Lecture 14: Network Traffic, Congestion, and Potential Games

Alexander Wolitzky

MIT

6.207/14.15: Networks, Spring 2022

Network Traffic

Having covered the basics of both graph theory and game theory, we're now ready to study strategic interactions on networks.

Start with a simple and important example: **network traffic** (also known as **routing games**).

- Multiple individuals need to get from point A to point B on a network.
  - Drivers on a road network; information packets under decentralized routing on a communication network.
- Each individual chooses a route to minimize its own travel time, given others' route choices. (Nash equilibrium.)
- What happens? Is the equilibrium outcome socially efficient? How inefficient can it be? What types of interventions can restore efficiency?
- Network traffic is important in its own right, and is also a point of entry into the study of **potential games**, an important general class of games with many engineering/CS applications.

A Simple Example

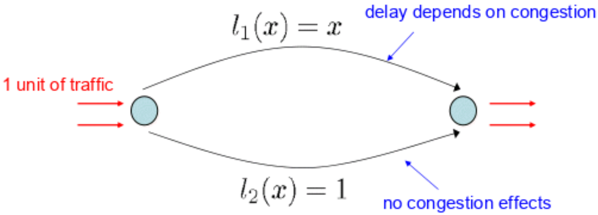
A unit mass of traffic must be routed over a network.

There are two routes.

- On route 1, the delay (or **latency**) depends on the mass of traffic taking that route: if mass x takes route 1, the latency is $l_1(x) = x$.
- On route 2, the delay is independent of the mass of traffic: for any mass of traffic x, the latency is $l_2(x) = 1$.

(Perhaps route 1 is a direct route on slow and easily congested local roads, while route 2 is an indirect route on a fast highway that's less congestible.)

Example: Diagram



Example: The Social Optimum

What is the **socially optimal** (total utility maximizing) routing, i.e. the routing pattern that minimizes average delay?

- If mass x takes route 1, average delay is

$$x \cdot x + (1 - x) \cdot 1 = x^2 + 1 - x$$

- This is minimized at $x = \frac{1}{2}$, for an average delay of $\frac{3}{4}$.

Note: At this socially optimal solution, different agents face different delays.

- Half the agents take route 1 and face delay $\frac{1}{2}$.
- Half the agents take route 2 and face delay 1.

The social optimum is not an equilibrium when each agent chooses her own route, as the agents who are "supposed" to take route 2 would deviate to taking route 1.

Example: The Nash Equilibrium

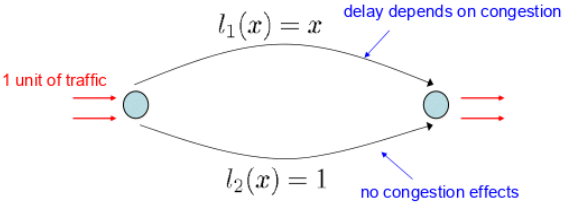
What is the (Nash) **equilibrium routing**, i.e. the routing pattern that results from each agent choosing the route that minimizes her own delay?

- For any $x < 1$, delay is less on route 1 than route 2.
- Hence, the only NE is for everyone to take route 1.
- This results in a delay of 1 for all agents.

Note: As compared to the social optimum, half the agents face the same delay (1) and half the agents face more delay (1 instead of $\frac{1}{2}$).

- No one is better off, and some people are strictly worse off!

Example: Intuition



Economic intuition for equilibrium inefficiency:

- Choosing the congestible route 1 rather than route 2 imposes a cost on other agents: a "negative externality."
- An individual agent does not take this externality into account when making her decision.
- Therefore, at equilibrium there are inefficiently many agents taking route 1.

Congestion vs. the Prisoner's Dilemma

The negative externality imposed on others by driving on a congestible road can be related to the prisoner's dilemma we saw last class:

	C	D
C	2,2	0,3
D	3,0	1,1

- Playing D rather than C always yields a selfish gain of 1, but imposes a negative externality of 2 on the other player.
- The unique Nash equilibrium outcome is (D, D), even though the socially optimal outcome is (C, C).
- Similarly, driving on a congestible road can save an individual agent time, but it imposes a negative externality on everyone else.
- Nash equilibrium in a congestion game involves overuse of congestible resources, relative to the social optimum.

The Price of Anarchy

In a game with negative payoffs ("costs" or "losses" that we want to minimize), the **price of anarchy** is the ratio of the total cost borne by all agents in the worst equilibrium to the total cost at the social optimum.

- $POA \geq 1$, because the social optimum minimizes costs.

In the example,

- There is a unique equilibrium with total cost 1.
- Total cost at the social optimum is $\frac{3}{4}$.
- Hence, the price of anarchy is $1/\frac{3}{4} = \frac{4}{3}$.
- In other words, total cost is $\frac{4}{3}$ times higher in the (worst) equilibrium as compared to the social optimum.

The Price of Anarchy

In a game with negative payoffs ("costs" or "losses" that we want to minimize), the **price of anarchy** is the ratio of the total cost borne by all agents in the worst equilibrium to the total cost at the social optimum.

- $POA \geq 1$, because the social optimum minimizes costs.

In the example,

- There is a unique equilibrium with total cost 1.
- Total cost at the social optimum is $\frac{3}{4}$.
- Hence, the price of anarchy is $1/\frac{3}{4} = \frac{4}{3}$.
- In other words, total cost is $\frac{4}{3}$ times higher in the (worst) equilibrium as compared to the social optimum.

If we consider the best equilibrium instead of the worst one, the corresponding ratio is called the **price of stability**.

- $1 \leq POS \leq POA$.

General Traffic Model

Directed network $G = (N, E)$.

Some given node is the origin, another node is the destination.

- We consider here a relatively simple model where everyone starts at the same origin and needs to get to the same destination.
- The analysis is similar in the more general case where different agents have different origins or destinations.

We normalize the total mass of traffic to 1.

General Traffic Model

Let P denote the set of paths between origin and destination.

Let x_p denote the **flow** on path $p \in P$.

- How many agents use path p.

Each link $i \in E$ has a **latency function** $l_i(x_i)$, where x_i is the total flow on link i, given by

$$x_i = \sum_{p \in P : i \in p} x_p.$$

- Here, $i \in p$ means link i is part of path p.

The latency function captures congestion effects.

- Assume $l_i(x_i)$ is non-negative and non-decreasing for each link i.
- The functions l_i can be different for different links i: some links can be more congestible than others.

General Traffic Model

A **routing pattern** (or **flow**) x is a probability distribution on P.

- A description of what fraction of the traffic takes each possible path from the origin to the destination.

The **total delay** (or **total latency**, or **total cost**) of a routing pattern x is

$$C(x) = \sum_{i \in E} x_i l_i(x_i).$$

This is simply the sum over links of the total delay (=mass of traffic times per-unit delay) incurred on each link.

We could also write this as

$$C(x) = \sum_{p \in P} x_p \sum_{i \in p} l_i(x_i),$$

where here we take the sum over paths of the total delay incurred on each path.

Socially Optimal Routing

A routing pattern x is **socially optimal** if it is a solution to the following problem:

$$\min_x \sum_{i \in E} x_i l_i \left(x_i \right)$$

subject to

$$\begin{aligned} x_i &= \sum_{p \in P : i \in p} x_p \text{ for all } i \in E, \\ \sum_{p \in P} x_p &= 1 \text{ and } x_p \geq 0 \text{ for all } p \in P. \end{aligned}$$

- First constraint: traffic on link i = mass of agents using a path that goes through link i.
- Second constraint: everyone must get from origin to destination.

Equilibrium Routing

A routing pattern x is an **equilibrium** if, for any path $p \in P$ with $x_p > 0$, there does not exist a path $p' \in P$ such that

$$\sum_{i \in p'} l_i(x_i) < \sum_{i \in p} l_i(x_i).$$

- Taking what everyone else is doing as given, no agent can switch to a faster route.

In other words, x is an equilibrium if

1. For all $p, p' \in P$ with $x_p, x_{p'} > 0$, we have

$$\sum_{i \in p} l_i(x_i) = \sum_{i \in p'} l_i(x_i), \text{ and}$$

2. For all $p, p' \in P$ with $x_p > 0$ and $x_{p'} = 0$, we have

$$\sum_{i \in p} l_i(x_i) \leq \sum_{i \in p'} l_i(x_i).$$

Equilibrium Routing: Comment

A routing pattern x is an **equilibrium** if, for any path $p \in P$ with $x_p > 0$, there does not exist a path $p' \in P$ such that

$$\sum_{i \in p'} l_i(x_i) < \sum_{i \in p} l_i(x_i).$$

This is simply the Nash equilibrium of the large-population game where no one individual's route affects overall traffic.

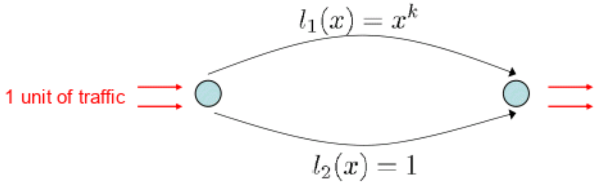
- In the context of routing games, this is also called **Wardrop equilibrium**.
- (Minor point: technically doesn't fit the usual definition of Nash equilibrium, since there is a continuum of players.)

Inefficiency of Equilibrium Routing

We've already seen that equilibrium routing can be inefficient.

In fact, this equilibrium inefficiency can be arbitrarily severe: $POA \approx \infty$.

Consider the same example as earlier but with a different latency function on route 1:



Note: we simply replaced $l_1(x) = x$ with $l_1(x) = x^k$.

- When k is large, congestion on route 1 only "gets bad" when almost everyone is using route 1.

Inefficiency of Equilibrium Routing (cntd.)

When k is large, x^k is close to 0 unless x is very close to 1.

- If 99% of agents take route 1, then when k is very large total delay is close to $.99 \cdot 0 + .01 \cdot 1 = .01$.
- As k increases, can have more and more agents take route 1 with incurring much delay.
- Socially optimal delay goes to 0 as $k \rightarrow \infty$.

But x^k is still less than 1 for all $x < 1$, so equilibrium again has everyone taking route 1, which yields equilibrium delay 1.

Therefore, the price of anarchy goes to $1/0 = \infty$ as $k \rightarrow \infty$.

How to Improve Efficiency?

There are different ways to reduce traffic or improve efficiency.

Leading contenders:

- Build new links or increase capacity on existing links.
- Introduce congestion pricing.

We'll see a classic example of how increasing capacity can backfire, while congestion pricing is a quite general solution.

- It can also be shown that a sufficiently large increase in capacity always reduces traffic, but we won't cover this result here. Basic idea: can show that equilibrium delay with 1 unit of traffic is less than optimal delay with 2 units of traffic, so delay must decrease if we "double the capacity of every edge."

Increasing Capacity

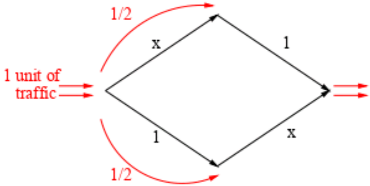
Can reducing the latency function $l_i(\cdot)$ on any link ever increase **socially optimal** delay.

No, because can always stick with the old routing pattern, which now involves less delay.

A special case of this observation: adding a new link always decreases optimal delay.

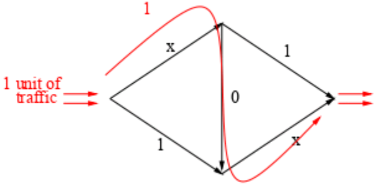
This raises the question, can adding a new link ever increase **equilibrium** delay?

Braess's Paradox



$C_{eq} = 1/2\ (1/2+1) + 1/2\ (1/2+1) = 3/2$

$C_{sys} = 3/2$



$C_{eq} = 1 + 1 = 2$

$C_{sys} = 3/2$

Braess's Paradox in the Real-World?

Braess's paradox shows that, in theory, closing a road can reduce commuting time, even if the number of commuters does not fall.

An interesting question: does this ever happen in real-world traffic networks?

There are several claimed cases, but evidence is mostly anecdotal.

Rather than trying to identify exogenous *actual* road closures, it's easier to run simulations about what *would* happen to real-world networks if some roads were closed.

- Debatable how convincing such simulations are.

Braess's Paradox in the Real-World? (cntd.)

One of the best-known papers doing this (Youn, Gastner, and Jeong, *Physical Review Letters* 2008) argues that closing Main Street would decrease traffic between Cambridge and Boston!
Possible explanation: Consider commuters going from Harvard Square to downtown Boston.

Three main routes that don't use Main Street:

1. Cross river at Harvard Sq, take Storrow Drive all the way.
2. Take Mass Ave through Cambridge and across river, then Storrow. (First part of the route is the most congestible: Mass Ave in Cambridge.)
3. Take Broadway/Hampshire through Cambridge, take Longfellow bridge. (Second part of the route is the most congestible: Longfellow bridge.)

Main St lets commuters use first part of Route 2 and second part of Route 3: causes traffic both on Mass Ave and Longfellow bridge.
If we closed Main St and forced commuters to choose between Routes 2 and 3, traffic could decrease on both routes.

Congestion Pricing

An alternative way to reduce traffic: congestion pricing.

Consider first example of inefficiency: $l_1(x) = x$, $l_2(x) = 1$.

- Efficient routing: $x = \frac{1}{2}$
- Equilibrium routing: $x = 1$

Why isn't efficient routing an equilibrium?

Each route 2 agent has an individual incentive to switch to route 1, as doesn't take into account that this increases delay for the mass $\frac{1}{2}$ agents on route 1.

Solution: impose a tax on route 1 ("congestion pricing").

Congestion Pricing (cntd.)

Suppose all drivers value their time at \$1 per unit.

Suppose the government declares that each driver must pay t dollars to use route 1, with the proceeds of $t \cdot x_1$ dollars distributed equally among all members of society.

Now an individual driver is indifferent between the two routes iff

$$x + t = 1.$$

(**Note:** she gets the proceeds $t \cdot x_1$ whichever route she takes, so this doesn't affect her decision.)

Given tax t, the equilibrium mass of drivers on route 1 equals $1 - t$.

Congestion Pricing (cntd.)

Given tax t, the equilibrium mass of drivers on route 1 equals $1 - t$.

To implement the socially efficient outcome of $x = \frac{1}{2}$, the government must set $t = \frac{1}{2}$.

What's the new equilibrium with this tax?

- Mass $\frac{1}{2}$ of drivers take route 1.
- This generates revenue $\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$, which is distributed equally among all drivers.
- Route 1 drivers face delay $\frac{1}{2}$, pay tax $\frac{1}{2}$, receive transfer $\frac{1}{4}$.
- Route 2 drivers face delay 1, pay no tax, receive transfer $\frac{1}{4}$.
- All drivers receive the same payoff of $-\frac{3}{4}$.
- Thus, the new equilibrium yields the socially optimal loss (and also eliminates inequality).

Congestion Pricing: General Analysis

Ability of congestion pricing to restore efficiency goes far beyond this example.

Key idea: set the toll on link i equal to the externality of using link i, evaluated at the social optimum x^*:

$$t_i = x_i^* l_i' \left(x_i^*\right).$$

- At the social optimum x^*, if you decide to use link i, this slows down x_i^* drivers by $l_i' \left(x_i^*\right)$ each.
- If you have to pay this externality to use link i, your incentives to use link i become perfectly aligned with social welfare.

This is called a **Pigouvian tax**
(or in the congestion pricing context, a **Pigouvian toll**).

Congestion Pricing: General Analysis

Pigouvian toll: impose a toll of $t_i = x_i^* l_i'(x_i^*)$ on each link i.

Theorem

With Pigouvian tolls, the socially optimal routing pattern x^* is also an equilibrium routing pattern.

We will see the same idea in a more general context when we discuss **Vickrey-Clarke-Groves** auctions later in the course.

- The idea that setting taxes equal to externalities restores efficiency is a key insight of economic theory.

General Analysis (cntd.)

Proof for 2-link case gives the intuition:

Two links with latency functions $l_1(x_1), l_2(x_2)$.

Socially optimal routing is given by solution to

$$\min x_1 l_1(x_1) + (1 - x_1) l_2(1 - x_1)$$

General Analysis (cntd.)

Proof for 2-link case gives the intuition:

Two links with latency functions $l_1(x_1), l_2(x_2)$.

Socially optimal routing is given by solution to

$$\min x_1 l_1(x_1) + (1 - x_1) l_2(1 - x_1)$$

Convex function with first-order condition for optimum x^*:

$$l_1(x_1^*) + x_1^* l_1'(x_1^*) = l_2(x_2^*) + x_2^* l_2'(x_2^*),$$

or equivalently

$$\underbrace{l_2(x_2^*) - l_1(x_1^*)}_{\text{net private benefit of taking route 1}} = \underbrace{x_1^* l_1'(x_1^*) - x_2^* l_2'(x_2^*)}_{\text{net externality of taking route 1}}.$$

General Analysis (cntd.)

Suppose we set tolls equal to externalities:

$$\begin{aligned} t_1 &= x_1^* l_1' (x_1^*) \\ t_2 &= x_2^* l_2' (x_2^*) . \end{aligned}$$

Then total cost of using route 1 is $l_1 (x_1^*) + x_1^* l_1' (x_1^*)$, total cost of using route 2 is $l_2 (x_2^*) + x_2^* l_2' (x_2^*)$.

The first-order condition for optimality on the previous slide says precisely that these two costs are equal.

Hence, in equilibrium x_1^* agents take route 1 and x_2^* agents take route 2.

Potential Games

Routing games are a special case of a general class of games called **potential games**.

Recognizing routing games as potential games will let us prove some important results about them, such as existence of pure-strategy equilibrium and an upper bound on the price of anarchy.

Intuitively, a potential game is one in which there exists a function $\phi : S \to \mathbb{R}$, called a **potential function**, such that, for any player i and any two strategies $s_i, s_i' \in S_i$, switching from s_i to s_i' has the same effect on player i's payoff as it has on the potential.

- The potential function thus reflects all players' incentives simultaneously.
- A key reason why this is important is that maxima of the potential function will correspond to equilibria of the game.
- Most games do not admit a potential function, but for those that do it's usually a very helpful object to work with.

Potential Games: Definition

Formally, a function $\phi : S \to \mathbb{R}$ is a **potential function** if, for all $i \in N$, $s_i, s_i' \in S_i$, and $s_{-i} \in S_{-i}$, we have

$$u_i\left(s_i', s_{-i}\right) - u_i\left(s_i, s_{-i}\right) = \phi\left(s_i', s_{-i}\right) - \phi\left(s_i, s_{-i}\right).$$

A game is a **potential game** if it admits a potential function.

Potential Games: Trivial Example

A trivial example of a potential game is a **common-interest game**, where all players have the same payoff function: there exists $u : S \to \mathbb{R}$ such that $u_i(s) = u(s)$ for all $i \in N$ and $s \in S$.

Claim: Every common-interest game is a potential game.

Proof: Simply let $\phi(s) = u(s)$ for all s.

Then, for all $i \in N$, $s_i, s_i' \in S_i$, and $s_{-i} \in S_{-i}$, we have

$$\begin{aligned} u_i\left(s_i', s_{-i}\right) - u_i\left(s_i, s_{-i}\right) &= u\left(s_i', s_{-i}\right) - u\left(s_i, s_{-i}\right) \\ &= \phi\left(s_i', s_{-i}\right) - \phi\left(s_i, s_{-i}\right). \end{aligned}$$

Unfortunately, finding a potential function is often not this easy.

Another Example

Recall the prisoner's dilemma:

	C	D
C	2,2	0,3
D	3,0	1,1

This is a potential game, with potential function given by

	C	D
C	0	1
D	1	2

- To see this, note that whenever a player switches her action from C to D, this increases her own payoff by 1, and also increases the potential by 1.

Note that the strategy profile that maximizes the potential is (D, D), which is also a Nash equilibrium.

- We will see that this is a general feature of potential games.

Potential Games: PSNE Existence

The following theorem is a simple and important example of the power of potential games:

Theorem

Every finite potential game has a pure strategy NE.

Proof:

- Since S is finite, ϕ has a maximizer s^*.
- Since s^* maximizes ϕ, there's no way to increase ϕ by changing any one coordinate s_i.
- Since $\phi\left(s_i, s_{-i}^*\right) - \phi\left(s_i^*, s_{-i}^*\right) = u\left(s_i, s_{-i}^*\right) - u\left(s_i^*, s_{-i}^*\right)$, there's no way to increase u_i by changing s_i.
- Hence, s^* is a PSNE.

Note: Conversely, every PSNE is either a local maximum or a saddle point of ϕ.

Routing Games and Potential Games

Theorem

Every routing game is a (convex) potential game, and therefore has a (unique) pure-strategy NE.

Note:

- We consider here routing games with a finite number n of players, rather than the continuum model we've considered thus far.
- In particular, x_j is now the **number** of agents using link j, not the fraction.
- We'll also now use j subscripts for links and i subscripts for players/numbers of players.

Routing Games and Potential Games (cntd.)

Theorem

Every routing game is a (convex) potential game, and therefore has a (unique) pure-strategy NE.

Proof:

We will show that a certain function $\phi(x)$ is a potential function.

We define $\phi(x)$ to be what total delay would be if the drivers arrived on the roads in sequence, and each driver only suffered the delay due to those drivers who arrived before her: that is,

$$\phi(x) = \sum_{j \in E} \sum_{i=1}^{x_j} l_j(i).$$

Note: this is **not** equal to total delay, which is

$$C(x) = \sum_{j \in E} x_j l_j(x_j).$$

Since each l_j is non-decreasing, we always have $\phi(x) \leq C(x)$.

Intuition for the Potential Function

The potential on link j is $\sum_{i=1}^{x_j} l_j(i)$.

- If a new agent i starts using link j, this increases her travel time by $l_j(x_j+1)$, and also increases the potential on link j by $l_j(x_j+1)$.

The total delay on link j is $x_j l_j(x_j)$.

- If a new agent i starts using link j, this increases the total delay on link j by $l_j(x_j+1) + x_j(l_j(x_j+1) - l_j(x_j))$.

Thus, the potential (but not the total delay) reflects individual agents' incentives to use the link.

- Another way of seeing this is that the increase in the potential equals the increase in total delay minus the externality $x_j(l_j(x_j+1) - l_j(x_j))$.
- If we want to capture incentives, we have to subtract off this externality, since individuals don't take it into account when making their choices.

Proof that We Have a Potential Function

Formally, if driver i switches from path p to path p', the effect on her delay is

$$\sum_{j \in p' \backslash p} l_j\left(x_j + 1\right) - \sum_{j \in p \backslash p'} l_j\left(x_j\right),$$

where $j \in p' \backslash p$ indicates that link j is in path p' but not path p, and similarly for $j \in p \backslash p'$.

By inspection, this is exactly the same as the effect on ϕ. Hence, ϕ is a potential function.

Note: To see that we could not have just taken C as the potential function, note that the effect of i switching from p to p' on C is something different:

$$\sum_{j \in p' \backslash p} \left[l_j\left(x_j + 1\right) + x_j\left(l_j\left(x_j + 1\right) - l_j\left(x_j\right)\right)\right]$$
$$- \sum_{j \in p \backslash p'} \left[l_j\left(x_j\right) + \left(x_j - 1\right)\left(l_j\left(x_j\right) - l_j\left(x_j - 1\right)\right)\right].$$

Potentials and the Price of Anarchy

We now use the potential game approach to prove the following important result: with affine latency functions, there is always an equilibrium that is not "too inefficient."

A latency function $l_j(x_j)$ is **affine** if it can be written as $l_j(x_j) = a_j x_j + b_j$ for constants $a_j, b_j \geq 0$.

Theorem

In any routing game with affine latency functions, $POA \leq 2$.

Price of Stability: Comments

Theorem

In any routing game with affine latency functions, $POA \leq 2$

- The simple example at the very beginning of lecture has affine latency and $POA = 4/3$.
- The theorem can actually be improved to say that $POA \leq 4/3$. So the simple example is actually the worst possible!
- Rough intuition: negative externalities are “as strong as possible” in the simple network.
- In contrast, we’ve seen that, with general polynomial latency functions, the price of anarchy can be arbitrarily high.

Price of Stability: Proof

Let x^* be a socially optimal routing (i.e., a routing that minimizes C), and let x^E be a routing that minimizes the potential ϕ (and hence is a PSNE).

We know that, for any routing x, we have $\phi(x) \leq C(x)$.

We will prove that, for any routing x, we have $\phi(x) \geq C(x)/2$.

Then we're done: we have

$$C\left(x^E\right) \underbrace{\leq}_{C \leq 2\phi} 2\phi\left(x^E\right) \underbrace{\leq}_{x^E \text{ minimizes } \phi} 2\phi(x^*) \underbrace{\leq}_{\phi \leq C} 2C(x^*).$$

That is, total delay at x^E is no more than twice the socially optimal delay.

Price of Stability: Proof (cntd.)

The fact that $\phi(x) \geq C(x)/2$ follows from the assumption of affine latency functions: $l_j(x_j) = a_j x_j + b_j$ for all j.

Recall that $C(x)$ is total delay, while $\phi(x)$ is total delay when drivers arrive in sequence and each driver only suffers the delay caused by earlier drivers.

- Total delay on link j equals

$$(a_j x_j + b_j)\, x_j = a_j x_j^2 + b_j x_j.$$

- Potential on link j equals

$$\sum_{i=1}^{x_j} l_j(i) = a_j (1 + 2 + \ldots + x_j) + b_j x_j.$$

- Note that $1 + 2 + \ldots + x_j = x_j (x_j + 1)/2 \geq x_j^2/2$.
- Hence, potential is at least (total delay)/2.

Summary

- Routing games are an important application of Nash equilibrium to networks, especially for understanding transportation and information routing.
- Equilibrium routing is typically inefficient, as individual agents do not take into account their contributions to congestion when making decisions.
- With non-affine latency functions, this inefficiency can be arbitrarily severe.
- With affine latency functions, inefficiency cannot be "too bad" (however, in reality a factor of 2 or $\frac{4}{3}$ is not great!).
- Increasing capacity does not always help, as shown by Braess's paradox.
- Congestion pricing presents a general solution.
- Routing games are an important example of potential games, an important general class of games with nice theoretical properties.

14.15 / 6.207 Networks
Spring 2022