Algorithmic Game Theory	Fall 2019, Week 5

Price of Stability
and
Introduction to Mechanism Design

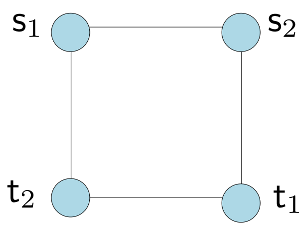
ETH Zürich Paolo Penna

This is the lecture where we start *designing* systems which involve selfish players. Roughly speaking, we consider two approaches:

- Suggest the players the **best** possible equilibrium, and measure how well this approach performs (Price of Stability);
- Provide **compensations** to the players to induce them to to behave in a desirable way (Mechanism Design).

1 Price of Stability

Example 1 (necessity of payment rules)**.** *In this graph, each edge can be built at a cost of* 1*. Each player i wants to go from s_i to t_i, and can pay a part of the cost for building the necessary edges.*



An edge exists if its cost is covered. If a player cannot connect to its target node t_i then he/she has a huge cost. The strategy of a player i is to specify, for each edge e, how much he/she contributes (the cost for the player is his/her total contribution).
*This game has **no pure Nash** equilibria (**Exercise!**).*

The above example suggests that we need rules on how the cost of an edge is shared among the players that use it. A natural scheme is to divide the cost of an edge equally among those that use it. This leads to so-called *fair cost sharing games*, which are congestion games with 'delays'

$$d_r(x) = c_r/x$$

where $c_r \geq 0$ is the cost for building resource r. We have n players and m resources. Player i selects some resources, i.e., his strategy set is $S_i \subseteq [1, \dots, m]$. The cost c_r is assigned in equal shares to the players allocating r (if any). That is, for strategy profile s denote by $n_r(s)$ the number of players using resource r. Then the cost $c_i(s)$ of player i

is $c_i(s) = \sum_{r \in s_i} d_r(n_r(S)) = \sum_{r \in s_i} c_i/n_r(s)$. As in the previous lectures, the *social cost* of a state s is the sum of all players' costs

$$cost(s) = \sum_i c_i(s).$$

Social cost = Costs of used resources

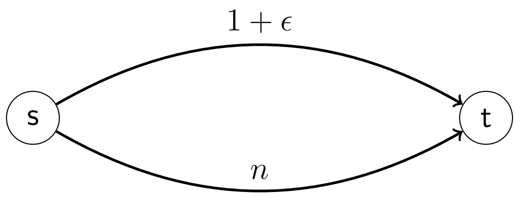
$$cost(s) = \sum_i c_i(s) = \sum_{r:\ n_r(s) \geq 1} c_r \quad . \tag{1}$$

Here is the simple proof:

$$cost(s) = \sum_i c_i(s) = \sum_i \sum_{r \in s_i} d_r(n_r(s)) = \sum_{r:\ n_r(s) \geq 1} n_r(s) \cdot c_r/n_r(s) = \sum_{r:\ n_r(s) \geq 1} c_r \quad .$$

The price of anarchy for pure Nash equilibria can be as big as the number of players n, even in a symmetric game.

Example 2 (Lower bound on price of anarchy)**.** *For $\epsilon > 0$, consider the following network cost sharing game, in which edge labels indicate the cost c_r of this resource:*



It is a pure Nash equilibrium if all players use the bottom edge, whereas the social optimum would be that all users use the top edge.

Although this is a very stylized example, there are indeed examples of such bad equilibria occurring in reality. A prime example are mediocre technologies, which win against better ones just because they are in the market early and get their share. As a concrete example consider social networks or messaging services. Even if you were to design and code-up a revolutionary new social network or messaging service, would people actually switch over from, say, Facebook or Whatsapp, if all their friends are using it?

1.1 Best Equilibria and Price of Stability

Let us first turn to the *price of stability*, which compares the **best** equilibrium with the social optimum. We can see this this concept in two ways:

- Try to reduce the negative effect of selfish behavior by suggesting a good equilibrium to the players (rather than let them compute an equilibrium by themselves);
- Understand if the inefficiency loss is just a question of equilibrium selection, or it is intrinsic in the fact that players will always play some equilibrium.

Definition 3 (**Price of Stability**). *For a cost-minimization game, the* price of stability for Eq *is defined as*

$$PoS_{\mathsf{Eq}} = \frac{\min_{p \in \mathsf{Eq}} cost(p)}{\min_{s \in S} cost(s)} \quad ,$$

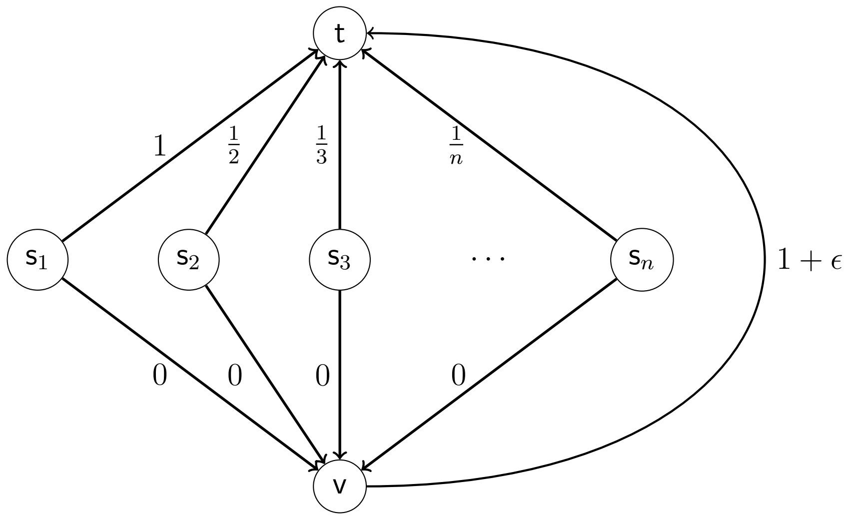
where Eq *is a set of probability distributions over the set of states* S*, and* $cost(p)$ *is the expected social cost of* p*.*[a]

[a] Recall definition from previous lectures: $cost(p) := \mathbf{E}_{s \sim p}[cost(s)] = \sum_{s \in S} p(s) cost(s)$.

Exercise 1. *Prove that in a symmetric cost sharing game every social optimum is a pure Nash equilibrium. Therefore, the price of stability for pure Nash equilibria is* 1*.*

In asymmetric games this is no longer true as the social optimum is not necessarily a pure Nash equilibrium.

Example 4 (Lower bound on price of stability)**.** *Consider the following game with* n *players. Each player* i *has source node* s_i *and destination node* t*.*



A player has two possible strategies: Either take the direct edge or take the detour via v*. The social optimum lets all players choose the indirect path, which leads to a social cost of* $1 + \epsilon$*. This, however, is not a Nash equilibrium. Player* n*, who currently faces a cost of* $(1 + \epsilon)/n$*, would opt out and take the direct edge, which would give him cost* $1/n$*.*

Therefore, the only pure Nash equilibrium lets all players choose their direct edge, yielding a social cost of $\mathcal{H}_n$*, where* $\mathcal{H}_n = \sum_{i=1}^{n} \frac{1}{i} = 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n}$ *is the* n*-th harmonic number. We have* $\mathcal{H}_n = \Theta(\log n)$*.*

Theorem 5. *The price of stability for pure Nash equilibria in fair cost sharing games is at most* $\mathcal{H}_n$*.*

Proof. Let us first derive upper and lower bounds on Rosenthal's potential function that apply to any state s, whether at equilibrium or not.

To obtain an upper bound we can use the definition of the potential function and the fact that each resource is used by no more than n players:

$$\begin{aligned}\Phi(s) = \sum_r \sum_{k=1}^{n_r(s)} \frac{c_r}{k} &= \sum_{r:\ n_r(s)\geq 1} c_r \cdot \left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n_r(s)}\right) \qquad (2)\\ &\leq \sum_{r:\ n_r(s)\geq 1} c_r \cdot \mathcal{H}_n \\ &= cost(s) \cdot \mathcal{H}_n \ .\end{aligned}$$

As for the lower bound, each term in the right hand side of (2) is at least c_r, thus

$$\Phi(s) \geq \sum_{r:\ n_r(s)\geq 1} c_r = cost(s) \ .$$

Now let s' be a state minimizing the potential, and s^* be a state minimizing the social cost. By definition, $\Phi(s') \leq \Phi(s^*)$, thus implying

$$cost(s') \leq \Phi(s') \leq \Phi(s^*) \leq cost(s^*) \cdot \mathcal{H}_n \ .$$

Note that s' must be a pure Nash equilibrium (because it is a local minimum for the potential). This proves that there is a pure Nash equilibrium s' that is only a factor of $\mathcal{H}_n$ more costly than s^*. □

We conclude that in both symmetric and asymmetric cost sharing games our assumption that players play an equilibrium has a rather mild impact. The possible source of inefficiency is the fact that players may not be able to select a particularly good equilibrium.

Another use of Potential: bound the price of stability

Theorem 6. *In a cost-minimization potential game, if s' is a state minimizing the potential, then*

$$PoS \leq \frac{cost(s')}{\min_{s\in S} cost(s)} \ .$$

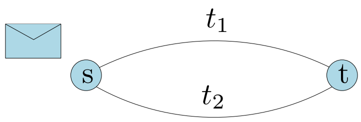
Proof. In potential games every state which is a local minimum for the potential is a pure Nash equilibrium (in particular, the global minimum s'). □

2 Introduction to Mechanism Design

In this section we introduce **mechanisms with money**. The main idea is that we provide compensations to the player in order to induce them to behave as we desire.

2.1 Examples and Main Intuitions

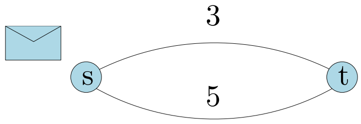
Problem (2 Links, 1 Packet)



We want to select the best (minimum cost), but the following complication arises:

- Each link is a player i and t_i is the time (latency) that this link 'suffers' if we use it for sending our message;
- We do not know t_i and we thus have to ask the player to report to us its cost;
- Players can **cheat** and report a cost which is different from the true one.

The reason for a player i to cheat is simple: if its link is selected, then the player has a cost t_i, while if its link is not selected, there is no cost for the player. For instance, if these were the true costs:



the top player has an incentive to declare a higher cost (say 6) so that our algorithm selects the other link. It is therefore natural to introduce compensations or **payments**.

First attempt: Pay the reported cost

Suppose we ask each player what is the cost of its link, and we provide that amount of money as compensation. Now the top link still has a reason to cheat: by reporting 4.99 he/she can get a higher payment (can speculate). Consider his **utility** in the two cases:

- Truth-telling (report 3): receives 3 as payment, and has cost 3 for being selected;

  Utility = Payment - Cost $= 3 - 3 = 0$.

- Cheating (report 4.99): receives 4.99 as payment, and has cost 3 for being selected.

  Utility = Payment - Cost $= 4.99 - 3 > 0$.

Surprisingly there are payments that always induce the players to tell the truth. The idea comes from the following type of auctions for selling one item to potential buyers.

2nd Price Auction (Vickrey)

1. Pick the highest offer (bid)
2. The winner pays the 2nd-highest offer

In this auction, each buyer sends his/her offer in a sealed envelope, and the winner is determined as above. Here buyers have a private valuation for the item, and they naturally try to get the item for a lower price (utility = valuation − price). For example:

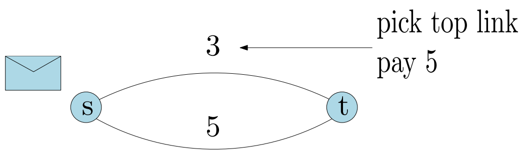
> Bids: 7, 10, 3
> Winner: second player
> Price: 7

Note for instance, that the second buyer cannot get the item for a better price, e.g., by bidding 8; bidding 6 would make this buyer to lose the auction.

This translates immediately into the following solution for our problem:

Second Attempt (Vickrey auction):

1. Select the best link (min reported cost)
2. Pay this link an amount equal to the 2nd-best cost

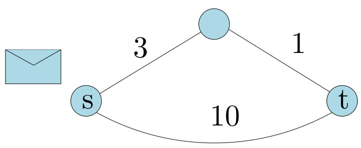


One can argue on this example that neither player can improve his/her utility by cheating. In fact this is true no matter which were the true costs of the two links. We prove this and more general results below. Now we want to build some intuition and generalize this problem.

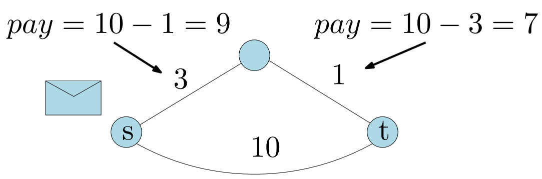
Observation 7. *In the example above, if we were paying the reported cost, the maximum this agent could speculate is 5 (the cost of the other link). To prevent him/her from cheating, we pay directly this amount to him/her.*

Let us consider the same problem, just slightly more general.

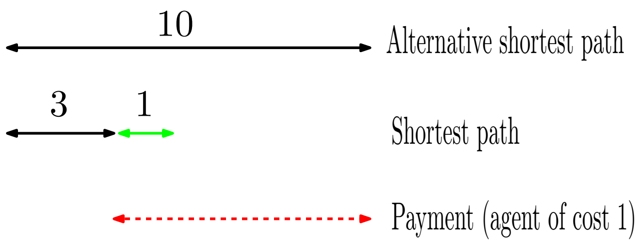
Problem (Shortest Path)



Imagine we pay each selected link an amount equal to the maximum he/she could speculate if the payment was the reported cost:



which can be also seen as follows,



This is perhaps the main intuition we use in a general construction described in Section 2.3.

2.2 Formal Setting

We consider the following setting.

- $\mathcal{A}$ is a set of feasible alternatives (or solutions)
- Each player i has private true cost function $t_i : \mathcal{A} \to \Re$ which gives his/her cost for every possible alternative,
$$t_i(a).$$
- Player i can misreport his/her true cost to some other cost function $c_i : \mathcal{A} \to \Re$.
- If agent i receives a payment p_i and alternative a is chosen, then his/her utility is
$$p_i - t_i(a).$$

A **mechanism** is a combination of an **algorithm** selecting a solution and a **payment** rule assigning payments to each agent. The mechanism takes in input the costs reported by the players.

A **mechanism** is a pair (A, P) which on input the costs $c = (c_1, \ldots, c_n)$ reported by the players, outputs

- A solution $A(c) \in \mathcal{A}$;
- A payment $P_i(c)$ for each player i.

The corresponding **utility** for each agent i is

$$u_i(c|t_i) := P_i(c) - t_i(A(c)).$$

What we want is that cheating is never convenient for a player. That is, **truth-telling** is a **dominant strategy**.

A mechanism (A, P) is **truthful** if

$$u_i(c_1, \ldots, c_{i-1}, t_i, c_{i+1}, \ldots, c_n | t_i) \geq u_i(c_1, \ldots, c_{i-1}, c_i, c_{i+1}, \ldots, c_n | t_i) \quad (3)$$

for all i, for all $\underbrace{t_i}_{\text{true cost}}$, for all $\underbrace{c_1, \ldots, c_{i-1}, c_{i+1}, \ldots, c_n}_{\text{costs reported by others}}$, and for all $\underbrace{c_i}_{\text{cost reported by } i}$.

2.3 VCG Mechanisms

We next present a very general construction of truthful mechanisms which will solve the shortest-path problem as a special case. Intuitively, the mechanism does the following:

- Compute the solution minimizing the social cost (sum of players costs) with respect to the reported costs (algorithm);
- Pay each agent his/her "marginal contribution" to this solution (payments). This part uses the intuition described at the end of Section 2.1.

Let us consider the **social cost** of a solution a with respect to costs $c = (c_1, \ldots, c_n)$ as

$$cost(a, c) := \sum_i c_i(a)$$

and let us denote the **optimum** as

$$opt(c) := \min_{a \in \mathcal{A}} cost(a, c).$$

A **VCG mechanism** is a pair (A, P) such that

- A in an optimal algorithm:

$$cost(A(c), c) = opt(c) \qquad \text{for all } c;$$

- P is of the following form:

$$P_i(c) = Q_i(c_{-i}) - \sum_{j \neq i} c_j(A(c))$$

where Q_i is an arbitrary function independent of c_i.

This construction yields truthful mechanisms:

Theorem 8 (Vickrey-Clarke-Groves)**.** *VCG mechanisms are truthful.*

Proof. Fix an agent i, and the costs c_{-i} reported by the others. Let

$$\tilde{t} := (c_1, \ldots, c_{i-1}, t_i, c_{i+1}, \ldots, c_n)$$

be the vector in which i is truth-telling.

Claim 9. *The utility of i for $\tilde{t}$ is*

$$u_i(\tilde{t}|t_i) = Q_i(c_{-i}) - opt(\tilde{t}).$$

Proof. By definition of the payments

$$u_i(\tilde{t}|t_i) = P_i(\tilde{t}) - t_i(A(\tilde{t})) = Q_i(c_{-i}) - \left(\sum_{j\neq i} c_j(A(c)) + t_i(A(\tilde{t}))\right)$$

Observe that for any solution $a \in \mathcal{A}$

$$\begin{aligned}\sum_{j\neq i} c_j(a) + t_i(a) =& c_1(a) + \cdots + c_{i-1}(a) + t_i(a) + c_{i+1}(a) + \cdots + c_n(a) \\ =& cost(a, \tilde{t}).\end{aligned}$$

Therefore

$$u_i(\tilde{t}|t_i) = Q_i(c_{-i}) - cost(A(\tilde{t}, \tilde{t})) = Q_i(c_{-i}) - opt(\tilde{t}).$$

This proves the claim. □

Claim 10. *The utility of i for c is*

$$u_i(c|t_i) = Q_i(c_{-i}) - cost(A(c), \tilde{t}).$$

Proof. The same proof of the previous claim with c_i in place of t_i yields the desired result. □

Now truthfulness is immediate from the two claims: since by definition of opt

$$opt(\tilde{t}) \leq cost(A(c), \tilde{t})$$

we get

$$u_i(\tilde{t}|t_i) = Q_i(c_{-i}) - opt(\tilde{t}) \geq Q_i(c_{-i}) - cost(A(c), \tilde{t}) = u_i(c|t_i).$$

□

Exercise 2. *Show that the mechanisms for shortest-path problem in Section 2.1 is VCG mechanism (and thus truthful no matter the costs of the edges). Describe the mechanism on general graph, assuming that the graph given is 2-connected (removing one edge never disconnects the graph). What is the function $Q_i()$?*

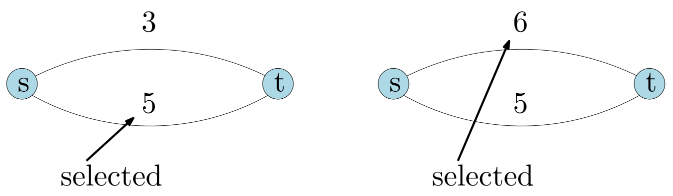
Exercise 3. *The above theorem is a weaker form of the original VCG theorem. Prove that the following weaker condition on the algorithm is sufficient. There exists a subset $R \subset \mathcal{A}$ of solutions such that A is optimal with respect to this fixed subset of solutions. That is, for every c*

- $A(c) \in R$*;*
- $cost(A(c), c) = \min_{a\in R} cost(a, c)$.

2.4 Longest Path

We conclude by describing a simple problem which does *not* have a truthful mechanism, no matter how clever or complicated we make the payments.

Longest Path: Find the longest path (instead of shortest-one)



Consider the top link, and any payment P_i for it. Truthfulness would require to deal with both these cases:

1. (true cost = 3, false cost = 6):

$$P_i(3,5) - 0 \geq P_i(6,5) - 3$$

2. (true cost = 6, false cost = 3):

$$P_i(6,5) - 6 \geq P_i(3,5) - 0$$

which cannot be both satisfied. So there is no truthful mechanism (A, P) for this problem.

Recommended Literature

For the Price of Stability see

- **Chapter 17.2.2** and **Chapter 19.3** in *Algorithmic Game Theory*, N. Nisan et al., pages 79–101, 2007.

For truthful mechanisms with money:

- **Chapter 2 (2*n*d price auction)** and **Chapter 7 (VCG)** in Tim Roughgarden, *Twenty Lectures on Algorithmic Game Theory*, Cambridge University Press, 2016 (both for problems with private valuations).
- **Chapter 9** in *Algorithmic Game Theory*, N. Nisan et al., pages 79–101, 2007.