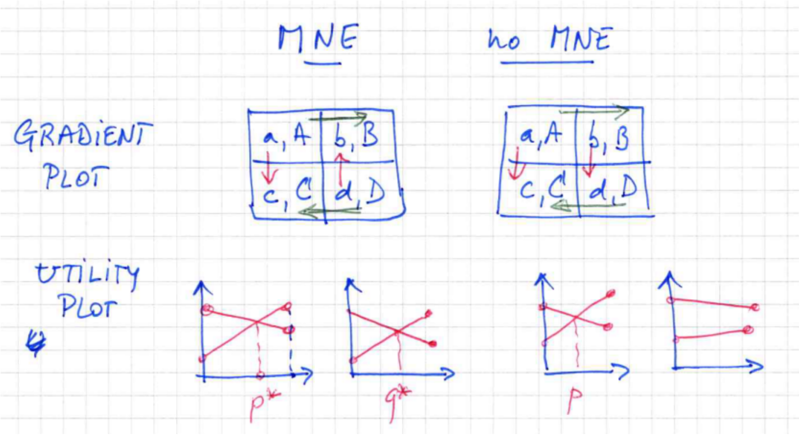
$$EU_1(T, q^*) = EU_1(B.q^*)$$

Mixed Nash equilibrium: "mutual stranglehold"

- P1 mixes (with prob p^*) to make P2 indifferent;
- Since P2 is indifferent, he's OK to mix too!
- This allows P2 to mix with prob q^* to make P1 indifferent;
- As a consequence P1 is willing to mix with prob p^*.



Some useful graphical representations



Nash equilibrium: Computation of mixed equilibrium

- **Battle of the Sexes** (two pure strategy NEs)
- Additional NE by **mixing pure strategies**?

$$s_1 = \{(A, p), (C, 1-p)\} \quad \text{and} \quad s_2 = \{(A, q), (C, 1-q)\}.$$

	Action (q)	Comedy ($1-q$)
Action (p)	2, 1	0, 0
Comedy ($1-p$)	0, 0	1, 2

- Determining the mixture parameters p and q
  - Ag 1 chooses p such that Ag 2 is **indifferent** btw actions A and C; if not Ag 2 would focus on the most lucrative option.

Condition for Ag 1: $u_2(s_1(p), A) = u_2(s_1(p), C)$

Nash equilibrium: Computation of mixed equilibrium

- **Battle of the Sexes** (mixed strategy NEs)

	Action (q)	Comedy ($1-q$)
Action (p)	2, 1	0, 0
Comedy ($1-p$)	0, 0	1, 2

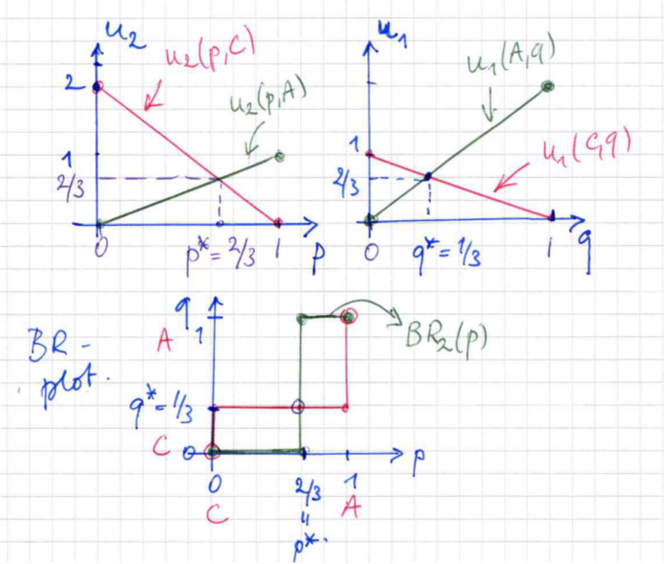
- Determining the mixture parameters p and q

$$\text{Ag 1:} \quad EU_2(s_1(p), A) = EU_2(s_1(p), C) \implies p = 2(1-p)$$

$$\text{Ag 2:} \quad EU_1(A, s_2(q)) = EU_1(C, s_2(q)) \implies 2q = (1-q)$$

Conclusion: $p = \frac{2}{3},\ q = \frac{1}{3} \qquad EU_1(s_1, s_2) = EU_2(s_1, s_2) = 2/3.$

Battle of the sexes: Mixed Nash Equilibrium



Nash equilibrium: Computation of mixed equilibrium

- **Prisoner's Dilemma** (does mixed strategy NE exist?)

	Quiet (q)	Confess ($1-q$)
Quiet (p)	$-1, -1$	$-12, 0$
Confess ($1-p$)	$0, -12$	$-8, -8$

- Determining the mixture parameters p and q

$$\text{Ag 1: } EU_2(s_1, Q) = EU_2(s_1, C) \Longrightarrow \quad -p - 12(1-p) = -8(1-p)$$

$$\text{Ag 2: } EU_1(Q, s_2) = EU_1(C, s_2) \Longrightarrow \quad -q - 12(1-q) = -8(1-q)$$

Conclusion: $p = q = \dfrac{4}{3}$ **impossible!**

Mixed NE for game with three actions

Rock-Paper-Scissors: no PURE NE!

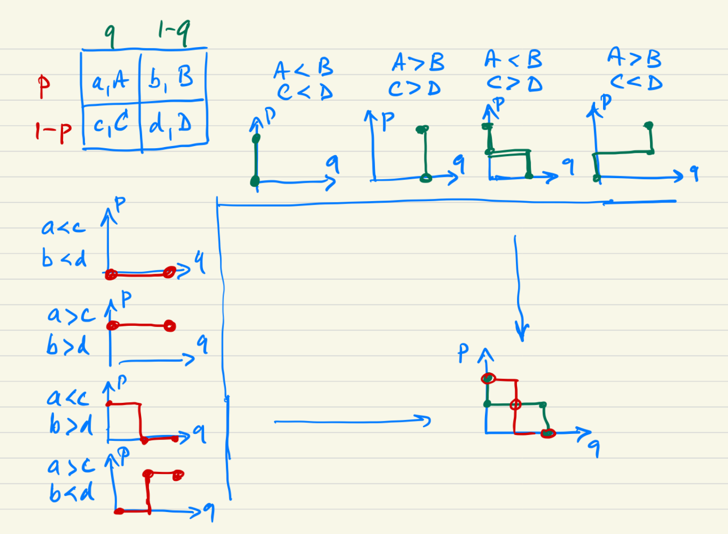
		v R	w P	$1-(v+w)$ S
p	R	$0,0$	$-1,1$	$1,-1$
q	P	$1,-1$	$0,0$	$-1,1$
$1-(p+q)$	S	$-1,1$	$1,-1$	$0,0$

Determine p, q by insisting that player 2 is indifferent btw actions:

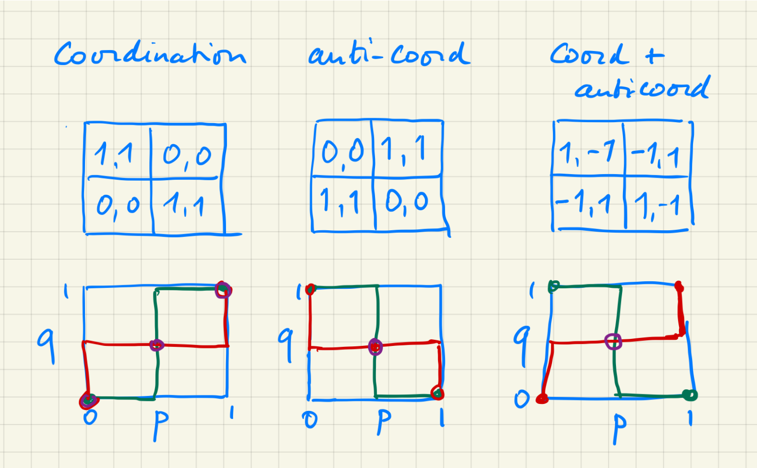
$$\begin{aligned} 0\cdot p+(-1)\cdot q+1\cdot(1-p-q) &= 1\cdot p+0\cdot q+(-1)\cdot(1-p-q) \\ 0\cdot p+(-1)\cdot q+1\cdot(1-p-q) &= -1\cdot p+1\cdot q+0\cdot(1-p-q) \end{aligned}$$

Result: $p=q=1/3$. Determine mixing parameters u, v similarly.

Nash equilibrium for 2 × 2 matrix game



Nash eq. for coordination and anti-coordination games



NE 2×3: Example of NE for 2×3 game

- Pay-off matrix and best response

		u	v	$1-u-v$
		L	C	R
p	T	3, 5	1, 1	2, 4
$1-p$	B	4, 1	0, 3	6, 2

- No pure Nash eq.
- There must be at least one mixed Nash eq.
  - Player 1: mixing: $\pi_1 = (p, 1-p)$
  - Player 2: mixing: $\pi_2 = (u, v, 1-(u+v))$

NE 2x3: Player 2 mixing to make Player 1 indifferent

Pl 2 mixing to make Pl. 1 indifferent:

$$EU_1(T,\pi_2) = 3u + v + 2(1-u-v) = u - v + 2$$

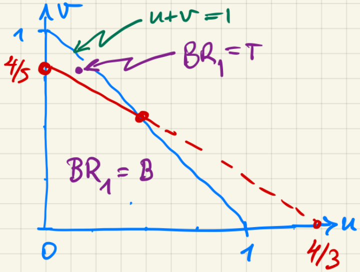
$$EU_1(B,\pi_2) = 4u + 0.v + 6(1-u-v) = -2u - 6v + 6$$

indifference

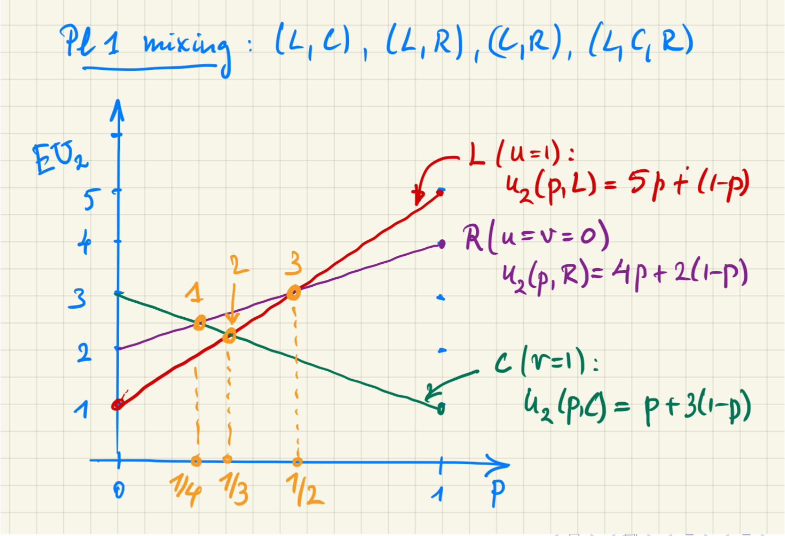
$\Downarrow$

$$u - v + 2 = -2u - 6v + 6$$

$$\Rightarrow \boxed{3u + 5v = 4}$$



NE 2x3: Player 1 mixing to make Player 2 indifferent



NE 2x3: Inspection of the Player 1 mixing options

- **M1: mixing R and C at** $p = 1/4$
  - Since L is not included in mix: $u = 0$
  - There is a unique point on indifference line for Player 1, corresponding to $u = 0$ and $v = 4/5$
  - Hence mixed NE at $p = 1/4$ and $u = 0, v = 4/5$
- **M2: mixing L and C at** $p = 1/3$
  - Not NE as deviating to R is better!
- **M3: Mixing L and R at** $p = 1/2$
  - Since C is not included: $v = 0$
  - No solution (point on indifference line outside feasible triangle!
- **Mixing all three strategies** dilutes pay-off and therefore is not optimal

NE 2x3: Expected utilities in Nash Eq.

3 5	1 1	2 4
4 1	0 3	6 2

π_2

π_1 \ π_2	0	4/5	1/5
1/4	0	4/20	1/20
3/4	0	12/20	3/20

$$u_1(\pi_1, \pi_2) = \frac{1}{20}\left[1 \cdot 4 + 0 \cdot 12 + 2 \cdot 1 + 6 \cdot 3\right]$$

$$= \frac{24}{20} = \frac{6}{5} = \underline{\underline{1.2}}$$

$$u_2(\pi_1, \pi_2) = \frac{1}{20}\left[1 \cdot 4 + 3 \cdot 12 + 4 \cdot 1 + 2 \cdot 3\right]$$

$$= \frac{50}{20} = \frac{5}{2} = \underline{\underline{2.5}}$$

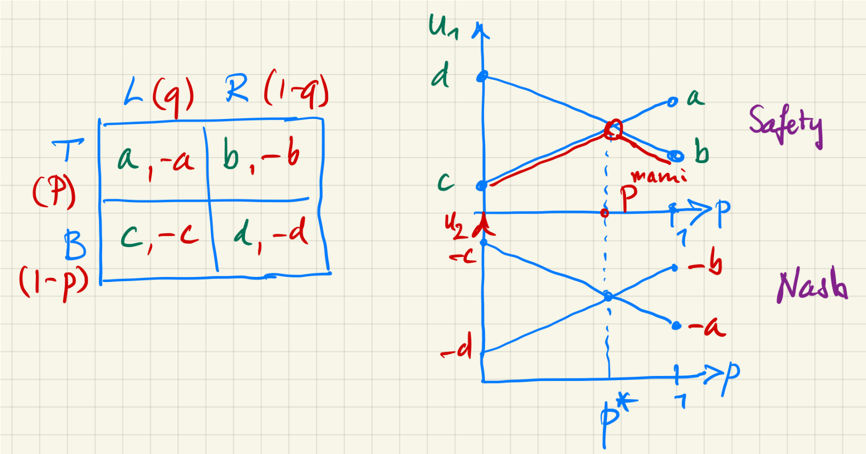
NE 2x3: Compare safety strategy to Nash Eq.

			min
3, 5	1, 1	2, 4	$\boxed{1}$
4, 1	0, 3	6, 2	0
min: 1	1	$\boxed{2}$ max	

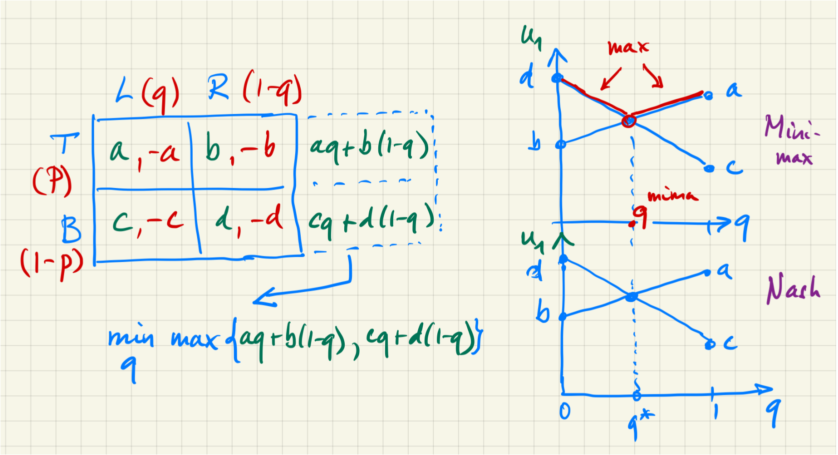
$$\boxed{v_1^{maxmin} = 1}$$

$$\boxed{v_2^{maxmin} = 2}$$

2 player zero-sum games: Nash = Maximin = Minimax



2 player zero-sum games: Nash = Maximin = Minimax



Fictitious Play (FP)

- FP uses simulation of many game iterations to learn about equilibria;
- FP refers to a dynamic process where at each stage, players play a (pure) best response to the empirical distribution of their opponent's play (which they interpret as a mixed strategy);
- **If FP converges (in distribution)**, the limit distribution coincides with mixed Nash strategies.
- Reminiscent of best response dynamics, but more general (includes not just last action, but whole history)
- **Form of learning**, works even opponent's utilities are unknown;

Fictitious Play: NE 2 × 3 example

	L	C	R
T	3,5	1,1	2,4
B	4,1	0,3	6,2

Suppose after $n = 40$ iterations (simulations)

$$N_1 = \begin{array}{cc} T & B \\ \boxed{30} & \boxed{10} \end{array} \qquad \hat{\pi}_1 = \boxed{3/4 \mid 1/4}$$

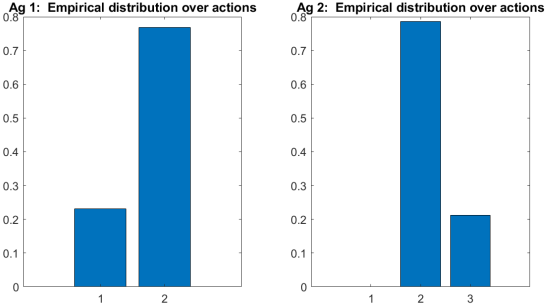
$$N_2 = \begin{array}{ccc} L & C & B \\ \boxed{15} & \boxed{5} & \boxed{20} \end{array} \qquad \hat{\pi}_2 = \boxed{3/8 \mid 1/8 \mid 4/8}$$

$$u_1(T, \hat{\pi}_2) = \frac{3}{8}\cdot 3 + \frac{1}{8}\cdot 1 + \frac{4}{8}\cdot 2 = \frac{9}{4}$$

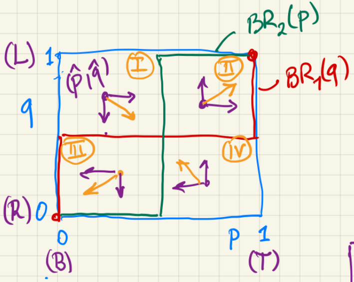
$$u_1(B, \hat{\pi}_2) = \frac{3}{8}\cdot 4 + \frac{1}{8}\cdot 0 + \frac{4}{8}\cdot 6 = \frac{18}{4} \longleftarrow \circledast$$

Fictitious Play: NE 2×3 example

- Recall: MNE at P1 (1/4, 3/4) and P2 (0,4/5, 1/5)
- Results for FP (300 iterations)



Caution: FP for 2 PNE and 1 MNE



After N iterations

Pl 1: $\nearrow n : T$, $\searrow N-n : B$

Pl 2: $\nearrow m : L$, $\searrow N-m : R$

$$\boxed{\hat p = \frac{n}{N}, \quad \hat q = \frac{m}{N}}$$

(I): $BR_1(\hat q) = T \ (p=1) \Rightarrow n \to n+1 \Rightarrow \hat p \uparrow$

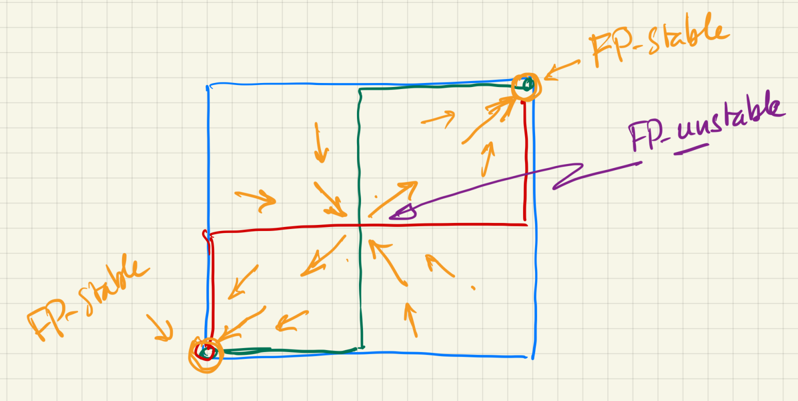
$BR_2(\hat p) = R \ (q=0) \Rightarrow m \to m \Rightarrow \hat q \downarrow$

(II) $BR_1(\hat q) = T \Rightarrow \hat p \uparrow \ ; \ BR_2(\hat p) = L \Rightarrow \hat q \uparrow$

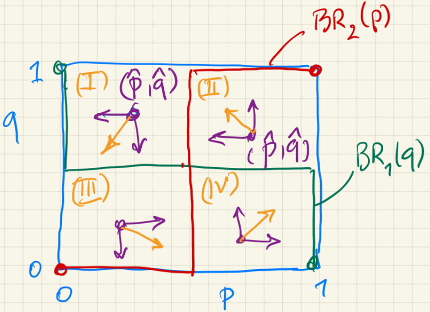
Mixed NE not found by FP.

FP No PNE, 1 MNE

PNE are stable but MNE is unstable under FP-dynamics



FP No PNE, 1 MNE



(I) $BR_1(\hat{q}) = (p=0) \Rightarrow (n \to n) \Rightarrow \hat{p} \downarrow$

$BR_2(\hat{p}) = (q=0) \Rightarrow (m \to m) \Rightarrow \hat{q} \downarrow$

(II) $BR_1(\hat{q}) = (p=0) \Rightarrow (n \to n) \Rightarrow \hat{p} \downarrow$

$BR_2(\hat{p}) = (q=1) \Rightarrow m \to m+1 \Rightarrow \hat{q} \uparrow$

FP No PNE, 1 MNE

Cycle yields approximate empirical distribution

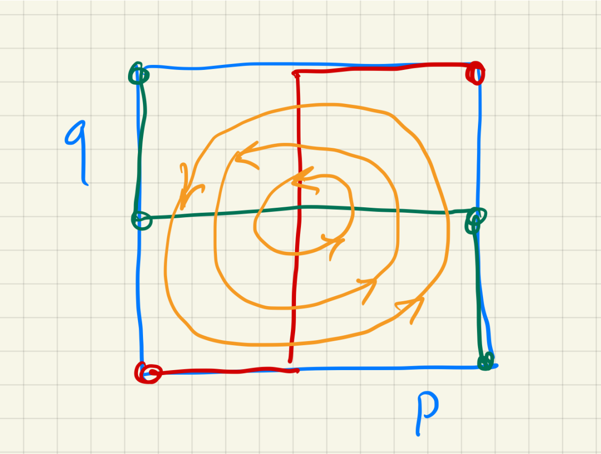
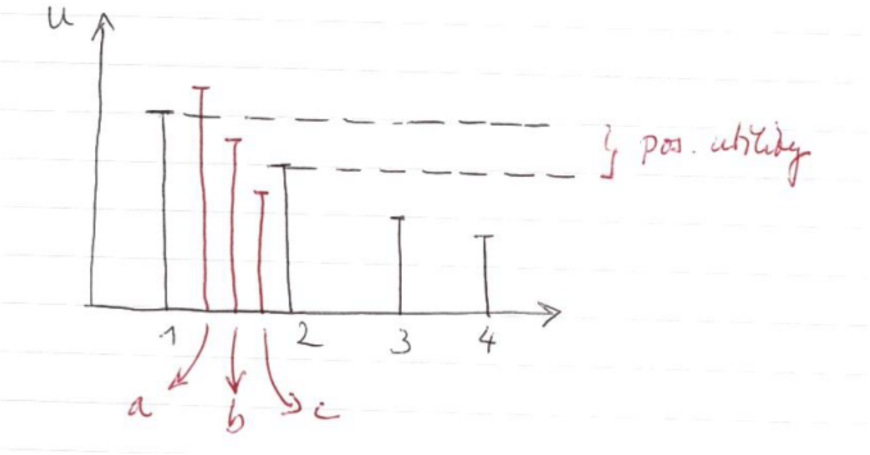


Table of Contents

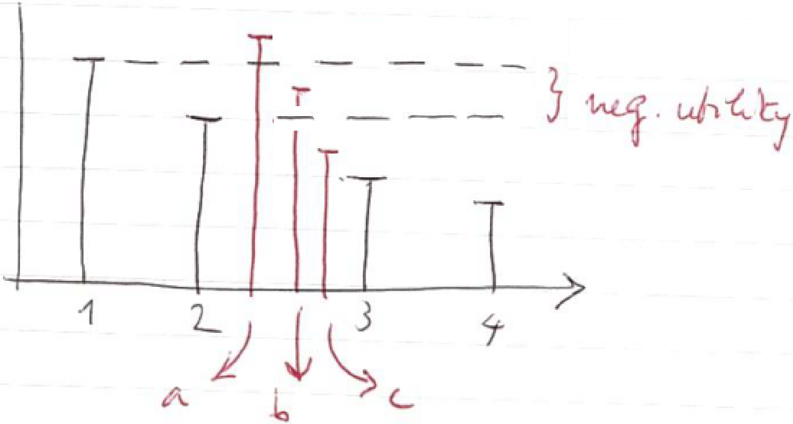
Vickrey auction: Second Price Auction

- **Vickrey Auction:**
  - n sealed bid auction for single item;
  - highest bid wins, but pays 2nd highest price;
- **Truth-telling** is (weakly) **dominant strategy;**
- NE: Neither winner nor loser(s) have incentive to deviate:
- **Winner:**
  - Higher: still winner, same price;
  - Lower: might lose, but if still winner, still paying 2nd price;
- **Loser:**
  - Higher: possibly winner, but at higher price;
  - Lower: still loser;
- Example of **mechanism design**.

Vickrey Auction: Alternatives for winner



Vickrey Auction: Alternatives for runner-up



Hawk or Dove

- Equilibria as a function of **exogenous pay-off variables**;
- **Exogenous variables** are imposed on the game (not by players);
- **Strategy: Hawk or dove:**
  - Two parties are in conflict over some good of value $2v > 0$;
  - Two doves share, each gets v;
  - Hawks fight, on average each gets half, but at a cost c (e.g. due to injury)
  - A dove is no match for a hawk and yields;
  - **Pay-off matrix:**

	Hawk	*Dove*
Hawk	$v-c,\ v-c$	$2v, 0$
Dove	$0, 2v$	$v,\ v$

- Different outcomes depending on cost of aggression c!

Hawk or Dove

- **Aggression is cheap** ($c < v$) H is dominant, hence: NE = (H,H) with utility $(v - c, v - c)$;
- **Aggression is risky** ($c > v$)
  - Two PNE = (H,D) and (D,H)
  - One MNE at $P(H) = p^* = v/c$.
  - If risk/cost increases $c \uparrow$, then $p^* \downarrow$.

Investment game

- Many agents (n) but only two strategies;
- **Strategies:** Each agent can either
  - I: invest 10 euro
  - N: not invest
- **Pay-offs:**
  - N: No investment: pay-off $= 0$;
  - I: Pay-off $= 15$:

$$\text{net pay-off } = \begin{cases} 5 & (= 15 - 10) & \text{if at least 90\% of ag. invest} \\ -10 & (= 0 - 10) & \text{otherwise} \end{cases}$$

- **Coordination game!**

Strategic Effects

- **Symmetric penalty kick game:**

		goal keeper	
		left	right
kicker	left	$0, 0$	$1, -1$
	right	$1, -1$	$0, 0$

- Suppose kicker has weak left kick: **direct and indirect effect!**

		goal keeper	
		left	right
kicker	left	$0, 0$	$a, -a$
	right	$1, -1$	$0, 0$

$$0 < a < 1$$

Strategic Effects

- Kicker has weak left side, so is inclined to kick less with left;
- Goalkeeper knows this, so anticipates less kicks to the left, so will tend to jump less to the left ...
- Since kicker knows this, he might reconsider and kick more with left, since goalie has tendency to jump to the right ...,
- but goalie knows this too, so might reconsider ...
- and so on ...
- Is there a way out of infinite regress???

Strategic Effects

$$\begin{array}{cc|c|c|}
 & & \text{goalie} & \\
 & & (q)\ L & (1-q)\ R \\
\hline
\text{Striker} & (p)\ L & 0,0 & a,-a \\
\hline
 & (1-p)\ R & 1,-1 & 0,0 \\
\hline
\end{array}$$

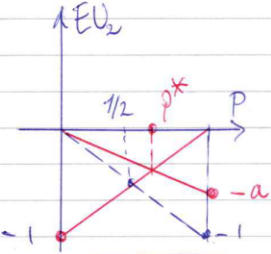
$0 < a < 1.$

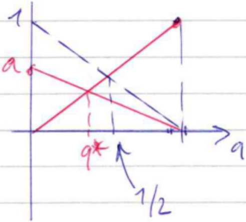
$$\left.\begin{aligned} EU_2(p,L) &= 0.p + (-1)(1-p) = p-1 \\ EU_2(p,R) &= -ap + 0(1-p) = -ap. \end{aligned}\right\} \text{goalie.}$$

$$\left.\begin{aligned} EU_1(L,q) &= 0.q + a(1-q) = a(1-q) \\ EU_1(R,q) &= q + 0 = q \end{aligned}\right\} \text{striker}$$

Strategic Effects

		q L	$1-q$ R
Striker	(p) L	$0,0$	$a,-a$
	$(1-p)$ R	$1,-1$	$0,0$





$\boxed{a\downarrow \Rightarrow p^*\uparrow}$ $\boxed{a\downarrow \Rightarrow q^*\downarrow}$

Strategic Effects

- Mixed Nash equilibrium at:

$$p^* = \frac{1}{1+a} = 1 - q^* \quad \Rightarrow \quad p^* > 1/2, \quad q^* < 1/2$$

- Zero-sum game: Utilities for kicker and goalie:

$$EU_1(p^*, q^*) = \frac{a}{1+a} = -EU_2(p^*, q^*)$$

- Notice that $p^* \to 1$ as $a \to 0$, i.e. if left kick is powerless ($a \approx 0$), make sure to kick left ($p^* \approx 1$).
- Is NE the right tool to think about this??

Homework dilemma: TA versus Wikipedia

- **Student:** either is honest (H), or cheats, copying Wikipedia (W), which means that the homework is solved with much less effort. However, there is a severe penalty if caught.
- **TA:** either trusts the students (no check: N) or conducts a thorough Wikpedia check (C). The latter requires considerably more effort from the TA!

Homework dilemma: TA versus Wikipedia

		Student	
		H (q)	W $(1-q)$
TA	N (p)	$\underline{2}$ 0	-1 $\underline{5}$
	C $(1-p)$	1 $\underline{1}$	$\underline{5}$ -10

Mixed NE:

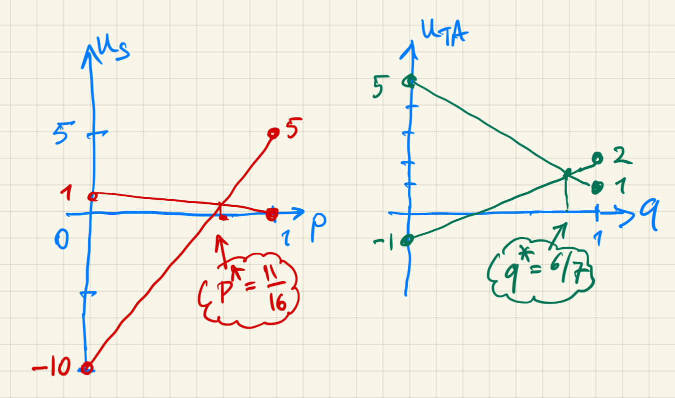
$$p^*: \quad 0p + 1(1-p) = 5p + (-10)(1-p)$$

$$1-p = 5p + 10p - 10 \Rightarrow \boxed{p^* = 11/16}$$

$$q^*: \quad 2q + (-1)(1-q) = 1q + 5(1-q) \Rightarrow \boxed{q^* = 6/7}$$

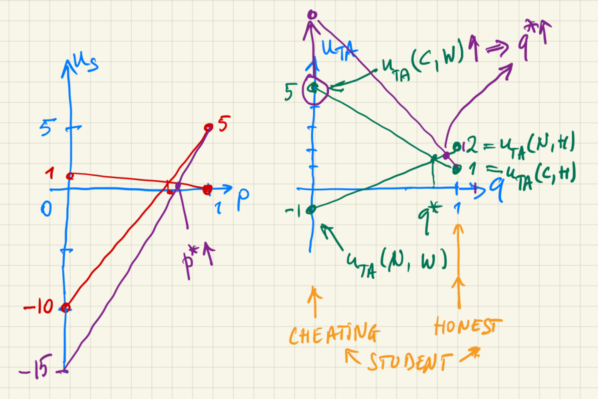
- Interpretation p^*: TA's actual randomisation probability.
- Interpretation q^*: Fraction of student population that cheats

Homework dilemma: TA versus Wikipedia



- How to increase the fraction q^* of honest students?
  - Increase penalty for cheating?
  - Increase pay-off for TA for catching cheater?

Homework dilemma: TA versus Wikipedia



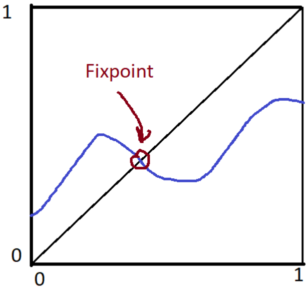
- Increase penalty for cheating $p^* \uparrow$ i.e. **less checking** by TA.
- Increase TA's pay-off for catching cheater: $q^* \uparrow$, i.e. **fewer students cheat!**

Table of Contents

Aside: Sperner's Lemma (1928) and fix-point theorems

Defintion: Fix-point

Point $a \in A$ is a **fix-point** for function $f : A \longrightarrow A$, iff $f(a) = a$.



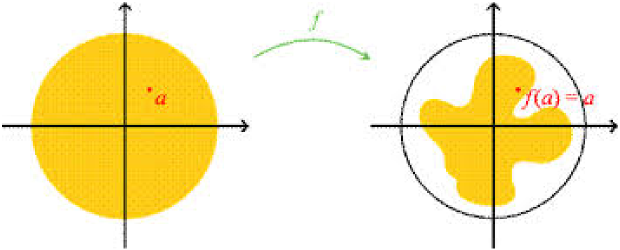
f: [0,1] ---> [0,1] continuous

Sometimes(!), fixpoints can be computed using **function iteration.**

Nash Theorem is based on Fixed-Point Theorem

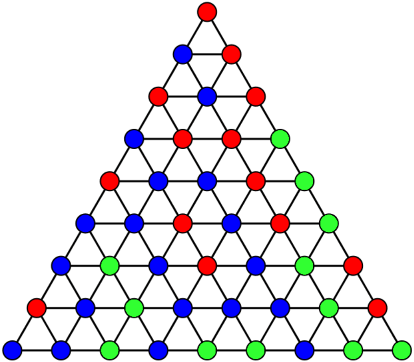
Brouwer's Fixed Point Thm

Let $K \subset \mathbb{R}^n$ be a compact and convex, and $f : K \longrightarrow K$ continuous. Then f has a fix-point in K, i.e. there exists a $x_0 \in K : f(x_0) = x_0$.



: **Non-constructive existence proof!**

Aside: Sperner's Lemma (1928) and fix-point theorems

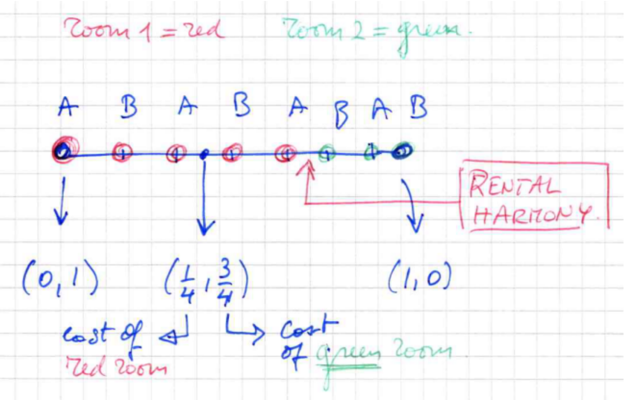


If in triangle:

- corner nodes have different colors
- nodes on outer edges have two possible colors (determined by corners)

Then there is a subtriangle with 3 differently colored corners.

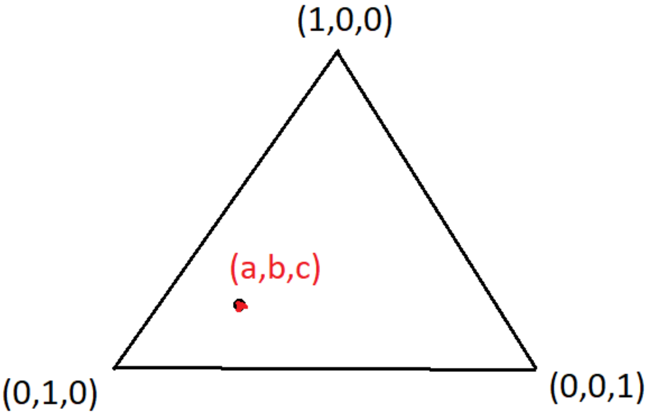
Sperner's lemma: Rental Harmony: Fair division of rent



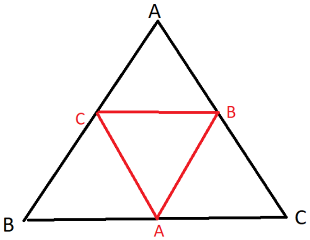
- Divide interval in segments by adding points
- Assign alternating decision makers (A and B)
- Decision maker decides which room he picks (for given rental division)

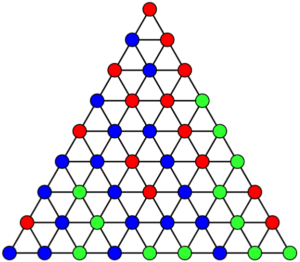
Sperner's lemma – Application 1: Rental Harmony

Flat sharing: three bedroom flat shared by tree friends but rooms are not of equal quality: how much should each contribute to rent?

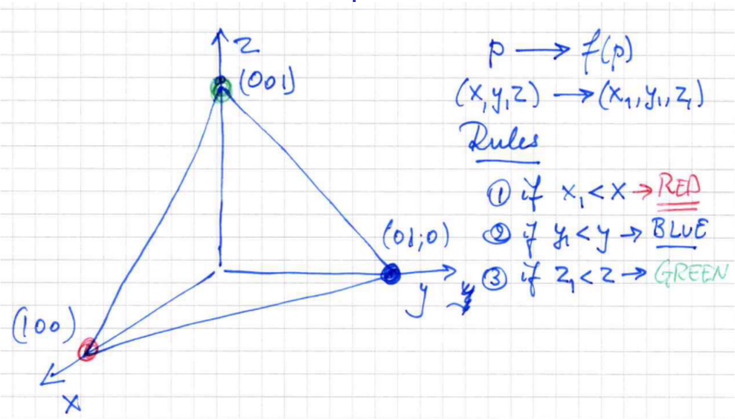


Sperner's lemma – Application 1: Rental Harmony





Sperner's lemma – Application 2: Proof of Brouwer's fixpoint thm



Youtube: Trefor Bazzett: A beautiful combinatorical proof of the Brouwer Fixed Point Theorem - Via Sperner's Lemma

Nash equilibrium: NE and IEWDS

Eliminating **weakly dominated** strategies might erase NEs!

	L	R
U	$2,3$	$4,3$
D	$3,3$	$1,1$

- Two NEs: (4,3) and (3,3)
- L weakly dominates R;
- Eliminating R would result in single solution (3,3);
- Notice that the Pareto-optimal NE would be eliminated.

Games with NO Nash Equilibrium

- **Dollar auction:** Sealed bid auction: highest bid gets item, 2nd highest pays this amount!
- More generally: One can construct games without NE by making sure that either
  - state space is **not compact**
  - utility function is **not continuous**
- **Ex. for 2-player games with continuous state spaces:**
  - Non-compact: $u_i(x, y) = xy \qquad \textit{for} \quad 0 \le x, y < 1$
  - Non-continuous: $0 \le x, y \le 1$ and

$$u_i(x, y) = \begin{cases} xy & \textit{if} \quad x, y < 1 \\ 0 & \textit{if} \quad x, y = 1 \end{cases}$$

Nash equilibrium: Computational aspects

- Finding a Nash equilibrium for 2-player zero-sum games can be done efficiently by formulating a linear program. Notice that in this case: NE = minimax = maximin.
- Finding a Nash equilibrium is not known to be NP-complete because it is not a decision problem
- PPAD (polynomial parity argument, directed version) is a class describing problems for which a solution always exists
- Daskalakis, Goldberg, and Papadimitriou showed that finding a sample Nash equilibrium of a general-sum finite game with two or more players is PPAD-complete (i.e. "difficult!").

Other solution concepts

- **Minimax equilibrium**: zero-sum special case of Nash
- **Trembling-hand perfect equilibrium**: each player's action is a best-response even if other players make small mistakes
- ϵ-**Nash equilibrium**: deviating benefits no agent more than ϵ
- **Correlated equilibrium**: agents can condition strategy on external signal: "If there is intelligent life on other planets, in a majority of them, they would have discovered correlated equilibrium before Nash equilibrium." -Myerson
- **Evolutionary stable state**: "A population is said to be in an evolutionarily stable state if its genetic composition is restored by selection after a disturbance, provided the disturbance is not too large." -Maynard Smith

Summary

- Game theory studies utility-based multiagent decision making.
- Solving a game means trying to predict its outcome.
- Rational agents never play strictly dominated actions.
- No agent has an incentive to **unilaterally deviate** from a **Nash equilibrium**.
- In finite games, there's always a NE (possibly in mixed strategies).
- Nash equilibria need not be Pareto optimal.

Some additional literature

- Presh Talwalkar: The Joy of Game Theory: An introduction to strategic thinking.
- Avinash K. Dixit, Barry J. Nalebuff: The Art of Strategy: A Game Theorist's Guide to Success in Business and Life
- William Poundstone: Prisoner's Dilemma: John von Neumann, Game Theory, and the Puzzle of the Bomb