
Best-Response Dynamics

This lecture segues into the third part of the course, where we ask: Do we expect strategic agents to reach an equilibrium? If so, which learning algorithms quickly converge to an equilibrium? Reasoning about these questions requires specifying *dynamics*, which describe how agents act when not at equilibrium. We consider dynamics where each agent's behavior is governed by an algorithm that attempts to, in some sense, learn the best response to how the other agents are acting. Ideally, we seek results that hold for multiple simple and natural learning algorithms. Then, even though agents may not literally follow such an algorithm, we can still have some confidence that our conclusions are robust and not an artifact of the particular choice of dynamics. This lecture focuses on variations of “best-response dynamics,” while the next two lectures study dynamics based on regret-minimization.

Section 16.1 defines best-response dynamics and proves convergence in potential games. Sections 16.2 and 16.3 introduce ϵ -best-response dynamics and prove that several variants of it converge quickly in atomic selfish routing games where all agents have a common origin and destination. Section 16.4 proves that, in the (λ, μ) -smooth games defined in Lecture 14, several variants of best-response dynamics quickly reach outcomes with objective function value almost as good as at an equilibrium.

16.1 Best-Response Dynamics in Potential Games

Best-response dynamics is a straightforward procedure by which agents search for a pure Nash equilibrium (PNE) of a game (Definition 13.2), using successive unilateral deviations.

Best-Response Dynamics

While the current outcome \mathbf{s} is not a PNE:
 pick an arbitrary agent i and an arbitrary
 beneficial deviation s'_i for agent i , and update
 the outcome to (s'_i, \mathbf{s}_{-i})

There might be many options for the deviating agent i and for the beneficial deviation s'_i . We leave both unspecified for the moment, specializing these choices later as needed.¹ We always allow the initial outcome to be arbitrary.

Best-response dynamics can be visualized as a walk in a graph, with vertices corresponding to strategy profiles, and outgoing edges corresponding to beneficial deviations (Figure 16.1). The PNE are precisely the vertices of this graph that have no outgoing edges. Best-response dynamics can only halt at a PNE, so it cycles in any game without one. It can also cycle in games that have a PNE (Exercise 16.1).

Best-response dynamics is a perfect fit for potential games (Section 13.3). Recall that a potential game admits a real-valued function Φ with the property that, for every unilateral deviation by some agent, the change in the potential function value equals the change in the deviator's cost (13.9). Routing games (Section 12.4), location games (Section 14.2), and network cost-sharing games (Section 15.2) are all potential games.

Theorem 13.7 notes that every potential game has at least one PNE, since the potential function minimizer is one. Best-response dynamics offers a more constructive proof of this fact.

Proposition 16.1 (Convergence of Best-Response Dynamics)

In a potential game, from an arbitrary initial outcome, best-response dynamics converges to a PNE.

Proof: In every iteration of best-response dynamics, the deviator's cost strictly decreases. By (13.9), the potential function strictly decreases. Thus, no cycles are possible. Since the game is finite by

¹This procedure is sometimes called “better-response dynamics,” with the term “best-response dynamics” reserved for the version in which s'_i is chosen to minimize i 's cost, given the strategies \mathbf{s}_{-i} of the other agents.

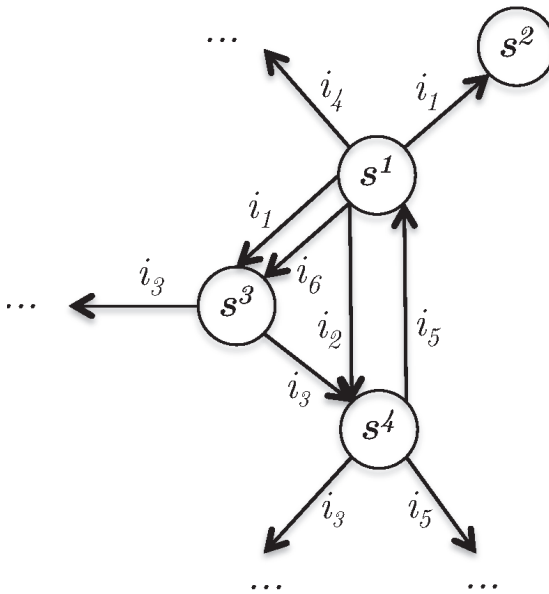


Figure 16.1: Best-response dynamics can be viewed as a walk in a graph. There is one vertex for each strategy profile. There is one edge for each beneficial unilateral deviation, labeled with the name of the deviating agent. PNE correspond to vertices with no outgoing edges, such as s^2 .

assumption, best-response dynamics eventually halts, necessarily at a PNE. ■

Translated to the graph in Figure 16.1, Proposition 16.1 asserts that every walk in a directed acyclic graph eventually stops at a vertex with no outgoing edges.

Proposition 16.1 shows that there is a natural procedure by which agents can reach a PNE of a potential game. How fast does this happen? One strong notion of “fast convergence” is convergence to a PNE in a reasonably small number of iterations. This occurs when, for example, the potential function Φ takes on only a small number of distinct values (Exercise 16.2). In general, best-response dynamics can decrease the potential function very slowly and require an exponential (in the number of agents k) number of iterations to converge (Lecture 19). This fact motivates the relaxed definitions of convergence studied in the rest of this lecture.

16.2 Approximate PNE in Selfish Routing Games

Our second notion of “fast convergence” settles for an approximate PNE.

Definition 16.2 (ϵ -Pure Nash Equilibrium) For $\epsilon \in [0, 1]$, an outcome \mathbf{s} of a cost-minimization game is an ϵ -pure Nash equilibrium (ϵ -PNE) if, for every agent i and deviation $s'_i \in S_i$,

$$C_i(s'_i, \mathbf{s}_{-i}) \geq (1 - \epsilon) \cdot C_i(\mathbf{s}). \quad (16.1)$$

Definition 16.2 is the same as Definition 14.5, reparametrized for convenience. An ϵ -PNE in the sense of Definition 16.2 corresponds to a $\frac{\epsilon}{1-\epsilon}$ -PNE under Definition 14.5.

We next study ϵ -best-response dynamics, in which we only permit moves that yield significant improvements.

ϵ -Best-Response Dynamics

While the current outcome \mathbf{s} is not an ϵ -PNE:

pick an arbitrary agent i who has an ϵ -move—a deviation s'_i with $C_i(s'_i, \mathbf{s}_{-i}) < (1 - \epsilon)C_i(\mathbf{s})$ —and an arbitrary such move for the agent, and update the outcome to (s'_i, \mathbf{s}_{-i})

ϵ -best-response dynamics can only halt at an ϵ -PNE, and it eventually converges in every potential game.

Our next result identifies a subclass of atomic selfish routing games (Section 12.4) in which a specialized variant of ϵ -best response dynamics converges quickly, meaning in a number of iterations that is bounded above by a polynomial function of all of the relevant parameters.²

²The number of outcomes of a game with k agents is exponential in k , so any polynomial bound on the number of iterations required is significant.

ϵ -Best-Response Dynamics (Maximum-Gain)

While the current outcome \mathbf{s} is not an ϵ -PNE:

among all agents with an ϵ -move, let i denote an agent who can obtain the largest cost decrease

$$C_i(\mathbf{s}) - \min_{\hat{s}_i \in S_i} C_i(\hat{s}_i, \mathbf{s}_{-i}),$$

and s'_i a best response to \mathbf{s}_{-i}

update the outcome to (s'_i, \mathbf{s}_{-i})

Theorem 16.3 (Convergence to an ϵ -PNE) *Consider an atomic selfish routing game where:*

1. *All agents have a common origin vertex and a common destination vertex.*
2. *For $\alpha \geq 1$, the cost function c_e of each edge e satisfies the α -bounded jump condition, meaning $c_e(x+1) \in [c_e(x), \alpha \cdot c_e(x)]$ for every edge e and positive integer x .*

Then, the maximum-gain variant of ϵ -best-response dynamics converges to an ϵ -PNE in at most $\frac{k\alpha}{\epsilon} \ln \frac{\Phi(\mathbf{s}_0)}{\Phi_{\min}}$ iterations, where \mathbf{s}_0 is the initial outcome and $\Phi_{\min} = \min_{\mathbf{s}} \Phi(\mathbf{s})$.

Analogs of Theorem 16.3 continue to hold for many different variants of ϵ -best-response dynamics (Problem 16.2); the only essential requirement is that every agent is given the opportunity to move sufficiently often. Even if we don't literally believe that agents will follow one of these variants of ϵ -best-response dynamics, the fact that simple and natural learning procedures converge quickly to approximate PNE in these games provides compelling justification for their study. Unfortunately, if either hypothesis of Theorem 16.3 is dropped, then all variants of ϵ -best-response dynamics can take an exponential (in k) number of iterations to converge (see the Notes).

*16.3 Proof of Theorem 16.3

The plan for proving Theorem 16.3 is to strengthen quantitatively the proof of Proposition 16.1 and show that every iteration of maximum-gain ϵ -best-response dynamics decreases the potential function by a lot. We need two lemmas. The first one guarantees the existence of an agent with a high cost; if this agent is chosen to move in an iteration, then the potential function decreases significantly. The issue is that some other agent might move instead. The second lemma, which is the one that needs the two hypotheses in Theorem 16.3, proves that the agent chosen to move has cost within an α factor of that of any other agent. This is good enough for fast convergence.

Lemma 16.4 *In every outcome \mathbf{s} , there is an agent i with $C_i(\mathbf{s}) \geq \Phi(\mathbf{s})/k$.*

Proof: In atomic selfish routing games, which have nondecreasing edge cost functions, the potential function can only underestimate the cost of an outcome. To see this, recall the definitions of the potential function (13.7) and objective function (11.3)–(11.4) of an atomic selfish routing game, and derive

$$\Phi(\mathbf{s}) = \sum_{e \in E} \sum_{i=1}^{f_e} c_e(i) \leq \sum_{e \in E} f_e \cdot c_e(f_e) = \sum_{i=1}^k C_i(\mathbf{s}) \quad (16.2)$$

for every outcome \mathbf{s} , where f_e denotes the number of agents that choose in \mathbf{s} a path including edge e . The inequality follows from the fact that cost functions are nondecreasing.

Since some agent must have cost at least as large as the average, we have

$$\max_{i=1}^k C_i(\mathbf{s}) \geq \frac{\sum_{i=1}^k C_i(\mathbf{s})}{k} \geq \frac{\Phi(\mathbf{s})}{k}$$

for every outcome \mathbf{s} , as claimed. ■

The next lemma relates the cost of the deviating agent in maximum-gain ϵ -best-response dynamics to those of the other agents.

Lemma 16.5 *Suppose agent i is chosen by maximum-gain ϵ -best-response dynamics to move in the outcome \mathbf{s} , and takes the ϵ -move s'_i . Then*

$$C_i(\mathbf{s}) - C_i(s'_i, \mathbf{s}_{-i}) \geq \frac{\epsilon}{\alpha} C_j(\mathbf{s}) \quad (16.3)$$

for every other agent j .

Proof: Fix the agent j . If j has an ϵ -move s'_j in \mathbf{s} , which by definition would decrease agent j 's cost by at least $\epsilon C_j(\mathbf{s})$, then

$$C_i(\mathbf{s}) - C_i(s'_i, \mathbf{s}_{-i}) \geq C_j(\mathbf{s}) - C_j(s'_j, \mathbf{s}_{-j}) \geq \epsilon C_j(\mathbf{s}).$$

The first inequality holds because i was chosen over j in maximum-game ϵ -best-response dynamics.

The trickier case is when the agent j has no ϵ -move available. We use here that all agents have the same set of available strategies. If s'_i is such a great deviation for agent i , why isn't it for agent j as well? That is, how can it be that

$$C_i(s'_i, \mathbf{s}_{-i}) \leq (1 - \epsilon)C_i(\mathbf{s}) \tag{16.4}$$

while

$$C_j(s'_i, \mathbf{s}_{-j}) \geq (1 - \epsilon)C_j(\mathbf{s})? \tag{16.5}$$

A key observation is that the outcomes (s'_i, \mathbf{s}_{-i}) and (s'_i, \mathbf{s}_{-j}) have at least $k - 1$ strategies in common. The strategy s'_i is used by i in the former outcome and by j in the latter outcome, and the $k - 2$ agents other than i and j use the same strategies in both outcomes. Since the two outcomes differ in only one chosen strategy, for every edge e of the network, the number of agents using e differs by at most one in the two outcomes. By the α -bounded jump hypothesis in Theorem 16.3, the cost of every edge differs by at most a factor of α in the two outcomes. In particular, the cost of agent j after deviating unilaterally to s'_i is at most α times that of agent i after the same unilateral deviation:

$$C_j(s'_i, \mathbf{s}_{-j}) \leq \alpha \cdot C_i(s'_i, \mathbf{s}_{-i}). \tag{16.6}$$

The inequalities (16.4)–(16.6) imply that $C_j(\mathbf{s}) \leq \alpha \cdot C_i(\mathbf{s})$. Combining this with (16.4) yields

$$C_i(\mathbf{s}) - C_i(s'_i, \mathbf{s}_{-i}) \geq \epsilon \cdot C_i(\mathbf{s}) \geq \frac{\epsilon}{\alpha} \cdot C_j(\mathbf{s}),$$

as required. ■

Lemma 16.4 guarantees that there is always an agent whose ϵ -move would rapidly decrease the potential function. Lemma 16.5

extends this conclusion to the agent who actually moves in maximum-gain ϵ -best-response dynamics. The bound on the number of iterations required for convergence now follows straightforwardly.

Proof of Theorem 16.3: In an iteration of maximum-gain ϵ -best-response dynamics where agent i makes an ϵ -move to the strategy s'_i ,

$$\Phi(\mathbf{s}) - \Phi(s'_i, \mathbf{s}_{-i}) = C_i(\mathbf{s}) - C_i(s'_i, \mathbf{s}_{-i}) \quad (16.7)$$

$$\geq \frac{\epsilon}{\alpha} \cdot \max_{j=1}^k C_j(\mathbf{s}) \quad (16.8)$$

$$\geq \frac{\epsilon}{\alpha k} \cdot \Phi(\mathbf{s}), \quad (16.9)$$

where equation (16.7) follows from the defining property (13.9) of a potential function, and inequalities (16.8) and (16.9) follow from Lemmas 16.5 and 16.4, respectively.

The derivation (16.7)–(16.9) shows that every iteration of maximum-gain ϵ -best-response dynamics decreases the potential function by at least a factor of $(1 - \frac{\epsilon}{\alpha k})$. Thus, every $\frac{k\alpha}{\epsilon}$ iterations decrease the potential function by at least a factor of $e = 2.718 \dots$ ³ Since the potential function begins with the value $\Phi(\mathbf{s}^0)$ and cannot drop lower than Φ_{\min} , maximum-gain ϵ -best-response dynamics converges in at most $\frac{k\alpha}{\epsilon} \ln \frac{\Phi(\mathbf{s}^0)}{\Phi_{\min}}$ iterations. ■

*16.4 Low-Cost Outcomes in Smooth Potential Games

This section explores our final notion of “fast convergence”: quickly reaching outcomes with objective function value *as good as if* agents had already converged to an approximate PNE. This guarantee does not imply convergence to an approximate PNE, but it is still quite compelling. When the primary reason for an equilibrium analysis is a price-of-anarchy bound, this weaker guarantee is a costless surrogate for convergence to an approximate equilibrium.

Weakening our notion of fast convergence enables positive results with significantly wider reach. The next result applies to all potential games that are (λ, μ) -smooth in the sense of Definition 14.2, including

³To see this, use that $(1 - x)^{1/x} \leq (e^{-x})^{1/x} = 1/e$ for $x \neq 0$.

all atomic selfish routing games (Section 12.4) and location games (Section 14.2). It uses the following variant of best-response dynamics, which is the analog of the variant of ϵ -best-response dynamics used in Theorem 16.3.

Best-Response Dynamics (Maximum-Gain)

While the current outcome \mathbf{s} is not a PNE:

among all agents with a beneficial deviation, let i
denote an agent who can obtain the largest cost
decrease

$$C_i(\mathbf{s}) - \min_{\hat{s}_i \in S_i} C_i(\hat{s}_i, \mathbf{s}_{-i}),$$

and s'_i a best response to \mathbf{s}_{-i}

update the outcome to (s'_i, \mathbf{s}_{-i})

We state the theorem for cost-minimization games; an analogous result holds for (λ, μ) -smooth payoff-maximization games (Remark 13.1).

Theorem 16.6 (Convergence to Low-Cost Outcomes)

Consider a (λ, μ) -smooth cost-minimization game with $\mu < 1$ that has a positive potential function Φ that satisfies $\Phi(\mathbf{s}) \leq \text{cost}(\mathbf{s})$ for every outcome \mathbf{s} . Let $\mathbf{s}^0, \dots, \mathbf{s}^T$ be a sequence of outcomes generated by maximum-gain best-response dynamics, \mathbf{s}^* a minimum-cost outcome, and $\eta \in (0, 1)$ a parameter. Then all but at most

$$\frac{k}{\eta(1-\mu)} \ln \frac{\Phi(\mathbf{s}^0)}{\Phi_{\min}}$$

outcomes \mathbf{s}^t satisfy

$$\text{cost}(\mathbf{s}^t) \leq \left(\frac{\lambda}{(1-\mu)(1-\eta)} \right) \cdot \text{cost}(\mathbf{s}^*), \quad (16.10)$$

where $\Phi_{\min} = \min_{\mathbf{s}} \Phi(\mathbf{s})$ and k is the number of agents.

Recall that in a (λ, μ) -smooth cost-minimization game, every PNE has cost at most $\frac{\lambda}{1-\mu}$ times the minimum possible (Section 14.4).

Thus Theorem 16.6 states that for all but a small number of outcomes in the sequence, the cost is almost as low as if best-response dynamics had already converged to a PNE.

Proof of Theorem 16.6: Fix $\eta \in (0, 1)$. The plan is to show that if \mathbf{s}^t is a *bad state*, meaning one that fails to obey the guarantee in (16.10), then the next iteration of maximum-gain best-response dynamics decreases the potential function significantly. This yields the desired bound on the number of bad states.

For an outcome \mathbf{s}^t , define $\delta_i(\mathbf{s}^t) = C_i(\mathbf{s}^t) - C_i(s_i^*, \mathbf{s}_{-i}^t)$ as the cost decrease that agent i would experience by switching her strategy to s_i^* , and $\Delta(\mathbf{s}^t) = \sum_{i=1}^k \delta_i(\mathbf{s}^t)$. The value $\delta_i(\mathbf{s}^t)$ is nonpositive when \mathbf{s}^t is a PNE, but in general it can be positive or negative. Using this notation and the defining property (14.9) of a (λ, μ) -smooth cost-minimization game, we can derive

$$\begin{aligned} \text{cost}(\mathbf{s}^t) &\leq \sum_{i=1}^k C_i(\mathbf{s}^t) \\ &= \sum_{i=1}^k [C_i(s_i^*, \mathbf{s}_{-i}^t) + \delta_i(\mathbf{s}^t)] \\ &\leq \lambda \cdot \text{cost}(\mathbf{s}^*) + \mu \cdot \text{cost}(\mathbf{s}^t) + \sum_{i=1}^k \delta_i(\mathbf{s}^t), \end{aligned}$$

and hence

$$\text{cost}(\mathbf{s}^t) \leq \frac{\lambda}{1 - \mu} \cdot \text{cost}(\mathbf{s}^*) + \frac{1}{1 - \mu} \Delta(\mathbf{s}^t). \quad (16.11)$$

This inequality implies that an outcome can be bad only when the amount $\Delta(\mathbf{s}^t)$ that agents have to gain by unilateral deviations to \mathbf{s}^* is large.

In a bad state \mathbf{s}^t , using inequality (16.11) and the assumption that $\Phi(\mathbf{s}) \leq \text{cost}(\mathbf{s})$ for all outcomes \mathbf{s} ,

$$\Delta(\mathbf{s}^t) \geq \eta(1 - \mu) \text{cost}(\mathbf{s}^t) \geq \eta(1 - \mu) \Phi(\mathbf{s}^t). \quad (16.12)$$

If an agent i switches her strategy to a best response in the outcome \mathbf{s}^t , then her cost decreases by at least $\delta_i(\mathbf{s}^t)$. (It could decrease by more, if her best response s_i' is better than s_i^* .) Inequality (16.12) implies

that, in a bad state \mathbf{s}^t , the cost of the agent chosen by maximum-gain best-response dynamics decreases by at least $\frac{\eta(1-\mu)}{k}\Phi(\mathbf{s}^t)$. Since Φ is a potential function and satisfies (13.9),

$$\Phi(\mathbf{s}^{t+1}) \leq \Phi(\mathbf{s}^t) - \max_{i=1}^k \delta_i(\mathbf{s}^t) \leq \left(1 - \frac{\eta(1-\mu)}{k}\right) \cdot \Phi(\mathbf{s}^t)$$

whenever \mathbf{s}^t is a bad state. This inequality, together with the fact that Φ can only decrease in each iteration of best-response dynamics, implies that the potential function decreases by a factor of at least $e = 2.718\dots$ for every sequence of $\frac{k}{\eta(1-\mu)}$ bad states. This yields the desired upper bound of $\frac{k}{\eta(1-\mu)} \ln \frac{\Phi(\mathbf{s}^0)}{\Phi_{\min}}$ on the total number of bad states. ■

The Upshot

- ☆ In each iteration of best-response dynamics, one agent unilaterally deviates to a better strategy.
- ☆ Best-response dynamics converges, necessarily to a PNE, in every potential game.
- ☆ Several variants of ϵ -best-response dynamics, where only moves that yield significant improvements are permitted, converge quickly to an approximate PNE in atomic selfish routing games where all agents have the same origin and destination.
- ☆ In (λ, μ) -smooth games, several variants of best-response dynamics quickly reach outcomes with objective function value almost as good as a PNE.

Notes

Proposition 16.1 and Exercises 16.3–16.4 are from Monderer and Shapley (1996). Theorem 16.3 and Problem 16.2

are due to Chien and Sinclair (2011). Skopalik and Vöcking (2008) show that, if either hypothesis of Theorem 16.3 is dropped, then ϵ -best-response dynamics can require an exponential number of iterations to converge, no matter how the deviating agent and deviation are chosen in each iteration. Approximation bounds for outcome sequences generated by best-response dynamics are first considered in Mirrokni and Vetta (2004). Theorem 16.6 is from Roughgarden (2015), inspired by results of Awerbuch et al. (2008). Problems 16.1 and 16.3 are from Even-Dar et al. (2007) and Milchtaich (1996), respectively.

Exercises

Exercise 16.1 (*H*) Exhibit a game with a PNE and an initial outcome from which best-response dynamics cycles forever.

Exercise 16.2 Consider an atomic selfish routing game (Section 12.4) with m edges and cost functions taking values in $\{1, 2, 3, \dots, H\}$. Prove that best-response dynamics converges to a PNE in at most mH iterations.

Exercise 16.3 A *generalized ordinal potential game* is a cost-minimization game for which there exists a *generalized ordinal potential function* Ψ such that $\Psi(s'_i, \mathbf{s}_{-i}) < \Psi(\mathbf{s})$ whenever $C_i(s'_i, \mathbf{s}_{-i}) < C_i(\mathbf{s})$ for some outcome \mathbf{s} , agent i , and deviation s'_i . Extend Proposition 16.1 to generalized ordinal potential games.

Exercise 16.4 (*H*) Prove the converse of Exercise 16.3: if best-response dynamics always converges to a PNE, for every choice of initial outcome and beneficial unilateral deviation at each iteration, then the game admits a generalized ordinal potential function.

Problems

Problem 16.1 Recall the class of cost-minimization games introduced in Problem 12.3, where each agent $i = 1, 2, \dots, k$ has a positive weight w_i and chooses one of m identical machines to minimize her load. Consider the following restriction of best-response dynamics:

Maximum-Weight Best-Response Dynamics

While the current outcome \mathbf{s} is not a PNE:

among all agents with a beneficial deviation, let i
 denote an agent with the largest weight w_i and
 s'_i a best response to \mathbf{s}_{-i}
 update the outcome to (s'_i, \mathbf{s}_{-i})

Prove that MaxWeight best-response dynamics converges to a PNE in at most k iterations.

Problem 16.2 (*H*) This problem considers another variant of ϵ -best-response dynamics.

ϵ -Best-Response Dynamics (Maximum-Relative-Gain)

While the current outcome \mathbf{s} is not an ϵ -PNE:

among all agents with an ϵ -move, let i denote an
 agent who can obtain the largest relative cost
 decrease

$$\frac{C_i(\mathbf{s}) - \min_{\hat{s}_i \in S_i} C_i(\hat{s}_i, \mathbf{s}_{-i})}{C_i(\mathbf{s})}$$

and s'_i a best response to \mathbf{s}_{-i}
 update the outcome to (s'_i, \mathbf{s}_{-i})

Prove that the iteration bound in Theorem 16.3 applies also to the maximum-relative-gain variant of ϵ -best-response dynamics.

Problem 16.3 This problem considers a variant of the cost-minimization games in Problem 16.1 where every agent has weight 1 but agents can have different individual cost functions. Formally, each agent i incurs a cost $c_j^i(\ell)$ on machine j if she is among ℓ agents using j . Assume that for each fixed i and j , $c_j^i(\ell)$ is nondecreasing in ℓ .

- (a) Prove that if there are only two machines, then best-response dynamics converges to a PNE.
- (b) *(H)* Prove that if there are three machines, then best-response dynamics need not converge.
- (c) *(H)* Prove that, no matter how many machines there are, a PNE always exists.