Lecture 15

Best-Case and Strong Nash Equilibria

This lecture has two purposes. The first is to introduce a simple model of network formation that resembles atomic selfish routing games but has positive externalities, meaning that an agent prefers to share the edges of her path with as many other agents as possible. Such games generally have multiple PNE with wildly varying costs. The second purpose of the lecture is to explain two approaches for confining attention to a subset of "reasonable" PNE. Ideally, better worst-case approximation bounds should hold for such a subset than for all PNE, and there should also be a plausible narrative as to why PNE in the subset are more worthy of study than the others.

Section 15.1 defines network cost-sharing games and considers two important examples. Section 15.2 proves an approximation bound for the best-case PNE of a network cost-sharing game. Sections 15.3 and 15.4 prove a bound on the POA of strong Nash equilibria, the subset of PNE for which no coalition of agents has a beneficial deviation.

15.1 Network Cost-Sharing Games

15.1.1 Externalities

The network formation model introduced next is a concrete example of a game with positive externalities. The *externality* caused by an agent in a game is the difference between her individual objective function value and her contribution to the social objective function value. The models studied in previous lectures have *negative* externalities, meaning that agents do not fully account for the harm that they cause. In a routing game, for example, an agent does not take into account the additional cost her presence creates for the other agents using the edges in her 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 also enriches the experience of other people in the same group. As an agent, you're generally bummed to see new agents show up in a game with negative externalities, and excited for the windfalls of new agents in a game with positive externalities.

15.1.2 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 agents. Agent i has an origin vertex $o_i \in V$ and a destination vertex $d_i \in V$, and her strategy set is the set of o_i-d_i paths of the graph. Outcomes of the game correspond to path vectors $\mathbf{P} = (P_1, \ldots, P_k)$, with the semantics that the subnetwork $(V, \cup_{i=1}^{k} P_i)$ gets formed.

We think of γ_e as the cost of building the edge e, for example of laying down high-speed Internet fiber to a neighborhood. This cost is independent of the number of agents that use the edge. Agents' costs are defined edge-by-edge, as in routing games (Lectures 11–12). If multiple agents use an edge e in their chosen paths, then they are jointly responsible for the edge's fixed cost γ_e, and we assume that they split it equally. In the language of cost-minimization games (Lecture 13), the cost $C_i(\mathbf{P})$ of agent i in the outcome $\mathbf{P}$ is

$$C_i(\mathbf{P}) = \sum_{e \in P_i} \frac{\gamma_e}{f_e}, \tag{15.1}$$

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

$$\mathrm{cost}(\mathbf{P}) = \sum_{e \in E \,:\, f_e \geq 1} \gamma_e. \tag{15.2}$$

Analogous to the objective function (11.3) and (11.4) in routing games, the function (15.2) can equally well be written as the sum $\sum_{i=1}^{k} C_i(\mathbf{P})$ of the agents' costs.

15.1.3 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 15.1, with k agents with a common origin o and destination d. One interpretation of this example is as a choice between two competing technologies. 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, and thus corresponds to the cheaper edge in Figure 15.1. VHS was the other technology, and it grabbed a larger market share early on. Coordinating on a single technology proved the primary driver in consumers' decisions—having the better technology is little consolation for being unable to rent anything from your corner store—and Betamax was eventually driven to extinction.

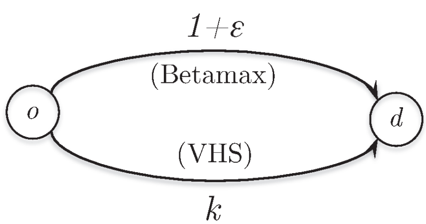

Figure 15.1: VHS or Betamax. The POA in a network cost-sharing game can be as large as the number k of agents. The parameter $\epsilon > 0$ can be arbitrarily small.

The optimal outcome in the network in Figure 15.1 is for all agents to pick the upper edge, for a total cost of $1 + \epsilon$. This is also a PNE (Definition 13.2). Unfortunately, there is a second PNE, in which all agents pick the lower edge. Since the cost of k is split equally, each agent pays 1. If an agent deviated unilaterally to the upper edge, she would pay the full cost $1+\epsilon$ of that edge and thus suffer a higher cost. This example shows that the POA in network cost-sharing games can be as high as k, the number of agents. Exercise 15.1 proves a matching upper bound.

The VHS-or-Betamax example is exasperating. We proposed a reasonable network model capturing positive externalities, and the POA—which has helped us reason about several models already—is distracted by an extreme equilibrium and yields no useful information. What if we focus only on the "nice" equilibria? We'll return to this question after considering another important example.

15.1.4 Example: Opting Out

Consider the network cost-sharing game shown in Figure 15.2. The k agents have distinct origins $o_1, \ldots, o_k$ but a common destination d. They have the option of meeting at the rendezvous point v and continuing together to d, incurring a joint cost of $1 + \epsilon$. Each agent can also "opt out," meaning take the direct o_i-d path solo. Agent i incurs a cost of $1/i$ for her opt-out strategy.

The optimal outcome is clear: if all agents travel through the rendezvous point, the cost is $1 + \epsilon$. Unfortunately, this is not a PNE: agent k can pay slightly less by switching to her opt-out strategy, which is a dominant strategy for her. Given that agent k does not use the rendezvous in a PNE, agent $k - 1$ does not either; she would have to pay at least $(1 + \epsilon)/(k - 1)$ with agent k absent, and her opt-out strategy is cheaper. Iterating this argument, there is no PNE in which any agent travels through v. Meanwhile, the outcome in which all agents opt out is a PNE.[1] The cost of this unique PNE is the kth harmonic number $\sum_{i=1}^{k} \frac{1}{i}$. This number lies between $\ln k$ and $\ln k + 1$, and we denote it by $\mathcal{H}_k$.

The POA in the opt-out example approaches $\mathcal{H}_k$ as ϵ tends to 0. Unlike the VHS-or-Betamax example, this inefficiency is not the result of multiple or unreasonable equilibria.

15.2 The Price of Stability

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

[1] This argument is an example of the iterated removal of strictly dominated strategies. When a unique outcome survives this procedure, it is the unique PNE of the game.

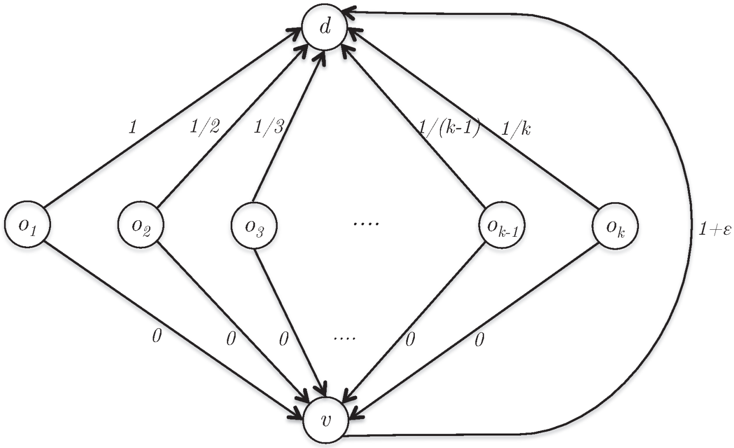

Figure 15.2: Opting out. There can be a unique PNE with cost $\mathcal{H}_k$ times that of an optimal outcome. The parameter $\epsilon > 0$ can be arbitrarily small.

Theorem 15.1 (Price of Stability) *In every network cost-sharing game with k agents, there exists a PNE 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 PNE. The opt-out example shows that the factor of $\mathcal{H}_k$ in Theorem 15.1 cannot be replaced by anything smaller.

The *price of stability* is the "optimistic" version of the POA, defined as the ratio

$$\frac{\text{cost of best equilibrium}}{\text{cost of optimal outcome}}.$$

Thus Theorem 15.1 states that the price of stability is at most $\mathcal{H}_k$ in every network cost-sharing game.

Proof of Theorem 15.1: Network cost-sharing games have the same form as atomic selfish routing games (Section 12.4), with each agent i picking an o_i-d_i path in a network. Moreover, an agent's cost (15.1) is the sum of the costs of the edges in her path, and each edge cost depends only on the number of agents using it. The "cost function" of an edge e can be thought of as $c_e(f_e) = \gamma_e/f_e$, where f_e is the number of agents using the edge.

Adapting the potential function (13.7) from the proof of Theorem 13.6 to network cost-sharing games yields

$$\Phi(\mathbf{P}) = \sum_{e \in E} \sum_{i=1}^{f_e} \frac{\gamma_e}{i} = \sum_{e \in E} \gamma_e \sum_{i=1}^{f_e} \frac{1}{i}. \tag{15.3}$$

As in that proof, the outcome that minimizes this function Φ is a PNE.[2] For instance, in the VHS-or-Betamax example, the low-cost PNE minimizes (15.3) while the high-cost PNE does not. While the minimizer of the potential function need not be the best PNE (Problem 15.1), we next prove that its cost is at most $\mathcal{H}_k$ times that of an optimal outcome.

The key observation is that the potential function (15.3), whose numerical value we don't care about per se, approximates well the objective function (15.2) that we do care about. Precisely, since

$$\gamma_e \le \gamma_e \sum_{i=1}^{f_e} \tfrac{1}{i} \le \gamma_e \cdot \mathcal{H}_k$$

for every edge e with $f_e \ge 1$, we can sum over such edges to derive

$$\text{cost}(\mathbf{P}) \le \Phi(\mathbf{P}) \le \mathcal{H}_k \cdot \text{cost}(\mathbf{P}) \tag{15.4}$$

for every outcome $\mathbf{P}$. The inequalities (15.4) state that PNE are inadvertently trying to minimize an approximately correct function Φ, so it makes sense that one PNE should approximately minimize the correct objective function.

To finish the proof, let $\mathbf{P}$ denote a PNE minimizing the potential function (15.3) and $\mathbf{P}^*$ an optimal outcome. We have

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

where the first and last inequalities follow from (15.4) and the middle inequality follows from the choice of $\mathbf{P}$ as a minimizer of Φ. ■

[2]Network cost-sharing games have decreasing per-agent cost functions, reflecting the positive externalities and contrasting with routing games. The proof of Theorem 13.6 holds for any edge cost functions, decreasing or otherwise.

How should we interpret Theorem 15.1? 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 POA. The price of stability is relevant for games where there is a third party who can propose an initial outcome—default behavior for the agents. 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 choosing 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 motivated to change them and, subject to this, to optimize performance. The price of stability quantifies the necessary degradation in the objective function value caused by the restriction to equilibrium outcomes.

The proof of Theorem 15.1 implies that every minimizer of the potential function (15.3) has cost at most $\mathcal{H}_k$ times that of an optimal outcome. There are plausible narratives for why such PNE are more relevant than arbitrary PNE; see the Notes for details. This gives a second interpretation of Theorem 15.1 that makes no reference to a third party and instead rests on the belief that potential function minimizers are in some sense the most important of the PNE.

15.3 The POA of Strong Nash Equilibria

This section gives an alternative approach to eluding the bad PNE of the VHS-or-Betamax example and proving meaningful bounds on the inefficiency of equilibria in network cost-sharing games. We once again argue about all (i.e., worst-case) equilibria, but first restrict attention to a well-motivated subset of PNE.[3]

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. In games with negative externalities, such as routing and location games, we zoomed all the way out to the set of coarse correlated equilibria (Lecture 14). The POA of PNE is reasonably close to 1 in

[3]When one equilibrium concept is only more stringent than another, the former is called an *equilibrium refinement* of the latter.

these games, so we focused on extending our worst-case bounds to ever-larger sets of equilibria. In network cost-sharing games, where worst-case PNE can be highly suboptimal, we need to zoom in to recover interesting POA bounds (Figure 15.3).

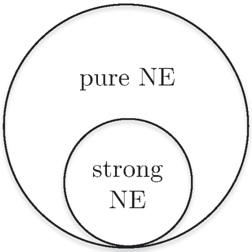

Figure 15.3: Strong Nash equilibria are a special case of pure Nash equilibria.

Recall the VHS or Betamax example (Section 15.1.3). The high-cost outcome is a PNE because an agent that deviates unilaterally would pay the full cost $1 + \epsilon$ of the upper edge. What if a coalition of *two* agents deviated jointly to the upper edge? Each deviating agent would then pay only $\approx \frac{1}{2}$, so this would be a profitable deviation for both of them. We conclude that the high-cost PNE does not persist when coalitional deviations are allowed.

Definition 15.2 (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 agents if

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

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

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

Every strong Nash equilibrium is a PNE, as beneficial deviations for singleton coalitions correspond to improving unilateral deviations. It is plausible that strong Nash equilibria are more likely to occur than other PNE.

To get a better feel for strong Nash equilibria, let's return to our two examples. As noted above, the high-cost PNE of the VHS or Betamax example is not a strong Nash equilibrium. The low-cost PNE is a strong Nash equilibrium. More generally, since the coalition of the entire agent set is allowed, intuition might suggest that strong Nash equilibria are always optimal outcomes. This is the case when all agents share the same origin and destination (Exercise 15.3), but not in general. In the opt-out example (Section 15.1.4), the same argument that proves that the all-opt-out outcome is the unique PNE also proves that it is a strong Nash equilibrium. This strong Nash equilibrium has cost arbitrarily close to $\mathcal{H}_k$ times that of an optimal outcome. Our next result states that no worse example is possible.

Theorem 15.3 (The POA of Strong Nash Equilibria) *In every network cost-sharing game with k agents, every strong Nash equilibrium has cost at most $\mathcal{H}_k$ times that of an optimal outcome.*

The guarantee in Theorem 15.3 differs from that in Theorem 15.1 in two ways. On the positive side, the guarantee holds for *every* strong Nash equilibrium, as opposed to just *one* PNE. Were it true that every network cost-sharing game has at least one strong Nash equilibrium, Theorem 15.3 would be a strictly stronger statement than Theorem 15.1. Unfortunately, a strong Nash equilibrium may or may not exist in a network cost-sharing game (see Figure 15.4 and Exercise 15.4), and so Theorems 15.1 and 15.3 offer incomparable guarantees.

*15.4 Proof of Theorem 15.3

The proof of Theorem 15.3 bears some resemblance to our previous POA analyses, but it has a couple of extra ideas. One nice feature is that the proof uses the potential function (15.3) in an interesting way. Our POA analyses of selfish routing and location games did not make use of their potential functions.

Fix a network cost-sharing game and a strong Nash equilibrium **P**. The usual first step in a POA analysis is to invoke the equilibrium hypothesis once per agent to generate upper bounds on agents' equilibrium costs. To use the strong Nash equilibrium assumption in the strongest-possible way, the natural place to start is with the most

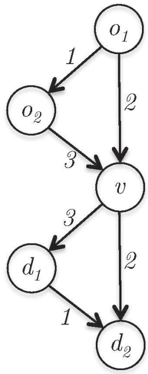

Figure 15.4: A network cost-sharing game with no strong Nash equilibrium.

powerful coalition $A_k = \{1, 2, \ldots, k\}$ of all k agents. Why doesn't this coalition collectively switch to the optimal outcome $\mathbf{P}^*$? It must be that for some agent i, $C_i(\mathbf{P}) \leq C_i(\mathbf{P}^*)$.[4] Rename the agents so that this is agent k.

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

By iterating this argument, we obtain a renaming of the agents as $\{1, 2, \ldots, k\}$ such that, for every i,

$$C_i(\mathbf{P}) \leq C_i(\mathbf{P}^*_{A_i}, \mathbf{P}_{-A_i}), \tag{15.5}$$

where $A_i = \{1, 2, \ldots, i\}$. Now that we have an upper bound on the equilibrium cost of every agent, we can sum (15.5) over the agents to

[4] This inequality is strict if at least one other agent is better off, but we don't need this stronger statement.

obtain

$$\begin{aligned}
\text{cost}(\mathbf{P}) &= \sum_{i=1}^{k} C_i(\mathbf{P}) \\
&\leq \sum_{i=1}^{k} C_i(\mathbf{P}^*_{A_i}, \mathbf{P}_{-A_i}) \qquad (15.6) \\
&\leq \sum_{i=1}^{k} C_i(\mathbf{P}^*_{A_i}). \qquad (15.7)
\end{aligned}$$

Inequality (15.6) is immediate from (15.5). Inequality (15.7) follows from the fact that network cost-sharing games have positive externalities; removing agents only decreases the number of agents using each edge and hence only increases the cost share of each remaining agent on each edge. The purpose of the inequality (15.7) is to simplify our upper bound on the equilibrium cost to the point that it becomes a telescoping sum.

Next we use the potential function Φ defined in (15.3). Letting f_e^i denote the number of agents of A_i that use a path in $\mathbf{P}^*$ that includes edge e, we have

$$C_i(\mathbf{P}^*_{A_i}) = \sum_{e \in P_i^*} \frac{\gamma_e}{f_e^i} = \Phi(\mathbf{P}^*_{A_i}) - \Phi(\mathbf{P}^*_{A_{i-1}}), \qquad (15.8)$$

with the second equation following from the definition of Φ.

Combining (15.7) with (15.8), we obtain

$$\begin{aligned}
\text{cost}(\mathbf{P}) &\leq \sum_{i=1}^{k} \left[\Phi(\mathbf{P}^*_{A_i}) - \Phi(\mathbf{P}^*_{A_{i-1}}) \right] \\
&= \Phi(\mathbf{P}^*) \\
&\leq \mathcal{H}_k \cdot \text{cost}(\mathbf{P}^*), \qquad (15.9)
\end{aligned}$$

where inequality (15.9) follows from our earlier observation (15.4) 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 15.3.

The Upshot

- ✰ In a network cost-sharing game, each agent picks a path from her origin to her destination, and the fixed cost of each edge used is split equally among its users.
- ✰ Different PNE of a network cost-sharing game can have wildly different costs, and the POA can be as large as the number k of agents. These facts motivate approximation bounds that apply only to a subset of PNE.
- ✰ The price of stability of a game is the ratio between the lowest cost of an equilibrium and the cost of an optimal outcome.
- ✰ The worst-case price of stability of network cost-sharing games is $\mathcal{H}_k = \sum_{i=1}^{k} \frac{1}{i} \approx \ln k$.
- ✰ A strong Nash equilibrium is an outcome such that no coalition of agents has a collective deviation that benefits at least one agent and harms no agent of the coalition.
- ✰ Every strong Nash equilibrium of a network-cost sharing game has cost at most $\mathcal{H}_k$ times that of an optimal outcome.
- ✰ Strong Nash equilibria are not guaranteed to exist in network cost-sharing games.

Notes

Network cost-sharing games and Theorem 15.1 are from Anshelevich et al. (2008a). The VHS or Betamax example is from Anshelevich et al. (2008b). Many other models of network formation have been proposed and studied; see Jackson (2008) for a textbook treatment. It is an open question to analyze the worst-case price of stability in undirected network cost-sharing games; see Bilò et al. (2016) for the latest progress. Experimental evidence that potential function minimizers are more commonly played than other PNE is given in Chen and Chen (2011); related theoretical results appear in Blume (1993) and Asadpour and Saberi (2009). The strong Nash equilibrium concept is due to Aumann (1959), and Andelman et al. (2009) propose studying the price of anarchy of strong Nash equilibria. Theorem 15.3, the example in Figure 15.4, and Problem 15.2 are from Epstein et al. (2009).

Exercises

Exercise 15.1 Prove that in every network cost-sharing game, the POA of PNE is at most k, the number of agents.

Exercise 15.2 If we modify the opt-out example (Section 15.1.4) so that all of the edges are undirected, and each agent i can choose an o_i-d path that traverses edges in either direction, what is the price of stability in the resulting network cost-sharing game?

Exercise 15.3 Prove that in every network cost-sharing game in which all agents have a common origin vertex and a common destination vertex, there is a one-to-one correspondence between strong Nash equilibria and minimum-cost outcomes. (Thus, in such games, strong Nash equilibria always exist and the POA of such equilibria is 1.)

Exercise 15.4 Prove that the network cost-sharing game shown in Figure 15.4 has no strong Nash equilibrium.

Exercise 15.5 Extend the model of network cost-sharing games by allowing each edge e to have a cost $\gamma_e(x)$ that depends on the number x of agents that use it. The joint cost $\gamma_e(x)$ is again split equally between the x users of the edge. Assume that each function γ_e is *concave*, meaning that

$$\gamma_e(i+1) - \gamma_e(i) \leq \gamma_e(i) - \gamma_e(i-1)$$

for each $i = 1, 2, \ldots, k-1$. Extend Theorems 15.1 and 15.3 to this more general model.

Exercise 15.6 *(H)* Continuing the previous exercise, suppose $\gamma_e(x) = a_e x^p$ for every edge e, where each $a_e > 0$ is a positive constant and the common exponent p lies in $(0, 1]$. For this special case, improve the upper bounds of $\mathcal{H}_k$ in Theorems 15.1 and 15.3 to $\frac{1}{p}$, independent of the number of agents k.

Problems

Problem 15.1 (a) Exhibit a network cost-sharing game in which the minimizer of the potential function (15.3) is not the lowest-cost PNE.

(b) Exhibit a network cost-sharing game with at least one strong Nash equilibrium in which the minimizer of the potential function is not a strong Nash equilibrium.

Problem 15.2 Suppose we weaken the definition of a strong Nash equilibrium (Definition 15.2) by requiring only that no coalition of at most ℓ agents has a beneficial deviation, where $\ell \in \{1, 2, \ldots, k\}$ is a parameter. Pure Nash equilibria and strong Nash equilibria correspond to the $\ell = 1$ and $\ell = k$ cases, respectively. What is the worst-case POA of ℓ-strong Nash equilibria in network cost-sharing games, as a function of ℓ and k? Prove the best upper and lower bounds that you can.

Problem 15.3 *(H)* Prove that in every atomic selfish routing network (Section 12.4) with edge cost functions that are polynomials with nonnegative coefficients and degree at most p, the price of stability is at most $p + 1$.