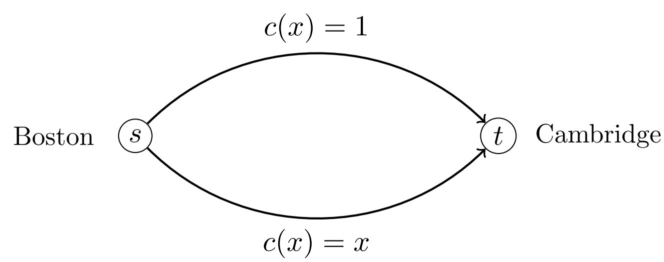
DS 574 Algorithmic Mechanism Design	**Congestion Pricing**
Fall 2026	Prof. Kira Goldner

Selfish Routing

You're driving from Boston to Cambridge and have two routes to choose from. The highway takes one hour regardless of how many drivers use it. The bridge is faster when it is uncrowded, but slows down as more drivers choose it: if a fraction x of the traffic takes the bridge, crossing takes x hours. We normalize the total traffic to one unit, so a fraction x is also an amount of traffic x. Each driver wants to get to Cambridge as quickly as possible.



When we defined Nash equilibria, we asked whether any player could benefit by changing their own action. Here a driver's action is a route. An equilibrium means that no driver can get to Cambridge faster by switching routes.

We model a large population in which each driver is an infinitesimal part of the traffic. One driver switching routes does not change either route's travel time. This is called *nonatomic* traffic. Thus a driver simply compares the travel times on the available routes at the current traffic levels.

A driver's objective. Each driver chooses their own route to minimize their own travel time, given the routes everyone else is taking. If fraction x is on the bridge, the driver compares x hours on the bridge with 1 hour on the highway. They choose a route, not the overall traffic split x. Their objective does not include the time other drivers spend on the road.

Definition 1. An *equilibrium flow*, or *Wardrop equilibrium*, is a routing of traffic in which no driver can reduce their travel time by changing routes. Every route that is used must be a fastest available route: used routes have equal travel times, and an unused route cannot be faster.

A *flow* just describes how the traffic is split among routes. In this example, putting x on the bridge leaves $1 - x$ on the highway. Write $C(f)$ for the total time spent driving under a flow f.

Pigou's example. Where will drivers go in equilibrium?

Claim 1. In Pigou's example, the unique equilibrium flow routes all of the traffic onto the bottom (bridge) edge. Every driver's travel time is 1, and $C(f) = 1$.

Proof. Write the flow as $(a, 1 - a)$, meaning a units on the top edge and $1 - a$ on the bottom. The top edge costs 1 and the bottom edge costs $1 - a$.

Nothing with $a > 0$ works. If $a > 0$ then $1 - a < 1$, so the bottom path is strictly cheaper than the top path, which is used. That violates the equilibrium condition.

And $a = 0$ does. There the bottom edge carries all one unit and costs 1, and the top edge costs 1. Both paths cost exactly 1, so no driver on a used path is paying more than the cheapest available path. Every driver takes an hour, so $C(f) = 1 \cdot 1 = 1$. $\square$

The designer's objective. The designer chooses how to split the traffic among routes to minimize the total time everyone spends driving. Unlike an individual driver, the designer can reroute a positive fraction of the traffic and accounts for how that changes travel times for everyone. An *optimal flow* is a traffic split that minimizes this total, whether or not drivers would choose to follow the assignment.

Write $C(f)$ for the *total travel time*, or social cost, of a flow f. To compute it, multiply the amount of traffic on each road by the time spent on that road, then add. Since total traffic is one unit, total and average travel time coincide.

In the highway-and-bridge example, the designer chooses $x \in [0, 1]$, the fraction assigned to the bridge, to minimize

$$C(x) = \underbrace{x \cdot x}_{\text{time spent by bridge drivers}} + \underbrace{(1 - x) \cdot 1}_{\text{time spent by highway drivers}}.$$

Sending fewer drivers over the bridge makes it faster for those who remain, but requires more drivers to spend an hour on the highway.

Claim 2. The optimal flow in Pigou's example splits the traffic evenly: 1/2 on each edge. Its total cost is 3/4.

Proof. Put x on the bottom and $1 - x$ on top. Then

$$C(x) = \underbrace{x \cdot x}_{\text{bridge}} + \underbrace{(1 - x) \cdot 1}_{\text{highway}} = x^2 - x + 1 = \left(x - \tfrac{1}{2}\right)^2 + \tfrac{3}{4}.$$

A square is nonnegative and equals zero only at $x = 1/2$, so $C(x) \geq 3/4$ with equality exactly at $x = 1/2$. There, $C = \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot 1 = \frac{3}{4}$. At this split, a highway driver takes an hour but could take the bridge in half an hour. So the optimal flow is not an equilibrium: the split minimizes total travel time, but the highway drivers are not minimizing their own travel time. $\square$

How much worse can selfish routing be than the best assignment of routes? We measure this using the price of anarchy.

Definition 2 (Koutsoupias–Papadimitriou '99)**.** The *price of anarchy* of an instance is

$$\text{PoA} = \frac{C(\text{worst equilibrium flow})}{C(\text{optimal flow})}.$$

What is the price of anarchy?

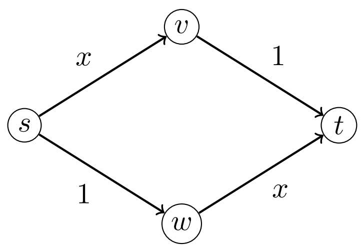
$$\text{PoA} = \frac{1}{3/4} = \frac{4}{3}.$$

Equilibrium travel time is one third larger than optimal travel time.

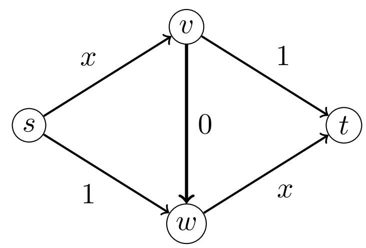
Braess's Paradox

Could adding a road make traffic worse? A planner could always leave the new road unused, so the best possible routing cannot get worse. But drivers choose their own routes. A new shortcut may change what they choose.

Consider two routes, each with one road whose travel time equals its traffic and one road that takes an hour. We still have one unit of traffic. First find the equilibrium without the middle road; then add a zero-cost road from v to w.



Before



After

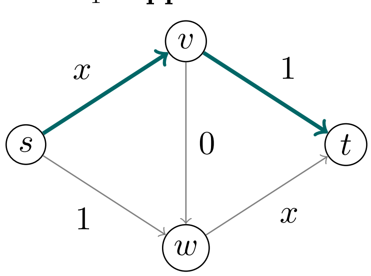
On a network, a route may use several roads. Write f_e for the traffic on road e and $c_e(f_e)$ for its travel time. The cost $c_P(f)$ of a route P is the sum of its roads' travel times. Total travel time is

$$C(f) = \sum_e \underbrace{f_e}_{\text{traffic on road } e} \underbrace{c_e(f_e)}_{\text{time per driver on } e}.$$

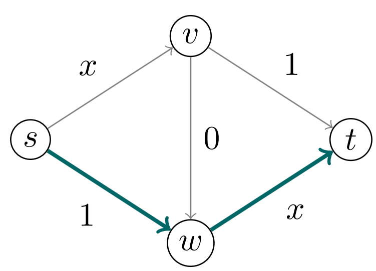
A driver using several roads contributes the time spent on each; adding road by road counts that driver's whole trip.

The highlighted roads show the three available routes. We use P_1, P_2, P_3 to refer to them in the calculations.

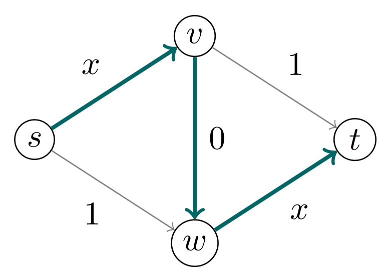
P_1: upper route



P_2: lower route



P_3: shortcut



Claim 3. In the "before" network, the unique equilibrium flow sends $1/2$ along each of P_1 and P_2. Every driver's travel time is $3/2$, and $C(f) = 3/2$.

Proof. Let a be the flow on P_1 and $1 - a$ the flow on P_2. Then $f_{sv} = a$ and $f_{wt} = 1 - a$, so

$$c_{P_1}(f) = a + 1, \qquad c_{P_2}(f) = 1 + (1 - a) = 2 - a.$$

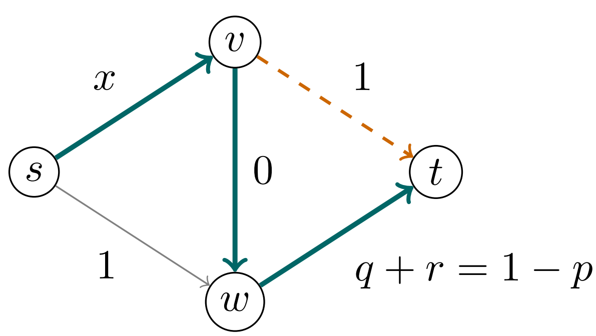
If $a > 1/2$, then $a + 1 > 2 - a$, so the drivers on P_1 are on a strictly more expensive path than P_2 and would switch; not an equilibrium. If $a < 1/2$, the same argument applies with the roles reversed. So $a = 1/2$ is the only candidate, and there $c_{P_1} = c_{P_2} = 3/2$, so both used paths cost the same and it is an equilibrium. Every driver takes $3/2$ hours, so $C(f) = 3/2$. $\square$

Theorem 1 (Braess's paradox)**.** *In the "after" network, the unique equilibrium flow routes* all *of the traffic along P_3. Every driver's travel time is 2, and $C(f) = 2$. Adding a free road made every single driver strictly worse off, from $3/2$ to 2.*

Proof. Write $p, q, r \geq 0$ for the flow on P_1, P_2, P_3, with $p + q + r = 1$. Then $f_{sv} = p + r$ and $f_{wt} = q + r$, so the three path costs are

$$c_{P_1} = (p + r) + 1, \qquad c_{P_2} = 1 + (q + r), \qquad c_{P_3} = (p + r) + 0 + (q + r).$$

Step 1: $r = 1$ is an equilibrium. Set $p = q = 0$, $r = 1$. Then $f_{sv} = f_{wt} = 1$, so $c_{P_1} = 2$, $c_{P_2} = 2$, and $c_{P_3} = 1 + 0 + 1 = 2$. All three paths cost exactly 2, so the only used path, P_3, is a cheapest path. It's an equilibrium, and every driver takes 2 hours.



The first road is unchanged. Replace the dashed one-hour road with the highlighted shortcut and final road, which together take $1 - p$ hours.

Step 2: why would anyone stay on an old route? Suppose $p > 0$ drivers use P_1. They take the road $s \to v$ and then spend an hour on $v \to t$. Instead, they could take the free middle road and finish on $w \to t$. That last road carries $q + r = 1 - p < 1$ units of traffic, so it takes less than an hour. Switching saves time:

$$c_{P_1} - c_{P_3} = 1 - (q + r) = p > 0.$$

Thus an equilibrium cannot have $p > 0$. Similarly, a driver on P_2 could replace the one-hour road $s \to w$ with $s \to v \to w$, saving $1 - (p + r) = q$ hours whenever $q > 0$. So $q = 0$ as well. All traffic must use P_3, and we already checked that this is an equilibrium. $\square$

Why is the old equilibrium no longer an equilibrium?

At $p = q = 1/2$ and $r = 0$, the old routes each cost $3/2$, but P_3 costs 1. Switching to the new path is a beneficial deviation.

Claim 4. The optimal flow in the "after" network has total cost $3/2$, achieved by $p = q = 1/2$, $r = 0$: route nothing at all through the new edge.

Proof. The total cost of the flow (p, q, r) is

$$C(p,q,r) \;=\; \underbrace{(p+r)^2}_{s\to v} + \underbrace{p\cdot 1}_{v\to t} + \underbrace{q\cdot 1}_{s\to w} + \underbrace{(q+r)^2}_{w\to t} + \underbrace{r\cdot 0}_{v\to w}.$$

Since $p+q+r = 1$, we have $p+r = 1-q$ and $q+r = 1-p$, so $C = (1-q)^2 + q + (1-p)^2 + p$, which has separated into two identical pieces. For any t,

$$(1-t)^2 + t \;=\; t^2 - t + 1 \;=\; \left(t - \tfrac{1}{2}\right)^2 + \tfrac{3}{4},$$

which is exactly the computation from Claim 2. Therefore

$$C \;=\; \left(p - \tfrac{1}{2}\right)^2 + \left(q - \tfrac{1}{2}\right)^2 + \tfrac{3}{2} \;\geq\; \tfrac{3}{2},$$

with equality exactly when $p = q = 1/2$, which forces $r = 0$. That flow is feasible, so the minimum is $3/2$. $\square$

What is the price of anarchy after adding the road?

$$\mathrm{PoA} = \frac{2}{3/2} = \frac{4}{3}.$$

The optimum is unchanged; the equilibrium is worse.

Both examples gave a price of anarchy of $4/3$. Could a larger network be worse, even if each road still has a linear travel-time function plus a constant?

Theorem 2 (Roughgarden–Tardos '02)**.** *Consider any selfish routing instance—any directed graph, any number of source-sink pairs, any traffic rates—in which every cost function is* affine*, i.e. of the form $c_e(x) = a_e x + b_e$ with $a_e, b_e \geq 0$. Then PoA $\leq 4/3$. Pigou's example shows this bound is tight.*

The general affine bound is stated without proof. Pigou and Braess attain it; other affine instances can have a smaller price of anarchy.

Interventions

In the optimal split for the highway and bridge, drivers on the highway would rather take the bridge. How could we make them willing to stay on the highway?

A toll on the bridge. Charge a toll equivalent to half an hour of travel time on the bridge and no toll on the highway. A driver then compares $x + 1/2$ on the bridge with 1 on the highway. At $x = 1/2$, both cost 1. If $x < 1/2$, highway drivers would switch to the bridge; if $x > 1/2$, bridge drivers would switch to the highway. Thus the optimal split is the unique equilibrium with this toll. The toll changes what drivers minimize without changing the physical travel times.

Acknowledgements

This lecture was developed in part using materials by Tim Roughgarden, and in particular, his book "Twenty Lectures on Algorithmic Game Theory" [Roughgarden, 2016].

References

Tim Roughgarden. *Twenty Lectures on Algorithmic Game Theory*. Cambridge University Press, 2016.