Lecture 12

Over-Provisioning and Atomic Selfish Routing

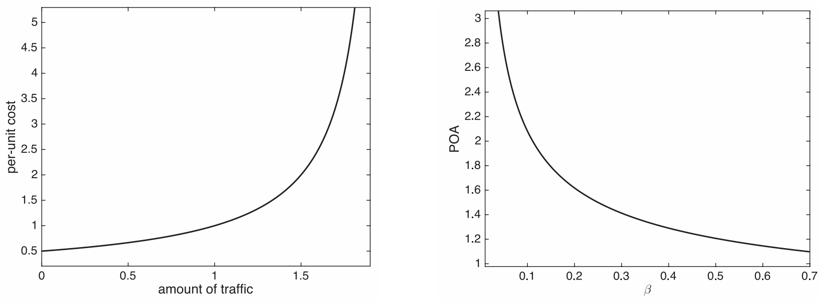
Last lecture proved generic tight bounds on the price of anarchy (POA) of selfish routing, parameterized by the edge cost functions. One particular instantiation of these bounds gives a rigorous justification for the common strategy of over-provisioning a communication network to achieve good performance (Section 12.1). A different result in the same vein states that a modest technology upgrade improves network performance more than implementing dictatorial control (Sections 12.2–12.3). Applications in which network users control a non-negligible fraction of the traffic are best modeled via an "atomic" variant of selfish routing (Section 12.4). The POA of atomic selfish routing is larger than in the "nonatomic" model, but remains bounded provided the network cost functions are affine (Sections 12.4–12.5), or more generally "not too nonlinear."

12.1 Case Study: Network Over-Provisioning

12.1.1 Motivation

The study of selfish routing provides insight into many different kinds of networks, including transportation, communication, and electrical networks. One big advantage in communication networks is that it is often relatively cheap to add additional capacity to a network. Because of this, a popular strategy to communication network management is to install more capacity than is needed, meaning that the network will generally not be close to fully utilized. One motivation for such network over-provisioning is to anticipate future growth in demand. Over-provisioning is also used for performance reasons, as it has been observed empirically that networks tend to suffer fewer packet drops and delays when they have extra capacity.

 https://doi.org/10.1017/CBO9781316779309.013



(a) M/M/1 delay function (b) Extra capacity vs. POA curve

Figure 12.1: Modest over-provisioning guarantees near-optimal routing. The first figure displays the per-unit cost $c(x) = 1/(u - x)$ as a function of the amount of traffic x for an edge with capacity $u = 2$. The second figure shows the worst-case POA as a function of the fraction of unused network capacity.

12.1.2 POA Bounds for Over-Provisioned Networks

The POA bounds for selfish routing developed in Lecture 11 are parameterized by the class of permissible network cost functions. In this section, we consider networks in which every cost function $c_e(x)$ has the form

$$c_e(x) = \begin{cases} \frac{1}{u_e - x} & \text{if } x < u_e \\ +\infty & \text{if } x \geq u_e. \end{cases} \tag{12.1}$$

The parameter u_e represents the capacity of edge e. A cost function of the form (12.1) is the expected per-unit delay in an M/M/1 queue, meaning a queue where jobs arrive according to a Poisson process with rate x and have independent and exponentially distributed services times with mean $1/u_e$. Such a function stays very flat until the amount of traffic nears the capacity, at which point the cost rapidly tends to $+\infty$ (Figure 12.1(a)). This is the simplest cost function used to model delays in communication networks.

For a parameter $\beta \in (0, 1)$, call a selfish routing network with cost functions of the form (12.1) *β-over-provisioned* if $f_e \leq (1 - \beta)u_e$ for every edge e, where f is some equilibrium flow. That is, at equilibrium, the maximum edge utilization in the network is at most $(1 - \beta) \cdot 100\%$.

Figure 12.1(a) suggests the following intuition: when β is not too close to 0, the equilibrium flow is not too close to the capacity on any edge, and in this range the edges' cost functions behave like low-degree polynomials with nonnegative coefficients. Theorem 11.2 implies that the POA is small in networks with such cost functions.

More formally, Theorem 11.2 reduces computing the worst-case POA in arbitrary β-over-provisioned selfish routing networks to computing the worst-case POA merely in β-over-provisioned Pigou-like examples. A computation (Exercise 12.2) then shows that the worst-case POA in β-over-provisioned networks is

$$\frac{1}{2}\left(1+\sqrt{\frac{1}{\beta}}\right), \tag{12.2}$$

an expression graphed in Figure 12.1(b).

Unsurprisingly, the bound in (12.2) tends to 1 as β tends to 1 and to $+\infty$ as β tends to 0. These are the cases where the cost functions effectively act like constant functions and like very high-degree polynomials, respectively. Interestingly, even relatively small values of β imply good POA bounds. For example, if $\beta = .1$, corresponding to a maximum edge utilization of 90%, then the POA is always at most 2.1. Thus a little over-provisioning is sufficient for near-optimal selfish routing, corroborating empirical observations.

12.2 A Resource Augmentation Bound

This section proves a guarantee for selfish routing in arbitrary networks, with no assumptions on the cost functions. What could such a guarantee look like? Recall that the nonlinear variant of Pigou's example (Section 11.1.3) shows that the POA in such networks is unbounded.

The key idea is to compare the performance of selfish routing to a handicapped minimum-cost solution that is forced to route extra traffic. For example, in Figure 11.2(b) with p large, with one unit of traffic, the equilibrium flow has cost 1 while the optimal flow has near-zero cost. If the optimal flow has to route *two* units of traffic, then there is nowhere to hide: it again routes $(1-\epsilon)$ units of traffic on the lower edge, with the remaining $(1+\epsilon)$ units of traffic routed on the upper edge. The cost of this flow exceeds that of the equilibrium flow with one unit of traffic.

This comparison between two flows at different traffic rates has an equivalent and easier-to-interpret formulation as a comparison between two flows with the same traffic rate but in networks with different cost functions. Intuitively, instead of forcing the optimal flow to route additional traffic, we allow the equilibrium flow to use a "faster" network, with each original cost function $c_e(x)$ replaced by the function $c_e(\frac{x}{2})/2$ (Exercise 12.3).

This transformation is particularly meaningful for cost functions of the form (12.1). If $c_e(x) = 1/(u_e - x)$, then the "faster" function is $1/(2u_e - x)$, corresponding to an edge with double the capacity. The next result, after this reformulation, gives a second justification for network over-provisioning: a modest technology upgrade improves performance more than implementing dictatorial control.

Theorem 12.1 (Resource Augmentation Bound) *For every selfish routing network and $r > 0$, the cost of an equilibrium flow with traffic rate r is at most the cost of an optimal flow with traffic rate $2r$.*

Theorem 12.1 also applies to selfish routing networks with multiple origins and destinations (Exercise 12.1).

*12.3 Proof of Theorem 12.1

Fix a network G with nonnegative, nondecreasing, and continuous cost functions, and a traffic rate r. Let f and f^* denote equilibrium and minimum-cost flows at the traffic rates r and $2r$, respectively.

The first part of the proof reuses the trick from the proof of Theorem 11.2 (Section 11.5) of employing fictitious cost functions, frozen at the equilibrium costs, to get a handle on the cost of the optimal flow f^*. Recall that since f is an equilibrium flow (Definition (11.3)), all paths P used by f have a common cost $c_P(f)$, call it L. Moreover, $c_P(f) \geq L$ for every path $P \in \mathcal{P}$. Analogously to (11.5)–(11.8), we have

$$\sum_{e \in E} f_e \cdot c_e(f_e) = \sum_{P \in \mathcal{P}} \underbrace{f_P}_{\text{sums to } r} \cdot \underbrace{c_P(f)}_{= L \text{ if } f_P > 0} = r \cdot L$$

and

$$\sum_{e \in E} f_e^* \cdot c_e(f_e) = \sum_{P \in \mathcal{P}} \underbrace{f_P^*}_{\text{sums to } 2r} \cdot \underbrace{c_P(f)}_{\geq L} \geq 2r \cdot L.$$

With respect to the fictitious frozen costs, we get an excellent lower bound on the cost of f^*, of twice the cost of the equilibrium flow f.

The second step of the proof shows that using the fictitious costs instead of the accurate ones overestimates the cost of f^* by at most the cost of f. Specifically, we complete the proof by showing that

$$\underbrace{\sum_{e \in E} f_e^* \cdot c_e(f_e^*)}_{\text{cost of } f^*} \geq \underbrace{\sum_{e \in E} f_e^* \cdot c_e(f_e)}_{\geq 2rL} - \underbrace{\sum_{e \in E} f_e \cdot c_e(f_e)}_{=rL}. \tag{12.3}$$

We prove that (12.3) holds term-by-term, with

$$f_e^* \cdot [c_e(f_e) - c_e(f_e^*)] \leq f_e \cdot c_e(f_e) \tag{12.4}$$

for every edge $e \in E$. When $f_e^* \geq f_e$, since the cost function c_e is nondecreasing and nonnegative, the left-hand side of (12.4) is nonpositive and there is nothing to show. Nonnegativity of c_e also implies that inequality (12.4) holds when $f_e^* < f_e$. This completes the proof of Theorem 12.1.

12.4 Atomic Selfish Routing

So far we've studied a *nonatomic* model of selfish routing, meaning that all agents have negligible size. This is a good model for cars on a highway or small users of a communication network. This section introduces *atomic* selfish routing networks, the more appropriate model for applications where each agent controls a significant fraction of the overall traffic. For example, an agent could represent an Internet service provider responsible for routing the data of a large number of end users.

An atomic selfish routing network consists of a directed graph $G = (V, E)$ with nonnegative and nondecreasing edge cost functions and a finite number k of agents. Agent i has an origin vertex o_i and a destination vertex d_i. Each agent routes 1 unit of traffic on a single o_i-d_i path, and seeks to minimize her cost.[1] Let $\mathcal{P}_i$ denote the o_i-d_i

[1]Two natural variants of the model allow agents to control different amounts of traffic or to split traffic over multiple paths. Tight worst-case POA bounds for both variants follow from ideas closely related to those of this and the next section. See the Notes for details.

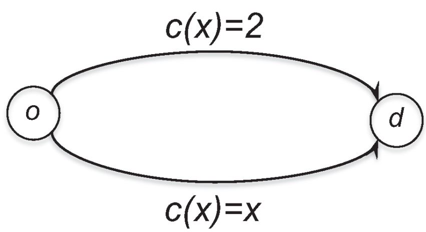


Figure 12.2: A Pigou-like network for atomic selfish routing.

paths of G. A *flow* can now be represented as a vector $(P_1, \ldots, P_k)$, with $P_i \in \mathcal{P}_i$ the path on which agent i routes her traffic. The *cost* of a flow is defined as in the nonatomic model, as (11.3) or (11.4). An *equilibrium* flow is one in which no agent can decrease her cost via a unilateral deviation.

Definition 12.2 (Equilibrium Flow (Atomic)) A flow $(P_1, \ldots, P_k)$ is an *equilibrium* if, for every agent i and path $\hat{P}_i \in \mathcal{P}_i$,

$$\underbrace{\sum_{e \in P_i} c_e(f_e)}_{\text{before deviating}} \leq \underbrace{\sum_{e \in \hat{P}_i \cap P_i} c_e(f_e) + \sum_{e \in \hat{P}_i \setminus P_i} c_e(f_e + 1)}_{\text{after deviating}}.$$

Definition 12.2 differs from Definition 11.3 because a deviation by an agent with non-negligible size increases the cost of the newly used edges.

To get a feel for the atomic model, consider the variant of Pigou's example shown in Figure 12.2. Suppose there are two agents, each controlling one unit of traffic. The optimal solution routes one agent on each edge, for a total cost of $1+2=3$. This is also an equilibrium flow, since neither agent can decrease her cost via a unilateral deviation. The agent on the lower edge does not want to switch, since her cost would jump from 1 to 2. More interestingly, the agent on the upper edge (with cost 2) has no incentive to switch to the lower edge, where her sudden appearance would drive the cost up to 2.

There is a second equilibrium flow in the network: if both agents take the lower edge, then both have a cost of 2 and neither can decrease her cost by switching to the upper edge. This equilibrium has cost 4. This example illustrates an important difference between nonatomic and atomic selfish routing: while different equilibria always have the same cost in the nonatomic model (see Lecture 13), they can have different costs in the atomic model.

Our definition of the POA in Lecture 11 assumes that all equilibria have the same cost. We next extend the definition to games with multiple equilibria using a worst-case approach.[2] Formally, the *price of anarchy (POA)* of an atomic selfish routing network is the ratio

$$\frac{\text{cost of worst equilibrium flow}}{\text{cost of optimal flow}}.$$

For example, in the network in Figure 12.2, the POA is $\frac{4}{3}$.[3]

A second difference between nonatomic and atomic selfish routing is that the POA can be larger in the latter model. To see this, consider the four-agent bidirected triangle network shown in Figure 12.3. Each agent has two options, a one-hop path and a two-hop path. In the optimal flow, every agent routes her traffic on her one-hop path. These one-hop paths are precisely the four edges with the cost function $c(x) = x$, so the cost of this flow is 4. This flow is also an equilibrium flow. On the other hand, if every agent routes her traffic on her two-hop path, then we obtain a second equilibrium flow (Exercise 12.5). Since the first two agents each incur three units of cost and the last two agents each incur two units of cost, the cost of this flow is 10. The POA of this network is $10/4 = 2.5$.

No atomic selfish routing network with affine cost functions has a larger POA.

Theorem 12.3 (POA Bound for Atomic Selfish Routing) *In every atomic selfish routing network with affine cost functions, the POA is at most* $\frac{5}{2}$.

Theorem 12.3 and its proof can be generalized to give tight POA bounds for arbitrary sets of cost functions; see the Notes.

*12.5 Proof of Theorem 12.3

The proof of Theorem 12.3 is a "canonical POA proof," in a sense made precise in Lecture 14. To begin, let's just follow our nose. We

[2]See Lecture 15 for some alternatives.

[3]The POA is well defined in every atomic selfish routing network, as every such network has at least one equilibrium flow (see Theorem 13.6).

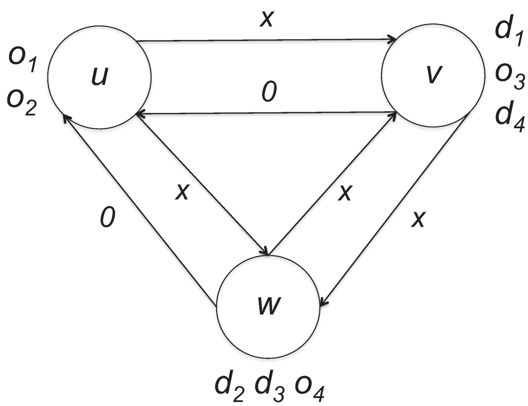


Figure 12.3: In atomic selfish routing networks with affine cost functions, the POA can be as large as 5/2.

need to bound from above the cost of every equilibrium flow; fix one f arbitrarily. Let f^* denote a minimum-cost flow. Write f_e and f_e^* for the number of agents in f and f^*, respectively, that pick a path that includes the edge e. Write each affine cost function as $c_e(x) = a_e x + b_e$ for $a_e, b_e \geq 0$.

The first step of the proof identifies a useful way of applying our hypothesis that f is an equilibrium flow. If we consider any agent i, using the path P_i in f, and any unilateral deviation to a different path $\hat{P}_i$, then we can conclude that i's equilibrium cost using P_i is at most what her cost would be if she switched to $\hat{P}_i$ (Definition 12.2). This looks promising: we want an upper bound on the cost of the equilibrium flow f, and hypothetical deviations give us upper bounds on the equilibrium costs of individual agents. Which hypothetical deviations should we single out for the proof? Given that f^* is the only other object referenced in the theorem statement, a natural idea is to use the optimal flow f^* to suggest deviations.

Formally, suppose agent i uses path P_i in f and path P_i^* in f^*. By Definition 12.2,

$$\sum_{e \in P_i} c_e(f_e) \leq \sum_{e \in P_i^* \cap P_i} c_e(f_e) + \sum_{e \in P_i^* \setminus P_i} c_e(f_e + 1). \qquad (12.5)$$

This completes the first step, in which we apply the equilibrium hypothesis to generate an upper bound (12.5) on the equilibrium cost of each agent.

The second step of the proof sums the upper bound (12.5) on individual equilibrium costs over all agents to obtain a bound on the total equilibrium cost:

$$\underbrace{\sum_{i=1}^{k} \sum_{e \in P_i} c_e(f_e)}_{\text{cost of } f} \leq \sum_{i=1}^{k} \left(\sum_{e \in P_i^* \cap P_i} c_e(f_e) + \sum_{e \in P_i^* \setminus P_i} c_e(f_e + 1) \right) \tag{12.6}$$

$$\leq \sum_{i=1}^{k} \sum_{e \in P_i^*} c_e(f_e + 1) \tag{12.7}$$

$$= \sum_{e \in E} f_e^* \cdot c_e(f_e + 1) \tag{12.8}$$

$$= \sum_{e \in E} \left[a_e f_e^* (f_e + 1) + b_e f_e^* \right], \tag{12.9}$$

where inequality (12.6) follows from (12.5), inequality (12.7) from the assumption that cost functions are nondecreasing, equation (12.8) from the fact that the term $c_e(f_e + 1)$ is contributed once by each agent i for which $e \in P_i^*$ (f_e^* times in all), and equation (12.9) from the assumption that cost functions are affine. This completes the second step of the proof.

The previous step gives an upper bound on a quantity that we care about—the cost of the equilibrium flow f—in terms of a quantity that we don't care about, the "entangled" version of f and f^* on the right-hand side of (12.9). The third and most technically challenging step of the proof is to "disentangle" the right-hand side of (12.9) and relate it to the only quantities that we care about for a POA bound, the costs of f and f^*.

We use the following inequality, which is easily checked (Exercise 12.6).

Lemma 12.4 *For every $y, z \in \{0, 1, 2, 3, \ldots\}$,*

$$y(z + 1) \leq \frac{5}{3} y^2 + \frac{1}{3} z^2.$$

We now apply Lemma 12.4 once per edge in the right-hand side of (12.9), with $y = f_e^*$ and $z = f_e$. Using the definition (11.4) of the cost $C(\cdot)$ of a flow, this yields

$$\begin{aligned} C(f) &\le \sum_{e\in E}\left[a_e\left(\frac{5}{3}(f_e^*)^2+\frac{1}{3}f_e^2\right)+b_ef_e^*\right] \\ &\le \frac{5}{3}\left[\sum_{e\in E} f_e^*(a_ef_e^*+b_e)\right]+\frac{1}{3}\sum_{e\in E}a_ef_e^2 \\ &\le \frac{5}{3}\cdot C(f^*)+\frac{1}{3}\cdot C(f). \end{aligned} \tag{12.10}$$

Subtracting $\frac{1}{3}C(f)$ from both sides and multiplying through by $\frac{3}{2}$ gives

$$C(f) \le \frac{5}{3}\cdot\frac{3}{2}\cdot C(f^*) = \frac{5}{2}\cdot C(f^*),$$

which completes the proof of Theorem 12.3.

The Upshot

☆ A selfish routing network with cost functions of the form $c_e(x) = 1/(u_e - x)$ is β-over-provisioned if the amount of equilibrium flow on each edge e is at most $(1-\beta)u_e$.

☆ The POA is small in β-over-provisioned networks even with fairly small β, corroborating empirical observations that a little over-provisioning yields good network performance.

☆ The cost of an equilibrium flow is at most that of an optimal flow that routes twice as much traffic. Equivalently, a modest technology upgrade improves performance more than implementing dictatorial control.

☆ In atomic selfish routing, where each agent controls a non-negligible fraction of the network traffic, different equilibrium flows can have different costs.

☆ The POA is the ratio between the objective function value of the worst equilibrium and that of an optimal outcome.

☆ The worst-case POA of atomic selfish routing with affine cost functions is exactly 2.5. The proof is "canonical" is a price sense.

Notes

Bertsekas and Gallager (1987) is a good reference for models of communication networks, and Olifer and Olifer (2005) for communication network management strategies such as over-provisioning. POA bounds for β-over-provisioned networks are discussed by Roughgarden (2010a). Theorem 12.1 is due to Roughgarden and Tardos (2002). Atomic selfish routing networks first appear in Rosenthal (1973), and the POA of such networks is first studied in Awerbuch et al. (2013) and Christodoulou and Koutsoupias (2005b). Our proof of Theorem 12.3 follows Christodoulou and Koutsoupias (2005a). Defining the POA via the worst-case equilibrium is the original proposal of Koutsoupias and Papadimitriou (1999). See Aland et al. (2011) for tight POA bounds for atomic selfish routing networks with polynomial cost functions, and Roughgarden (2015) for general cost functions. Tight POA bounds for agents controlling different amounts of traffic are given by Awerbuch et al. (2013), Christodoulou and Koutsoupias (2005b), Aland et al. (2011), and Bhawalkar et al. (2014). For agents who can split traffic over multiple paths, POA bounds appear in Cominetti et al. (2009) and Harks (2011), and tight bounds in Roughgarden and Schoppmann (2015). In all of these atomic models, when edges' cost functions are polynomials with nonnegative coefficients and degree at most p, the POA is bounded by a constant that depends on p. The dependence on p is exponential, in contrast to the sublinear dependence in nonatomic selfish routing networks. Problems 12.1–12.4 are from Chakrabarty (2004), Christodoulou and Koutsoupias (2005b), Koutsoupias and Papadimitriou (1999), and Awerbuch et al. (2006), respectively.

Exercises

Exercise 12.1 Multicommodity selfish routing networks are defined in Exercise 11.5. Generalize Theorem 12.1 to such networks.

Exercise 12.2 This exercise outlines the proof that the worst-case POA in β-over-provisioned networks is at most the expression in (12.2).

(a) Prove that, in a Pigou-like network (Section 11.3) with traffic rate r and cost function $1/(u-x)$ with $u > r$ on the lower edge, the POA is the expression in (12.2), where $\beta = 1 - \frac{r}{u}$.

(b) Adapt the Pigou bound (11.2) to β-over-provisioned networks by defining

$$\alpha_\beta = \sup_{u>0} \sup_{r \in [0,(1-\beta)u]} \sup_{x \geq 0} \left\{ \frac{r \cdot c_u(r)}{x \cdot c_u(x) + (r-x) \cdot c_u(r)} \right\},$$

where c_u denotes the cost function $1/(u-x)$. Prove that, for every $\beta \in (0,1)$, α_β equals the expression in (12.2).

(c) *(H)* Prove that the POA of every β-over-provisioned network is at most the expression in (12.2).

Exercise 12.3 Prove that the following statement is equivalent to Theorem 12.1: If f^* is a minimum-cost flow in a selfish routing network with cost functions c and f is an equilibrium flow in the same network with cost functions $\tilde{c}$, where $\tilde{c}_e(x)$ is defined as $c_e(x/2)/2$, then

$$\widetilde{C}(f) \leq C(f^*).$$

The notation $\widetilde{C}$ and C refers to the cost of a flow (11.3) with the cost functions $\tilde{c}$ and c, respectively.

Exercise 12.4 Prove the following generalization of Theorem 12.1: for every selfish routing network and $r, \delta > 0$, the cost of an equilibrium flow with traffic rate r is at most $\frac{1}{\delta}$ times the cost of an optimal flow with traffic rate $(1+\delta)r$.

Exercise 12.5 Verify that if each agent routes her traffic on her two-hop path in the network in Figure 12.3, then the result is an equilibrium flow.

Exercise 12.6 *(H)* Prove Lemma 12.4.

Problems

Problem 12.1 *(H)* For selfish routing networks with affine cost functions, prove the following stronger version of Theorem 12.1: for every such network and $r > 0$, the cost of an equilibrium flow with traffic rate r is at most that of an optimal flow with traffic rate $\frac{5}{4}r$.

Problem 12.2 Recall the four-agent atomic selfish routing network in Figure 12.3, where the POA is 2.5.

(a) Using a different network, show that the POA of atomic selfish routing with affine cost functions can be 2.5 even when there are only three agents.

(b) How large can the POA be with affine cost functions and only two agents?

Problem 12.3 This problem studies a scenario with k agents, where agent i has a positive weight w_i. There are m identical machines. Each agent chooses a machine, and wants to minimize the *load* of her machine, defined as the sum of the weights of the agents who choose it. This problem considers the objective of minimizing the *makespan*, defined as the maximum load of a machine. A *pure Nash equilibrium* is an assignment of agents to machines so that no agent can unilaterally switch machines and decrease the load she experiences.

(a) *(H)* Prove that the makespan of a pure Nash equilibrium is at most twice that of the minimum possible.

(b) Prove that, as k and m tend to infinity, pure Nash equilibria can have makespan arbitrarily close to twice the minimum possible.

Problem 12.4 This problem modifies the model in Problem 12.3 in two ways. First, every agent has unit weight. Second, each agent i must choose from a restricted subset S_i of the m machines.

(a) Prove that, for every constant $a \geq 1$, with sufficiently many agents and machines, the makespan of a pure Nash equilibrium can be more than a times the minimum makespan of a feasible schedule (assigning each agent i to a machine in her set S_i).

(b) Prove that there is a constant $a > 0$ such that the makespan of every pure Nash equilibrium is at most $a \ln m$ times the minimum possible. Can you obtain an even tighter dependence on the number m of machines?