

CS364A: Algorithmic Game Theory

Lecture #15: Best-Case and Strong Nash Equilibria*

Tim Roughgarden[†]

November 11, 2013

1 Network Cost-Sharing Games

1.1 Lecture Themes

The *externality* caused by a player in a game is the difference between its own objective function value and its contribution to the social objective function value. The models we've looked at thus far have *negative* externalities, meaning that players cause more harm to the system than they realize (or choose to care about). In a routing game, for example, a player does not take into account the additional cost its presence causes for the other players using the edges in its path.

There are also important applications that exhibit *positive* externalities. You usually join a campus organization or a social network to derive personal benefit from it, but your presence (hopefully) also enriches the experience of other people in the same group. As a player, you're generally bummed to see new players show up in a game with negative externalities, and excited for the windfalls of new players in a game with positive externalities. The first theme of this lecture is the study of positive externalities, in a concrete model of network formation.

In the model we study, there will generally be multiple pure Nash equilibria. We're used to that, from routing and location games, but here different equilibria can have wildly different costs. This motivates confining attention to a subset of "reasonable" Nash equilibria that hopefully possesses two properties: first, better worst-case inefficiency bounds should hold for the subset than for all equilibria; second, there should be a plausible narrative as to why equilibria in the subset are more worthy of study than the others. No fully satisfactory approach to this problem is known in the model we study, but we'll cover two partially successful approaches.

*©2013, Tim Roughgarden. These lecture notes are provided for personal use only. See my book *Twenty Lectures on Algorithmic Game Theory*, published by Cambridge University Press, for the latest version.

[†]Department of Computer Science, Stanford University, 462 Gates Building, 353 Serra Mall, Stanford, CA 94305. Email: tim@cs.stanford.edu.

1.2 The Model

A *network cost-sharing game* takes place in a graph $G = (V, E)$, which can be directed or undirected, and each edge $e \in E$ carries a fixed cost $\gamma_e \geq 0$. There are k players. Player i has a source vertex $s_i \in V$ and a sink vertex $t_i \in V$, and its strategy set is the s_i - t_i paths of the graph. Outcomes of the game correspond to path vectors (P_1, \dots, P_k) , with the semantics that the subnetwork $(V, \cup_{i=1}^k P_i)$ gets formed.

We think of γ_e as the fixed cost of building the edge e — laying down high-speed Internet fiber to a neighborhood, for example — and this cost is independent of the number of players that use the edge. Players' costs are defined edge-by-edge, like in routing games. If $f_e \geq 1$ players use an edge e in their chosen paths, then they are jointly responsible for the edge's fixed cost γ_e . In this lecture, we assume that this cost is split equally amongst the players. That is, in the language of cost-minimization games (Lecture 13), the cost $C_i(\mathbf{P})$ of player i in the outcome \mathbf{P} is

$$C_i(\mathbf{P}) = \sum_{e \in P_i} \frac{\gamma_e}{f_e}, \quad (1)$$

where $f_e = |\{j : e \in P_j\}|$ is the number of players that choose a path that includes e . The global objective is simply to minimize the total cost of the formed network:

$$\text{cost}(\mathbf{P}) = \sum_{e \in E: f_e \geq 1} \gamma_e. \quad (2)$$

Note that, analogous to routing games, the objective function (2) can equally well be written as the sum $\sum_{i=1}^k C_i(\mathbf{P})$ of the players' costs.

Remark 1.1 This is a very stylized game-theoretic model of how a network might form. Many such models have been studied, but it is quite difficult to capture observed properties of real-world networks with an analytically tractable model. See Jackson [5] for a textbook treatment of network formation games.

1.3 Example: VHS or Betamax

Let's build our intuition for network cost-sharing games through a couple of examples. The first example demonstrates how tragic miscoordination can occur in games with positive externalities.

Consider the simple network in Figure 1, with k players with a common source s and sink t . One can interpret an edge as the adoption of a particular technology. For example, back in the 1980s, there were two new technologies enabling home movie rentals. Betamax was lauded by technology geeks as the better one — corresponding to the lower-cost edge in Figure 1 — and VHS was the other one. VHS grabbed a larger market share early on. Since coordinating on a single technology proved the primary driver in consumers' decisions — have the better technology is little consolation for being unable to rent anything from your corner store — Betamax was eventually driven to extinction.

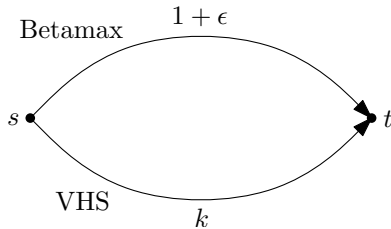


Figure 1: VHS or Betamax. The price of anarchy in a network cost-sharing game can be as large as the number k of players.

The optimal solution in Figure 1 is for all players to pick the upper edge, for a total cost of $1 + \epsilon$. This is also a Nash equilibrium. Unfortunately, there is a second Nash equilibrium, in which all players pick the lower edge. Since the cost of k is split equally, each player pays 1. If a player deviated unilaterally to the upper edge, it would pay the full cost $1 + \epsilon$ of that edge and thus suffer a higher cost. This example shows that the price of anarchy in network cost-sharing games can be as high as k , the number of players. (For a matching upper bound, see the Exercises.)

The VHS or Beta example is exasperating. We proposed a reasonable network model capturing positive externalities, and the price of anarchy — which helped us reason about several models already — is distracted by a manifestly unreasonable Nash equilibrium and yields no useful information. Can we salvage our approach by focusing only on the “reasonable” equilibria? We’ll return to this question after considering another important example.

1.4 Example: Opting Out

Consider the network cost-sharing game shown in Figure 2. The k players have distinct sources s_1, \dots, s_k but a common destination t . They have the option of meeting at the rendezvous point v and continuing together to t , incurring a joint cost of $1 + \epsilon$. Each player can also “opt out,” meaning take the direct s_i - t path solo. (Insert your own joke about public transportation in California here.) Player i incurs a cost of $1/i$ for its opt-out strategy.

The optimal solution is clear: if all players travel through the rendezvous point, the overall cost is $1 + \epsilon$. Unfortunately, this is not a Nash equilibrium: player k would pay slightly less by switching to its opt-out strategy (which is a dominant strategy for this player). Given that player k does not use the rendezvous in a Nash equilibrium, player $k - 1$ does not either — it would have to pay at least $(1 + \epsilon)/(k - 1)$ with player k absent, and its opt-out strategy is cheaper. Iterating this argument, there is no Nash equilibrium in which any player travels through v . Meanwhile, the outcome in which all players opt out is a Nash equilibrium.¹ The cost of this (unique) Nash equilibrium is the k th Harmonic number $\sum_{i=1}^k \frac{1}{i}$, which lies between $\ln k$ and $\ln k + 1$.

The price of anarchy in the opt-out example is \mathcal{H}_k , which is much smaller than in the VHS

¹We just performed a procedure called the iterated deletion of dominated strategies. When a unique outcome survives this procedure, it is the unique Nash equilibrium.

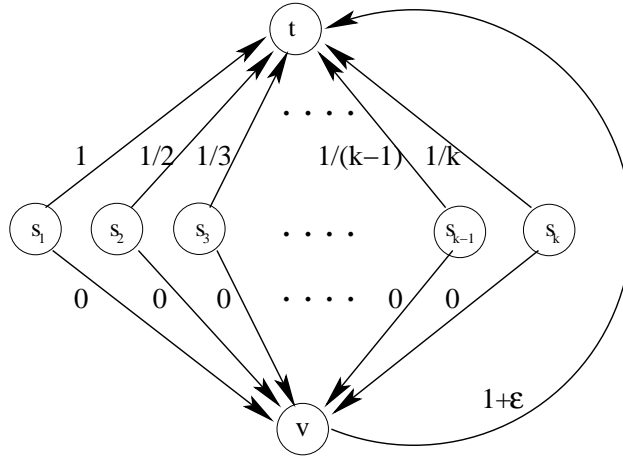


Figure 2: Opting out. There can be a unique Nash equilibrium with cost \mathcal{H}_k times that of an optimal outcome.

or Betamax example, but it is also a much more compelling example. This inefficiency is not the result of multiple or unreasonable equilibria, and it captures the observable phenomenon that games with positive externalities can suffer efficiency loss from underparticipation.

2 The Price of Stability

The two examples in the previous section limit our ambitions: we cannot hope to prove anything interesting about worst-case Nash equilibria of network cost-sharing games, and even when there is a unique Nash equilibrium, it can cost \mathcal{H}_k times that of an optimal solution. This section proves the following guarantee on *some* Nash equilibrium of every network cost-sharing game.

Theorem 2.1 (Price of Stability of Network Cost-Sharing Games [1]) *In every network cost-sharing game with k players, there exists a pure Nash equilibrium with cost at most \mathcal{H}_k times that of an optimal outcome.*

The theorem asserts in particular that every network cost-sharing game possesses at least one pure Nash equilibrium, which is already a non-trivial fact. The opt-out example shows that the factor of \mathcal{H}_k cannot be replaced by anything smaller.

The *price of stability* is the “optimistic” version of the price of anarchy, defined as the ratio between the cost of the best Nash equilibrium and that of an optimal outcome. Thus Theorem 2.1 asserts that the price of stability in every network cost-sharing game is at most \mathcal{H}_k . In terms of Figure 3, we are working with the entire set of pure Nash equilibria, but arguing only about one of them, rather than about all of them. This is the first occasion we’ve argued about anything other than the worst of a set of equilibria.

A bound on the price of stability, which only ensures that one equilibrium is approximately optimal, provides a significantly weaker guarantee than a bound on the price of

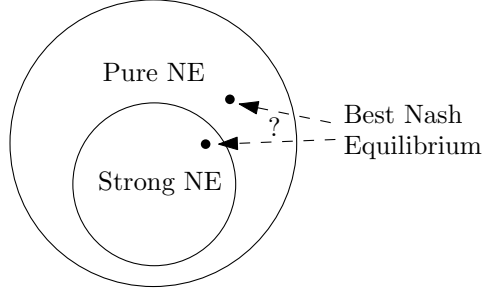


Figure 3: The best Nash equilibrium may be a strong NE, or not.

anarchy. It is well motivated in games where there is a third party that can propose an initial outcome — “default behavior” for the players. It’s easy to find examples in real life where an institution or society effectively proposes one equilibrium out of many — even just in choosing which side of the road everybody drives on. For a computer science example, consider the problem of designing the default values of user-defined parameters of software or a network protocol. One sensible approach is to set default parameters so that users are not incentivized to change them and, subject to this, to optimize performance. The price of stability quantifies the necessary degradation in solution quality caused by restricting solutions to be equilibria.

Proof of Theorem 2.1: The proof of Theorem 2.1 goes through Rosenthal’s potential function, introduced in Lecture 13. Recall the definition that we gave for atomic selfish routing games

$$\Phi(\mathbf{P}) = \sum_{e \in E} \sum_{i=1}^{f_e} c_e(i),$$

where c_e denotes the per-player cost incurred on edge e . Network cost-sharing games have exactly the same form as atomic selfish routing games — each of k players picks an s_i - t_i path in a network, and the player cost (1) is a sum of the edge costs, each of which depends only on the number of players using it — with the per-player cost of an edge e with f_e users being γ_e/f_e . Positive externalities are reflected by decreasing per-player cost functions, in contrast to the nondecreasing cost functions that were appropriate in routing games. Thus the potential function specializes to

$$\Phi(\mathbf{P}) = \sum_{e \in E} \sum_{i=1}^{f_e} \frac{\gamma_e}{i} = \sum_{e \in E} \gamma_e \sum_{i=1}^{f_e} \frac{1}{i} \quad (3)$$

in a network cost-sharing game.

In Lecture 13, we proved that the outcome that minimizes the potential function Φ is a Nash equilibrium, and we noted at the time that the proof worked for any cost functions, nondecreasing or not. That is, the strategic players are inadvertently striving to minimize Φ . This argument proves that every network cost-sharing game has a pure Nash equilibrium — the outcome that minimizes (3). For instance, in the VHS or Betamax example, the low-cost

Nash equilibrium minimizes (3) while the high-cost Nash equilibrium does not. While the minimizer of the potential function need not be the lowest-cost Nash equilibrium (see the Problems), we can prove that it has cost at most \mathcal{H}_k times that of an optimal outcome.

The key observation is that the potential function in (3), whose numerical value we don't care about per se, approximates well the objective function (2) that we do care about. Precisely, since $\gamma_e \leq \gamma_e \sum_{i=1}^{f_e} \frac{1}{i} \leq \gamma_e \cdot \mathcal{H}_k$ for every $i \in \{1, 2, \dots, k\}$, we have

$$\text{cost}(\mathbf{P}) \leq \Phi(\mathbf{P}) \leq \mathcal{H}_k \cdot \text{cost}(\mathbf{P}) \quad (4)$$

for every outcome \mathbf{P} . The inequalities (4) state that Nash equilibria are effectively trying to minimize an approximately correct function Φ , so it makes sense that one such equilibrium should approximately minimize the correct objective function.

To be precise, let \mathbf{P} minimize Φ (a Nash equilibrium) and let \mathbf{P}^* be an outcome of minimum cost. We have

$$\begin{aligned} \text{cost}(\mathbf{P}) &\leq \Phi(\mathbf{P}) \\ &\leq \Phi(\mathbf{P}^*) \\ &\leq \mathcal{H}_k \cdot \text{cost}(\mathbf{P}^*), \end{aligned}$$

where the first and last inequalities follow from (4) and the middle inequality follows from the choice of \mathbf{P} as the minimizer of Φ . This completes the proof of Theorem 2.1. ■

Open Question 1 (POS in Undirected Networks) In the VHS or Betamax example, it doesn't matter whether the network is directed or undirected. The opt-out example, on the other hand, makes crucial use of a directed network (see the Exercises). An interesting open question is whether or not the price of stability of every undirected network cost-sharing game is bounded above by a constant; see [2] for the latest progress.

3 Strong Nash Equilibria and Their POA

This section gives an alternative approach to eluding the bad Nash equilibrium of the VHS or Betamax example and proving meaningful bounds on the inefficiency of equilibria in network cost-sharing games. The plan is to once again argue about all (i.e., worst-case) equilibria, but to first identify a strict subset of the pure Nash equilibria that we care about.

In general, when studying the inefficiency of equilibria in a class of games, one should zoom out (i.e., enlarge the set of equilibria) as much as possible subject to the existence of meaningful POA bounds. We zoomed out in games with negative externalities, such as routing and location games. The POA of pure Nash equilibria is close to 1 in these games, so we focused on extending worst-case bounds to ever-larger sets of equilibria. This lecture, where worst-case Nash equilibria can be highly suboptimal, we need to zoom in to recover interesting inefficiency bounds. In terms of Figure 3, we aim to bound the cost of all Nash equilibria that fall into the smaller set.

To motivate the subclass of Nash equilibria that we study, let's return to the VHS or Betamax example. The high-cost Nash equilibrium is an equilibrium because a player that deviates unilaterally would pay the full cost $1 + \epsilon$ of the upper edge. What if a coalition of *two* players deviated jointly to the upper edge? Each deviating player would be responsible for a cost of only $\approx \frac{1}{2}$, so this would be a profitable deviation. Thus the high-cost Nash equilibrium does not persist when coalitional deviations are allowed.

Definition 3.1 (Strong Nash Equilibrium) Let \mathbf{s} be an outcome of a cost-minimization game.

- (a) Strategies $\mathbf{s}'_A \in \prod_{i \in A} S_i$ are a *beneficial deviation* for a subset A of players if

$$C_i(\mathbf{s}'_A, \mathbf{s}_{-A}) \leq C_i(\mathbf{s})$$

for every player $i \in A$, with the inequality holding strictly for at least one player of A .

- (b) The outcome \mathbf{s} is a *strong Nash equilibrium* if there is no coalition of players with a beneficial deviation.

Nash equilibria can be thought of as strong Nash equilibria in which only singleton coalitions are allowed. Every strong Nash equilibrium is thus a Nash equilibrium — that is, the former concept is an *equilibrium refinement* of the latter.

To get a better feel for strong Nash equilibria, let's return to our two examples. As noted above, the high-cost Nash equilibrium of the VHS or Betamax example is not strong. The low-case Nash equilibrium is strong. In fact, since a coalition of the entire player set is allowed, intuition might suggest that strong Nash equilibria have to be optimal. This is the case when all players share the same source and destination (see the Exercises), but not in general. In the opt-out example, the same argument that proves that the outcome in which everybody “opt outs” is the unique Nash equilibrium also proves that it is a strong Nash equilibrium. Thus, the opt-out example has a strong Nash equilibrium with cost \mathcal{H}_k times that of the minimum-cost outcome. Our next result states that no worse example is possible.

Theorem 3.2 (POA of Strong Nash Equilibria in Network Cost-Sharing Games [4])

In every network cost-sharing game with k players, every strong Nash equilibrium has cost at most \mathcal{H}_k times that of an optimal outcome.

The guarantee in Theorem 3.2 differs from that in Theorem 2.1 in two ways. On the positive side, the guarantee holds for *every* strong Nash equilibrium, as opposed to just *one* Nash equilibrium. If it were the case that every network cost-sharing game has at least one strong Nash equilibrium, then Theorem 3.2 would be a strictly stronger statement than Theorem 2.1. The second difference, however, is that Theorem 3.2 does not assert existence, and for good reason (see Figure 4 below). These two differences render Theorems 2.1 and 3.2 incomparable guarantees.

Proof of Theorem 3.2: The proof bears some resemblance to our previous POA analyses, but has a couple of extra ideas. One nice feature is that — perhaps unsurprisingly given

the bound that we're trying to prove — the proof uses Rosenthal's potential function in an interesting way. Our previous POA analyses for classes of potential games (selfish routing, location games) did not make use of the potential function.

The first step in our previous POA analyses was to invoke the Nash equilibrium hypothesis once per player to generate upper bounds on players' equilibrium costs. Here, we're making a strong hypothesis — we begin by letting \mathbf{P} be an arbitrary strong Nash equilibrium rather than an arbitrary Nash equilibrium — and aspire to a stronger conclusion. After all, what we're trying to prove is false for arbitrary Nash equilibria (as in the VHS or Betamax example).

The natural place to start is with the most powerful coalition $A_k = \{1, 2, \dots, k\}$ of all k players. Why doesn't this coalition collectively switch to the optimal outcome \mathbf{P}^* ? It must be that for some player i , $C_i(\mathbf{P}) \leq C_i(\mathbf{P}^*)$.² Rename the players so that this is player k .

We want an upper bound on the equilibrium cost of *every* player, not just that of player k . To ensure that we get an upper bound for a new player, we next invoke the strong Nash equilibrium hypothesis for the coalition $A_{k-1} = \{1, 2, \dots, k-1\}$ — why don't these $k-1$ players collectively deviate to $\mathbf{P}_{A_{k-1}}^*$? There must be a player $i \in \{1, 2, \dots, k-1\}$ with $C_i(\mathbf{P}) \leq C_i(\mathbf{P}_{A_{k-1}}^*, P_k)$. We rename the players of A_{k-1} so that this is true for player $k-1$ and continue.

Iterating the argument yields a renaming of the players as $\{1, 2, \dots, k\}$ such that, for every i ,

$$C_i(\mathbf{P}) \leq C_i(\mathbf{P}_{A_i}^*, \mathbf{P}_{-A_i}), \quad (5)$$

where $A_i = \{1, 2, \dots, i\}$. Now that we have an upper bound on the equilibrium cost of every player, we can sum (5) over the players to obtain

$$\begin{aligned} \text{cost}(\mathbf{P}) &= \sum_{i=1}^k C_i(\mathbf{P}) \\ &\leq \sum_{i=1}^k C_i(\mathbf{P}_{A_i}^*, \mathbf{P}_{-A_i}) \end{aligned} \quad (6)$$

$$\leq \sum_{i=1}^k C_i(\mathbf{P}_{A_i}^*). \quad (7)$$

Inequality (6) is immediate from (5). Inequality (7) follows from the fact that network cost-sharing games have positive externalities — deleting players only decreases the number f_e of players using a given edge and hence only increases the cost share of each remaining player on each edge. The motivation for the second inequality is to simplify our upper bound on the equilibrium cost to the point that it becomes a telescoping sum (cf., the location game analysis in Lecture 14).

Next we use the fact that network cost-sharing games are potential games. Recalling the definition (3) of the potential function Φ , we see that the decrease in potential function

²This inequality is strict if at least one other player is better off, but we won't need this stronger statement.

value from deleting a player is exactly the cost incurred by that player. Formally:

$$C_i(\mathbf{P}_{A_i}^*) = \sum_{e \in P_i^*} \frac{\gamma_e}{f_e^i} = \Phi(\mathbf{P}_{A_i}^*) - \Phi(\mathbf{P}_{A_{i-1}}^*), \quad (8)$$

where f_e^i denotes the number of players of A_i that use a path in \mathbf{P}^* that includes edge e . This equation is the special case of the Rosenthal potential function condition (see Lecture 14) in which a player deviates to the empty-set strategy.

Combining (7) with (8), we obtain

$$\begin{aligned} \text{cost}(\mathbf{P}) &\leq \sum_{i=1}^k \left[\Phi(\mathbf{P}_{A_i}^*) - \Phi(\mathbf{P}_{A_{i-1}}^*) \right] \\ &= \Phi(\mathbf{P}^*) \\ &\leq \mathcal{H}_k \cdot \text{cost}(\mathbf{P}^*), \end{aligned}$$

where the inequality follows from our earlier observation (4) that the potential function Φ can only overestimate the cost of an outcome by an \mathcal{H}_k factor. This completes the proof of Theorem 3.2. ■

4 Epilogue

Network cost-sharing games can have “unreasonable” bad Nash equilibria, and this motivates the search for a subset of Nash equilibria with two properties: better worst-case bounds than for arbitrary Nash equilibria, and a plausible narrative justifying restricting the analysis to this subset. Both of our two solutions — best-case Nash equilibria and worst-case strong Nash equilibria — meet the first criterion, admitting an approximation bound of \mathcal{H}_k rather than k . The justification for focusing on best-case Nash equilibria is strongest in settings where a third party can propose an equilibrium, although there is additional experimental evidence that potential function optimizers (as in Theorem 2.1) are more likely to be played than other Nash equilibria [3]. Worst-case bounds for strong Nash equilibria are attractive when such equilibria exist, as it is plausible that such equilibria are more likely to persist than regular Nash equilibria. While strong Nash equilibria are guaranteed to exist in classes of network cost-sharing games with sufficiently simple network structure [4], they do not, unfortunately, exist in general. Figure 4 gives a concrete example; we leave verifying the non-existence of strong Nash equilibria as an exercise.

References

- [1] E. Anshelevich, A. Dasgupta, J. Kleinberg, É. Tardos, T. Wexler, and T. Roughgarden. The price of stability for network design with fair cost allocation. *SIAM Journal on Computing*, 38(4):1602–1623, 2008.

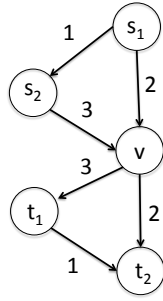


Figure 4: A network cost-sharing game with no strong Nash equilibrium.

- [2] V. Bilò, M. Flammini, and L. Moscardelli. The price of stability for undirected broadcast network design with fair cost allocation is constant. In *Proceedings of the 54rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 638–647, 2013.
- [3] R. Chen and Y. Chen. The potential of social identity for equilibrium selection. *American Economic Review*, 101(6):2562–2589, 2011.
- [4] A. Epstein, M. Feldman, and Y. Mansour. Strong equilibrium in cost sharing connection games. *Games and Economic Behavior*, 67(1):51–68, 2009.
- [5] M. O. Jackson. *Social and Economic Networks*. Princeton, 2008.