Lecture 02: Potential Games and Best Response Dynamics

Jun Wang
UCL

Outline of MAAI

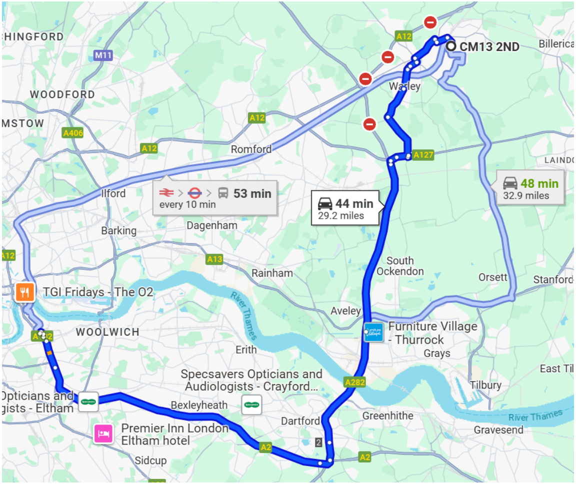
1. Basic Game Theory and Nash Equilibrium
2. Potential Games and Best Respond Dynamics
3. Repeated Games and the Theory of Cooperation
4. Minimax Games
5. General-Sum Games and L-H Algorithms
6. Single-agent Reinforcement Learning and MDP
7. Learning Stochastic Games
8. Non-regret Learning and Correlated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

- Subject to change

Outline of This Lecture

Table of Contents

Traffic Routing



Congestion Games: Multiple agents (e.g., Google Route Planners) collectively seek optimal routes in transportation networks to minimise congestion.

Motivation for Potential Games in Multi-Agent AI

Understanding Potential Games:

- A potential game is one where the incentive of all players to change their strategy can be captured by a single global potential function.
- Ensure the existence of pure Nash equilibria, facilitating predictable and stable outcomes in multi-agent systems.

Applications:

- **Traffic Routing:** Optimising paths in transportation networks to reduce congestion.
- **Network Bandwidth Allocation:** Distributing resources in communication networks efficiently.
- **Load Balancing:** Equitably distributing tasks across servers in computing environments.

Table of Contents

Cournot Model of Duopoly: Theory of Competition[1]

- The Cournot model, formulated by French economist Antoine Augustin Cournot in 1838, serves as a fundamental framework in industrial organisation and microeconomics.
- It examines how two firms in a duopoly (two firms dominating a market), compete by independently deciding the quantities they produce.
- The single market price is influenced by the total production from both firms.
  - Each firm seeks to maximise its profit, assuming the other firm's output remains unchanged.
- This strategic interplay results in a Nash equilibrium, where neither firm can increase its profit by altering its production unilaterally.

[1] Antoine Augustin Cournot. *Recherches sur les Principes Mathématiques de la Théorie des Richesses*. Hachette, 1838.

Model Assumptions

As an example, let us consider two Internet Service Providers (ISPs) competing to provide bandwidth in a shared market.

- Players: Two ISPs competing to provide bandwidth.
- Actions: Each ISP determines its output, q_1 and q_2.
- Costs: Cost for ISP i is $C_i(q_i)$, where C_i is increasing. A specific case assumes $C_i(q_i) = cq_i$, with $c \geq 0$.
- Market Demand:

$$P(Q) = \begin{cases} a - Q, & \text{if } Q \leq a, \\ 0, & \text{if } Q > a, \end{cases}$$

where $Q = q_1 + q_2$ and $a > 0$.

Formal Game Definition

Definition of the Game

The Cournot duopoly is modelled as a strategic game:

$$\Gamma = (N, S, u), \text{ where:}$$

- $N = \{1, 2\}$: The set of players (the two ISPs).
- $q = (q_1, q_2)$: The strategy space, with $q_i = [0, \infty)$ representing feasible outputs for ISP i.
- $u = (u_1, u_2)$: The payoff functions, where $u_i : S \to \mathbb{R}$ gives the profit of ISP i, i.e.,

$$u_i(q_1, q_2) = q_i P(q_1 + q_2) - C_i(q_i).$$

  Substituting the linear cost function and demand:

$$u_i(q_1, q_2) = q_i(a - c - q_1 - q_2).$$

Nash Equilibrium

Definition:

A Nash equilibrium is a strategy profile $(q_1^*, q_2^*) \in S$ such that:

$$u_1(q_1^*, q_2^*) \geq u_1(q_1, q_2^*) \quad \forall q_1 \in S_1,$$

and

$$u_2(q_1^*, q_2^*) \geq u_2(q_1^*, q_2) \quad \forall q_2 \in S_2.$$

The first-order conditions: To maximise profit, each ISP solves:

$$\frac{\partial u_1}{\partial q_1} = a - c - 2q_1 - q_2 = 0,$$

and

$$\frac{\partial u_2}{\partial q_2} = a - c - 2q_2 - q_1 = 0.$$

Solving the Equilibrium

Best-response functions:

$$q_1^* = \frac{1}{2}(a - c - q_2), \quad q_2^* = \frac{1}{2}(a - c - q_1).$$

Solving the best-response functions simultaneously gives the equilibrium:

$$q_1^* = q_2^* = \frac{a - c}{3}.$$

Market outcomes:

$$\begin{aligned} \text{Total Output:} \quad & Q = q_1^* + q_2^* = \frac{2(a-c)}{3}, \\ \text{Market-Clearing Price:} \quad & P(Q) = a - Q = \frac{a + 2c}{3}. \end{aligned}$$

Equilibrium Payoffs

Profit calculation: Substituting equilibrium quantities:

$$u_1^* = u_2^* = q_1^*(P(Q) - c).$$

Simplifying:

$$u_1^* = u_2^* = \frac{(a-c)^2}{9}.$$

Comparison to Monopoly (single firm):

- Monopoly output: $q_m = \frac{a-c}{2}$.
- Monopoly profit: $u_m = \frac{(a-c)^2}{4}$.
- Total duopoly profit: $u_1^* + u_2^* = \frac{2(a-c)^2}{9}$.

  Duopoly profits are lower due to competition.

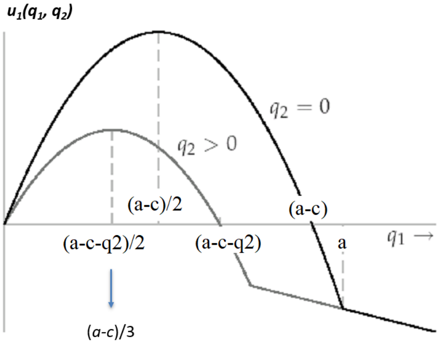
Equilibrium Payoffs: Visualisation

- Each ISP wants to be a monopolist (the black curve).

$$q_2 = 0; R_1(q_2) = \frac{1}{2}(a - q_2 - c) \implies q_1 = \frac{(a-c)}{2}.$$

- In equilibrium, the output of ISP 1 reduces to:

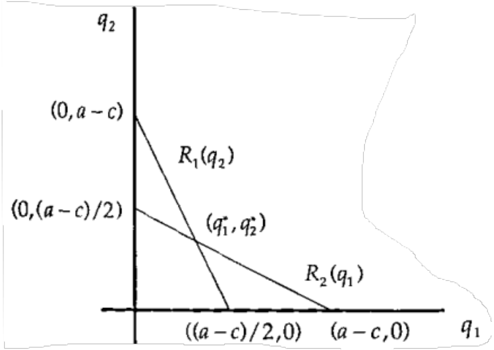
$$q_1 = \frac{(a-c)}{3} \quad \text{(the grey curve).}$$



Duopoly profits are lower due to competition.

Solution by Best Response

- ISP 1's best response: $R_1(q_2) = \frac{1}{2}(a - q_2 - c)$
- ISP 2's best response: $R_2(q_1) = \frac{1}{2}(a - q_1 - c)$



These two best response functions intersect only once at the equilibrium point (q_1^*, q_2^*).

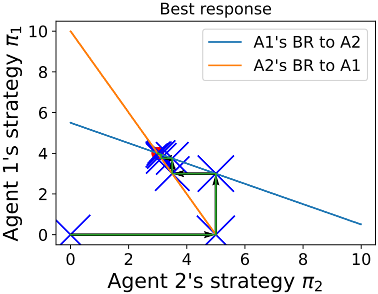
Solution as a Result of Learning

- ISP 1's best response:

$$R_1(q_2) = \frac{1}{2}(a - q_2 - c)$$

- ISP 2's best response:

$$R_2(q_1) = \frac{1}{2}(a - q_1 - c)$$



Learning-type adjustment processes:

- Starting from a random solution, firms react simultaneously to the opponent's most recent output:

$$q^t \equiv (q_1^t, q_2^t) = (R_1(q_2^{t-1}), R_2(q_1^{t-1})) \equiv f(q^{t-1}).$$

- Alternatively, firms take turns best-responding or respond to the average of past plays.

Discussion about Adjustment Processes

Measurements

- Asymptotically stable: If it converges to a particular steady state (a single Nash equilibrium) for all initial states close to it.
- Globally stable: If it converges from every starting state.

- Simultaneous Adjustment: ISPs update quantities simultaneously, converging to an equilibrium under specific conditions. $\longrightarrow$ best response dynamics.
- Fictitious Play: A more plausible solution that ISPs base actions on historical averages of opponents' choices, potentially leading an equilibrium. $\longrightarrow$ non-regret learning.
- But not always converge: May exhibit a circling effect.

Limitation: Ignores how the current action influences the future.

- Repeated Interaction: Cooperation may emerge through repeated interactions if credible punishment strategies are established. $\longrightarrow$ repeated game.

Problem Statement

Payoff Matrix for a Two-Player, Three-Action Game:

$$\begin{bmatrix} (0,0) & (5,4) & (4,5) \\ (4,5) & (0,0) & (5,4) \\ (5,4) & (4,5) & (0,0) \end{bmatrix}$$

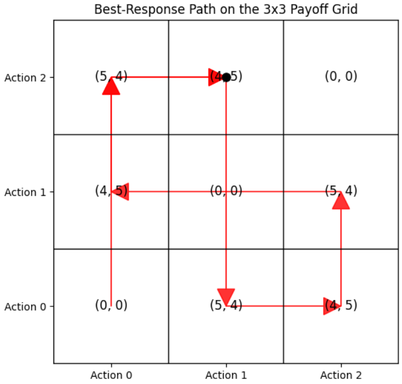
Objective:

- Implement a best-response process in Python for this game.
- Simulate the evolution of strategies in the pure strategy space.
- Analyse the outcomes from different initial joint actions.

Problem Statement

Payoff Matrix for a Two-Player, Three-Action Game:

$$\begin{bmatrix} (0,0) & (5,4) & (4,5) \\ (4,5) & (0,0) & (5,4) \\ (5,4) & (4,5) & (0,0) \end{bmatrix}$$



best-response process.

Table of Contents

Recall: Cournot Competition Game Setup

The Game Setup

- I: Number of firms.
- Each firm i chooses a production quantity $q_i \in (0, \infty)$.
- Total quantity: $Q = q_1 + q_2 + \cdots + q_I$.
- Profit (payoff) of firm i:

$$u_i(q_i, q_{-i}) = q_i(P(Q) - c),$$

where $P(Q)$ is the price function, c is the marginal cost, and q_{-i} are the quantities of other firms.

Potential Function in Cournot Competition

Potential Function:

- Define the function:

$$\Phi(q_1, \ldots, q_I) = \prod_{i=1}^{I} u_i(q_i, q_{-i}) = \prod_{i=1}^{I} q_i(P(Q) - c),$$

where $Q = \sum_{i=1}^{I} q_i$.

Note that for all i firms and all $q_i > 0$, we have:

$$u_i(q_i, q_{-i}) - u_i(q_i', q_{-i}) > 0 \iff \Phi(q_i, q_{-i}) - \Phi(q_i', q_{-i}) > 0.$$

Proof?

Ordinal Potential Games

- A game is an ordinal potential game if there exists a function Φ such that:

$$u_i(s_i, s_{-i}) - u_i(s_i', s_{-i}) > 0 \iff \Phi(s_i, s_{-i}) - \Phi(s_i', s_{-i}) > 0.$$

- In the Cournot competition, Φ serves as an ordinal potential function.
- The set of pure strategy Nash equilibria of the original game corresponds to that of the game where profits are given by Φ.

Exact Potential Games

- A strategic form game $G = \langle I, S_i, u_i \rangle$ is an exact potential game if there exists a function $\Phi : S \to \mathbb{R}$ such that $\Phi(s)$ reflects the exact changes in $u_i(s)$:

$$u_i(x, s_{-i}) - u_i(z, s_{-i}) = \Phi(x, s_{-i}) - \Phi(z, s_{-i}),$$

  for all i, s_{-i}, $x, z \in S_i$.

Note: The potential function has a natural analogy to potential energy in physical systems (e.g., how forces act in a system). It will be useful both for locating pure strategy Nash equilibria and also for the analysis of myopic (short sighted) dynamics.

Example of Potential Games: Prisoner's Dilemma

- Game matrix:

Player 1/Player 2	Cooperate	Defect
Cooperate	(3, 3)	(0, 5)
Defect	(5, 0)	(1, 1)

- Corresponding potential function matrix (A potential function assigns a real value for every strategy profile):

Player 1/Player 2	Cooperate	Defect
Cooperate	4	6
Defect	6	7

Pure Strategy Nash Equilibria in Ordinal Potential Games

Theorem: Every finite ordinal potential game has at least one pure strategy Nash equilibrium.

Proof

- The global maximum of an **ordinal potential function** is a pure strategy Nash equilibrium.
- Suppose s^* corresponds to the global maximum. Then, for any $i \in I$, we have: $\Phi(s_i^*, s_{-i}^*) - \Phi(s_i, s_{-i}^*) \geq 0 \quad$ for all $s_i \in S_i$.
- Since Φ is a potential function, for all i and all $s_i \in S_i$:

$$u_i(s_i^*, s_{-i}^*) - u_i(s_i, s_{-i}^*) \geq 0 \iff \Phi(s_i^*, s_{-i}^*) - \Phi(s_i, s_{-i}^*) \geq 0.$$

- Thus: $u_i(s_i^*, s_{-i}^*) - u_i(s_i, s_{-i}^*) \geq 0 \quad$ for all $s_i \in S_i$ & all $i \in I$.
- Hence, s^* is a pure strategy Nash equilibrium.

Note: There may also be other pure strategy Nash equilibria corresponding to local maxima of Φ.

Example: Cournot Competition Again

- Suppose $P(Q) = a - bQ$ and costs $c_i(q_i)$ are arbitrary.
- Define the function:

$$\Phi^*(q_1, \ldots, q_I) = a \sum_{i=1}^{I} q_i - b \sum_{i=1}^{I} q_i^2 - b \sum_{i} \sum_{i \neq j} q_i q_j - \sum_{i=1}^{I} c_i(q_i).$$

- It can be shown that for all i and all q_{-i}:

$$u_i(q_i, q_{-i}) - u_i(q_i', q_{-i}) = \Phi^*(q_i, q_{-i}) - \Phi^*(q_i', q_{-i}),$$

for all $q_i, q_i' > 0$.

Simple Dynamics in Finite Ordinal Potential Games

Definition

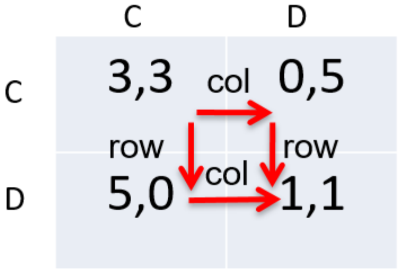
- A path in strategy space S is a sequence of strategy vectors $(s_0, s_1, \ldots)$ such that every two consecutive strategies differ in one coordinate (i.e., exactly in one player's strategy).
- An improvement path is a path $(s_0, s_1, \ldots)$ such that:

$$u_{i_k}(s^k) < u_{i_k}(s^{k+1}),$$

where s^k and s^{k+1} differ in the i_k-th coordinate, i.e., the payoff improves for the player who changes their strategy.

Prisoner's Dilemma

An improvement path can be thought of as generated dynamically by myopic players.



Simple Dynamics in Finite Ordinal Potential Games

Proposition: In every finite ordinal potential game, every improvement path is finite.

Proof

- Suppose $(s_0, s_1, \dots)$ is an improvement path. Therefore, we have:

$$\Phi(s^0) < \Phi(s^1) < \cdots ,$$

  where Φ is the ordinal potential.
- Since the game is finite, i.e., it has a finite strategy space, the potential function takes on finitely many values, and the above sequence must end in finitely many steps.

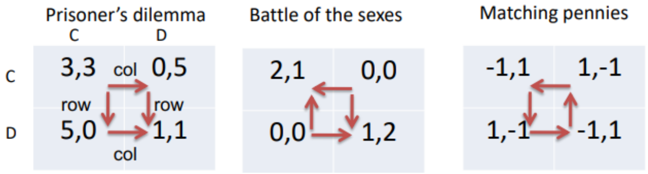
- This implies that in finite ordinal potential games, every maximal improvement path terminates in an equilibrium point.
- The *simple myopic learning process* converges to equilibrium.

Characterisation of Finite Exact Potential Games[2]

- For a finite path $\gamma = (s^0, \dots, s^N)$, let:

$$I(\gamma) = \sum_{i=1}^{N} u^{m_i}(s^i) - u^{m_i}(s^{i-1}),$$

  where m_i denotes the player changing its strategy in the i-th step of the path.
- The path $\gamma = (s^0, \dots, s^N)$ is **closed** if $s^0 = s^N$. It is a **simple closed path** if $s^l \neq s^k$ for every $0 \leq k < l \leq N-1$.



[2] Dov Monderer and Lloyd S Shapley. "Potential games". In: *Games and economic behavior* 14.1 (1996), pp. 124–143.

Characterisation of Finite Exact Potential Games[3]

Theorem:

A game G is an exact potential game if and only if for all finite simple closed paths γ, $I(\gamma) = 0$. Moreover, it is sufficient to check simple closed paths of length 4.

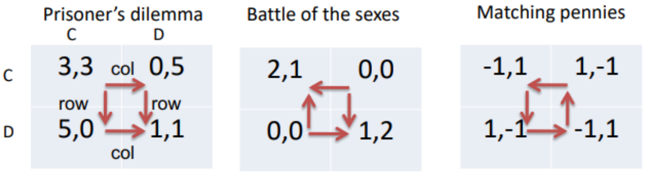


Figure 2: True for the first two, but not the last one.

- Arrows are edges connecting each possible pure strategy profile with a directed graph.
- Directed edge (u, v) means v (which differs from u only in the strategy of a single player, i) is a (strictly) better action for i, given the strategies of the other players.

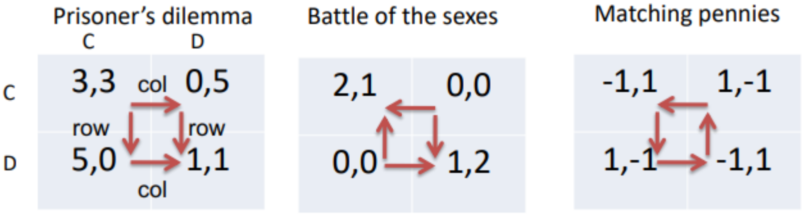
[3] Dov Monderer and Lloyd S Shapley. "Potential games". In: *Games and economic behavior* 14.1 (1996), pp. 124–143.

Characterisation of Finite Exact Potential Games

Lemma: A game is a potential game if and only if local improvements always terminate.

Intuition:

- A potential function exists if and only if the graph does not contain cycles:
  - If cycles exist, no potential function (e.g., (a, b, c, a) means $f(a) < f(b) < f(c) < f(a)$) $\rightarrow$ not a function.
  - If no cycles exist, can easily define an ordinal potential function. **Why?**



Best-response Dynamics

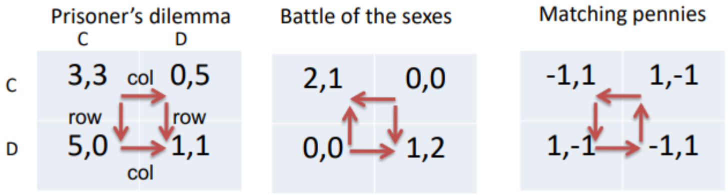
Routine for Best-response Dynamics

while $s = (s_i, s_{-i})$ is not a PNE do:

1. Pick an arbitrary agent i.
2. Identify a beneficial deviation s_i' from s_i for agent i.
3. Update the outcome to (s_i', s_{-i}).

end

Theorem: In a Finite Potential Game, any iterative best response process will terminate and eventually converge to a Pure Strategy Nash Equilibrium (PSNE). **Examples:**
Direction of local improvement, indicated by arrows.



Best-response Dynamics

- Convergence Guaranteed: Dynamics always reach Nash Equilibrium.
- Cost Function: Convergence occurs with any cost functions, not necessarily monotonic.
- Initial State Irrelevance: Starting action profile does not affect convergence.
- Agent Selection Independence: Any agent can be chosen to update their strategy without impacting the outcome.
- Better Response Suffices: Agents need only to choose strategies that improve their current payoffs, not necessarily the best one.

Infinite Potential Games[4]

Proposition:

- Let G be a continuous potential game with compact strategy sets. Then G has at least one Pure Strategy Nash equilibrium (PSNE).

Proposition:

- Let G be a game such that $S_i \subseteq \mathbb{R}$ and the payoff functions $u_i : S \to \mathbb{R}$ are continuously differentiable.
- Let $\Phi : S \to \mathbb{R}$ be a function. Then, Φ is a potential for G if and only if Φ is continuously differentiable and:

$$\frac{\partial u_i(s)}{\partial s_i} = \frac{\partial \Phi(s)}{\partial s_i}, \quad \text{for all } i \in I \text{ and all } s \in S.$$

[4] David H Mguni et al. "Learning in Nonzero-Sum Stochastic Games with Potentials". In: *ICML*. 2021.

Q Learning with Potentials[5]

- Multiple stages potential games:

$$R_i(s,(a^i,a^{-i})) - R_i(s',(a^i,a^{-i})) = \phi(s,(a^i,a^{-i})) - \phi(s',(a^i,a^{-i})).$$

- Bellman optimality backup:

$$[T_\Phi F](s) := \sup_{a\in\mathcal{A}} \left[g(s,a) + \gamma \int_{s'\in\mathcal{S}} ds' P(s';a,s)F[s']\right].$$

Multiagent learning converted into a single update:

$$\lim_{k\to\infty} T_\Phi^k V^\pi = \sup_{\pi\in\Pi} V^\pi.$$

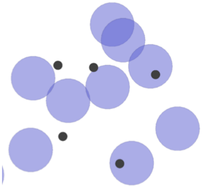


Figure 3: Coordination Navigation: selfish agents (purple) seek to reach rewards (black) whilst minimising contact with each other.

[5] David H Mguni et al. "Learning in Nonzero-Sum Stochastic Games with Potentials". In: *ICML*. 2021.

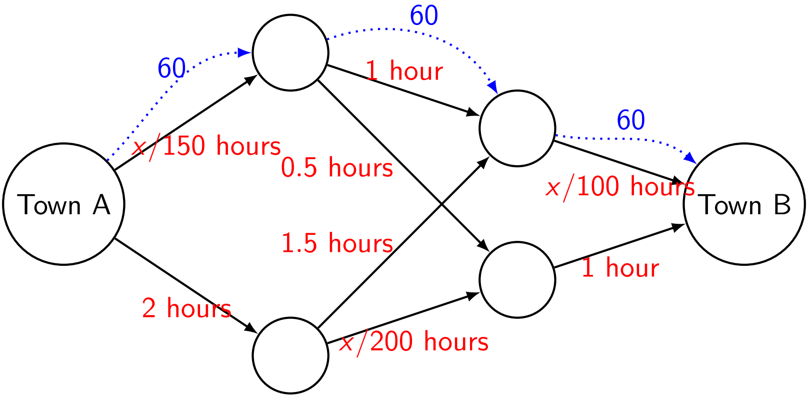
Table of Contents

Network Congestion Games

Definition:

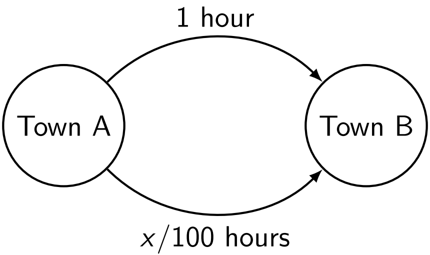
- A directed graph $G = (V, E)$ consists of vertices (nodes) and edges.
- Each edge $e \in E(G)$ has a delay function f_e.
- We have n users (agents); the strategy of user i is to choose a path A_j from a source node s_i to a destination node t_i.
- The delay of a path is the sum of delays of edges on the path.
- Each user wants to minimise their own delay by choosing the best path.

An Example



- Traffic splits dynamically, with drivers minimising travel time.
- Each edge has a delay function (red) based on traffic load x.
- A sample path (dotted blue) shows 60 agents (load) using it.

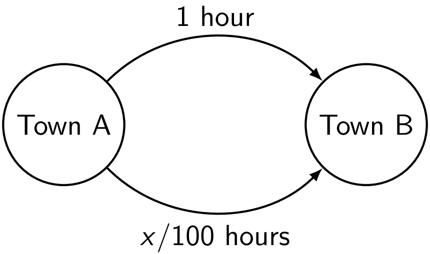
Traffic Routing (Pigou's example)[6]



- 100 units of traffic (representing 100 drivers but generalisable to a real number) travel from town A to town B.
- Each driver aims to minimise their travel time.
- In unbalanced traffic, drivers on the most loaded path have incentive to switch paths.

[6] Arthur Pigou. *The economics of welfare*. Routledge, 1920.

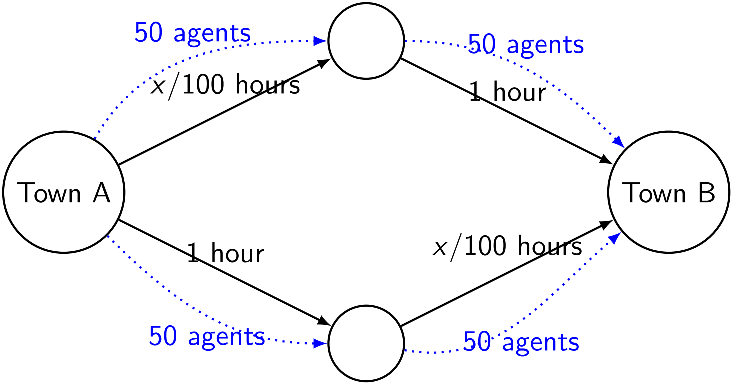
Traffic Routing (Pigou's example)[7]



- If both paths have 50, average delay is 0.75 hours.
- In a Nash equilibrium, everyone goes bottom and the average delay is 1 hour.
- NE leads to slower travel times.

[7] Arthur Pigou. *The economics of welfare*. Routledge, 1920.

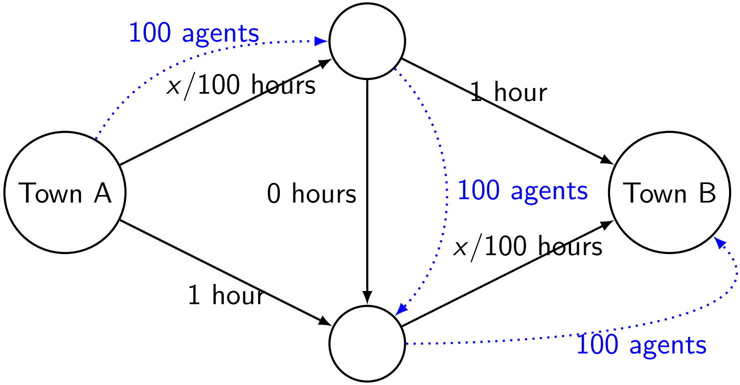
Traffic Routing: Braess's Paradox[8]



Delay is 1.5 hours for everybody at the unique Nash equilibrium.

[8] Dietrich Braess. "Über ein Paradoxon aus der Verkehrsplanung". In: *Unternehmensforschung* 12 (1968), pp. 258–268.

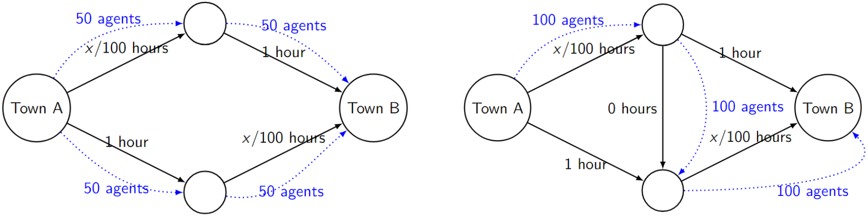
Traffic Routing: Braess's Paradox



- A 0-delay superhighway is built between the middle nodes.
- Always optimal to follow the zig-zag path, regardless of others.

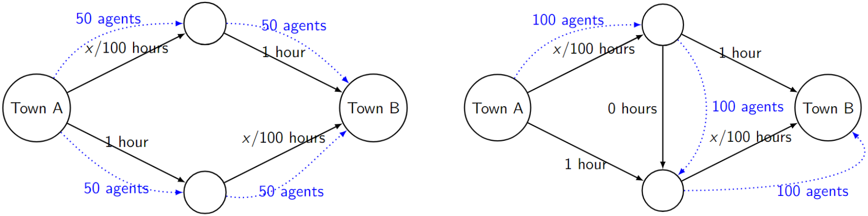
The new road increases delay to 2 hours for all at Nash equilibrium.

Traffic Routing: Braess's Paradox



- Adding a fast road to a network isn't always beneficial.
- In the left network, a traffic pattern exists where all players have a delay of 1.5 hours, whereas on the right 2 hours.
- Question: How does the Nash Equilibrium compare to the optimal solution?

Price of Anarchy[9]



- The Price of Anarchy (PoA) quantifies efficiency loss due to selfish agent behaviour in the worst case.
- Defined as the ratio of the worst Nash equilibrium's average travel time to the minimum possible average travel time.
    - Here, PoA $= \frac{2}{1.5} = \frac{4}{3}$.

[9] Tim Roughgarden. *Twenty lectures on algorithmic game theory*. Cambridge University Press, 2016.

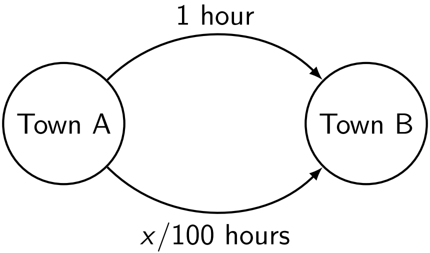
Price of Anarchy: Prisoner's Dilemma

- 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 $(1+1)/2 = 1$.
    - PoA $= 4/1 = 4$.

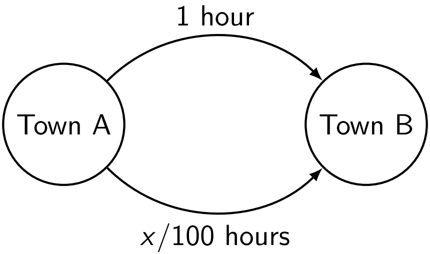
Price of Anarchy: Pigou's Example[10]



- In equilibrium, the average travel time is 1 hour when 100 drivers use the lower road.
    - Question: What is the PoA here?

[10] Arthur Pigou. *The economics of welfare*. Routledge, 1920.

Price of Anarchy: Pigou's Example

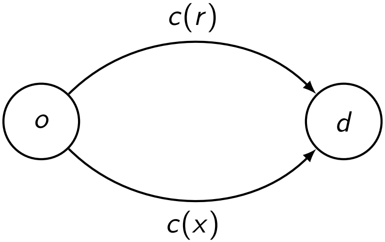


- a drivers use the upper road; $100 - a$ use the lower road.
- Average travel time:

$$\frac{a}{100}1 + \frac{100-a}{100}\frac{100-a}{100}$$

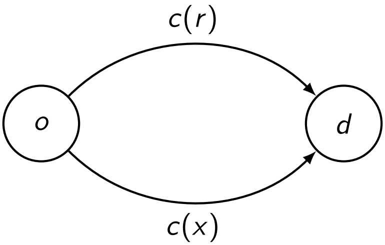
- Optimal occurs when $a = 50$, giving the average time of $3/4$.
  - $\text{PoA} = \frac{1}{3/4} = \frac{4}{3}$.

Pigou-like Network: A General Setting



- Two vertices: origin o and destination d, connected by an *upper* and *lower* edge.
- A total traffic rate $r \in \mathbb{R}^+$ flows from o to d.
- The lower edge has cost $c(x) \in C$ (traffic-dependent), where $x \in [0, r]$; and the upper edge has constant cost $c(r)$.

Pigou Bound

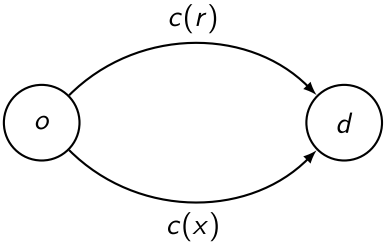


- By construction, the lower edge is a dominant strategy for all traffic r, as it has no higher cost than the constant cost $c(r)$.
- At equilibrium, all traffic flows on the lower edge, with a total travel time of $rc(r)$.
- Nonetheless, the minimum total travel time is:

$$\inf_{0 \le x \le r} \{xc(x) + (r - x)c(r)\},$$

where the infimum represents the greatest lower bound.

Pigou Bound



The PoA, given c and r, is: PoA $= \sup_{0 \leq x \leq r} \frac{rc(r)}{xc(x)+(r-x)c(r)}$.

- The Pigou bound for C, an arbitrary set of nonnegative, continuous, and nondecreasing cost functions:

$$\alpha(C) = \sup_{c \in C, 0 \leq x \leq r} \frac{rc(r)}{xc(x) + (r - x)c(r)}.$$

- Q1: Prove that if C is the set of cost functions $c(x) = ax + b$, $a, b \geq 0$, then $\alpha(C) = 4/3$.
- Q2: What is $\alpha(C)$ when $c(x) = x^p$, p is large, and $r = 1$?

Note: supremum looks at the least upper-bound.

Answer for Question 1

Prove that if C is the set of cost functions of the form $c(x) = ax + b$ with $a, b \geq 0$, then $\alpha(C) = 4/3$.

- To minimise:

$$xc(x) + (r - x)c(r)$$

Expanding:

$$x(ax + b) + (r - x)(ar + b) = ax^2 + ar^2 + br - arx.$$

- Differentiating with respect to x and setting to 0:

$$2ax - ar = 0 \quad \Rightarrow \quad x = r/2.$$

- Substituting $x = r/2$:

$$\alpha(C) = \sup_{c \in C} \frac{rc(r)}{\left(\frac{r}{2}\right) c\left(\frac{r}{2}\right) + \left(\frac{r}{2}\right) c(r)} = \sup_{c \in C} \frac{ar + b}{\frac{3}{4}ar + b} = 4/3,$$

where $b = 0$ gives the upper bound.

PoA is independent of the network topology[11]

Theorem:

For every set C of cost functions and every selfish routing network with cost function in C, the PoA is at most $\alpha(C)$.

Description	**Typical Form**	**Price of Anarchy**
Linear	$ax + b$	$4/3$
Quadratic	$ax^2 + bx + c$	$3\sqrt{3}/2 \approx 1.6$
Cubic	$ax^3 + bx^2 + cx + d$	$4\sqrt{4}/3 \approx 1.9$
Quartic	$ax^4 + \dots$	$5\sqrt{5}/4 \approx 2.2$
Poly. of degree $\leq d$	$\sum_{i=0}^{d} a_i x^i$	$\frac{d+1}{d+1\sqrt{d+1-d}} \approx \frac{d}{\ln d}$

[11] Tim Roughgarden. *Twenty lectures on algorithmic game theory.* Cambridge University Press, 2016.

Congestion Games: Refined Definition

$C = \langle I, M, (S_i)_{i \in I}, (c_j)_{j \in M} \rangle$, where:

- $I = \{1, 2, \ldots, I\}$: the set of players.
- $M = \{1, 2, \ldots, m\}$: the set of resources.
- S_i: the set of resource combinations (e.g., links or shared resources) that player i can use. A strategy for player i is $s_i \in S_i$, the subset of resources this player is using.
- Utilities:

$$u_i(s_i, s_{-i}) = \sum_{j \in s_i} c^j(k_j),$$

where $c^j(k)$ is the cost (negative of the reward) to each user using resource j, if k_j users are using it. k_j is the number of users of resource j under the joint strategy s.

Every Congestion Game is a Potential Game[12]

Theorem:

Every congestion game is a potential game and thus has a pure strategy Nash equilibrium.

Proof (Step 1):

- For each resource j, define:

$$k_j^i = \sum_{i' \neq i} \mathbb{I}[j \in s_{i'}],$$

  where k_j^i is the number of players excluding i using j, and $\mathbb{I}[j \in s_{i'}]$ is the indicator for the event $j \in s_{i'}$.

- The utility difference for player i between two strategies s_i, s_i':

$$u_i(s_i, s_{-i}) - u_i(s_i', s_{-i}) = \sum_{j \in s_i} c^j(k_j^i + 1) - \sum_{j \in s_i'} c^j(k_j^i + 1).$$

[12] Robert W Rosenthal. "A class of games possessing pure-strategy Nash equilibria". In: *International Journal of Game Theory* 2 (1973), pp. 65–67.

Every Congestion Game is a Potential Game

Proof (Step 2):

- Define the potential function:

$$\Phi(s) = \sum_{j \in \bigcup_{i \in I} s_i} \sum_{k=1}^{k_j} c^j(k),$$

where:

- j is a resource (it is from a set of resources that has been used by any $i \in I$); i is a player.
- s_i: the set of resources used by player i (their strategy).
- $c^j(k)$: the utility of j when used by k players.
- k_j: the number of players using j.

Every Congestion Game is a Potential Game

Proof (Step 3):

- We can write the potential function as:

$$\Phi(s_i, s_{-i}) = \sum_{j \in \bigcup_{i' \neq i} s_{i'}} \left[\sum_{k=1}^{k_j} c^j(k) \right] + \sum_{j \in s_i} c^j(k_j^i + 1),$$

where:

  - The first term enumerates resources j used by all other players $i' \neq i$,
  - If any resource j has been additionally used by player i, the second term adds this additional cost.

Every congestion game is a potential game

Proof (Step 4):

- It is easy to see that:

$$\Phi(s_i, s_{-i}) - \Phi(s_i', s_{-i}) = \sum_{j \in \bigcup_{i' \neq i} s_{i'}} \left[\sum_{k=1}^{k_j} c^j(k) \right] + \sum_{j \in s_i} c^j(k_j^i + 1)$$

$$- \sum_{j \in \bigcup_{i' \neq i} s_{i'}} \left[\sum_{k=1}^{k_j} c^j(k) \right] - \sum_{j \in s_i'} c^j(k_j^i + 1).$$

  Simplifying:

$$\Phi(s_i, s_{-i}) - \Phi(s_i', s_{-i}) = \sum_{j \in s_i} c^j(k_j^i + 1) - \sum_{j \in s_i'} c^j(k_j^i + 1).$$

- This is equivalent to:

$$u_i(s_i, s_{-i}) - u_i(s_i', s_{-i}).$$

Thus, every finite congestion game has a pure strategy equilibrium due to the property of potential games.

Conclusions

Potential and Congestion Games:

- Provide a robust framework for modeling strategic interactions in multi-agent systems.
- Agents select resources, with each resource's cost depending on its congestion (number of agents choosing it).
- Applications: Traffic networks, data routing, and load balancing.

What's Next?

- Understanding the nature of cooperation, improving PoA.
- Examining situations where we need to repeatedly play a game.

Table of Contents

References I

[Bra68] Dietrich Braess. "Über ein Paradoxon aus der Verkehrsplanung". In: *Unternehmensforschung* 12 (1968), pp. 258–268.

[Cou38] Antoine Augustin Cournot. *Recherches sur les Principes Mathématiques de la Théorie des Richesses*. Hachette, 1838.

[Mgu+21] David H Mguni et al. "Learning in Nonzero-Sum Stochastic Games with Potentials". In: *ICML*. 2021.

[MS96] Dov Monderer and Lloyd S Shapley. "Potential games". In: *Games and economic behavior* 14.1 (1996), pp. 124–143.

[Pig20] Arthur Pigou. *The economics of welfare*. Routledge, 1920.

References II

[Ros73] Robert W Rosenthal. "A class of games possessing pure-strategy Nash equilibria". In: *International Journal of Game Theory* 2 (1973), pp. 65–67.

[Rou16] Tim Roughgarden. *Twenty lectures on algorithmic game theory*. Cambridge University Press, 2016.