Basics of Game Theory

Daniel Khashabi

Fall 2015
Last Update: November 12, 2015

1 Introduction

The file contains some example implementations to demonstrate the implementation of ideas in practice [1].

2 Basics

Mixed strategy: The average payoff of player i is:

$$U_i(\mu_i, \mu_{-i}) = \sum_{a_1 \in A_1} \sum_{a_2 \in A_2} \dots \sum_{a_N \in A_N} u_i(a_1, a_2, \dots, a_N) \mu_1(a_1) \dots \mu_N(a_N)$$

Nash equilibrium point: $\mu^* = (\mu_1^*, \dots, \mu_N^*)$ is a equilibrium point, iff:

$$U_i(\mu_i, \mu_{-i}^*) \leq U_i(\mu_i^*, \mu_{-i}^*)$$

Best response of player:

$$BR_i(\mu_{-i}) \in \arg\max_{\mu_i} U_i(\mu_i, \mu_{-i})$$

Alternate view of NE: It can be thought of as, fixed points of the best responses:

$$\begin{bmatrix} \mu_1 \\ \vdots \\ \mu_N \end{bmatrix} \in \begin{bmatrix} BR_1(\mu_{-1}^*) \\ \vdots \\ BR_N(\mu_{-N}^*) \end{bmatrix}$$

Example 1. *Consider the following partnership game:*

$p1 \setminus p2$	Work hard	Be lazy
Work hard	$(10, 10)$	$(-5, 5)$
Be lazy	$(-5, 5)$	$(0, 0)$

[1] Codes available at `https://github.com/danyaljj/gameTheory`

- *(W, W) and (L, L) are pure-strategy NEs.*

To find mixed strategy NEs we use the best responses.

- *Best response for player 1:* $\max_{x\in[0,1]} 10xy - 5x(1-y) + 5(1-x)y = \max_{x\in[0,1]} 5x(2y-1) + 5y$, *which can be simplified as*

$$\begin{cases} x^* \in [0,1] & y = 1/2 \\ x^* = 1 & y > 1/2 \\ x^* = 0 & y < 1/2 \end{cases}$$

- *Best response for player 2:* $\max_{y\in[0,1]} 10xy + 5x(1-y) - 5(1-x)y = \max_{y\in[0,1]} 5y(2x-1) + 5x$, *which similar to the previous result:*

$$\begin{cases} y^* \in [0,1] & x = 1/2 \\ y^* = 1 & x > 1/2 \\ y^* = 0 & x < 1/2 \end{cases}$$

This results in a mixed-strategy NE of $(x^*, y^*) = (1/2, 1/2)$, *with corresponding payoff* $5/2, 5/2$.

Theorem 1 (Nash)**.** *For any game, a NE exists.*

The proof is using fixed-point theorems.

Here are some theorems and definitions we need to know outside the scope of Game Theory.

Definition 1 (Closed graph)**.** *f has closed graph if for any sequence $\{x_n\} \in C$ and $\{y_n\}$, s.t.*

- $x_n \to x$ *as* $n \to \infty$.
- $y_n \in f(x_n)$ $\forall y_n \to y$ *as* $n \to \infty$.

we have $y \in f(x)$.

Theorem 2 (Brouwer)**.** *Let $C \subseteq \mathbb{R}^n$ be a convex, closed, bounded set. If the function $f : C \to C$ be continuous, then f has a fixed point in C (i.e. $\exists x \in C, \quad s.t. \quad f(x) = x$).*

Theorem 3 (Kukutani)**.** *If C is a closed, bounded and convex subset of $\mathbb{R}^n$. Consider f to be a mapping from each element of C to a subset of C (i.e. $f : C \to 2^C$). Suppose $f(x) \neq \emptyset, \quad \forall x$. f has a closed graph, and $f(x)$ is a convex set, $\forall x$. Then f has a fixed-point in C.*

With Brouwer theorem. Suppose $(\mu_1, \ldots, \mu_N)$ are a set of strategies. If instead of these strategies, $\square$

With Kukutani theorem. $\square$

3 Zero-sum games

In general the payoff can be represented as $U_1(x,y) = \sum_{i,j} A(i,j)x_i y_j = x^\top Ay$, and specifically for zero-sum games we have $U_1(x,y) = -U_2(x,y)$.

For zero-sum games a mixed-strategy NE (x^*, y^*) is a saddle-point, since it satisfies:

$$\begin{cases} x^{*\top} Ay^* \geq x^\top Ay^* \\ x^{*\top} Ay^* \leq x^{*\top} Ay \end{cases}$$

Minimax Theorem (von Neumann):

$$\min_y \max_x x^\top Ay = \max_x \min_y x^\top Ay$$

and the solution to this problem could be found via either of the following LPs:

$$LP1 : \begin{cases} \max_{x,v_1} v_1 \\ \text{s.t. } v_1 \leq (x^\top A)_j, \quad \forall j \\ x \geq 0 \\ 1^\top x = 1 \end{cases} \qquad LP2 : \begin{cases} \min_{y,v_2} v_2 \\ \text{s.t. } v_2 \geq (Ay)_i, \quad \forall i \\ y \geq 0 \\ 1^\top y = 1 \end{cases}$$

These two LPs are duals of each other.

Example 2. *We can easily write numerical programs*

Symmetric zero-sum games: A game is called symmetric zero-sum if the payoff matrix is skew-symmetric, i.e. $A = -A^\top$ (or $a_{ij} = -a_{ji}$, and a_{ii} for all i). The average pay-off (value of the game) in such games is zero. Consider distributions x and y over actions of the first and second players; the value of the game is:

$$x^\top Ay = y^\top A^\top x = -y^\top Ax$$

And in the special $x = y$, $x^\top Ax = -x^\top Ax$, which implies that $x^\top Ax = 0$. In addition, we know that

$$\min_y x^\top Ay \leq x^\top Ay \leq \max_x x^\top Ay$$

This holds for the special case $x = y$ as well:

$$\begin{cases} \min_y x^\top Ay \leq x^\top Ax = 0 \\ \max_x x^\top Ay \geq y^\top Ay = 0 \end{cases}$$

Based on von Neumann's minimax theorem we know that:

$$\min_y \max_x x^\top Ay = \max_x \min_y x^\top Ay = 0$$

which would imply that the value of this game is always zero. In addition, $x = y$ for any value of x results in a zero value which clearly is a solution to this problem. Therefore if x is a saddle-point strategy for player 1, x is also a saddle-point mixed strategy for player 2 as well.

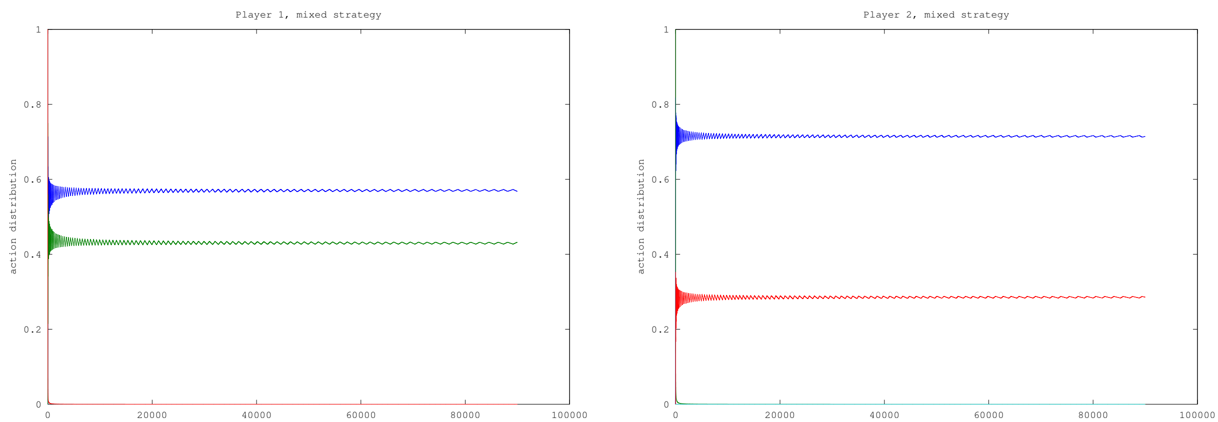


Figure 1: Convergence of the empirical mixed strategy to its optimal point, with primal-dual iterations.

Example 3. *Instead of solving the LP programs directly, we can have an iterative primal-dual method to approximate the objective. Essentially at each iteration each agent plays its pure-strategy, and we replace x and y with their empirical definition (the number of repetitions for a pure-strategy over the total number of actions). The pure action is chosen based on other player's empirical mixed strategy. For example player 2 choses the action j with the smallest of $(x^\top A)_j$. The outputs of such strategy for an example game are shown in Figure 1.*

4 Continuous action spaces

The assumption is that each action is continuous $a_i \in A_i \subseteq \mathbb{R}^n$. In this setting each action has its own constrained space of possible actions. Similarly, the actions spaces can be coupled. In other words, the stacked vector for all actions is in $\Omega \subseteq \mathbb{R}^{n_1+\ldots+n_N}$.

Theorem 4. *In the case uncoupled continuous actions $\Omega = A_1 \times \ldots \times A_N$. Assuming that each action space A_i is closed and bounded, and each payoff $u_i(a_i, a_{-i})$ is continuous on Ω, then there exists a mixed-strategy NE.*

Idea behind the proof: Discretize the action space; by Nash's theorem, there exist a mixed strategy NE for the discretized game. The rest is showing NE converges to the NE of the continuous space as the size of the discretization becomes smaller.

Theorem 5 (Rosen)**.** *Let Ω be a coupled constrained set, convex, closed and bounded set. Further, if the payoff $u_i(a_i, a_{-i})$ is concave in a_i, for any a_{-i}. Then there exist a pure-strategy NE.*

5 Optimal Auctions

The theorem is due to Myerson [?]. Here is the scenario:

- A seller is selling one item to one of the N buyers.
- The buyer's valuations are i.i.d. A buyer i has valuation distributed with pdf f_i.
- Each bidder selects a bid $b_i \in \mathcal{B}_i$.
- Allocation rule: The probability that the bidder i gets the object is $\Pi_i(b_i, b_{-i})$ ($\Pi_i \geq 0, \forall i$ and $\sum_i \Pi_i = 1$).
- Payment rule: If the bidder i gets the object, it pays $q_i(b_i, b_{-i})$. The valuations are in the bounded range $v_i \in [0, \theta_{i,\max}]$.
- Mechanism: A set of rules announced by the seller.

Problem: The mechanism design problem for the seller is to select the space $\mathcal{B}$, the allocation function $\{\Pi_i\}$ and the payment rule $\{q_i\}$, to maximize the expected revenue of the seller.

Define the following notations:

- $\alpha_i(\theta_i) = \mathbb{P}\left(\text{bidder } i \text{ gets the object}|\theta_i\right) = \mathbb{E}_{\theta_{-i}}\left[\Pi_i(\theta_i, \theta_{-i})\right]$
- $m_i(\theta_i) = \mathbb{E}\left[\text{ Payment of the bidder } i|\theta_i\right] = \mathbb{E}_{\theta_{-i}}\left[q_i(\theta_i, \theta_{-i})\Pi_i(\theta_i, \theta_{-i})\right]$

Revelation principle: each bidder truthfully reveal their bids.
Incentive Compatibility (IC): Each bidder will bid their true valuation (in other words, it is not beneficial for each user to lie about their valuation). The payoff for bidder is $\theta_i\alpha_i(\theta_i) - m_i(\theta_i)$. If the bidder reports its valuation as $\hat{\theta}_i$, IC entails:

$$\theta_i\alpha_i(\theta_i) - m_i(\theta_i) \geq \theta_i\alpha_i(\tilde{\theta}_i) - m_i(\tilde{\theta}_i), \quad \forall \theta, \tilde{\theta}, i$$

Individual Rationality (IR): Says that the bidders will participate voluntarily:

$$\theta_i\alpha_i(\theta_i) - m_i(\theta_i) \geq 0 \quad \forall i, \theta_i$$

Lemma 1. *IC is equivalent to the following two:*

$$m_i(\theta_i) = m_i(0) + \theta_i\alpha_i(\theta_i) - \int_0^{\theta_i} \alpha_i(\theta)d\theta \tag{1}$$

$$\alpha_i \textit{ is a non-decreasing function} \tag{2}$$

Proof. We prove each direction separately. First lets prove that IC entails the Equation 1 and Equation 2.

Based on IC the utility function $\theta_i\alpha_i(\tilde{\theta}_i) - m_i(\tilde{\theta}_i)$, where $\tilde{\theta}_i$ is the reported valuation, reaches its maximum when $\tilde{\theta}_i = \theta_i$. Therefore

$$\frac{\partial}{\partial\tilde{\theta}_i}\left[\theta_i\alpha_i(\tilde{\theta}_i) - m_i(\tilde{\theta}_i)\right]\Bigg|_{\tilde{\theta}_i=\theta_i} = 0$$

$$\begin{aligned}\Rightarrow \theta_i \alpha_i'(\theta_i) - m_i'(\theta_i) = 0 \Rightarrow m_i(\theta_i) &= m_i(0) + \int_0^{\theta_i} z\alpha_i'(z)dz \\ &= m_i(0) + \left[z\alpha_i(z)|_0^{\theta_i} - \int_0^{\theta_i} \alpha_i(z)dz \right] \\ &= m_i(0) + \theta_i \alpha_i(\theta_i) - \int_0^{\theta_i} \alpha_i(z)dz\end{aligned}$$

Which finishes the proof of "IC $\Rightarrow$ Equation 1".

Now we prove Equation 2. We know, for two arbitrary valuations $\tilde{\theta}$ and $\hat{\theta}$:

$$\tilde{\theta}_i \alpha_i(\tilde{\theta}_i) - m_i(\tilde{\theta}_i) \geq \tilde{\theta} \alpha_i(\hat{\theta}_i) - m_i(\hat{\theta}_i)$$
$$\hat{\theta}_i \alpha_i(\hat{\theta}_i) - m_i(\hat{\theta}_i) \geq \hat{\theta} \alpha_i(\tilde{\theta}_i) - m_i(\tilde{\theta}_i)$$

Adding these two we would get

$$\left(\tilde{\theta} - \hat{\theta}_i\right) \alpha_i(\tilde{\theta}_i) \geq \left(\tilde{\theta} - \hat{\theta}_i\right) \alpha_i(\hat{\theta}_i) \Rightarrow \left(\tilde{\theta} - \hat{\theta}_i\right) \left(\alpha_i(\tilde{\theta}_i) - \alpha_i(\hat{\theta}_i)\right) \geq 0$$

This α_i is non-decreasing.

Now we prove the other direction: "Equation 1 and Equation 2 $\Rightarrow$ IC". We want to prove that:

$$\theta_i \alpha_i(\theta_i) - m_i(\theta_i) - (\theta_i \alpha_i(\tilde{\theta}_i) - m_i(\tilde{\theta}_i)) \geq 0$$

We simplify the LHS:

$$\theta_i \alpha_i(\theta_i) - m_i(\theta_i) - (\theta_i \alpha_i(\tilde{\theta}_i) - m_i(\tilde{\theta}_i)) = \theta_i \alpha_i(\theta_i) - m_i(\theta_i) - (\tilde{\theta}_i \alpha_i(\tilde{\theta}_i) - m_i(\tilde{\theta}_i) + (\theta_i - \tilde{\theta}_i)\alpha_i(\tilde{\theta}_i)) \geq 0$$

Also we know:

$$\begin{cases} m_i(\theta_i) = m_i(0) + \theta_i \alpha_i(\theta_i) - \int_0^{\theta_i} \alpha_i(\theta)d\theta \Rightarrow \int_0^{\theta_i} \alpha_i(\theta)d\theta - m_i(0) = \theta_i \alpha_i(\theta_i) - m_i(\theta_i) \\ m_i(\tilde{\theta}_i) = m_i(0) + \tilde{\theta}_i \alpha_i(\tilde{\theta}) - \int_0^{\tilde{\theta}} \alpha_i(\theta)d\theta \Rightarrow \int_0^{\tilde{\theta}_i} \alpha_i(\theta)d\theta - m_i(0) = \tilde{\theta}_i \alpha_i(\tilde{\theta}_i) - m_i(\tilde{\theta}_i) \end{cases}$$

Therefore:

$$\theta_i \alpha_i(\theta_i) - m_i(\theta_i) - (\theta_i \alpha_i(\tilde{\theta}_i) - m_i(\tilde{\theta}_i)) = \int_0^{\theta_i} \alpha_i(\theta)d\theta - \int_0^{\tilde{\theta}_i} \alpha_i(\theta)d\theta - (\theta_i - \tilde{\theta}_i)\alpha_i(\tilde{\theta}_i)$$

which is essentially greater or equal to zero, size α_i is non-decreasing.

□

Lemma 2. *IR and IC are equivalent to Equation 1, Equation 2, and*

$$m_i(0) \leq 0 \tag{3}$$

Proof. We know that IR means

$$\theta_i \alpha_i(\theta_i) - m_i(\theta_i) \geq 0 \quad \forall i, \theta_i$$

We use the result of the previous lemma, and rewrite the LHS of above:

$$\int_0^{\theta_i} \alpha_i(\theta)d\theta - m_i(0) \geq 0$$

If we replace $\theta_i = 0$, we would get the desired result that $m_i(0) \leq 0$. □

The goal of the optimal auction is to maximize the seller's revenue. In other words, we desire to solve:

$$\begin{cases} \max_{\{\Pi_i\},\{q_i\}} \sum_{i=1}^{N} \mathbb{E}\left(m_i(\theta_i)\right) \\ \text{s.t. Equation 1, Equation2, Equation3} \end{cases} \tag{4}$$

We expand the objective using Equation 1:

$$\begin{aligned}
\mathbb{E}\left(m_i(\theta_i)\right) &= m_i(0) + \mathbb{E}\left[\theta_i \alpha_i(\theta_i)\right] - \mathbb{E}\left[\int_0^{\theta_i} \alpha_i(\theta) d\theta\right] \\
&= m_i(0) + \int_0^{\theta_{i,\max}} \theta_i \alpha_i(\theta_i) f_i(\theta_i) d\theta_i - \int_0^{\theta_{i,\max}} \left[\int_0^{\theta_i} \alpha_i(\theta) d\theta\right] f_i(\theta_i) d\theta_i \\
&= m_i(0) + \int_0^{\theta_{i,\max}} \theta_i \alpha_i(\theta_i) f_i(\theta_i) d\theta_i - \int_0^{\theta_{i,\max}} \left[\int_\theta^{\theta_{i,\max}} f_i(\theta_i) d\theta_i\right] \alpha_i(\theta) d\theta \\
&= m_i(0) + \int_0^{\theta_{i,\max}} \theta_i \alpha_i(\theta_i) f_i(\theta_i) d\theta_i - \int_0^{\theta_{i,\max}} \left(1 - F_i(\theta_i)\right) \alpha_i(\theta) d\theta \\
&= m_i(0) + \int_0^{\theta_{i,\max}} \alpha_i(\theta_i) f_i(\theta_i) \left[\theta_i - \frac{1 - F_i(\theta_i)}{f_i(\theta_i)}\right] d\theta_i
\end{aligned}$$

Let $\psi_i(\theta_i) = \theta_i - \frac{1-F_i(\theta_i)}{f_i(\theta_i)}$. We replace $\alpha_i(\theta_i)$ with its definition and convert the integration over θ_i to $\theta = \theta_1 \times ... \times \theta_N$, and also $f(\theta) = \prod_i f_i(\theta)$.

$$\begin{aligned}
\mathbb{E}\left(m_i(\theta_i)\right) &= m_i(0) + \int_0^{\theta_{i,\max}} \alpha_i(\theta_i) f_i(\theta_i) \psi_i(\theta_i) d\theta_i \\
&= m_i(0) + \int_0^{\theta_{i,\max}} \Pi_i(\theta_i, \theta_{-i}) f(\theta) \psi_i(\theta_i) d\theta
\end{aligned}$$

Therefore the overall objective function is:

$$\sum_{i=1}^{N} \mathbb{E}\left(m_i(\theta_i)\right) = \sum_{i=1}^{N} m_i(0) + \int_0^{\theta_{i,\max}} \left[\sum_{i=1}^{N} \Pi_i(\theta_i, \theta_{-i}) \psi_i(\theta_i)\right] f(\theta) d\theta$$

The objective in Equation 4 is over the set of all $\{\Pi_i\}, \{q_i\}$. The maximizer of this objective, ignoring the constraints, is achieved when:

$$\Pi_i(\theta_i, \theta_{-i}) > 0 \iff \psi_i(\theta_i) = \max_j \psi_j(\theta_j)$$

Essentially the object will be assigned to the bidder which has the highest virtual bid. To satisfy the objective in Equation 1, choose:

$$q_i(\theta_i, \theta_{-i}) = \theta_i \Pi_i(\theta_i, \theta_{-i}) - \int_0^{\theta_i} \Pi_i(\theta, \theta_{-i}) d\theta$$

To verify that Equation 2 is satisfied we need the following assumption:

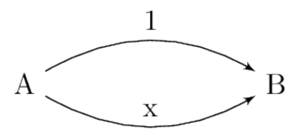


Figure 2: Pigou's Example

Assumption 1. *ψ_i is a strictly increasing function.*

Since $\psi_i(\theta_i) = \theta_i - \frac{1-F_i(\theta_i)}{f_i(\theta_i)}$, $\frac{1-F_i(\theta_i)}{f_i(\theta_i)}$ should be non-increasing $\iff$ $\frac{f_i(\theta_i)}{1-F_i(\theta_i)}$ should be non-decreasing.

Consider two scenarios:

- If $\Pi_i(\theta_i, \theta_{-i}) > 0$, then for any $\tilde{\theta}_i \geq \theta_i$ $\Pi_i(\tilde{\theta}_i, \theta_{-i}) = 1$, therefore $\Pi_i(\tilde{\theta}_i, \theta_{-i}) \geq \Pi_i(\theta_i, \theta_{-i})$ for any $\tilde{\theta}_i$.

- If $\Pi_i(\theta_i, \theta_{-i}) = 0$, therefore trivially $\Pi_i(\tilde{\theta}_i, \theta_{-i}) \geq \Pi_i(\theta_i, \theta_{-i})$ for any $\tilde{\theta}_i$.

6 Price of Anarchy

Suppose there are two routes between A and B. Each pas has its delay (cost) with we denote with $c(x)$, and it is a function of its flow x:

- The delay of the first path: $c_1(x_1)$

- The delay of the second path: $c_2(x_2)$

In usual scenarios we assume that the total flow is constant and fixed, say $x_1 + x_2 = 1$.

The average cost of the routing can be calculated as

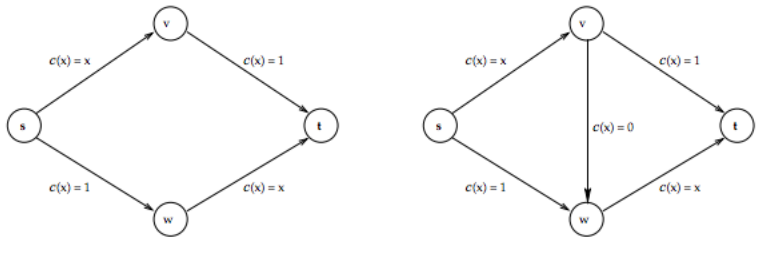
$$\frac{x_1 c(x_1) + x_2 c(c_2)}{x_1 + x_2}$$

Definition 2 (Selfish routing)**.** *At equilibrium if both path 1 and 2 are used, then $c_1(x_1) = c_2(x_2)$. If only path 1 is used, $c_1(x_1) \leq c_2(x_2)$. If only path 2 is used $c_2(x_2) \leq c_1(x_1)$. Such an (x_1, x_2) are said to be Wardrop equilibrium.*

Example 4 (Pigou's example)**.** *] Consider the costs based on Figure 2. The costs are $c_1(x_1) = 1$ and $c_2(x_2) = x_2$. The Wardrop equilibrium in these example is either $(x_1, x_2) = (1, 0)$ or $(x_1, x_2) = (0, 1)$.*

Definition 3 (Price of Anarchy (PoA))**.** *PoA is defined as*

$$\frac{\textit{Optimal cost}}{\textit{Cost under selfish routing}}$$



(a) Initial network (b) Augmented network

Figure 3: Braess's Network.

Example 5 (Braess's Paradox)**.** *Consider the network shown in Figure 3 (a). The cost of each path is* $1 + x$*. for a flow of* x*. The equilibrium flow for each path is* $x_1 = x_2 = 1/2$*. Note the scenario that* $(x_1, x_2) = (1, 0)$ *or* $(0, 1)$ *is contradictory with the definition of the Wardrop equilibrium. The total average cost under this strategy is* $0.5(1 + 0.5) \times 2 = 1.5$*.*

Suppose we add a zero-cost link between the two path, as shown in Figure 3 (b). How do we expect the flows in each route to change? It is easy to verify that in this case all of the flow will go in the path $s \to v \to w \to t$ *(why?). In this case, the total cost becomes* 2 *which is more than the previous case.*

This relatively surprising behavior is called Braess's Paradox, which is result of selfish behavior of individuals; individual elements (flow) chose to go through the edges which has zero delay (cost), although the overall cost increased.

Example 6 (PoA for linear latency)**.** *Consider a two-node network with linear latency. If the input flow is* r *the costs of the two edges are:*

$$c_1(x_1) = ar + b, \quad c_2(x_2) = ax_2 + b$$

The socially optimal answer is

$$\min_{x_2 \in [0,r]} \frac{(r - x_2)(ar + b) + (ax_2 + b)x_2}{r}$$

$$\Rightarrow -(ar + b) + 2ax_2 + b = 0 \Rightarrow x_2 = r/2$$

This would result in average delay of

$$\frac{(ar + b)r/2 + (ar/2 + b)r/2}{r} = 3a/4 + b$$

The Wardop optimal answer is $(x_1, x_2) = (0, r)$*, which results in average delay of* $(ar + b)$*.*

The Price of Anarchy is:

$$\alpha = \max_{a,b} \frac{ar + b}{3a/4 + b} \leq 4/3$$

Example 7 (PoA for quadratic latency)**.** *Consider a two-node network with linear latency. If the input flow is r the costs of the two edges are:*

$$c_1(x_1) = ar^2 + br + c, \quad c_2(x_2) = ax_2^2 + bx_2 + c$$

The socially optimal answer is

$$\min_{x_2 \in [0,r]} \frac{(r - x_2)(ar^2 + br + c) + (ax_2^2 + bx_2 + c)x_2}{r}$$

$$\Rightarrow -(ar^2 + br + c) + 3ax_2^2 + 2bx_2 + c = 0 \Rightarrow x_2 = r/2$$

This would result in average delay of

$$\frac{(ar + b)r/2 + (ar/2 + b)r/2}{r} = 3a/4 + b$$

The Wardop optimal answer is $(x_1, x_2) = (0, r)$, which results in average delay of $(ar + b)$.

The Price of Anarchy is:

$$\alpha = \max_{a,b} \frac{ar + b}{3a/4 + b} \leq 4/3$$

Theorem 6. *The pure PoA of any generalized routing problem (G, L) with linear latencies is less than* 4/3*.*

7 Blackwell Approachability

Define the reward to player for choosing action i to be a vector $r(i, j)$, when adversary has chosen action j. Define a mixed strategy reward to be

$$\sum_{i,j} r(i, j)p(i)q(j) = R(p, q)$$

In the decision making p_t is allowed to be a function of $H_t = \{p_1, ..., p_{t-1}, q_1, ..., q_{t-1}\}$, and q_t is a function of $H_t \cup \{p_t\}$.

Approachability: Given a set S it is called approachable, if it is possible for the player to choose a sequence of mixed strategies $\{p_t\}$ s.t.

$$d\left(\frac{1}{T}\sum_{t=1}^{T} R(p_t, q_t), S\right) \to 0 \text{ as } T \to +\infty$$

for some distance measure $d(x, S) = \min_{y \in S} \|x - y\|^2$.

Proposition 1. *Let S be a half-space*

$$S = \{x : w^\top x \geq b\}$$

then S is approachable if and only if the zero-sum game with payoff $w^\top r(i, j)$ to player has a value $\geq b$.

Theorem 7 (Blackwell approachability)**.** *Let S be a compact convex set. Then S is approachable if and only if every half-space containing S is approachable.*

Proof. It is easy to observe that if S is approachable, any superset of S is also approachable.

Now we show that if every half-space containing S is approachable, S is approachable too. The proof is constructive, i.e. we show it by providing an algorithm.

Algorithm: At time T suppose $\bar{R}_T = \frac{1}{T}\sum_{t=1}^{T} R(p_t, q_t) \in S$, then pick an arbitrary distribution p_t

□

Example 8. *Player chooses action i and adversary chooses action j. The cost of the player is $c(i,j)$. Denote the mixed strategies of the player and adversary at time t with p_t and q_t. p_t is chosen based on H_t, and q_t is chosen based on $\{H_t, p_t\}$. We use Blackwell approachability to show that for any valid choice of $\{q_t\}$:*

$$\lim_{T\to+\infty}\left(\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{m} c(i,j)p_t(i)q_t(j) - \min_{i\in\{1,2,\ldots,n\}}\frac{1}{T}\sum_{t=1}^{T} c(i,j)q_t(j)\right) \le 0 \tag{5}$$

Suppose we denote the minimizer with i^*; we can rewrite the above objective function in the following way:

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{m}\left(c(i,j)p_t(i)q_t(j) - c(i^*,j)p_t(i)q_t(j)\right)$$

We define the following short-hand notation $C_t(p,q) \in \mathbb{R}^n$:

$$l_t(i^*,p,q) = \sum_{i=1}^{n}\sum_{j=1}^{m}\left(c(i,j)p_t(i)q_t(j) - c(i^*,j)p_t(i)q_t(j)\right)$$

$$C_t(p,q) = \left[l_t(1,p,q),\ldots,l_t(n,p,q)\right]^\top$$

We will show the average cost of $C_t(p,q)$ (average over time), is approachable to the set $S = \{(x_1,...,x_n)|x_1,...,x_n \le 0\}$ (negative orthant), for any distribution of q_t. In other words, for any choice of q_t, there always exists sequence of distributions p_t such that for any sequence of distribution q_t, the average cost will converge to S for big enough T. This would result in the desired in Equation 5.

Now we prove the approachability of the average cost vectors. We can use Blackwell's theorem here; the average cost is approachable to S, if and only if it is approachable to any half-space containing S (i.e. $\{x|a^\top x \le b\}$, for arbitrary a and $b \ge 0$).

For any choice of a, we can choose p_t to be $p_t = a/\|a\|$. Now we can verify that for any choice of a, $a^\top C_t\left(a/\|a\|, q\right) = 0$. This shows that the there exists an algorithm for any choice of q_t and any choice of a, the average cost is inside the set S.

8 Bibliographical notes

Preliminary version mostly based on R. Srikant's Game Theory course in UIUC.