Lecture 3: Price of Anarchy (PoA): Routing

Lecturer: Yishay Mansour *Scribe: Shai Vardi, Naama Ben-Aroya*[1]

1 Lecture Overview

In this lecture we consider the problem of routing traffic to optimize the performance of a congested and unregulated network. We are given a network, a rate of traffic between each pair of nodes and a latency function specifying the time needed to traverse each edge given its congestion. The goal is to route traffic while minimizing the *total latency*. In many situations, network traffic cannot be regulated, thus each user minimizes his latency by choosing among the available paths with respect to the congestion caused by other users. We will see that this "selfish" behavior does not perform as well as an optimized regulated network.
We investigate the *price of anarchy* by exploring characteristics of *Nash Equilibrium* and *minimal latency optimal flow*.
We prove that if the latency of each edge is a linear function, then the PoA is at most 4/3, while in unsplittable (atomic) routing the PoA is bounded by 2.6. We also show that if the latency function is only known to be continuous, nondecreasing and differentiable, then there is no bounded coordination ratio.

2 Introduction

Last lecture we observed the problem of Job Scheduling. We can look at job scheduling as a private case of routing, where each player wants to send a particular amount of traffic along a path from source to destination, and has to choose exactly one line to pass his traffic along. We call this problem Parallel Lines Routing.

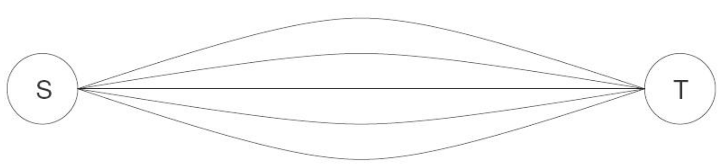


Figure 1: Routing on Parallel Lines

Today, we shall investigate the problem of routing traffic in a network. The problem is defined as follows: Given a rate of traffic between pairs of nodes in the network, find an assignment of the traffic to paths so that the total latency is minimized. Each

[1]This scribe is based in part on the scribe notes of Maya Ben-Ari and Lotem Kaplan 2006

link in the network is associated with a latency function which is typically load-dependent, i.e. the latency increases as the link becomes more congested.
In many domains (such as the internet or road networks) it is impossible to impose regulation of traffic, and therefore we are interested in those settings where each user acts according to his own selfish interests. We assume that each user will always select the minimum latency path to its destination. In other words, we assume all users are rational and nonmalicious. This can actually be viewed as a noncooperative game where each user plays the best response given the state of all other users, and thus we expect the chosen routes to form a Nash equilibrium.

2.1 Motivation for the Model

Each edge in the network is assigned a latency function, that specifies the delay of the edge as a function of the congestion on that edge.

- **The Player Model**
  We will examine two models:
  - **Splittable (non-atomic):** We will treat traffic similarly to flow in a network in which we can split up the load to several paths. Analogously, we can look at it as (infinitely) many users, where each user controls an infinitisimal portion of the total traffic.
  - **Unsplittable (atomic) :** A finite number of users; each user chooses a path on which he transports all of his load. (He is not allowed to split the load)
- **The Global Target Function** is to minimize the total latency suffered by all users.

Reminder: The *price of anarchy* (PoA) is defined as follows:

$$PoA = \max_{x \in PNE} \frac{C(x)}{OPT},$$

where OPT denote the minimum latency among all feasible flows, and $C(x)$ is the total cost of flow x.

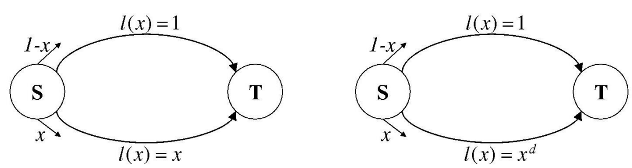
Our goal is to bound the PoA.

Example: Routing on Parallel Lines We use nearly the same model as in the last lecture (see figure 1): a set of n players with weights $w_i,\ i = 1, ..., n$ $(w_i > 0)$. m lines (machines) with speeds $s_i,\ i = 1, ..., m$.
Our goal is to minimize the congestion on the lines. The players are allowed split their flow between different lines.

Nash Equilibrium is achieved when the Load on each line is $L_i(a) = \frac{\sum_{i=1}^{n} w_i}{\sum_{j=1}^{m} s_j} = \frac{W}{S}$ (this is NE because if there is a line with more than $\frac{W}{S}$ then there is a line with less than $\frac{W}{S}$ so any player using the more loaded line will benefit from passing flow to the less loaded line).

The *Optimum* is achieved (as seen in the last lecture) by dividing the flow equally between the lines. Therefore we achieve $PoA = 1$.

We will now look at a more interesting example:



(a) (b)

Figure 2: bounded and unbounded PoA

Figure 2(a) shows an example for which Nash flow will only traverse in the lower path, and reach the cost 1, while OPT will divide the flow equally among the two paths. The target function is $1 \cdot (1-x) + x \cdot x$ and it reaches minimum with value $\frac{3}{4}$ when x = $\frac{1}{2}$, giving a Price of Anarchy of $\frac{4}{3}$ for this example. Combining the example with the tighter upper bound to be shown, we demonstrate a tight bound of $\frac{4}{3}$ for linear latency functions.

In Figure 2(b) the flow at Nash will continue to use only the lower path. Let $(1-x)$ be the flow on the upper path in the optimum solution, and x - the flow on the lower path. The overall cost is $(1-x) \cdot 1 + x \cdot x^d = 1 - x + x^{d+1}$.

Lemma 2.1 $\lim_{d\to\infty} OPT = 0$

Proof. We can choose $(1-x) = \frac{\log d}{d+1}$

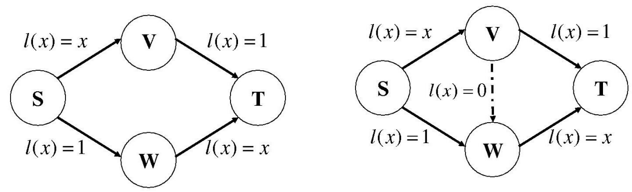
$$x^{d+1} \approx e^{-\log d} = \frac{1}{d}$$

$$\lim_{d\to\infty} OPT = \lim_{d\to\infty} 1 - x + x^{d+1} \approx \lim_{d\to\infty} \frac{\log d}{d+1} + \frac{1}{d} = 0$$

□

It is clear that in this case $\lim_{d\to\infty} PoA = \infty$, meaning that the PoA cannot be bounded from above in some cases when nonlinear latency functions are allowed.

Before we continue, let's examine an example setting which has inspired much of the work in this traffic model. Consider the network in Figure 3(a). There are



(a) (b)

Figure 3

two disjoint paths from S to T. Each path follows exactly two edges. The latency functions are labeled on the edges. Suppose one unit of traffic needs to be routed from S to T. The optimal flow coincides with the Nash equilibrium such that half of the traffic takes the upper path and the other half takes the lower path. In this manner, the latency perceived by each user is $\frac{3}{2}$. In any other nonequal distribution of traffic among the two paths, there will be a difference in the total latency of the two paths and users will be motivated to reroute to the less congested path.

Note Incidentally, we will soon realize that in any scenario in which the flow at Nash Equilibrium is split over more than a single path, the latency of all the chosen paths must be equal.

Now, consider Figure 3(b) where a fifth edge of latency zero is added to the network. While the optimum flow has not been affected by this augmentation and stays $\frac{3}{2}$, Nash will only occur by routing the entire traffic on the single S $\rightarrow$ V $\rightarrow$ W $\rightarrow$ T path, hereby increasing the latency each user experiences to 2 (because if we split the flow to the upper and lower paths, then the user will be motivated to reroute to the less congested path, using the new edge. However, if the entire traffic is routed trough S $\rightarrow$ V $\rightarrow$ W $\rightarrow$ T no user will benefit from a change, and therefore this is a Nash Equilibrium).
Amazingly, adding a new zero latency link had a negative effect for all agents. This counter-intuitive impact is known as *Braess's paradox.*

Anecdote 1 *Two live and well known examples of Braess's paradox occurred when 42nd street was closed in New York City and instead of the predicted traffic gridlock, traffic flow actually improved. In the second case, traffic flow worsened when a new road was constructed in Stuttgart, Germany, and only improved after the road was closed.*

2.2 Formal Definition of the Problem

The problem of routing flow in a network with flow dependent latencies is defined as follows:

- We consider a directed graph $G = (V, E)$.

- Input
  - k pairs of source and destination vertices (s_i, t_i).
  - Demand r_i (the amount of required flow between s_i and t_i).
    We may assume that $r_i > 0$.
  - Each edge $e \in E$ is given a load-dependent latency function $\ell_e : \mathbb{R}^+ \to \mathbb{R}^+$.
    We restrict our discussion to nonnegative, differentiable and nondecreasing latency functions.
- Output
  Flow f - A function that defines for each path p a flow f_p.
  f induces *flow on edge* e, $f_e = \sum_{p:e\in p} f_p$ (a flow on an edge is the sum of flows of all the paths that contains that edge).
- We denote the set of simple paths connecting the pair (s_i, t_i) by $\mathcal{P}_i$.
  And let $\mathcal{P} = \bigcup_i \mathcal{P}_i$.
- Solution is *feasible* if $\forall i,\ \sum_{p\in\mathcal{P}_i} f_p = r_i$ (for all i - the sum of the flow over all paths between s_i to t_i is equal to the demand r_i).
- The latency of a path ℓ_p is defined as the sum of latencies of all edges in the path. $\ell_p(f) = \sum_{e\in p} \ell_e(f_e)$.
- The *total (social) cost* of a flow f, $C(f) \triangleq \sum_{p\in\mathcal{P}} \ell_p(f) \cdot f_p = \sum_{e\in E} \ell_e(f_e) \cdot f_e$ (the two formulas express the same value, since they differ only in the accumulation order).
  The cost of player i, denoted by $c_i(f)$, is $\sum_{p\in\mathcal{P}_i} \ell_p(f) \cdot f_p$.
- (G, r, ℓ) - A triple which defines an *instance* of the routing problem.
- Our goal is to find a feasible flow f that will minimize the total cost $C(f) = \sum_e \ell_e(f_e) \cdot f_e$.

We denote this problem by (G, r, ℓ).

3 Characterizations of Nash & OPT Flows

3.1 Flows at Nash Equilibrium

Lemma 3.1 *A feasible flow f for instance (G, r, ℓ) is Nash Equilibrium if for every $i \in \{1, ..., k\}$ and p, $p' \in \mathcal{P}_i$*

- *if $f_p > 0$ then $\ell_p(f) \le \ell_{p'}(f)$*

From the lemma it follows that flow at Nash Equilibrium will be routed only through best response (BR) paths. Consequently, all paths assigned with a positive flow between (s_i, t_i) have equal latency.

Definition 3.2 *Define* $L_i(f) = min_{p \in P_i} \ell_p(f)$. *(The minimum over all paths, including those with no traffic).*

Corollary 3.1 f *is a flow at a Nash Equilibrium for instance* (G, r, ℓ) *if and only if* $C(f) = \sum_i L_i(f) \cdot r_i$.

Proof. $\Rightarrow$
Follows directly from the definition of Nash Equilibrium
$\Leftarrow$
If $C(f) = \sum_i L_i(f) \cdot r_i$ then for every load of i on any path with $load > 0$, the cost is $L_i(f)$, otherwise there will be a "surplus". □

3.2 Optimal Solution - Flow

Our goal is to find the optimal solution, that is to find a feasible flow f that will minimize the total cost $C(f) = \sum_{e \in E} \ell_e(f_e) f_e$.

Let $c_e(x) = \ell_e(x) \cdot x$. Clearly, it follows that $C(f) = \sum_{e \in E} c_e(f_e)$.

To find the optimum flow, ${c^*}_e(x)$, we will look at ${c'}_e(x)$ (the derivative of c). $c'_e(x) = \ell_e(x) + x \cdot \ell'_e(x)$

In our case of network routing, we assume that for each edge $e \in E$ the function $c_e(x) = \ell_e(x) \cdot x$ is a convex function, and therefore, our target function $C(f)$ is also convex. Our assumption on $\ell_e(x)$ implies that $c_e(x)$ is differentiable for every x.
Let $c'_e(x) = \frac{d}{dx} c_e(x)$. Let $c'_p(x) = \sum_{e \in p} c'_e(x)$.

Lemma 3.3 *(The optimality condition) Let* (G, r, ℓ) *be a splittable game. For each edge* $e \in E$ *the function* $c_e(x) = \ell_e(x) \cdot x$ *is a convex, continuous and differentiable function. A flow* f *is optimal for* (G, r, ℓ) *iff*

$$\forall p, p' \in \mathcal{P}_i,\ f_p > 0 \Rightarrow\ c'_p(f) \leq c'_{p'}(f)$$

Proof. The formal proof is given using convex programming (see appendix A). The intuition behind the proof is that if there is a path p' such that ${c'}_p(f^*) > {c'}_{p'}(f^*)$ then we can transfer some of the load from p to p' and gain the difference between the derivatives. □

Notice the resemblance between the characterization of optimality conditions (Lemma 3.3), and Nash Equilibrium (Lemma 3.1). In fact, an optimal flow can be interpreted as a Nash equilibrium with respect to a different edge latency functions.
We will use this resemblance to reach the bound on PoA.

Let

$$\begin{aligned} \ell^*_e(x) &\triangleq c'_e(x) = (\ell_e(x) \cdot x)' = \ell_e(x) + x \cdot \ell'_e(x) \\ \ell^*_p(x) &\triangleq \textstyle\sum_{e \in p} \ell^*_e(x) \end{aligned}$$

Corollary 3.2 *f is an optimal flow for (G, r, ℓ) iff it is Nash Equilibrium for the instance (G, r, ℓ^*).*

Proof. f is OPT for $\ell \Leftrightarrow$ (optimality conditions) $\forall i, \forall p, p' \in P_i, \forall f_p > 0,\ c'_p(f) \leq c'_{p'}(f) \Leftrightarrow$ (by def.) $\forall p, p' \in P_i, \forall f_p > 0,\ \ell^*_p(f) \leq \ell^*_{p'}(f) \Leftrightarrow f$ is Nash Eqi. for ℓ^* $(\forall i\ \forall p, p' \in \mathcal{P}_i)$. □

3.3 Existence of Flows at Nash Equilibrium

We exploit the similarity between the characterizations of Nash and OPT flows to establish that a Nash equilibrium indeed exists and its cost is unique.

The optimum flow is characterized by the derivative:

$$c'_p(f^*) \leq c'_{p'}(f^*)$$

The definition of Nash equilibrium requires the same condition, but for different functions:

$$c_p(f) \leq c_{p'}(f)$$

We will define a function h_e such that $h'_e(x) = \ell_e(x)$.

Definition 3.4 $h_e(x) \triangleq \int_o^x \ell_e(y)dy$

Clearly, $h'_e(x) = \ell_e(x)$. We further define a potential function:

Definition 3.5 $\Phi(f) = \sum_e h_e(f_e)$

Informally, we can say that the function h_e "replaces" c_e as the cost function.

Claim 3.3 *h_e is non-negative, monotonous increasing and differentiable. Furthermore, Φ is a convex function.*

Theorem 3.4 *For every splittable (non-atomic) game (G, r, ℓ):*

1. *There exists at least one Nash equilibrium.*
2. *If f and f' are Nash equilibria then for every $e, c_e(f) = c_e(f')$*

Proof.

1. Every flow that minimizes Φ is a Nash equilibrium.
2. Let f and f' be Nash equilibria. Define $g = \lambda f + (1-\lambda)f'$ for $\lambda \in [0, 1]$ Because Φ is convex:

$$\Phi(g) \leq \lambda\Phi(f) + (1 - \lambda)\Phi(f')$$

Because f and f' both minimize Φ, we get $\Phi(f) = \Phi(f')$ and so, of course, $\Phi(g) = \Phi(f) = \Phi(f')$. Because $\Phi(g)$ is the sum of convex functions, this is possible only if there is an equality for all the member in the sum (otherwise we could differentiate between them for some λ).

Therefore, $c_e(g) = c_e(f) = c_e(f')$. □

3.4 Bounding the Price of Anarchy

The relationship between Nash and OPT characterizations provide a general method for bounding the price of anarchy $PoA = \frac{Nash}{OPT}$.
We notice that because $\ell_e(\cdot)$ is non-decreasing, we get $h_e(x) = \int_0^x \ell_e(t)dt \leq x\ell_e(x) = c_e(x)$

Theorem 3.5 *If there exists a constant $\alpha > 1$ such that $\forall x,\ \alpha h_e(x) \geq c_e(x)$ then $PoA \leq \alpha$.*

Proof.

$$\begin{aligned} Nash = C(f) &= \textstyle\sum_{e\in E} c_e(f_e) \\ &\leq \alpha \textstyle\sum_{e\in E} h_e(f_e) \\ &\leq \alpha \textstyle\sum_{e\in E} h_e(f_e^*) \\ &\leq \alpha \textstyle\sum_{e\in E} c_e(f_e^*) \\ &= \alpha \cdot C(f^*) = \alpha \cdot OPT \end{aligned}$$

The first inequality follows from the hypothesis, the second follows from the fact that Nash flow f is OPT for the function $h_e(f_e)$ and the final inequality follows from $h_e(x) \leq c_e(x)$. □

Corollary 3.6 *If the latency function $\ell_e(\cdot)$ is a polynomial function of degree d, $\ell_e(x) = \sum_{i=0}^d a_{e,i}x^i$, then $PoA \leq d+1$.*

Proof.
$h_e(x) = \int_0^x \ell_e(y)dy = \sum_{i=0}^d \frac{1}{i+1}a_i x^{i+1} \geq \frac{1}{d+1}\sum_{i=0}^d a_i x^{i+1} = \frac{1}{d+1}x \cdot \ell_e(x) = \frac{1}{d+1}c_e(x)$ □

Note From the corollary, an immediate coordination ratio of 2 is established for linear latency functions. We will now show a tighter bound of $\frac{4}{3}$.

4 A Tight Bound for Linear Latency Functions

We will now focus on a scenario where all edge latency functions are linear $\ell_e(x) = a_e x + b_e$, for constants $a_e, b_e \geq 0$. A fairly natural example for such a model is a network employing a congestion control protocol such as TCP. We have already seen in Figure 2(a) an example where the coordination ratio was $\frac{4}{3}$. We have also established an upper bound of 2 according to Corollary 3.6. We shall now show that the $\frac{4}{3}$ ratio is also a tight upper bound.

Prior to this result, we examine a simple case where $\ell_e(x) = b_e$. In this case both OPT and Nash will route all the flow to the shortest paths. Thus, $Nash = OPT$.

Lemma 4.1

$$xy \leq x^2 + \frac{y^2}{4}$$

Proof. See appendix. □

Theorem 4.1 *If the latency functions are all of the form $\ell_e(x) = a_e x + b_e$ then $PoA \leq \frac{4}{3}$.*

Proof. Let f be a flow at Nash equilibrium and f^* an optimal flow. Given a flow f, we define $\ell_e^f = a_e f_e + b_e$, and $C^f(x) = \sum \ell_e^f x_e$.

$$\begin{aligned}
C^f(x) &= \sum (a_e f_e + b_e) x_e \\
&= \sum (a_e f_e x_e + b_e x_e) && (xy \leq x^2 + \tfrac{y^2}{4}) \\
&\leq \sum (a_e x_e^2 + b_e x_e) + \sum a_e f_e^2 \tfrac{1}{4} \\
&\leq C(x) + \tfrac{1}{4} C(f) && (\forall x\ C^f(f) \leq C^f(x))
\end{aligned}$$

$$\begin{aligned}
C^f(f) = C(f) &\leq C(f^*) + \tfrac{1}{4} C(f) \\
\tfrac{3}{4} C(f) &\leq C(f^*) \\
C(f) &\leq \tfrac{4}{3} C(f^*) \\
PoA &\leq \tfrac{4}{3}
\end{aligned}$$

□

5 Unsplittable (Atomic) Routing

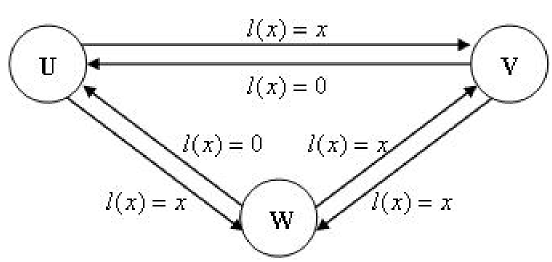


Figure 4: Unsplittable (atomic) routing example 1

Examples

1. Consider the graph in Figure 4. There are 4 players with the following requests (s_i, t_i): $\{(U, V); (U, W); (V, W); (W, V)\}$, and $\forall i : r_i = 1$.
   The optimal solution will route the players through the edges with $l(x) = x$, i.e., player 1 will use the (only) path $U \to V$, player 2: $U \to W$, player 3: $V \to W$ and player 4: $W \to V$. The total cost of the optimal solution is $1 \cdot 1 + 1 \cdot 1 + 1 \cdot 1 + 1 \cdot 1 = 4$. This flow is a Nash equilibrium, but it is not unique. Another NE is generated from routing player 1 through $U \to W \to V$, player 2: $U \to V \to W$, player 3: $V \to U \to W$ and player 4: $W \to U \to V$. The total cost of that flow is $2 \cdot 2 + 2 \cdot 2 + 1 \cdot 1 + 1 \cdot 1 = 10$. In this NE, we have $PoA = \frac{10}{4} = 2.5$.

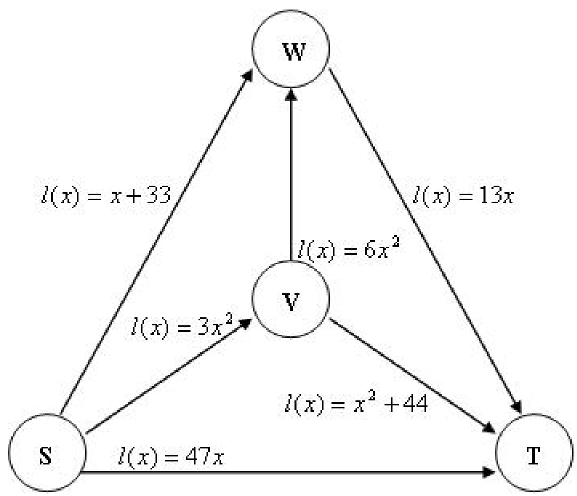


Figure 5: Unsplittable (atomic) routing example 2

2. Consider the graph in Figure 5. There are 2 players with $s_i = S, t_i = T, r_i = i$ $(i = 1, 2)$. There are four available paths from S to T in this graph:

$$\begin{array}{ll} p_1 & S \to T \\ p_2 & S \to V \to T \\ p_3 & S \to W \to T \\ p_4 & S \to V \to W \to T \end{array}$$

   We will show that there is no pure equilibrium in this game using the following lemmas:

   (a) If player 2 choose either path p_1 or p_2 then player 1 will choose path p_4.
   (b) If player 1 choose path p_4 then player 2 will choose path p_3.
   (c) If player 2 choose either path p_3 or p_4 then player 1 will choose path p_1.
   (d) If player 1 choose path p_1 then player 2 will choose path p_2.

Giving these four lemmas, one can easily see that there is no equilibrium in this game.
Proof of (a): Assuming that player 2 chose path p_2, the cost of player 1 will be

path	*cost*
p_1	47
p_2	$27 + 53 = 80$
p_3	$34 + 13 = 47$
p_4	$27 + 6 + 13 = 46$

Therefore, player 1 will choose p_4. Similarly, if player 2 choose path p_1, the cost of player 1 will be

path	*cost*
p_1	141
p_2	$3 + 47 = 50$
p_3	$34 + 13 = 47$
p_4	$3 + 6 + 13 = 22$

Again, player 1 would prefer p_4. □

(Proofs for (b), (c) and (d) are similar.)

We showed that in an unsplittable game there is does not necessarily exist a Nash equilibrium.

5.1 Existence of a Nash equilibrium

Theorem 5.1 *If (G, r, ℓ) is an unsplittable game with $\forall i \in N, r_i = 1$ then there exists a Nash equilibrium.*

Proof. We define a potential function Φ_a:

$$\Phi_a(f) = \sum_e \sum_{i=1}^{f(e)} \ell_e(i)$$

(We can think of Φ_a as a discrete version of Φ).
We observe the effect of player i when he moves (his load) from path p to path p'.

$$\ell_{p'}(f') - \ell_p(f) = \sum_{e \in p' - p} \ell(f_e + 1) - \sum_{e \in p' - p} \ell(f_e)$$

We now look at the effect this has on the potential Φ_a:

$$\Phi_a(f') - \Phi_a(f) = \begin{cases} +\ell_e(f_e + 1) & e \in p' - p \\ -\ell_e(f_e) & e \in p - p' \\ 0 & otherwise (e \in p \cap p' \vee e \notin p \cup p') \end{cases}$$

Therefore

$$\ell_{p'}(f') - \ell_p(f_) = \Phi_a(f') - \Phi_a(f)$$

And so if we let the players play "best response" the potential decreases. As the potential is non-negative, a series of best reponses will converge to a Nash equilibrium.
□

5.2 Quality of Nash equilibria in unsplittable linear games

We will show how to bound the PoA for linear functions, assuming a Nash equilibrium exists.

Let f be a Nash equilibrium in an unsplittable game. Let f^* be the optimal flow in the same game.

Assume player i uses path p_i in f and p_i^* in f^*. For linear cost functions, we have:

$$c_{p_i}(f) = \sum_{e\in p_i}(a_e f_e + b_e) \le \sum_{e\in p_i^*}(a_e(f_e + r_i) + b_e)$$

(Otherwise player i would switch to path p_i^*, in contradiction to the definition of Nash equilibrium).

Lemma 5.1 $C(f) \le C(f^*) + \sum_e a_e f_e f_e^*$

Proof.

$$\begin{aligned}
C(f) &\le \sum_i r_i(\sum_{e\in p_i^*} a_e(f_e + r_i) + b_e) \\
&\le \sum_i r_i(\sum_{e\in p_i^*} a_e(f_e + f_e^*) + b_e) \\
&= \sum_{e\in E}[a_e(f_e + f_e^*) + b_e)]f_e^* \\
&= \sum_{e\in E}[a_e(f_e^*)^2 + b_e f_e^*] + \sum_{e\in E} a_e f_e f_e^* \\
&= C(f^*) + \sum_{e\in E} a_e f_e f_e^*
\end{aligned}$$

The transition from line 2 to line 3 is due to changing the order of the sums. f^* is the optimal flow and we summed over all edges according to the optimum flow.

□

We now show a better bound on the price of anarchy for unsplittable (atomic) routing games.

Theorem 5.2 *Let (G, r, ℓ) be an unsplittable game. Then PoA $\le \frac{3+\sqrt{5}}{2} \sim 2.618$*

Proof.
Let f be a Nash equilibrium and f^* be an optimal flow.

$$\sum_{e\in E} a_e f_e f_e^* \le \sqrt{\sum_{e\in E} a_e f_e^2}\sqrt{\sum_{e\in E} a_e (f_e^*)^2} \le \sqrt{C(f)}\sqrt{C(f^*)}$$

The first inequality stems from Cachy-Schwartz on $\sqrt{a_e} f_e + \sqrt{a_e} f_e^*$.
We have shown that

$$\begin{aligned} C(f) &\leq C(f^*) + \sum a_e f_e f_e^* \\ &\leq C(f^*) + \sqrt{C(f)C(f^*)} \\ \Rightarrow \frac{C(f)}{C(f^*)} &\leq 1 + \sqrt{\frac{C(f)}{C(f^*)}} \end{aligned}$$

Assigning $x = \frac{C(f)}{C(f^*)}$, get $x \leq 1 + \sqrt{x} \Rightarrow (x-1)^2 \leq x$.
Solving for x, we get $\frac{C(f)}{C(f^*)} = x \leq \frac{3+\sqrt{5}}{2} \simeq 2.618$. □

6 APPENDIX A

Convex Set

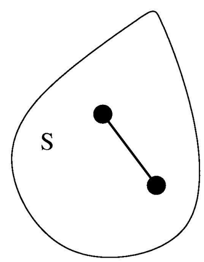


Figure 6: Convex Set

Definition 6.1 *A set S is called a* convex set *if* $\forall A, B \in S$, $0 \leq \lambda \leq 1$, $\lambda A + (1 - \lambda)B \in S$.

Intuitively, a set S is convex if the linear segment connecting two points in the set, is entirely in the set.
(see Figure 6)

Convex Function

Definition 6.2 *Function f is called a* convex function *if* $\forall x, y$, $0 \leq \lambda \leq 1$, $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$.

(see Figure 7)

Convex Programming

Let $F(x)$ be a convex function, and S a convex set.
A convex Programming is of the form: $\min F(x)$, $s.t.\ x \in S$.

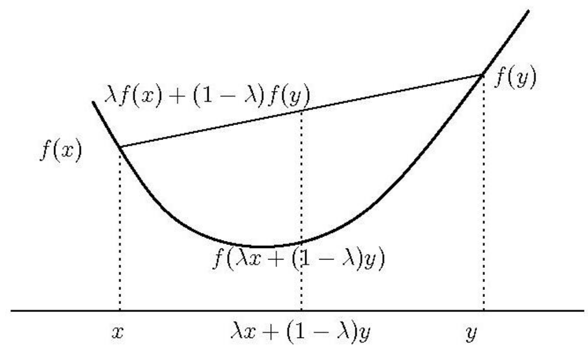


Figure 7: Convex Set

Lemma 6.3 *If $F(\cdot)$ is strictly convex then the solution is unique.*

Proof. Assume that $x \neq y$ are both minimum solutions. Let $z = \frac{1}{2}x + \frac{1}{2}y$. Because S is a convex set, $z \in S$.
Since $F(\cdot)$ is strictly convex: $F(z) < \frac{1}{2}F(x) + \frac{1}{2}F(y)$, contradicting the minimality of $F(x)$ and $F(y)$. □

Lemma 6.4 *If $F(\cdot)$ is convex then the solution set U is convex.*

Lemma 6.5 *If $F(\cdot)$ is convex and y is not optimal ($\exists x : F(x) < F(y)$) then y is not a local minimum. Consequently, any local minimum is also a global minimum.*

Proof. Assume that y is not optimal, i.e. $\exists x : F(x) < F(y)$. Let $Z = \lambda x + (1 - \lambda)y$. Since $F(\cdot)$ is convex $F(z) \leq \lambda F(x) + (1 - \lambda)F(y) < F(y)$, for every $0 < \lambda < 1$.
□

Note Lemma 6.5 implies that the "gradient method" converges to an optimal solution in convex programming.

7 APPENDIX B

Lemma 7.1

$$xy \leq x^2 + \frac{y^2}{4}$$

Proof.

$$
\begin{aligned}
xy &\leq x^2 + \tfrac{y^2}{4} \\
4xy &\leq 4x^2 + y^2 \\
0 &\leq 4x^2 - 4xy + y^2 \\
0 &\leq (2x - y)^2
\end{aligned}
$$

□