Lecture 01: Introduction of Multiagent AI and the Existence of Nash Equilibrium

Jun Wang
UCL

Outline of MAAI

1. Basic Game Theory and Nash Equilibrium
2. Potential Games and Best Respond Dynamics
3. Repeated Games and the Theory of Cooperation
4. Minimax Games
5. General-Sum Games and L-H Algorithms
6. Single-agent Reinforcement Learning and MDP
7. Learning Stochastic Games
8. Non-regret Learning and Correlated Equilibrium
9. Counterfactual Regret Minimisation (AI Poker)
10. Learning with a Population of Agents

- Subject to change

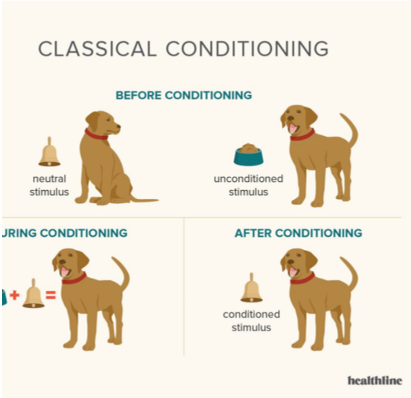
Main References of This Lecture

- Drew Fudenberg and Jean Tirole. "Game theory". In: *Cambridge, Massachusetts* 393.12 (1991), p. 80
- John F Nash Jr. "Equilibrium points in n-person games". In: *Proceedings of the national academy of sciences* 36.1 (1950), pp. 48–49
- Albert Xin Jiang and Kevin Leyton-Brown. "A Tutorial on the Proof of the Existence of Nash Equilibria". In: *University of British Columbia Technical Report TR-2007-25. pdf* 14 (2009)

Outline

Table of Contents

Background: What is Learning?[1]

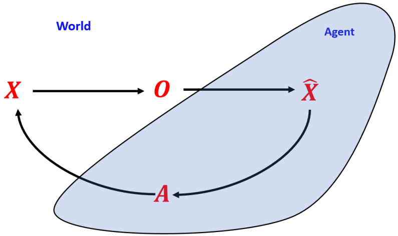


- In biology, learning is a behavioural change driven by experience.
- Classical conditioning: Learning to associate stimuli. For example:
  - Pairing bell (O), food (X), a dog salivates when the bell rings.
  - Representation: $X \rightarrow O \rightarrow \hat{X}$
- Learned behaviours enhance survival and reflect intelligence (e.g., avoiding harmful foods).

[1] Ivan Petrovitch Pavlov and W Gantt. "Lectures on conditioned reflexes: Twenty-five years of objective study of the higher nervous activity (behaviour) of animals.". In: (1928).

Background: Generalist AI Agents

- Since its inception, AI has aimed to develop generalist agents that adapt to complex tasks.
- Such agents act effectively in uncertain environments, achieving subgoals that support their ultimate objectives[2].
- They require:
  - Sensors to perceive their environment: $P(X|O)$.
  - Actuator to act upon it: $P(A|\hat{X})$.

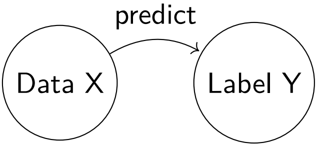


The variable X is the (unknown) environmental state, O its observation, $\hat{X}$ the estimation, and A the agent's action, which influences X.

[2] James S Albus. "Outline for a theory of intelligence". In: *IEEE transactions on systems, man, and cybernetics* 21.3 (1991), pp. 473–509.

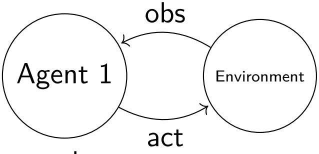
Motivations: Strategic Interactions

(Sensoring) classic machine learning and pattern recognition:

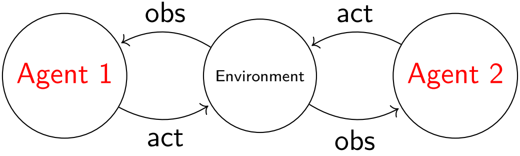


Examples: Classification, Regression, Clustering, Generative models etc.

(Acting upon it) decision making (Single-agent and multiagent systems):



Examples: Reinforcement learning, Bandit etc.



Examples: Online Auction, Route planning, Multiagent learning, GAN, Self-driving cars, StarCraft games, Poker etc.

Standard Machine Learning Formulation

- Predictions are obtained through a predictive model f_θ parameterised by $\theta \in \Theta$:

  $$\hat{y} = f_\theta(x) \in Y, x \in X,$$

  where the input feature space: $X \subseteq \mathbb{R}^d$; Output space: $Y \subseteq \mathbb{R}$

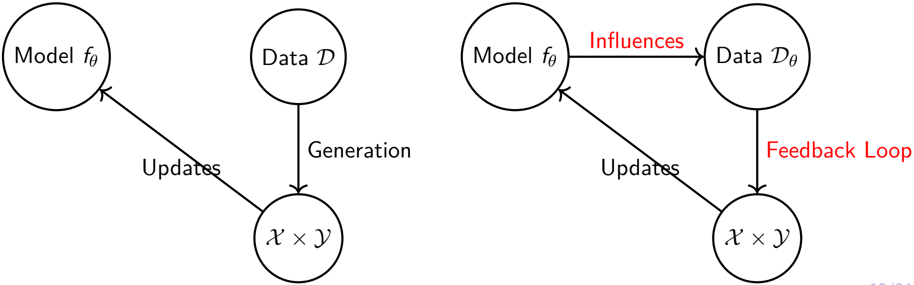
- The quality of a model is assessed by its ability to predict y from x over a target population (Standard ML Objective):

  $$\theta_{ML} \in \arg\min_{\theta \in \Theta} \text{Risk}(\theta, D) = \mathbb{E}_{z \sim D}[\ell(\theta; z)], \tag{1}$$

  where $D \in \Delta(X \times Y)$ denotes a probability distribution over $X \times Y$; $z \equiv \{x, y\}$ and ℓ is a loss function, e.g., squared errors.

Prediction Needs to be Strategic

- However, a deployed prediction model θ might influence the data distribution D too. Thus, $\theta \Leftrightarrow D$. Examples include:
  1. **Loan Approval:** Predicting creditrating alters borrower behaviour.
  2. **Traffic Routing:** Traffic predictions affect driver routes.
  3. **Content Moderation:** Flagging posts changes user content.
  4. **Energy Pricing:** Demand forecasts shift consumption.
  5. **Job Rec.:** Job suggestions influence applications.
  6. **Stock Predictions:** Market forecasts shape trading.
  7. **Health Alerts:** Outbreak warnings modify public behavior.



Strategic Prediction[3]

- To model this, mathematically, we define $D : \Theta \to \Delta(X \times Y)$ mapping model parameters $\theta \in \Theta$ to a data-generating distribution $D(\theta)$.

A **Strategic** (calculated and forward-looking approach) Prediction should consider:

- Performative Stability: A model θ_{PS} is stable if it satisfies:

$$\theta_{PS} \in \arg\min_{\theta \in \Theta} \mathrm{Risk}(\theta, D(\theta_{PS})). \tag{2}$$

- Performative Optimality: A model θ_{PO} is optimal if it satisfies:

$$\theta_{PO} \in \arg\min_{\theta \in \Theta} \mathrm{Risk}(\theta, D(\theta)). \tag{3}$$

- Performative optima minimise risk under distribution shifts; stable points reflect fixed point equilibrium under current data.

[3] Moritz Hardt and Celestine Mendler-Dünner. "Performative prediction: Past and future". In: *arXiv preprint arXiv:2310.16608* (2023).

An Example (1/3)[4]

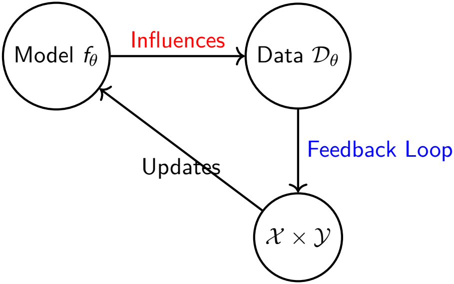
Banks build predictive models to assess creditworthiness:

1. The initial model predicts repayment probabilities based on credit scores, i.e., $\hat{y} = f_\theta(x) \in Y$.
2. Predictions influence loan approval decisions.
3. Loan approval/rejection decisions shape applicants' behaviour (e.g., saving habits, credit management), i.e., $D(\theta)$.
4. Changed applicant behaviour alters the repayment probability distribution. $D(\theta) \longrightarrow X \times Y$.
5. The optimal model must account for this dynamic feedback loop, seeking a balance between prediction accuracy and performative effects.

[4] Moritz Hardt and Celestine Mendler-Dünner. "Performative prediction: Past and future". In: *arXiv preprint arXiv:2310.16608* (2023).

An Example (2/3)

- Game Setup:
  - Institution deploys classifier f_θ to predict outcomes (e.g., loan repayment).
  - Applicants adapt features strategically to improve outcomes.
- Stackelberg (Bi-level) Structure:
  - Institution (Leader Agent) deploys a classifier.
  - Applicants (Follower Agent) adjust features after observing f_θ.



Performative effects: the model influences the data-generating process.

An Example (3/3)[5]

$\theta \to D(\theta)$: Modified pairs after applicants respond to f_θ.

Algorithm 1 Sampling procedure for $\mathcal{D}(\theta)$

1: Sample $(x, y) \sim \mathcal{D}$: Feature-outcome pairs before classifier deployment.
2: Compute $x_{\text{BR}} \leftarrow \arg\max_{x'} u(x', \theta) - c(x', x)$, where $u(x', \theta)$: applicant's utility, $c(x', x)$: cost of deviation.
3: Output sample (x_{BR}, y)

$D(\theta) \to f_\theta$: e.g., Performative Optimality:

$$f_\theta^{\text{SE}} \text{ is optimal } \iff \theta_{\text{PO}} \in \arg\min_\theta Risk(\theta; D(\theta))$$

$Risk(\theta; D(\theta))$: Performative risk on $D(\theta)$.

[5] Juan Perdomo et al. "Performative prediction". In: *International Conference on Machine Learning*. PMLR. 2020, pp. 7599–7609.

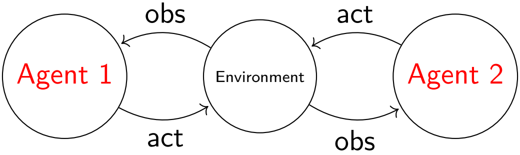
Multiagent Learning vs. Classic Machine Learning

- Classic ML typically optimises a single objective (e.g., minimising loss).
- Multiagent learning represents a multi-stakeholder, sometime adversarial setup with two or more competing objectives.
- Equilibrium states (stable points) replace traditional ML optimal solutions.

Table of Contents

What is Game Theory?

- Game theory provides a mathematical framework for understanding how AI agents interact strategically, influencing each other's outcomes in multi-agent systems.
- It offers insights into decision-making in competitive, cooperative, and mixed scenarios, enabling the design of AI systems that can effectively interact with other AIs and humans, extending beyond traditional human-to-human interactions.



Historical Milestones in Game Theory

- 1713: Waldegrave introduces the mixed strategy for *Le Her* card game [BF15].
- 1928: von Neumann formalises game theory for strictly competitive games [Neu28].
- 1950s: Nash develops the Nash equilibrium, a cornerstone of modern game theory [Nas50].
- 1970s: Applications expand to evolutionary biology and experimental economics [May82].
- 1990s: Emergence of algorithmic game theory, integrating computational complexity with strategic decision-making [Nis+07].

Many recent applications of game theory focus on machine learning and artificial intelligence.

Why Study Game Theory in Machine Learning?

- Multi-Agent Systems: Models interactions in environments like autonomous vehicles, recommender systems, and online markets.
- Strategic Decision-Making: Analyses dynamic settings where agents adapt and learn strategies.
- Reinforcement Learning (RL): Foundations for self-play and multi-agent RL, with concepts like Nash equilibrium.
- Adversarial Learning: Underpins adversarial settings, including GANs and model robustness.
- Mechanism Design: Aligns incentives for fairness, stability, and efficiency in distributed systems and marketplaces.
- Applications: Bidding algorithms, AI playing Games (Chess, Poker, Video Games), resource allocation, pricing, scalable AI systems, and LLM agents.

The Scenario (1/2)

- You are a student with two assignments due tomorrow:
  - An individual exam (solo work)
  - A group presentation (with partner)
- Time constraints:
  - Can only Prepare for one: Exam or Presentation
  - Goal: Maximise average grade (equal weight)
  - Your partner faces the same choice
- Payoff Structure:
  - Exam: 92 (if studied), 80 (if not)
  - Presentation:
    - Both prepare: 100 each
    - One prepares: 92 for preparer
    - Neither prepares: 84 each

Key constraint: No communication & simultaneous decision.

Payoff Matrix (2/2)

- You have two choices either Exam or Presentation; the same case for your partner.
- Calculations shown (from your perspective):
  - (Presentation,Presentation): $0.5 \times 80 + 0.5 \times 100 = 90$
  - (Exam, Presentation): $0.5 \times 92 + 0.5 \times 92 = 92$
  - (Exam, Exam): $0.5 \times 92 + 0.5 \times 84 = 88$
  - (Presentation, Exam): $0.5 \times 92 + 0.5 \times 80 = 86$

A Finite Strategic Game can be described conveniently in a table (matrix) below (aka a Matrix game or Normal-Form Game):

		Your Partner	
		Presentation	Exam
You	Presentation	(90,90)	(86,92)
	Exam	(92,86)	(88,88)

Key Insight: Each player considers studying for the exam as the "optimal" strategy $\longrightarrow$ we can predict the behaviour and pattern.

How are the Players Likely to Behave?

To reason about how players behave in a game, we make a few assumptions:

- Payoffs reflect all the rewards: The payoff for a player should capture all the rewards the player receives from the game.
- Complete knowledge: Players know all possible strategies and payoffs for other players.
- Rationality: Players are rational:
  - Each player aims to maximise their own payoff.
  - Each player succeeds in picking the optimal strategy given the payoffs.

However, it is important to note that not all of these assumptions always hold true in every situation, e.g., bounded rationality[6].

[6] "Modelling Bounded Rationality in Multi-Agent Interactions by Generalized Recursive Reasoning". In: *IJCAI* (2020).

The Definition of Dominant Strategies

Definition: Strictly Dominant Strategy

A strategy that is strictly better than all other options, regardless of what the other player does.

		Your Partner	
		Presentation	Exam
You	Presentation	(90,90)	(86,92)
	Exam	(92,86)	(88,88)

- The "Exam" strategy is strictly better for both players, regardless of the other's choice.
- Rational players always choose dominant strategies, leading to a predictable outcome.
- Paradoxically, cooperation could yield a better outcome: (90,90), but rational self-interest prevents it.

Another Example: the Prisoner's Dilemma

- Two suspects are arrested and suspected of committing a robbery.
- They are interrogated separately.
- The payoffs are as follows:

$Player1/Player2$	Don't Confess	Confess
Don't Confess	$-1, -1$	$-10, 0$
Confess	$0, -10$	$-4, -4$

- Confessing is a strictly dominant strategy: For each player, confessing always yields a higher payoff than remaining silent, regardless of what the other player does.
- The dilemma: Even though not confessing would give both players a better outcome (1 year each), rational play leads both players to confess, resulting in a worse social outcome (4 years each).

Prisoner's Dilemma: General Cases

The Prisoner's Dilemma illustrates the tension between cooperation and defection.

Payoff Matrix:

$$\begin{bmatrix} (R, R) & (S, T) \\ (T, S) & (P, P) \end{bmatrix}$$

where $T > R > P > S$, highlighting incentive misalignment and challenges in achieving mutual benefits.Defection often dominates, despite better outcomes from cooperation.

Applications:

- *Sports*: Athletes face incentives to use performance-enhancing drugs, harming fairness.
- *Arms races*: Nations invest in costly armaments despite mutual disarmament being preferable.

The solution lies in designing mechanisms to align incentives and foster cooperation (repeated games).

Another Example: Performance-Enhancing Drugs

Consider the following game where two players decide whether to use performance-enhancing drugs:

Player 1 / Player 2	Don't Use Drugs	Use Drugs
Don't Use Drugs	$3, 3$	$1, 4$
Use Drugs	$4, 1$	$2, 2$

- Using Drugs is a dominant strategy for both players, despite the fact that mutual cooperation (not using drugs) would yield a better outcome (3, 3).

The Definition of Best Response

Definition: Best Response

A strategy is a best response for a player if it maximises their payoff given the strategy of the other player.

- Let player 1 choose strategy S and player 2 choose strategy T.
- The payoff to player i for strategies (S, T) is denoted by $U_i(S, T)$.
- S is a best response to T if:

$$U_1(S, T) \geq U_1(S', T) \quad \text{for all strategies } S' \text{ of player 1.}$$

- S is a strict best response if:

$$U_1(S, T) > U_1(S', T) \quad \text{for all strategies } S' \text{ of player 1.}$$

Dominant Strategy and Best Response

- A strategy is dominant if it is the best response to every strategy of the other player.
- A strictly dominant strategy is a strategy that is strictly better than all other strategies, no matter what the opponent does.
- In the Prisoner's Dilemma, both players have strictly dominant strategies:
  - Confess is strictly better than Not Confessing, regardless of the other player's choice.

Not All Games Have Dominant Strategies

- Marketing Game: Only Firm 1 has a dominant strategy.
- Two firms compete: Firm 2 is established. 60% of consumers prefer low-priced products; 40% prefer upscale options.
- Direct competition: Firm 1 captures 80% of sales, Firm 2 captures 20%.
- Avoiding competition: Each firm dominates its own segment.

Form 1 /Form 2	Low-Priced	Upscale
Low-Priced	0.48, 0.12	0.60, 0.40
Upscale	0.40, 0.60	0.32, 0.08

Sales Calculation:

- 60% of the market prefers low-priced items: $60\% \times 80\% = 48\%$ for Firm 1.
- 40% of the market prefers upscale items: $40\% \times 80\% = 32\%$ for Firm 1.

Not All Games Have Dominant Strategies

Let us reason:

- Firm 1 has a dominant strategy: Low-priced.
- Firm 2 does not have a dominant strategy and thus assumes that Firm 1 will play Low-priced.
- Therefore, the stable prediction (we say in the equilibrium) for the game is (Low-Priced, Upscale), resulting in Firm 1 capturing 60% and Firm 2 capturing 40% of the market.

Firm 1/ Firm 2	Low-Priced	Upscale
Low-Priced	0.48, 0.12	0.60, 0.40
Upscale	0.40, 0.60	0.32, 0.08

What If Neither Player Has a Dominant Strategy?

What is the best response to each strategy in a three-action game?

	A	B	C
A	4, 4	0, 2	0, 2
B	0, 0	1, 1	0, 2
C	0, 0	0, 2	1, 1

Player 1's Responses:

- A is the best response to A (Player 2's choice).
- B is the best response to B.
- C is the best response to C.
- There is no single dominant strategy for either player.

Thus, our best prediction is that players will choose strategies that are best responses to each other's choices $\Rightarrow$Equilibrium Concepts.

Best Responses to Each Other: Nash Equilibrium[7]

Definition: Nash Equilibrium

A strategy pair (S, T) is a Nash Equilibrium if:

- S is a best response to T, and
- T is a best response to S.

At a NE, no player wants to unilaterally deviate to an alternative strategy.

Example: Three-action Game

- Suppose Firm 1 chooses A and Firm 2 also chooses A.
- These strategies are mutual best responses, and neither player wants to deviate.
- Hence, (A, A) is a Nash Equilibrium.
- It is also the unique equilibrium in this case.

[7] John F Nash Jr. "Equilibrium points in n-person games". In: *Proceedings of the national academy of sciences* 36.1 (1950), pp. 48–49.

Coordination Game

Scenario: Preparing a presentation with a partner, but unsure which software they will use.
Incompatibility between software choices incurs effort.

Player 1/Player 2	PowerPoint	Keynote
PowerPoint	1, 1	0, 0
Keynote	0, 0	1, 1

Both (PowerPoint, PowerPoint) and (Keynote, Keynote) are Nash Equilibria.
This is a coordination game where the players prefer to coordinate but have multiple equilibria to choose from.

What is a Game[8]?

Definition of a Game

A finite strategic (normal-form) game is defined as:

$$\Gamma = \langle \mathcal{I}, (S_i)_{i \in \mathcal{I}}, (u_i)_{i \in \mathcal{I}} \rangle$$

where:

1. A set of players $\mathcal{I} = \{1, \ldots, n\}$.
2. Each player $i \in \mathcal{I}$ has a strategy set S_i, where $s_i \in S_i$ represents a specific strategy.
3. A payoff function $u_i : S_1 \times \cdots \times S_n \to \mathbb{R}$, which gives the payoff for player i based on the strategy profile $(s_1, \ldots, s_n)$.

[8] Drew Fudenberg and Jean Tirole. "Game theory". In: *Cambridge, Massachusetts* 393.12 (1991), p. 80.

Solution Concepts in Strategic Games

Definition (Math): Dominant Strategy

A strategy $s_i \in S_i$ is dominant for player i if it maximises the payoff for i regardless of the strategies chosen by other players:

$$u_i(s_1, \ldots, s_{i-1}, s_i, s_{i+1}, \ldots, s_n) \geq u_i(s_1, \ldots, s_{i-1}, s_i', s_{i+1}, \ldots, s_n)$$

for all $s_i' \in S_i$ and all $(s_1, \ldots, s_n) \in S_1 \times \cdots \times S_n$.

Definition (Math): Nash Equilibrium

A strategy profile $(s_1^*, \ldots, s_n^*)$ is a Nash Equilibrium if no player can improve their payoff by unilaterally changing their strategy:

$$u_i(s_1^*, \ldots, s_{i-1}^*, s_i, s_{i+1}^*, \ldots, s_n^*) \leq u_i(s_1^*, \ldots, s_i^*, \ldots, s_n^*)$$

for all $s_i \in S_i$ and for each $i \in \mathcal{I}$.

Important Games (1): Battle of the Sexes

Scenario: Two people want to rent a movie, but they prefer different genres (e.g., romance vs thriller).
There are two Nash Equilibria: (Romance, Romance) and (Thriller, Thriller), but which one will be played (predicted) depends on the players' social conventions, culture, or past experiences.

Player 1/Player 2	Romance	Thriller
Romance	1, 2	0, 0
Thriller	0, 0	2, 1

Important Games (2): Stag Hunt

Scenario: Two hunters can either hunt a stag together or hunt hares separately.
If both hunt stag, they get the best outcome, but if one chooses stag and the other chooses hare, the stag hunter gets nothing.

Hunter 1/Hunter 2	Hunt Stag	Hunt Hare
Hunt Stag	4, 4	0, 3
Hunt Hare	3, 0	3, 3

It illustrates the tension between pursuing individual security versus engaging in potentially more rewarding cooperation.

Important Games (3): Hawk-Dove (or Game of Chicken)

Scenario: Two players compete for a shared resource. Each player can choose to be aggressive (Hawk) or passive (Dove).
If both are passive, they share the resource. If both are aggressive, it results in a disastrous outcome. The aggressor wins if the other is passive.

	Dove	Hawk
Dove	3, 3	1, 5
Hawk	5, 1	0, 0

A Classification of Norma-Form Games can be found in Appendix 2 in slide 82.

Table of Contents

Do Nash Equilibria Always Exist?

Player 1/Player 2	Heads	Tails
Heads	-1, +1	+1, -1
Tails	+1, -1	-1, +1

- Matching Pennies Game:
  - Player 2 wins when the outcome matches.
  - Player 1 wins when the outcome differs.
- Zero-Sum Game Example:
  - A zero-sum game: one player's gain is the other player's loss.
  - Example: Allied landing in Europe on June 6, 1944.
- How Would You Play This Game?

What about Mixed Strategies?

- Players randomise their strategies using probabilities instead of deterministic actions.
- Let Player 2 commit to playing Heads with probability q.
- Expected payoffs for Player 1:
  - If Player 1 chooses Heads: $1 - 2q$
  - If Player 1 chooses Tails: $2q - 1$

Player 1/Player 2	Heads (q)	Tails (1-q)	Expected Payoffs
Heads	-1, +1	+1, -1	(1-2q)
Tails	+1, -1	-1, +1	(2q-1)

What if Player 1 also randomises their strategy by choosing Heads with probability p?

Nash Equilibrium for Mixed Strategies

- In Matching Pennies:
  - No pure strategy Nash equilibrium exists.
  - Nash equilibrium requires both players to randomise: $p = q = \frac{1}{2}$.

	Heads	Tails
Heads	-1, +1	+1, -1
Tails	+1, -1	-1, +1

Why?

Intuition Behind Mixed Strategies

- Mixed strategies make it difficult for opponents to predict the next move.
- In sports, players randomise actions to avoid predictability (e.g., penalty kicks in soccer).
- Nash Equilibrium is an equilibrium of beliefs:
  - Players believe the other is using a Nash equilibrium strategy.
  - This leads both players to adopt the same strategy, reinforcing the equilibrium.

Mixed Strategies Example 1: American Football

American Football

- Offense can either run with the ball or pass forward.

Offense/Defense	Defend Pass	Defend Run
Pass (p)	$0, 0$	$10, -10$
Run ($1-p$)	$5, -5$	$0, 0$

Table 1: Payoff Matrix (Offense, Defense)

- Suppose Offense passes with probability p (Defense with prob. q defending a pass).
- Defense is indifferent when $p = 1/3$ (calculation next slide).
- Equally, we can get $q = 2/3$, so $(1/3, 2/3)$ is a Nash equilibrium.
- Expected payoff to Offense: $10/3$ (yard gain).

American Football: Step-by-Step Calculation

Step 1: Defense's Expected Payoffs

- Defend Pass: $p0 + (1-p)(-5) = -5 + 5p$
- Defend Run: $p(-10) + (1-p)0 = -10p$

Step 2: Indifference Condition for Defense (also the condition for a mixed strategy NE, why?)

$$-5 + 5p = -10p$$

Solve for p: $\quad 15p = 5 \quad \Rightarrow \quad p = \frac{1}{3}$

Step 3: Offense's Expected Payoff (at $p = \frac{1}{3}$)

$$\text{Payoff}_{\text{Offense}} = \frac{10}{3}\,(\text{yard gain}).$$

Mixed Strategies Example 2: Penalty-Kick Game[9][10]

Penalty-Kick Game

- Real Data: economists analysed 1,417 penalty kicks from five years of professional soccer matches among European clubs.
- Success rates of penalty kickers based on decisions by the kicker and goalkeeper:

Kicker/Goal-keeper	Defend Left	Defend Right
Left	58%	95%
Right	93%	70%

Table 2: Left and right are from kicker's perspective.

[9] P-A Chiappori, Steven Levitt, and Timothy Groseclose. "Testing mixed-strategy equilibria when players are heterogeneous: The case of penalty kicks in soccer". In: *American Economic Review* 92.4 (2002), pp. 1138–1151.

[10] Ignacio Palacios-Huerta. "Professionals play minimax". In: *The Review of Economic Studies* 70.2 (2003), pp. 395–415.

Penalty-Kick Game (Continued)

Penalty-kick game

Kicker/Goal-keeper	Defend left	Defend right
Left	0.58, -0.58	0.95, -0.95
Right	0.93, -0.93	0.70, -0.70

- As a kicker, what is your strategy of kicking left or right?
- As a goalkeeper, what is your strategy of defending left or right?

Penalty Kick Game (Continued)

Kicker/Goal-keeper	Defend left(q)	Defend right($1-q$)
Left (p)	0.58, -0.58	0.95, -0.95
Right($1-p$)	0.93, -0.93	0.70, -0.70

- The kicker's strategy p depends on the goalkeeper's probability of defending left (q).
- Kicker is indifferent (enable nonzero p) when:

$$(0.58)(q) + (0.95)(1-q) = (0.93)(q) + (0.70)(1-q)$$

  Solving this gives $q = 0.42$ for the goalkeeper and similarly $p = 0.39$ for the kicker.
- This theoretical model closely matches real-world data.

Mixed Nash Equilibrium in Finite Strategic Games

Definition: Mixed Nash Equilibrium

In a finite strategic (normal-form) game $\Gamma = \langle \mathcal{I}, (S_i)_{i \in \mathcal{I}}, (u_i)_{i \in \mathcal{I}} \rangle$, a Mixed Nash Equilibrium is a profile of mixed strategies

$$(\sigma_1^*, \sigma_2^*, \ldots, \sigma_n^*),$$

where each σ_i^* is a probability distribution over S_i, such that:

$$u_i(\sigma_i^*, \sigma_{-i}^*) \geq u_i(\sigma_i, \sigma_{-i}^*) \quad \forall \sigma_i \in \Delta(S_i), \forall i \in \mathcal{I}.$$

- σ_i^*: The mixed strategy for player i, where $\sigma_i^*(s_i)$ is the probability assigned to pure strategy $s_i \in S_i$.
- $u_i(\sigma_i, \sigma_{-i})$: The expected payoff for player i under mixed strategy profile (σ_i, σ_{-i}).
- $\Delta(S_i)$: The set of all possible probability distributions over the strategy set S_i.

Table of Contents

The Existence of Nash Equilibrium (Nash, 1950)[11]

The Existence Theorem of Nash Equilibrium:

Every finite game (with a finite number of players and action profiles) has at least one Mixed Nash equilibrium.

[Nas50] proved the theorem directly using Kakutani's fixed-point theorem, whereas [Nas51] provided a better proof by proposing a function below:

Proof Outline

- Define a function $f : \{\Delta(S_i)\}_i \to \{\Delta(S_i)\}_i$ that adjusts strategy profiles.
- Show that f has a fixed point using Brouwer's Fixed-Point Theorem (for functions).
- Prove that fixed points of f correspond to Nash equilibria.

[11] John F Nash Jr. "Equilibrium points in n-person games". In: *Proceedings of the national academy of sciences* 36.1 (1950), pp. 48–49; John Nash. "Non-Cooperative Games". In: *Annals of Mathematics* 54.2 (1951), pp. 286–295.

Brouwer Fixed Point Theorem[12]

Brouwer Fixed Point Theorem:

Let $X \subset \mathbb{R}^n$ be a convex and compact set. If $T : X \to X$ is continuous, then there exists a fixed point $x^* \in X$ such that $T(x^*) = x^*$.

- Convex set: A set X is convex if, for any $x, y \in X$, the line segment joining x and y lies entirely in X.
- Compact set: A set X is compact if it is closed (i.e., includes its boundaries) and bounded (i.e., fits within a finite region).
- Continuous function: A function $T : X \to X$ is continuous if for any $\epsilon > 0$, there exists $\delta > 0$ such that $\|x - y\| < \delta$ implies $\|T(x) - T(y)\| < \epsilon$.

[12] Luitzen Egbertus Jan Brouwer. "Über abbildung von mannigfaltigkeiten". In: *Mathematische annalen* 71.1 (1911), pp. 97–115.

Visualisation of Fixed Point Theorem

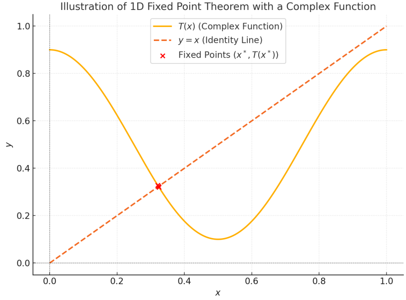


Figure 1: $T(x)$ and $y = x$. The red dots are fixed points where $T(x) = x$.

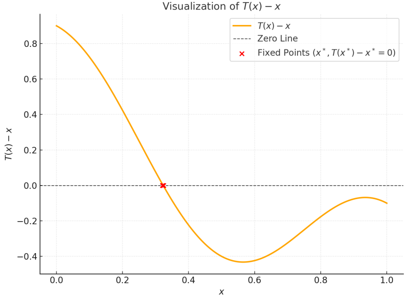


Figure 2: $T(x) - x$. The red dots mark zero crossings corresponding to fixed points.

Brouwer Fixed Point Theorem (1D Case)

Brouwer Fixed Point Theorem (1D Case):

Let $[a, b] \subset \mathbb{R}$ be a closed interval. If $T : [a, b] \to [a, b]$ is continuous, then there exists $x^* \in [a, b]$ such that $T(x^*) = x^*$.

Proof

1. If $T(a) = a$ or $T(b) = b$, then $x^* = a$ or $x^* = b$ is the fixed point.
2. Suppose $T(a) > a$ and $T(b) < b$. Define a function $g(x) = T(x) - x$.
3. Observe: $g(a) = T(a) - a > 0, \quad g(b) = T(b) - b < 0.$
4. Since g is continuous (as T is continuous), by the Intermediate Value Theorem[a], there exists $x^* \in (a, b)$ such that: $g(x^*) = 0$.
5. Therefore, $T(x^*) - x^* = 0$, implying $T(x^*) = x^*$. $\square$

[a] Intuitively, if a continuous function starts at $f(a)$ and ends at $f(b)$, it must take every value between $f(a)$ and $f(b)$ at least once.

Step 1: Constructing the Adjustment Function f[13]

Setup: Let $\{\sigma_i\}_i \in \{\Delta(S_i)\}_i$ be a strategy profile, and for each player $i \in N$, define a gain from a pure strategy a_i:

$$\phi_{i,a_i}(\sigma) = \max\{0, u_i(a_i, \sigma_{-i}) - u_i(\sigma)\},$$

where:

- $u_i(a_i, \sigma_{-i})$ is player i's utility when choosing a deterministic action $a_i \in S_i$ while others play σ_{-i}.
- $u_i(\sigma)$ is player i's utility under the profile s.

Adjustment Function:

$$f(\sigma) = \sigma', \quad \text{where} \quad \sigma'_i(a_i) = \frac{\sigma_i(a_i) + \phi_{i,a_i}(\sigma)}{1 + \sum_{b_i \in S_i} \phi_{i,b_i}(\sigma)}. \tag{4}$$

Interpretation: $f(\sigma)$ redistributes probability mass to increase the weight of better responses for each player.

[13] Albert Xin Jiang and Kevin Leyton-Brown. "A Tutorial on the Proof of the Existence of Nash Equilibria". In: *University of British Columbia Technical Report TR-2007-25. pdf* 14 (2009).

Step 2: Continuity of f and Existence of a Fixed Point

Claim: The function $f : \{\Delta(S_i)\}_i \to \{\Delta(S_i)\}_i$ is continuous.

Proof

- Each $\phi_{i,a_i}(\sigma)$ is continuous as it depends on the continuous utility function u_i.
- The numerator $\sigma_i(a_i) + \phi_{i,a_i}(\sigma)$ and denominator $1 + \sum_{b_i \in S_i} \phi_{i,b_i}(\sigma)$ are continuous, making $\sigma_i'(a_i)$ continuous.

$\square$

By Brouwer's Fixed-Point Theorem: Since $\{\Delta(S_i)\}_i$ is convex, compact, and $f : \{\Delta(S_i)\}_i \to \{\Delta(S_i)\}_i$ is continuous, f has at least one fixed point.

Step 3: Fixed Points of f are Nash Equilibria

Claim: A strategy profile σ is a Nash equilibrium if and only if σ is a fixed point of f (equivalency).

Proof: (i) Nash Equilibrium $\Rightarrow$ Fixed Point

If σ is a Nash equilibrium, no player can improve their utility by deviating. Thus:

$$\phi_{i,a_i}(\sigma) = 0, \quad \forall i \in N,\ a_i \in S_i.$$

Substituting $\phi_{i,a_i}(\sigma) = 0$ into $f(\sigma)$, we find $\sigma_i'(a_i) = \sigma_i(a_i)$. Hence, σ is a fixed point. $\square$

Step 3 (Continued): Fixed Points of f are Nash Equilibria

Proof: **(ii) Fixed Point $\Rightarrow$ Nash Equilibrium**

- Consider an arbitrary fixed point σ of f.
- By the linearity of expectation (weighted average), there exists at least one action a_i' in the support (i.e., $s(a_i') > 0$) of s such that: $u_i(\sigma) \geq u_i(a_i', \sigma_{-i})$[a]; it follows: $\phi_{i,a_i'}(\sigma) = 0$.
- Since σ is a fixed point of f, we have: $\sigma_i'(a_i') = \sigma_i(a_i')$.
- Consider the adjustment function, which defines $\sigma_i'(a_i')$. The numerator simplifies to $\sigma_i(a_i')$, which is positive because a_i' is in i's support. Hence, the denominator must equal 1 (as a fixed point).
- Consequently, for any i and $b_i \in S_i$, $\phi_{i,b_i}(\sigma) = 0$.
- This implies no player can improve their expected payoff by switching to a pure strategy.
- **Conclusion:** σ is a Nash equilibrium. $\square$

[a] The average is always between the min and max values of the terms.

The Second Way for Proof in Step 3[14]

In a fixed point, the adjustment function (Eq. (4)) can be rewritten as:

$$\sigma_i(a_i)\left(1+\sum_{b_i\in S_i}\phi_{i,b_i}(\sigma)\right)=\sigma_i(a_i)+\phi_{i,a_i}(\sigma)$$

Simplify and rearrange:

$$\sigma_i(a_i)\sum_{b_i\in S_i}\phi_{i,b_i}(\sigma)=\phi_{i,a_i}(\sigma).$$

We can argue that $\sum_{b_i\in S_i}\phi_{i,b_i}(\sigma)$ has to be zero (therefore NE) in the fixed point. Otherwise (if $\sum_{b_i\in S_i}\phi_{i,b_i}(\sigma)>0$)

1. $\sigma_i(a_i)=0 \Leftrightarrow \phi_{i,a_i}(\sigma)=0$ in NE already;
2. $\sigma_i(a_i)>0 \Leftrightarrow \phi_{i,a_i}(\sigma)>0$ contradicting the fixed-point condition (cannot improve $\sigma_i(a_i)$ by selecting $\sigma_i(a_i)$). $\square$

[14] Lecture 5 Note on Computational Game Theory, Yishay Mansour.

The Third Method for Proof in Step 3[15]

Summation over $a_i : \phi_{i,a_i}(\sigma) > 0$ from Eq. (4) gives:

$$\sum_{a_i:\phi_{i,a_i}(\sigma)>0} \sigma_i'(a_i) = \frac{\sum_{a_i:\phi_{i,a_i}(\sigma)>0} \sigma_i(a_i) + \sum_{a_i:\phi_{i,a_i}(\sigma)>0} \phi_{i,a_i}(\sigma)}{1 + \sum_{b_i \in S_i} \phi_{i,b_i}(\sigma)} \tag{5}$$

$$\geq \sum_{a_i:\phi_{i,a_i}(\sigma)>0} \sigma_i(a_i), \tag{6}$$

where the inequality holds because:

$$\frac{\sum_{a_i:\phi_{i,a_i}(\sigma)>0} \sigma_i(a_i) + \sum_{a_i:\phi_{i,a_i}(\sigma)>0} \phi_{i,a_i}(\sigma)}{1 + \sum_{b_i \in S_i} \phi_{i,b_i}(\sigma)} - \sum_{a_i:\phi_{i,a_i}(\sigma)>0} \sigma_i(a_i)$$

$$= \frac{\sum_{a_i:\phi_{i,a_i}(\sigma)>0} \phi_{i,a_i}(\sigma) - \left(\sum_{a_i:\phi_{i,a_i}(\sigma)>0} \sigma_i(a_i)\right) \sum_{b_i \in S_i} \phi_{i,b_i}(\sigma)}{1 + \sum_{b_i \in S_i} \phi_{i,b_i}(\sigma) \text{ (red is the same)}}$$

$$= \frac{\sum_{b_i \in S_i} \phi_{i,b_i}(\sigma) \left(1 - \sum_{a_i:\phi_{i,a_i}(\sigma)>0} \sigma_i(a_i)\right)}{1 + \sum_{b_i \in S_i} \phi_{i,b_i}(\sigma)} \geq 0 \quad \text{(green} > 0) \tag{7}$$

[15] https://cs.brown.edu/courses/cs295-8/nash.pdf

The Third Method for Proof in Step 3 (Continued)

- Eq. (7) indicates that equality holds only when $\sum_{b_i \in S_i} \phi_{i,b_i}(\sigma) = 0$, i.e., $\phi_{i,a_i}(\sigma) = 0$ for all $a_i \in S_i$.
- In the fixed point, we know that the above equality holds because every fixed point $\sigma' = \sigma$, namely:

$$\sum_{a_i : \phi_{i,a_i}(\sigma) > 0} \sigma_i'(a_i) = \sum_{a_i : \phi_{i,a_i}(\sigma) > 0} \sigma_i(a_i).$$

- As such $\phi_{i,a_i}(\sigma) = 0$ for all $a_i \in S_i$, implying $u_i(a_i, \sigma_{-i}) \leq u_i(\sigma)$ for all a_i. Thus, no player i has an incentive to deviate, satisfying the definition of a Nash equilibrium. $\square$

Best Responses as an Adjustment Function[16]

- A natural approach to strategy adjustment is the best response process.
- However, typically, the best response is not a well-defined function, as it may lack continuity (we will see next).
- To overcome this, redefine the objective for player i as:

$$\max_q u_i(q, p_{-i}) - ||p_i - q||^2,$$

  adding a penalty term $||p_i - q||^2$ to deviations.
- This ensures a continuous mapping.
- The updated best response for player i is q_i, forming the adjusted strategy profile:

$$f(p) = (q_1, q_2, \ldots, q_n).$$

- A detailed proof can be found in Appendix 3 (slide 88)

[16] https://www.cs.cornell.edu/courses/cs6840/2012sp/lecture39.pdf

Table of Contents

Best Response Correspondence

- The best response for player i to a strategy σ_{-i} (strategies of all other players) is:

$$B_i(\sigma_{-i}) = \arg\max_{\sigma_i} u_i(\sigma_i, \sigma_{-i}),$$

where $u_i(\sigma_i, \sigma_{-i})$ is player i's payoff.

- For a joint strategy σ, the joint best response correspondence is (it is not a function see slide 79 for the definition):

$$B(\sigma) = \prod_i B_i(\sigma_{-i}).$$

This maps a strategy profile σ to the set of all joint best responses: $B : \Sigma \mapsto 2^{\Sigma}$.

- A Nash equilibrium (NE) is a mixed strategy σ^* such that:

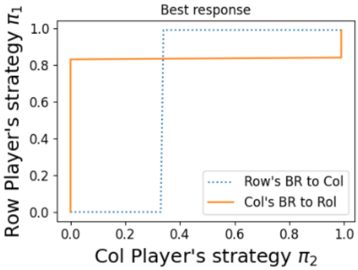
$$\sigma^* \in B(\sigma^*), \text{or alternatively } \sigma_i^* \in B_i(\sigma_{-i}^*) \quad \forall i.$$

- The existence of a NE is guaranteed by Kakutani's Fixed Point Theorem (see slide 80).

An Example: Best Response Correspondence

Payoff Matrix

ROW / COL	$L\,(\pi_2)$	$R\,(1-\pi_2)$
$U\,(\pi_1)$	$(2,1)$	$(0,0)$
$D\,(1-\pi_1)$	$(0,0)$	$(1,5)$



- Define π_1 as the probability of ROW player (player 1) playing strategy U.
- π_2 is the probability that COL player (player 2) plays L.
- Nash Equilibrium (NE):

$$\{(\pi_1, 1-\pi_1), (\pi_2, 1-\pi_2)\} = \{(5/6, 1/6), (1/3, 2/3)\}$$

An Example: Best Response Correspondence (Continued)

COL's Best Response Correspondence:

$$\pi_2 \in B_C(\pi_1) = \begin{cases} 0 & [0, \frac{5}{6}) \\ [0,1] & \frac{5}{6} \\ 1 & (\frac{5}{6}, 1] \end{cases}$$

ROW's Best Response Correspondence:

$$\pi_1 \in B_R(\pi_2) = \begin{cases} 0 & [0, \frac{1}{3}) \\ [0,1] & \frac{1}{3} \\ 1 & (\frac{1}{3}, 1] \end{cases}$$

- COL chooses L ($\pi_2 = 1$) when: $\pi_1 1 + (1-\pi_1)0 > \pi_1 0 + (1-\pi_1)5$ $\longrightarrow 1 \geq \pi_1 > \frac{5}{6}$
- Row chooses U ($\pi_1 = 1$) when: $\pi_2 2 + (1-\pi_2)0 > \pi_2 0 + (1-\pi_2)1$ $\longrightarrow 1 \geq \pi_2 > \frac{1}{3}$

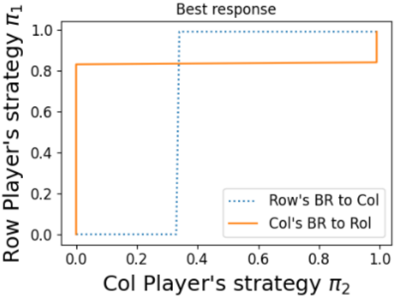


Table of Contents

Pareto Optimality

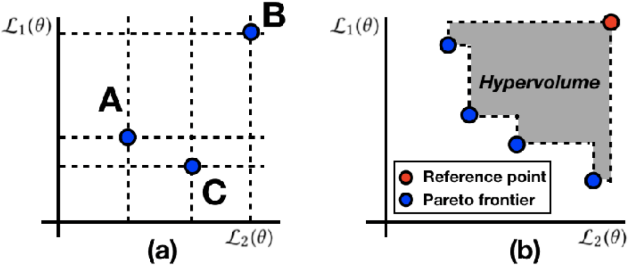
Even playing the best responses does not always lead to a good group outcome, e.g., Prisoner's dilemma.

Definition: Pareto Optimality

A choice of strategies is Pareto optimal if:

- No other choice of strategies gives all players a payoff at least as high,
- and at least one player achieves a strictly higher payoff.

In other words, no player can achieve a higher payoff without reducing others' payoffs. Everyone must do at least as well.



Social Optimality

Definition: Social Welfare Maximiser

A choice of strategies is a social welfare maximiser (or socially optimal) if it maximises the sum of all players' payoffs.

Example:

- The unique Nash equilibrium in this game is socially optimal.

	Exam	Presentation
Exam	98, 98	94, 96
Presentation	96, 94	92, 92

Prisoner's Dilemma: How Bad a NE is?

- Consider a cost-based game (years in prison as the cost):

Player 1/Player 2	Don't confess	Confess
Don't confess	1, 1	10, 0
Confess	0, 10	4, 4

- The worst Nash equilibrium occurs when both confess, with a cost of $(4+4)/2 = 4$.
- The highest social welfare occurs when both don't confess, giving $(1+1)/2 = 1$.
- We shall next study how to measure the efficiency of a system degrades due to selfish behaviour, e.g., Price of Anarchy (PoA) $= 4/1 = 4$.

Table of Contents

Conclusions and Next Steps

We have explored:

- The importance of strategic thinking in machine learning and AI.
- Nash equilibrium as a stable solution concept for multi-agent systems.
- The existence of Nash equilibrium in finite games.

Next, we will focus on:

- Investigating the learning process using best-response dynamics to reach Nash equilibrium.
- Evaluating how selfish behavior impacts system efficiency.

Table of Contents

References I

[20] "Modelling Bounded Rationality in Multi-Agent Interactions by Generalized Recursive Reasoning". In: *IJCAI* (2020).

[Alb91] James S Albus. "Outline for a theory of intelligence". In: *IEEE transactions on systems, man, and cybernetics* 21.3 (1991), pp. 473–509.

[BF15] David R Bellhouse and Nicolas Fillion. "Le her and other problems in probability discussed by Bernoulli, Montmort and Waldegrave". In: *Statistical Science* (2015), pp. 26–39.

[Bro11] Luitzen Egbertus Jan Brouwer. "Über abbildung von mannigfaltigkeiten". In: *Mathematische annalen* 71.1 (1911), pp. 97–115.

References II

[CLG02] P-A Chiappori, Steven Levitt, and Timothy Groseclose. "Testing mixed-strategy equilibria when players are heterogeneous: The case of penalty kicks in soccer". In: *American Economic Review* 92.4 (2002), pp. 1138–1151.

[FT91] Drew Fudenberg and Jean Tirole. "Game theory". In: *Cambridge, Massachusetts* 393.12 (1991), p. 80.

[HM23] Moritz Hardt and Celestine Mendler-Dünner. "Performative prediction: Past and future". In: *arXiv preprint arXiv:2310.16608* (2023).

[JL09] Albert Xin Jiang and Kevin Leyton-Brown. "A Tutorial on the Proof of the Existence of Nash Equilibria". In: *University of British Columbia Technical Report TR-2007-25. pdf* 14 (2009).

References III

[May82] John Maynard Smith. *Evolution and the Theory of Games*. Cambridge University Press, 1982. ISBN: 9780521288842.

[Nas50] John F Nash Jr. "Equilibrium points in n-person games". In: *Proceedings of the national academy of sciences* 36.1 (1950), pp. 48–49.

[Nas51] John Nash. "Non-Cooperative Games". In: *Annals of Mathematics* 54.2 (1951), pp. 286–295.

[Neu28] John von Neumann. "Zur Theorie der Gesellschaftsspiele". In: *Mathematische Annalen* 100 (1928), pp. 295–320. DOI: 10.1007/BF01448847.

[Nis+07] Noam Nisan et al., eds. *Algorithmic Game Theory*. Cambridge University Press, 2007. ISBN: 9780521872829.

References IV

[Pal03] Ignacio Palacios-Huerta. "Professionals play minimax". In: *The Review of Economic Studies* 70.2 (2003), pp. 395–415.

[Per+20] Juan Perdomo et al. "Performative prediction". In: *International Conference on Machine Learning*. PMLR. 2020, pp. 7599–7609.

[PG28] Ivan Petrovitch Pavlov and W Gantt. "Lectures on conditioned reflexes: Twenty-five years of objective study of the higher nervous activity (behaviour) of animals.". In: (1928).

Table of Contents

A.1 Background: Correspondence

Definition

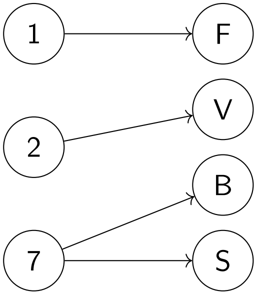
A correspondence between sets A and A' is any subset of the Cartesian product $A \times A'$, allowing:

- One-to-many relationships.
- Many-to-many relationships.

Example

Examples of correspondences:

- (1:00 p.m., Football)
- (2:00 p.m., Volleyball)
- (7:00 p.m., Soccer)
- (7:00 p.m., Basketball)



For the discussion on best response correspondence, see slide 64.

Background: Kakutani's Fixed Point Theorem

Sufficient Conditions for a Fixed Point

A correspondence $B : \Sigma \mapsto \Sigma$ has a fixed point if:

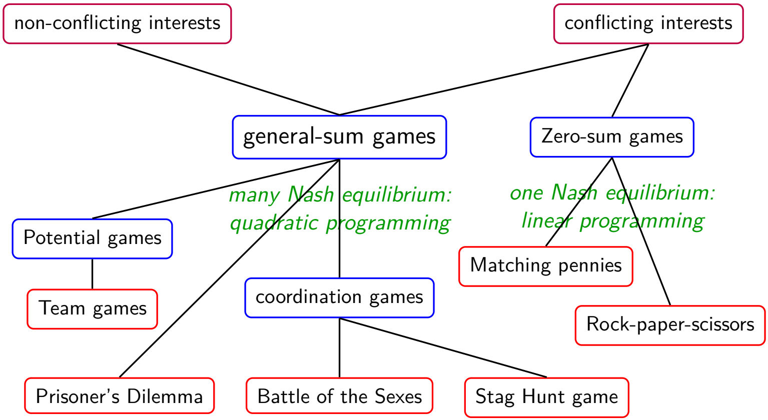
1. Σ is a compact, convex, nonempty subset of a finite-dimensional Euclidean space.
2. $B(\sigma) \neq \emptyset$ for all σ.
3. $B(\sigma)$ is convex for all σ.
4. $B(\sigma)$ has a closed graph:

$$\text{If } x_n \to x, y_n \to y, y_n \in B(x_n), \text{ then } y \in B(x).$$

This has been utilised in the proof of NE existence in slide 64.

Table of Contents

A.2 Classification of Normal-Form Games



Normal-Form Games (1/4)

- **Prisoner's Dilemma**
  A classic example of cooperation vs. defection dilemmas.

$$\text{Payoff Matrix} = \begin{bmatrix} (R, R) & (S, T) \\ (T, S) & (P, P) \end{bmatrix}$$

  where $T > R > P > S$, illustrating incentive misalignment.

- **Matching Pennies**
  Models adversarial settings, useful in studying competitive learning.

$$\text{Payoff Matrix} = \begin{bmatrix} (1, -1) & (-1, 1) \\ (-1, 1) & (1, -1) \end{bmatrix}$$

  Mixed Strategy Nash Equilibrium due to opposing objectives.

Normal-Form Games (2/4)

- **Coordination Game**
  Illustrates multi-agent consensus and coordination challenges.

  $$\text{Payoff Matrix} = \begin{bmatrix} (a, a) & (0, 0) \\ (0, 0) & (b, b) \end{bmatrix}$$

  where $a, b > 0$; multiple equilibria emphasise equilibrium selection issues.

- **Battle of the Sexes**
  Highlights asymmetric preferences in agent collaboration.

  $$\text{Payoff Matrix} = \begin{bmatrix} (2, 1) & (0, 0) \\ (0, 0) & (1, 2) \end{bmatrix}$$

  with two *Pure Strategy Nash Equilibria* that reflect coordination trade-offs.

Normal-Form Games (3/4)

- **Rock-Paper-Scissors**
  Used to model cyclic dominance in competitive environments.

$$\text{Payoff Matrix} = \begin{bmatrix} (0,0) & (-1,1) & (1,-1) \\ (1,-1) & (0,0) & (-1,1) \\ (-1,1) & (1,-1) & (0,0) \end{bmatrix}$$

  No Pure Strategy Nash Equilibrium, emphasising the need for randomised strategies.

- **Team Coordination Game**
  Models scenarios where agents must coordinate to achieve a common goal.

$$\text{Payoff Matrix} = \begin{bmatrix} (3,3) & (0,0) \\ (0,0) & (2,2) \end{bmatrix}$$

  Multiple equilibria highlight the importance of strategy alignment in team settings.

Normal-Form Games (4/4)

- **Stag Hunt**
  Demonstrates the tension between individual security and rewarding cooperation.

$$\text{Payoff Matrix} = \begin{bmatrix} (4,4) & (0,3) \\ (3,0) & (3,3) \end{bmatrix}$$

  Cooperation leads to higher payoffs, but requires mutual trust.

- **Hawk-Dove (Chicken) Game**
  Models conflict over shared resources with strategies of aggression and passivity.

$$\text{Payoff Matrix} = \begin{bmatrix} (3,3) & (1,5) \\ (5,1) & (0,0) \end{bmatrix}$$

  Highlights the risks of mutual aggression and the benefits of strategic compromise.

Table of Contents

Setup

Setting:

- We have n players. Each player i has a strategy set $\Delta(S_i)$ (the simplex over pure strategies S_i).
- Let $p = (p_1, \ldots, p_n)$ denote a profile of mixed strategies.
- Each player i has a payoff function $u_i(q_i, p_{-i})$.

Goal: Show that the following best-response map

$$f_i(p) \; = \; \arg \max_{q_i \in \Delta(S_i)} \left\{ u_i(q_i, p_{-i}) - \|q_i - p_i\|^2 \right\}$$

has a fixed point, which implies the existence of a Nash equilibrium.

Step 1: Uniqueness of the Best Response

Lemma (Uniqueness). For each player i, the maximiser

$$\max_{q_i \in \Delta(S_i)} \left\{ u_i(q_i, p_{-i}) - \|q_i - p_i\|^2 \right\}$$

is unique.

Proof Outline:

1. $u_i(q_i, p_{-i})$ is linear in q_i.
2. $-\|q_i - p_i\|^2$ is *strictly concave* in q_i.
3. The sum of a linear and a strictly concave function is *strictly concave*.
4. A strictly concave function on a convex compact set has a *unique* global maximiser.

Hence each $f_i(p)$ is well-defined and single-valued.

Step 2: Continuity of the Best-Response Map

Claim. If the optimum of a parametrised, strictly concave optimisation problem is unique, then the optimal solution depends *continuously* on the parameters.

Why?

- This is a standard result in convex/concave analysis.
- Uniqueness prevents "jumps" in the maximiser.
- Hence each f_i is continuous in (p_{-i}, p_i).

Therefore the combined best-response map

$$f = (f_1, f_2, \ldots, f_n) : \prod_{i=1}^{n} \Delta(S_i) \;\to\; \prod_{i=1}^{n} \Delta(S_i)$$

is *continuous*.

Step 3 (Part 1): Fixed Point Implies NE

Lemma. If p is a fixed point of f, i.e. $p_i = f_i(p)$ for all i, then p is a *Nash equilibrium (NE)*.

Restate the NE Definition:

- A profile $p = (p_1, \dots, p_n)$ is a *Nash equilibrium* if

$$\forall i,\ \forall q_i \in \Delta(S_i), \quad u_i(p_i,\, p_{-i}) \ \geq\ u_i(q_i,\, p_{-i}).$$

- We must show: if $p_i = f_i(p)$ (i.e. each p_i maximises $u_i(\cdot, p_{-i}) - \| \cdot - p_i \|^2$), then no player has a profitable deviation.

Proof Outline (Contradiction):

1. Assume p is *not* a NE. Then $\exists i$, $\exists q_i$ s.t. $u_i(q_i, p_{-i}) > u_i(p_i, p_{-i})$.
2. We'll show this contradicts the uniqueness of p_i as the maximiser of $\phi_i(q) = u_i(q, p_{-i}) - \|q - p_i\|^2$.

Step 3 (Part 2): Objective ϕ_i & Small Deviation

Best-Response Objective:

$$\phi_i(q) \;=\; u_i(q,\, p_{-i}) \;-\; \|\, q - p_i\|^2.$$

If p is a fixed point, then

$$p_i \;=\; f_i(p) \;=\; \arg\max_{q \in \Delta(S_i)} \phi_i(q).$$

Since the maximiser is *unique*, for any other $q \neq p_i$, we should have $\phi_i(q) \leq \phi_i(p_i)$.

Contradiction Setup:

- We assumed $\exists q_i$ with $u_i(q_i, p_{-i}) > u_i(p_i, p_{-i})$.
- Even though q_i improves *payoff*, it might not maximise ϕ_i because of the penalty $\|q - p_i\|^2$.

Construct a Small Deviation:

$$r_i(\varepsilon) \;=\; (1-\varepsilon)\, p_i \;+\; \varepsilon\, q_i, \quad 0 < \varepsilon \ll 1.$$

We will compare $\phi_i(r_i(\varepsilon))$ to $\phi_i(p_i)$.

Step 3 (Part 3): Analysis of $\delta_i(\varepsilon)$

Compare:

$\delta_i(\varepsilon) \;=\; \phi_i(r_i(\varepsilon)) \;-\; \phi_i(p_i); \phi_i(q) = u_i(q, p_{-i}) - \|q - p_i\|^2.$

Expansion:

$$u_i(r_i(\varepsilon)) = (1-\varepsilon)\, u_i(p_i) + \varepsilon\, u_i(q_i), \quad \|\, r_i(\varepsilon) - p_i\|^2 = \varepsilon^2 \,\|\, q_i - p_i\|^2.$$

Hence

$$\delta_i(\varepsilon) = \underbrace{(1-\varepsilon)\, u_i(p_i) + \varepsilon\, u_i(q_i) - u_i(p_i)}_{\text{linear part}} \;-\; \underbrace{\varepsilon^2 \|\, q_i - p_i\|^2}_{\text{quadratic penalty}} \tag{8}$$

$$= \varepsilon\,[u_i(q_i) - u_i(p_i)] \;-\; \varepsilon^2 \|q_i - p_i\|^2. \tag{9}$$

Dominance of Linear Term for Small ε:

- Since $u_i(q_i) > u_i(p_i)$, for sufficiently small ε,

$$\varepsilon[u_i(q_i) - u_i(p_i)] \;>\; \varepsilon^2 \|q_i - p_i\|^2.$$

- Thus $\delta_i(\varepsilon) > 0$. So $\phi_i(r_i(\varepsilon)) > \phi_i(p_i)$.

Step 3 (Part 4): Conclusion of the Contradiction

Contradiction and Conclusion:

- We found a strategy $r_i(\varepsilon)$ that *increases* ϕ_i beyond $\phi_i(p_i)$.
- This violates the assumption that p_i was the unique maximiser of ϕ_i.
- Therefore, our assumption that p is not a NE is false.

Hence: If p is a fixed point ($p_i = f_i(p)$), then p *must* be a Nash equilibrium.

Final Step: By Brouwer's theorem, f has a fixed point. Combining with this result shows a *Nash equilibrium exists*.
Go back to the main slide (slide 62).