Potential functions and the inefficiency of equilibria

Tim Roughgarden*

Abstract. We survey one area of the emerging field of algorithmic game theory: the use of approximation measures to quantify the inefficiency of game-theoretic equilibria. Potential functions, which enable the application of optimization theory to the study of equilibria, have been a versatile and powerful tool in this area. We use potential functions to bound the inefficiency of equilibria in three diverse, natural classes of games: selfish routing networks, resource allocation games, and Shapley network design games.

Mathematics Subject Classification (2000). 68Q25, 68W25, 90B35, 91A65.

Keywords. Game theory, inefficiency of equilibria, Nash equilibrium, network design, potential functions, price of anarchy, price of stability, resource allocation, selfish routing.

1. Introduction

The interface between theoretical computer science and microeconomics, often called *algorithmic game theory*, has been an extremely active research area over the past few years. Recent points of contact between the two fields are diverse and include, for example, increased attention to computational complexity and approximation in combinatorial auctions (e.g. [9]); a new focus on worst-case analysis in optimal auction design (e.g. [17]); and a renewed emphasis on the computability and learnability of equilibrium concepts (e.g. [14], [18], [26]). This survey touches on just one connection between theoretical computer science and game theory: the use of approximation measures to quantify the inefficiency of game-theoretic equilibria.

1.1. Quantifying the inefficiency of equilibria. Even in very simple settings, selfish behavior can lead to highly inefficient outcomes [11]. A canonical example of this phenomenon is provided by the "Prisoner's Dilemma" [28], in which strategic behavior by two captured and separated prisoners inexorably draws them into the worst-possible outcome. We will see several concrete examples of the inefficiency of selfish behavior in networks later in the survey.

Must more recently, researchers have sought to *quantify* the inefficiency of selfish behavior. Koutsoupias and Papadimitriou [23] proposed a framework to systematically study this issue. The framework presupposes a strategic environment (a *game*), a definition for the outcome of selfish behavior (an *equilibrium concept*), and a real-

*Supported in part by ONR grant N00014-04-1-0725, DARPA grant W911NF-04-9-0001, and an NSF CAREER Award.

Proceedings of the International Congress
of Mathematicians, Madrid, Spain, 2006
© 2006 European Mathematical Society

valued, nonnegative objective function defined on the possible outcomes of the game. The *price of anarchy* [23], [26] is then defined as the ratio between the objective function value of an equilibrium and that of an optimal solution. (For the moment, we ignore the question of whether or not equilibria exist and are unique.) If the price of anarchy of a game is 1, then its equilibria are fully efficient. More generally, bounding the price of anarchy in a class of games provides a guarantee on the worst-possible inefficiency of equilibria in these games.

The price of anarchy is directly inspired by other popular notions of approximation in theoretical computer science [23]. One example is the *approximation ratio* of a heuristic for a (typically NP-hard) optimization problem, defined as the largest ratio between the objective function value of the solution produced by the heuristic and that of an optimal solution. While the approximation ratio measures the worst-case loss in solution quality due to insufficient computational effort, the price of anarchy measures the worst-case loss arising from insufficient ability (or willingness) to control and coordinate the actions of selfish individuals. Much recent research on the price of anarchy is motivated by optimization problems that naturally occur in the design and management of large networks (like the Internet), in which users act selfishly, but implementing an optimal solution is not practical.

1.2. Potential functions. The price of anarchy has been successfully analyzed in a diverse array of game-theoretic models (see e.g. [32], [33] and the references therein). This survey discusses three of these models, with the goal of illustrating a single mathematical tool for bounding the price of anarchy: *potential functions*. The potential function technique is by no means the only one known for bounding the inefficiency of equilibria, but (so far) it has been the most versatile and powerful.

Potential functions enable the application of optimization theory to the study of equilibria. More precisely, a potential function for a game is a real-valued function, defined on the set of possible outcomes of the game, such that the equilibria of the game are precisely the local optima of the potential function. This idea was first used to analyze selfish behavior in networks by Beckmann, McGuire, and Winsten [4], though similar ideas were used earlier in other contexts.

When a game admits a potential function, there are typically consequences for the existence, uniqueness, and inefficiency of equilibria. For example, suppose a game admits a potential function and either: (1) there are a finite number of distinct outcomes; or (2) the set of outcomes is compact and the potential function is continuous. In either case, the potential function achieves a global optimum, which is also a local optimum, and hence the game has at least one equilibrium. This is a much more elementary approach to establishing the existence of equilibria than traditional fixed-point proofs (e.g. [25]). Moreover, if the potential function has a unique local optimum, then the game has a unique equilibrium. Finally, if the potential function is "close to" the true objective function, then the equilibria that are global optima of the potential function have nearly-optimal objective function value, and are thus approximately efficient.

The power of the potential function approach might suggest that its applicability is limited. Fortunately, many important and natural classes of games admit well-behaved potential functions. To suggest what such functions look like, we briefly interpret some classical results about electric networks in terms of potential functions. Consider electrical current in a two-terminal network of resistors. By Kirchhoff's equations and Ohm's law, we can interpret this current as an "equilibrium", in the sense that it equalizes the voltage drop along all paths in the network between the two terminals. (View current as a large population of "selfish particles", each seeking out a path with minimum voltage drop.) On the other hand, Thomson's principle states that electrical current also minimizes the dissipated energy over all flow patterns that achieve the same total current. In other words, energy dissipation serves as a potential function for current in an electrical network. For further details and discussion, see Kelly [21] and Doyle and Snell [10].

1.3. Survey overview. Each of the next three sections introduces a model of selfish behavior in networks, and uses a potential function to bound the inefficiency of equilibria in the model. We focus on these three examples because they are simple, natural, and diverse enough to illustrate different aspects of potential function proof techniques. In order to emphasize the most important concepts and provide a number of self-contained proofs, we often discuss only special cases of more general models and results.

Section 2 discusses selfish routing networks, a model that generalizes the electrical networks of Subsection 1.2 and has been extensively studied by the transportation, networking, and theoretical computer science communities. Section 3 analyzes the performance of a well-studied distributed protocol for allocating resources to heterogeneous users. Section 4 bounds the inefficiency of equilibria in a model of selfish network design. Section 5 concludes.

2. Selfish routing and the price of anarchy

2.1. The model. In this section, we study the inefficiency of equilibria in the following model of noncooperative network routing. A *multicommodity flow network*, or simply a *network*, is a finite directed graph $G = (V, E)$, with vertex set V and (directed) edge set E, together with a set $(s_1, t_1), \ldots, (s_k, t_k)$ of source-sink vertex pairs. We also call such pairs *commodities*. We denote the set of simple s_i-t_i paths by $\mathcal{P}_i$, and always assume that this set is non-empty for each i. We allow the graph G to contain parallel edges, and a vertex can participate in multiple source-sink pairs.

A *flow* in a network G is a nonnegative vector indexed by the set $\mathcal{P} = \bigcup_{i=1}^{k} \mathcal{P}_i$. For a flow f and a path $P \in \mathcal{P}_i$, we interpret f_P as the amount of traffic of commodity i that chooses the path P to travel from s_i to t_i. We use r to denote a nonnegative vector of *traffic rates*, indexed by the commodities of G. A flow f is *feasible* for r if it routes all of the prescribed traffic: for each $i \in \{1, 2, \ldots, k\}$, $\sum_{P \in \mathcal{P}_i} f_P = r_i$.

We model the negative consequences of network congestion in the following simple way. For a flow f in a network G and an edge e of G, let $f_e = \sum_{P \in \mathcal{P} : e \in P} f_P$ denote the total amount of traffic employing edge e. Each edge e then has a nonnegative, continuous, and nondecreasing *cost function* c_e, which describes the cost incurred by traffic using the edge e as a function of f_e. We call a triple of the form (G, r, c) a *selfish routing network* or simply an *instance*.

Next we describe a notion of equilibrium in selfish routing networks – the expected outcome of "selfish routing". Define the cost of a path P with respect to a flow f as the sum of the costs of the constituent edges: $c_P(f) = \sum_{e \in P} c_e(f_e)$. Assuming that selfish traffic attempts to minimize its incurred cost, we obtain the following definition of a *Wardrop equilibrium* [38].

Definition 2.1 ([38]). Let f be a feasible flow for the instance (G, r, c). The flow f is a *Wardrop equilibrium* if, for every commodity $i \in \{1, 2, \ldots, k\}$ and every pair $P, \tilde{P} \in \mathcal{P}_i$ of s_i-t_i paths with $f_P > 0$, $c_P(f) \leq c_{\tilde{P}}(f)$.

In Definition 2.1, we are implicitly assuming that every network user controls a negligible portion of the overall traffic, so that the actions of an individual user have essentially no effect on the network congestion. In the game theory literature, games with this property are called *nonatomic* [35]. Atomic variants of selfish routing have also been extensively studied (see e.g. [32]). We will study other types of atomic games in Sections 3 and 4.

Example 2.2 (Pigou's example [27]). Consider the two-vertex, two-edge network shown in Figure 1. There is one commodity and the traffic rate is 1. Note that the lower edge is cheaper than the upper edge if and only if less than one unit of traffic uses it. There is thus a unique Wardrop equilibrium, with all traffic routed on the lower edge. In this flow, all traffic incurs one unit of cost.

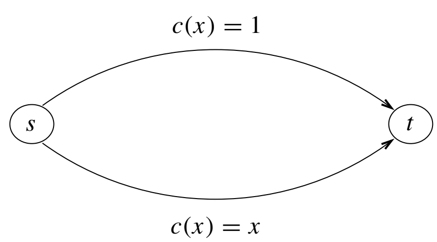

Figure 1. Pigou's example (Example 2.2).

Pigou's example already illustrates that equilibria can be inefficient. More specifically, note that routing half of the traffic on each of the two edges would produce a "better" flow: all of the traffic would incur at most one unit of cost, while half of the traffic would incur only $1/2$ units of cost.

The inefficiency of the Wardrop equilibrium in Example 2.2 arises from what is often called a *congestion externality* – a selfish network user accounts only for its own cost, and not for the costs that its decision imposes on others. The "better" routing of traffic in Example 2.2 is not a Wardrop equilibrium because a selfish network user routed on the upper edge would switch to the lower edge, indifferent to the fact that this switch (slightly) increases the cost incurred by a large portion of the population.

In Example 2.2, there is a unique Wardrop equilibrium. In Subsection 2.2 we will use a potential function to prove the following theorem, which states that Wardrop equilibria exist and are "essentially unique" in all selfish routing networks.

Theorem 2.3 ([4]). *Let (G, r, c) be an instance.*

(a) *The instance (G, r, c) admits at least one Wardrop equilibrium.*

(b) *If f and $\tilde{f}$ are Wardrop equilibria for (G, r, c), then $c_e(f_e) = c_e(\tilde{f}_e)$ for every edge e.*

The Wardrop equilibrium in Example 2.2 is intuitively inefficient; we next quantify this inefficiency. We define our objective function, the *cost* of a flow, as the sum of the path costs incurred by traffic:

$$C(f) = \sum_{P \in \mathcal{P}} c_P(f) f_P = \sum_{e \in E} c_e(f_e) f_e. \tag{1}$$

The first equality in (1) is a definition; the second follows easily from the definitions. An *optimal flow* for an instance (G, r, c) is feasible and minimizes the cost. Since cost functions are continuous and the set of feasible flows is compact, every instance admits an optimal flow. In Pigou's example (Example 2.2), the Wardrop equilibrium has cost 1, while routing half of the traffic on each edge yields an optimal flow with cost $3/4$.

Definition 2.4 ([23], [26]). The *price of anarchy* $\rho(G, r, c)$ of an instance (G, r, c) is

$$\rho(G, r, c) = \frac{C(f)}{C(f^*)},$$

where f is a Wardrop equilibrium and f^* is an optimal flow. The *price of anarchy* $\rho(\mathcal{I})$ of a non-empty set $\mathcal{I}$ of instances is $\sup_{(G,r,c)\in\mathcal{I}} \rho(G, r, c)$.

Definition 2.1 and Theorem 2.3(b) easily imply that all Wardrop equilibria have equal cost, and thus the price of anarchy of an instance is well defined unless there is a flow with zero cost. In this case, all Wardrop equilibria also have zero cost, and we define the price of anarchy of the instance to be 1.

Example 2.5 (Nonlinear Pigou's example [34]). The inefficiency of the Wardrop equilibrium in Example 2.2 can be amplified with a seemingly minor modification to the network. Suppose we replace the previously linear cost function $c(x) = x$ on the lower edge with the highly nonlinear one $c(x) = x^p$ for p large (Figure 2). As in Example 2.2, the cost of the unique Wardrop equilibrium is 1. The optimal flow routes a small ε fraction of the traffic on the upper edge and has cost $\varepsilon + (1-\varepsilon)^{p+1}$. Since this approaches 0 as ε tends to 0 and p tends to infinity, the price of anarchy of this selfish routing network grows without bound as p grows large.

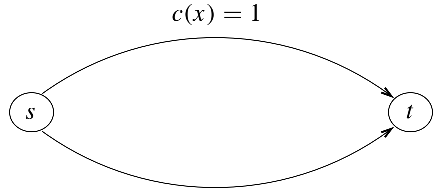

Figure 2. A nonlinear variant of Pigou's example (Example 2.5).

Example 2.5 demonstrates that the price of anarchy can be large in (very simple) networks with nonlinear cost functions. In Subsection 2.2 we use a potential function to show the converse: the price of anarchy is large *only* in networks with "highly nonlinear" cost functions.

2.2. A potential function for wardrop equilibria. We now show that Wardrop equilibria can be characterized as the minima of a potential function, and use this characterization to prove both Theorem 2.3 and upper bounds on the price of anarchy of selfish routing. To motivate this potential function, we first characterize the optimal flows of an instance.

Optimal flows for an instance (G, r, c) minimize the cost (1) subject to linear flow feasibility constraints. Assume for the moment that for every edge e, the function $x \cdot c_e(x)$ is convex. The cost (1) is then a convex (separable) function, and we can apply the Karush–Kuhn–Tucker conditions (see e.g. [5]) to characterize its global minima. To state this characterization cleanly, assume further that all cost functions are differentiable, and let $c^*_e(x) = (x \cdot c_e(x))' = c_e(x) + x \cdot c'_e(x)$ denote the *marginal cost function* for the edge e. The KKT conditions then give the following.

Proposition 2.6 ([4]). *Let (G, r, c) be an instance such that, for every edge e, the function $x \cdot c_e(x)$ is convex and differentiable. Let c^*_e denote the marginal cost function of the edge e. Then f^* is an optimal flow for (G, r, c) if and only if, for every commodity $i \in \{1, 2, \ldots, k\}$ and every pair $P, \tilde{P} \in \mathcal{P}_i$ of s_i-t_i paths with $f_P > 0$, $c^*_P(f) \le c^*_{\tilde{P}}(f)$.*

Comparing Definition 2.1 and Proposition 2.6, we discover that Wardrop equilibria and optimal flows are essentially the same thing, just with different sets of cost functions.

Corollary 2.7. *Let (G, r, c) be an instance such that, for every edge e, the function $x \cdot c_e(x)$ is convex and differentiable. Let c_e^* denote the marginal cost function of the edge e. Then f^* is an optimal flow for (G, r, c) if and only if it is a Wardrop equilibrium for (G, r, c^*).*

To construct a potential function for Wardrop equilibria, we need to "invert" Corollary 2.7: of what function do Wardrop equilibria arise as global minima? The answer is simple: to recover Definition 2.1 as an optimality condition, we seek a function $h_e(x)$ for each edge e – playing the previous role of $x \cdot c_e(x)$ – such that $h'_e(x) = c_e(x)$. Setting $h_e(x) = \int_0^x c_e(y)\,dy$ for each edge e thus yields the desired potential function.

Precisely, call

$$\Phi(f) = \sum_{e \in E} \int_0^{f_e} c_e(x)\,dx \tag{2}$$

the *potential function* for an instance (G, r, c). Analogously to Corollary 2.7, the following proposition holds.

Proposition 2.8 ([4]). *Let (G, r, c) be an instance. A flow feasible for (G, r, c) is a Wardrop equilibrium if and only if it is a global minimum of the corresponding potential function Φ given in* (2).

Remark 2.9. Thomson's principle for electrical networks (Subsection 1.2) can be viewed as the special case of Proposition 2.8 for single-commodity flow networks with linear cost functions (of the form $c_e(x) = a_e x$).

Theorem 2.3 now follows easily.

Proof of Theorem 2.3 (Sketch). Since cost functions are continuous and the set of feasible flows is compact, part (a) of the theorem follows immediately from Proposition 2.8 and Weierstrass's Theorem. Since cost functions are nondecreasing, the potential function Φ in (2) is convex; moreover, the set of feasible flows is convex. Part (b) of the theorem now follows from routine convexity arguments. □

Much more recently, the potential function (2) has been used to upper bound the price of anarchy of selfish routing. The intuition behind this connection is simple: if Wardrop equilibria exactly optimize a potential function (2) that is a good approximation of the objective function (1), then they should also be approximately optimal. Formally, we have the following.

Theorem 2.10 ([34]). *Let (G, r, c) be an instance, and suppose that $x \cdot c_e(x) \le \gamma \cdot \int_0^x c_e(y)\,dy$ for all $e \in E$ and $x \ge 0$. Then the price of anarchy $\rho(G, r, c)$ is at most γ.*

Proof. Let f and f^* be a Wardrop equilibrium and an optimal flow for (G, r, c), respectively. Since cost functions are nondecreasing, the cost of a flow (1) is always at least its potential function value (2). The hypothesis ensures that the cost of a flow is at most γ times its potential function value. The theorem follows by writing

$$C(f) \leq \gamma \cdot \Phi(f) \leq \gamma \cdot \Phi(f^*) \leq \gamma \cdot C(f^*),$$

with the second inequality following from Proposition 2.8. □

Theorem 2.10 implies that the price of anarchy of selfish routing is large only in networks with "highly nonlinear" cost functions. For example, if c_e is a polynomial function with degree at most p and nonnegative coefficients, then $x \cdot c_e(x) \leq (p+1)\int_0^x c_e(y)\,dy$ for all $x \geq 0$. Applying Theorem 2.10, we find that the price of anarchy in networks with cost functions that are polynomials with nonnegative coefficients grows at most linearly with the degree bound p.

Corollary 2.11 ([34]). *If (G, r, c) is an instance with cost functions that polynomials with nonnegative coefficients and degree at most p, then $\rho(G, r, c) \leq p + 1$.*

This upper bound is nearly matched by Example 2.5. (The upper and lower bounds differ by roughly a $\ln p$ multiplicative factor.) Qualitatively, Example 2.5 and Corollary 2.11 imply that a large price of anarchy can be caused by highly nonlinear cost functions, but not by complex network topologies or by a large number of commodities.

2.3. An optimal bound on the price of anarchy. We have established that the price of anarchy of selfish routing depends on the "degree of nonlinearity" of the network cost functions. However, even in the simple case of polynomial cost functions, there is gap between the lower bound on the price of anarchy provided by Example 2.5 and the upper bound of Theorem 2.10. We conclude this section by showing how a different analysis, which can be regarded as a more "global" application of potential function ideas, provides a tight bound on the price of anarchy for essentially every set of allowable cost functions.

We first formalize a natural lower bound on the price of anarchy based on "Pigou-like examples".

Definition 2.12 ([7], [31]). Let $\mathcal{C}$ be a nonempty set of cost functions. The *Pigou bound* $\alpha(\mathcal{C})$ for $\mathcal{C}$ is

$$\alpha(\mathcal{C}) = \sup_{c \in \mathcal{C}} \sup_{x, r \geq 0} \frac{r \cdot c(r)}{x \cdot c(x) + (r - x)c(r)}, \tag{3}$$

with the understanding that $0/0 = 1$.

The point of the Pigou bound is that it lower bounds the price of anarchy in instances with cost functions in $\mathcal{C}$.

Proposition 2.13. *Let $\mathcal{C}$ be a set of cost functions that contains all of the constant cost functions. Then $\rho(\mathcal{C}) \geq \alpha(\mathcal{C})$.*

Proof. Fix a choice of $c \in \mathcal{C}$ and $x, r \geq 0$. We can complete the proof by exhibiting a selfish routing network with cost functions in $\mathcal{C}$ and price of anarchy at least $c(r)r/[c(x)x + (r-x)c(r)]$. Since c is nondecreasing, this expression is at least 1 if $x \geq r$; we can therefore assume that $x < r$.

Let G be a two-vertex, two-edge network as in Figure 1. Give the lower edge the cost function $c_1(y) = c(y)$ and the upper edge the constant cost function $c_2(y) = c(r)$. By assumption, both of these cost functions lie in $\mathcal{C}$. Set the traffic rate to r. Routing all of the traffic on the lower edge yields a Wardrop equilibrium with cost $c(r)r$. Routing x units of traffic on the lower edge and $r - x$ units of traffic on the upper edge gives a feasible flow with cost $[c(x)x + (r-x)c(r)]$. The price of anarchy in this instance is thus at least $c(r)r/[c(x)x + (r-x)c(r)]$, as desired □

Proposition 2.13 holds more generally for every set $\mathcal{C}$ of cost functions that is *inhomogeneous* in the sense that $c(0) > 0$ for some $c \in \mathcal{C}$ [31].

We next show that, even though the Pigou bound is based only on Pigou-like examples, it is also an upper bound on the price of anarchy in general multicommodity flow networks. The proof requires the following *variational inequality* characterization of Wardrop equilibria, first noted by Smith [36].

Proposition 2.14 ([36]). *A flow f feasible for (G, r, c) is a Wardrop equilibrium if and only if*

$$\sum_{e \in E} c_e(f_e) f_e \leq \sum_{e \in E} c_e(f_e) f_e^*$$

for every flow f^* feasible for (G, r, c).

Proposition 2.14 can be derived as an optimality condition for minimizers of the potential function (2), or can be proved directly using Definition 2.1.

We now show that the Pigou bound is tight.

Theorem 2.15 ([7], [31]). *Let $\mathcal{C}$ be a set of cost functions and $\alpha(\mathcal{C})$ the Pigou bound for $\mathcal{C}$. If (G, r, c) is an instance with cost functions in $\mathcal{C}$, then*

$$\rho(G, r, c) \leq \alpha(\mathcal{C}).$$

Proof. Let f^* and f be an optimal flow and a Wardrop equilibrium, respectively, for an instance (G, r, c) with cost functions in the set $\mathcal{C}$. The theorem follows by writing

$$\begin{aligned} C(f^*) &= \sum_{e \in E} c_e(f_e^*) f_e^* \\ &\geq \frac{1}{\alpha(\mathcal{C})} \sum_{e \in E} c_e(f_e) f_e + \sum_{e \in E} (f_e^* - f_e) c_e(f_e) \\ &\geq \frac{C(f)}{\alpha(\mathcal{C})}, \end{aligned}$$

where the first inequality follows from Definition 2.12 and the second from Proposition 2.14. □

Different, more recent proofs of Theorem 2.15 can be found in [8], [37].

Proposition 2.13 and Theorem 2.15 establish the qualitative statement that, for essentially every fixed restriction on the allowable network cost functions, the price of anarchy is maximized by Pigou-like examples. Determining the largest-possible price of anarchy in Pigou-like examples (i.e., the Pigou bound) is a tractable problem in many cases. For example, it is precisely 4/3 when $\mathcal{C}$ is the affine functions [34], and more generally is $[1 - p \cdot (p+1)^{-(p+1)/p}]^{-1} \approx p/\ln p$ when $\mathcal{C}$ is the set of polynomials with degree at most p and nonnegative coefficients [31]. In these cases, the maximum price of anarchy (among all multicommodity flow networks) is achieved by the instances in Examples 2.2 and 2.5. For further examples, see [7], [31].

For much more on topics related to the price of anarchy of selfish routing, including many extensions and generalizations of the results described in this section, see [32], [33] and the references therein.

3. Efficiency loss in resource allocation protocols

We next study the performance of a protocol for allocating resources to heterogeneous users. While there are a number of conceptual differences between this model and the selfish routing networks of Section 2, the inefficiency of equilibria in these models can be analyzed in a similar way.

3.1. The model. We consider a single divisible resource – the capacity of a single network link, say – to be allocated to a finite number $n > 1$ of competing users. These users are *heterogeneous* in the sense that different users can have different values for capacity. We model this by giving each user i a nonnegative real-valued *utility function* U_i that expresses this user's value for a given amount of capacity. We assume that each U_i is concave, strictly increasing, and continuously differentiable. A *resource allocation game* is defined by the n utility functions $U_1, \dots, U_n$ and the link capacity $C > 0$.

An *allocation* for a resource allocation game is a nonnegative vector $(x_1, \dots, x_n)$ with $\sum_{i=1}^n x_i = C$. We study the following protocol for allocating capacity. Each user i submits a nonnegative *bid* b_i for capacity. The protocol allocates capacity in proportion to bids, with

$$x_i = \frac{b_i}{\sum_{j=1}^n b_j} \cdot C \tag{4}$$

units of capacity allocated to user i. The *payoff* Q_i to a user i is defined as its value for the capacity it receives, minus the bid that it made (and presumably now has to pay):

$$Q_i(b_1, \dots, b_n) = U_i(x_i) - b_i = U_i\left(\frac{b_i}{\sum_{j=1}^n b_j} \cdot C\right) - b_i.$$

Assume that if all users bid zero, then all users receive zero payoff.

An *equilibrium* is then a bid vector in which each user bids optimally, given the bids of the other users. To state this precisely, we use the notation $b_{-i} = (b_1, b_2, \ldots, b_{i-1}, b_{i+1}, \ldots, b_n)$ to denote the bids of the users other than i, and sometimes write (b_i, b_{-i}) for a bid vector $(b_1, \ldots, b_n)$.

Definition 3.1. A bid vector $(b_1, \ldots, b_n)$ is an *equilibrium* of the resource allocation game $(U_1, \ldots, U_n, C)$ if for every user $i \in \{1, 2, \ldots, n\}$,

$$Q_i(b_i, b_{-i}) = \sup_{\tilde{b}_i \geq 0} Q_i(\tilde{b}_i, b_{-i}). \tag{5}$$

One easily checks that in every equilibrium, at least two users submit strictly positive bids.

While equilibria are most naturally defined for bid vectors, we will be interested in the quality of the corresponding allocations. An *equilibrium allocation* is an allocation $(x_1, \ldots, x_n)$ induced by an equilibrium bid vector – i.e., there is an equilibrium $(b_1, \ldots, b_n)$ such that (4) holds for each user i. We next give a characterization of equilibrium allocations that will be crucial for designing a potential function for resource allocation games.

First, a simple calculation shows that concavity of the utility function U_i (in x_i) implies strict concavity of the payoff function Q_i (in b_i) for every fixed vector b_{-i} with at least one strictly positive component. Similarly, the latter function is continuously differentiable for each such fixed b_{-i}. We can therefore characterize solutions to (5) via standard first-order optimality conditions, which yields the following.

Proposition 3.2 ([16], [20]). *Let $(U_1, \ldots, U_n, C)$ be a resource allocation game and $(b_1, \ldots, b_n)$ a bid vector with at least two strictly positive bids. Let $B = \sum_{j=1}^{n} b_j$ denote the sum of the bids. This bid vector is an equilibrium if and only if*

$$U_i'\left(\frac{b_i}{B} \cdot C\right)\left(1 - \frac{b_i}{B}\right) \leq \frac{B}{C}$$

for every user $i \in \{1, 2, \ldots, n\}$, with equality holding whenever $b_i > 0$.

Reformulating Proposition 3.2 in terms of allocations gives the following corollary (cf., Definition 2.1).

Corollary 3.3 ([16], [20]). *Let $(U_1, \ldots, U_n, C)$ be a resource allocation game. An allocation $(x_1, \ldots, x_n)$ is an equilibrium if and only if for every pair $i, j \in \{1, 2, \ldots, n\}$ of users with $x_i > 0$,*

$$U_i'(x_i)\left(1 - \frac{x_i}{C}\right) \geq U_j'(x_j)\left(1 - \frac{x_j}{C}\right).$$

Proof. The "only if" direction follows easily from Proposition 3.2 and equation (4). For the "if" direction, suppose $(x_1, \ldots, x_n)$ satisfies the stated condition. There is then a scalar $\lambda \geq 0$ such that $U_i'(x_i)[1 - (x_i/C)] \leq \lambda$ for all users i, with equality holding whenever $x_i > 0$. Setting $b_i = \lambda x_i$ for each i yields a bid vector that meets the equilibrium condition in Proposition 3.2. □

Example 3.4 ([20]). Consider a resource allocation game in which the capacity C is 1, one user has the utility function $U_1(x_1) = 2x_1$, and the other $n-1$ users have the utility function $U_i(x_i) = x_i$. Corollary 3.3 implies that in the unique equilibrium allocation, the first user receives $\frac{1}{2} + \varepsilon$ units of capacity, while each of the other $n-1$ users receive δ units of capacity (with $\varepsilon, \delta \to 0$ as $n \to \infty$). In this allocation, $U_i'(x_i)(1-x_i)$ is the same for each user i, and is slightly less than 1. The corresponding equilibrium bid vector is roughly the same as the equilibrium allocation vector.

In the next subsection, we use a potential function to show that every resource allocation game has a unique equilibrium allocation.

We claim that the equilibrium allocation in Example 3.4 is suboptimal. As in the previous section, we formalize this claim by introducing an objective function and studying the price of anarchy. We define the *efficiency* of an allocation $(x_1, \dots, x_n)$ of a resource allocation game to be the sum of the users' utilities:

$$\mathcal{E}(x_1, \dots, x_n) = \sum_{i=1}^{n} U_i(x_i). \tag{6}$$

An *optimal* allocation has the maximum-possible efficiency.

The *price of anarchy* of a resource allocation game is the ratio $\mathcal{E}(x)/\mathcal{E}(x^*)$, where x is the equilibrium allocation and x^* is an optimal allocation. Note that the price of anarchy of such a game is at most 1. In Example 3.4, the optimal allocation gives all of the capacity to the first user and has efficiency 2. The equilibrium allocation has efficiency approaching $3/2$ as $n \to \infty$; the price of anarchy can therefore be arbitrarily close to $3/4$ in this family of examples.

Why does inefficiency arise in Example 3.4? First, note that if the first user is the only one bidding a strictly positive amount (leading to the optimal allocation), then the bid vector cannot be an equilibrium: the first user can bid a smaller positive amount and continue to receive all of the capacity. A similar argument holds whenever the first user's bid comprises a sufficiently large fraction of the sum of the users' bids: if the first user lowers its bid, its allocation diminishes, but the price it pays per unit of bandwidth decreases by a large enough amount to increase its overall payoff. This intuition is mathematically reflected in Corollary 3.3 in the term $U_i'(x_i)(1-x_i)$ – the marginal benefit of increased capacity to a user becomes increasing tempered as its allocation grows. Inefficiency thus arises in Example 3.4 because of "market power" – the fact that the actions of a single user have significant influence over the effective price of capacity. Indeed, resource allocation games were initially studied by Kelly [22] under the assumption that no users have nontrivial market power. Under this assumption, equilibria are fully efficient – i.e., the price of anarchy is always 1 [22]. See [19, §1.3–1.4] for further discussion.

Remark 3.5. Selfish routing networks and resource allocation games differ in a number of ways. In the former, there is a continuum of selfish network users that each have a finite set of strategies (paths); in the latter, there is a finite set of users, each with a continuum of strategies (bids). In selfish routing, the objective is cost minimization; in resource allocation, it is efficiency maximization. Finally, and perhaps most fundamentally, inefficiency appears to arise for different reasons in the two models. Recall that in selfish routing networks, inefficiency stems from congestion externalities (see the discussion following Example 2.2). Example 3.4 shows that market power is the culprit behind inefficient equilibria in resource allocation games. Despite all of these conceptual differences, the next two subsections show that the inefficiency of equilibria can be quantified in the two models via remarkably similar analyses.

3.2. A potential function for equilibria. As in Subsection 2.2, our first step toward constructing a potential function for equilibrium allocations is to characterize optimal allocations. Since efficiency (6) is a separable concave function, a straightforward application of first-order optimality conditions yields the following.

Proposition 3.6. *Let* $(U_1, \dots, U_n, C)$ *be a resource allocation game. An allocation* $(x_1, \dots, x_n)$ *is optimal if and only if for every pair* $i, j \in \{1, 2, \dots, n\}$ *of users with* $x_i > 0$, $U_i'(x_i) \geq U_j'(x_j)$.

Given the near-identical characterizations of equilibrium and optimal allocations in Corollary 3.3 and Proposition 3.6, respectively, we again ask: of what function does an equilibrium allocation arise as the global maximum? To recover Corollary 3.3 as an optimality condition, we seek a function H_i for each user i such that $H_i'(x_i) = U_i'(x_i)[1 - (x_i/C)]$ for all $x_i \geq 0$. Setting $H_i(x_i) = U_i(x_i)[1 - (x_i/C)] + \left[\int_0^{x_i} U_i(y)\,dy\right]/C$ thus yields the desired potential function. Precisely, for the resource allocation game $(U_1, \dots, U_n, C)$, define

$$\Phi_{RA}(x_1, \dots, x_n) = \sum_{i=1}^{n} \hat{U}_i(x_i), \tag{7}$$

where

$$\hat{U}_i(x_i) = \left(1 - \frac{x_i}{C}\right) \cdot U_i(x_i) + \frac{x_i}{C} \cdot \left(\frac{1}{x_i}\int_0^{x_i} U_i(y)\,dy\right). \tag{8}$$

A simple calculation shows that each function $\hat{U}_i$ is strictly concave, increasing, and continuously differentiable. Regarding $(\hat{U}_1, \dots, \hat{U}_n, C)$ as a resource allocation game, applying Proposition 2.6 to it, and appealing to Corollary 3.3 formalizes the fact that Φ_{RA} is a potential function.

Proposition 3.7 ([16], [20]). *An allocation of the game* $(U_1, \dots, U_n, C)$ *is an equilibrium allocation if and only if it is a global maximum of the corresponding potential function* Φ_{RA}.

Existence and uniqueness of equilibrium allocations follow immediately.

Proposition 3.8 ([16], [20]). *In every resource allocation game, there is a unique equilibrium allocation.*

Proof. Existence follows from Proposition 3.7 and the facts that the potential function (7) is continuous and the set of all allocations is compact. Uniqueness follows from Proposition 3.7 and the fact that the potential function (7) is strictly concave. □

Proposition 3.7 also has consequences for the price of anarchy in resource allocation games. To see why, note that the value of $\hat{U}_i(x_i)$ in (8) can be viewed as a weighted average of two quantities – the "true utility" $U_i(x_i)$ and the "average utility" $\left[\int_0^{x_i} U_i(y)\,dy\right]/x_i$. Since U_i is increasing, the latter quantity can only underestimate the utility $U_i(x_i)$, and hence $\hat{U}_i(x_i) \leq U_i(x_i)$ for all i and $x_i \geq 0$. On the other hand, since U_i is nonnegative and concave, the average utility between 0 and x_i is at least half of the utility $U_i(x_i)$ at x_i. Thus $\hat{U}_i(x_i) \geq U_i(x_i)/2$ for all i and $x_i \geq 0$. It follows that

$$\mathcal{E}(x_1, \ldots, x_n) \geq \Phi_{RA}(x_1, \ldots, x_n) \geq \mathcal{E}(x_1, \ldots, x_n)/2$$

for every allocation $(x_1, \ldots, x_n)$. Following the proof of Theorem 2.10 now gives a lower bound of $1/2$ on the price of anarchy in resource allocation games.

Theorem 3.9 ([20]). *In every resource allocation game, the price of anarchy is at least* $1/2$.

3.3. An optimal bound on the price of anarchy. There is a gap between the lower bound of $1/2$ on the price of anarchy given in Theorem 3.9 and the upper bound of $3/4$ that is achieved (in the limit) in Example 3.4. As in Subsection 3.3, an optimal (lower) bound can be obtained by leveraging the potential function characterization of equilibria (Proposition 3.7) in a less crude way. Our argument will again be based on a "variational inequality", which can be derived directly from Corollary 3.3 or viewed as a first-order optimality condition for the potential function (7).

Proposition 3.10. *Let* $(U_1, \ldots, U_n, C)$ *be a resource allocation game. For each user* i*, define the modified utility function* $\hat{U}_i$ *as in* (8)*. An allocation* $\hat{x}$ *is an equilibrium for* $(U_1, \ldots, U_n, C)$ *if and only if*

$$\sum_{i=1}^{n} \hat{U}_i'(\hat{x}_i)\hat{x}_i \geq \sum_{i=1}^{n} \hat{U}_i'(\hat{x}_i)x_i$$

for every feasible allocation x.

Next is the analogue of the Pigou bound (Definition 2.12) for resource allocation games. This definition is primarily motivated by the upper bound on the price of anarchy provided by Example 3.4; we state it in a form that also permits easy application of Proposition 3.10 in the proof of Lemma 3.13 below.

Definition 3.11. Let $\mathcal{U}$ denote the set of real-valued, nonnegative, strictly increasing, continuously differentiable, and concave (utility) functions. Define the *JT bound* β by

$$\beta = \inf_{U \in \mathcal{U}} \inf_{C>0} \inf_{0 \leq \hat{x}, x^* \leq C} \frac{U(\hat{x}) + \hat{U}'(\hat{x})(x^* - \hat{x})}{U(x^*)}, \tag{9}$$

where $\hat{U}$ is defined as in (8), as a function of U and C.

In the rest of this section, we show that the JT bound is exactly the worst price of anarchy occurring in resource allocation games, and explicitly compute the bound.

Lemma 3.12. *For every $\varepsilon > 0$, there is a resource allocation game with price of anarchy at most $\beta + \varepsilon$, where β is the JT bound.*

Lemma 3.13. *In every resource allocation game, the price of anarchy is at least the JT bound β.*

Lemma 3.14. *The JT bound β is exactly* $3/4$.

Lemmas 3.12–3.14 give an explicit optimal bound on the price of anarchy in resource allocation games.

Theorem 3.15 ([20]). *In every resource allocation game, the price of anarchy is at least* $3/4$. *Moreover, this bound is tight.*

We now prove Lemmas 3.12–3.14 in turn.

Proof of Lemma 3.12. Fix a choice of a utility function U, a capacity $C > 0$, and values for $\hat{x}, x^* \in [0, C]$. We aim to exhibit a resource allocation game with price of anarchy (arbitrarily close to)

$$\frac{U(\hat{x}) + \hat{U}'(\hat{x})(x^* - \hat{x})}{U(x^*)}. \tag{10}$$

Recall from (8) that $\hat{U}'(\hat{x}) = U'(\hat{x}) \cdot [1 - (\hat{x}/C)]$. A calculation shows that (10) is at least 1 if $\hat{x} \geq x^*$, so we can assume that $\hat{x} < x^*$. Since (10) is decreasing in C, we can assume that $C = x^*$.

Define a resource allocation game in which the capacity is C, the first user has the utility function $U_1(x_1) = U(x_1)$, and the other $n - 1$ users each have the linear utility function $U_i(x_i) = \hat{U}'(\hat{x}) \cdot x_i$. Giving all of the capacity to the first user is a feasible allocation with efficiency $U_1(C) = U(x^*)$. Arguing as in Example 3.4, the equilibrium allocation has efficiency approaching $U_1(\hat{x}) + (C - \hat{x}) \cdot \hat{U}'(\hat{x}) = U(\hat{x}) + \hat{U}'(\hat{x})(x^* - \hat{x})$ as the number n of users tends to infinity. The price of anarchy in this family of instances thus tends to (at most) the expression in (10) as $n \to \infty$, completing the proof. □

Proof of Lemma 3.13. Let $(U_1, \ldots, U_n, C)$ be a resource allocation game. Let x^* and $\hat{x}$ denote optimal and equilibrium allocations, respectively. Define the modified utility function $\hat{U}_i$ as in (8). The lemma follows by writing

$$\begin{aligned}\sum_{i=1}^{n} U_i(x_i^*) &\leq \sum_{i=1}^{n} \left[\frac{1}{\beta}\big(U_i(\hat{x}_i) + \hat{U}_i'(\hat{x}_i)(x_i^* - \hat{x}_i)\big)\right] \\ &\leq \frac{1}{\beta} \sum_{i=1}^{n} U_i(\hat{x}_i),\end{aligned}$$

where the first inequality follows from Definition 3.11 and the second from Proposition 3.10. □

Proof of Lemma 3.14. Setting U to the identity function, $\hat{x} = 1/2$, and $C = x^* = 1$ shows that the JT bound is at most 3/4. Now fix arbitrary choices of U, C, and $\hat{x}, x^* \in [0, C]$. We need to show that (10) is at least 3/4. As in the proof of Lemma 3.12, we can assume that $\hat{x} < x^* = C$. We can then write

$$\begin{aligned} U(\hat{x}) + \hat{U}'(\hat{x})(x^* - \hat{x}) &= U(\hat{x}) + \left(1 - \frac{\hat{x}}{x^*}\right) U'(\hat{x})(x^* - \hat{x}) \\ &\geq U(\hat{x}) + \left(1 - \frac{\hat{x}}{x^*}\right)\big(U(x^*) - U(\hat{x})\big) \\ &= \left(\frac{\hat{x}}{x^*}\right) \cdot U(\hat{x}) + \left(1 - \frac{\hat{x}}{x^*}\right) \cdot U(x^*) \\ &\geq \left(\frac{\hat{x}}{x^*}\right)^2 \cdot U(x^*) + \left(1 - \frac{\hat{x}}{x^*}\right) \cdot U(x^*) \\ &\geq \frac{3}{4} \cdot U(x^*), \end{aligned}$$

where the first equality follows from the definition of $\hat{U}$ in (8), the first and second inequalities follow from the concavity and nonnegativity of U, and the final inequality follows from the fact that the function $y^2 - y + 1$ is uniquely minimized when $y = 1/2$. The proof is complete. □

Remark 3.16. The original proof of Theorem 3.15 is fairly different than the one given here. Specifically, Johari and Tsitsiklis [20] first show that the price of anarchy is minimized in games in which all users have linear utility functions, and then explicitly determine a worst-case example (the same as Example 3.4) by analyzing a linear program. We instead presented the proof above to further highlight the connections between resource allocation games and selfish routing networks.

Despite the numerous common features in our analyses of the price of anarchy in selfish routing networks and in resource allocation games, the precise relationship between the two models is not completely understood. In particular, we lack a unifying analysis of the price of anarchy in the two models.

Open Question 1. Find a compelling generalization of selfish routing networks and resource allocation games in which the price of anarchy can be analyzed in a uniform way. Ideally, such a generalization would unify Theorems 2.15 and 3.15, and would also apply to several of the more general classes of games described in [19], [32].

As with selfish routing networks, we have only scratched the surface of the literature on the price of anarchy in resource allocation games. For much more on the subject, including generalizations of these games to general networks, see Johari and Tsitsiklis [20] and Johari [19].

4. The price of stability in network design games

Our final class of games is a model of network design with selfish users. These games share some features with selfish routing networks, but also differ in a few fundamental respects.

4.1. The model. In this section we study *Shapley network design games*, first proposed by Anshelevich et al. [1]. The game occurs in a directed graph $G = (V, E)$, in which each edge $e \in E$ has a fixed nonnegative cost c_e. There is a finite set of k selfish players, and each player $i \in \{1, 2, \dots, k\}$ is identified with a source-sink vertex pair (s_i, t_i). Let $\mathcal{P}_i$ denote the set of simple s_i-t_i paths.

Each player i chooses a path $P_i \in \mathcal{P}_i$ from its source to its destination. This creates a network $\left(V, \bigcup_i P_i\right)$, and we define the *cost* of this outcome as

$$c(P_1, \dots, P_k) = \sum_{e \in \bigcup_i P_i} c_e. \tag{11}$$

We assume that this cost is shared among the players in the following way. First, if edge e lies in f_e of the chosen paths, then each player choosing such a path pays a proportional share $\pi_e = c_e/f_e$ of the cost. The overall cost $c_i(P_1, \dots, P_k)$ to player i is then the sum $\sum_{e \in P_i} \pi_e$ of these proportional shares. Selfish players naturally attempt to minimize their incurred cost.

We next define our notion of equilibria for Shapley network design games. In contrast to selfish routing networks and resource allocation games, these network design games are finite games – there is a finite set of players, each with a finite set of strategies. This is the classical setting for *Nash equilibria* [25]. As in Definition 3.1, we use P_{-i} to denote the vector of strategies chosen by the players other than i.

Definition 4.1. An outcome $(P_1, \dots, P_k)$ of a Shapley network design game is a (*pure-strategy*) *Nash equilibrium* if for every player i,

$$c_i(P_i, P_{-i}) = \min_{\tilde{P}_i \in \mathcal{P}_i} c_i(\tilde{P}_i, P_{-i}).$$

In a pure-strategy Nash equilibrium, every player chooses a single strategy. In a *mixed-strategy* Nash equilibrium, a player can randomize over several strategies. We will not discuss mixed-strategy Nash equilibria in this survey, though the price of anarchy of such equilibria has been studied in different models (see e.g. [3], [23]).

Example 4.2 ([2]). Consider the network shown in Figure 3. There are k players, each with the same source s and sink t. The edge costs are k and $1 + \varepsilon$, where $\varepsilon > 0$ is arbitrarily small. In the minimum-cost outcome, all players choose the lower edge. This outcome is also a Nash equilibrium. On the other hand, suppose all of the players choose the upper edge. Each player i then incurs cost 1, and if player i deviates to the lower edge it pays the full cost of $1 + \varepsilon$. This outcome is thus a second Nash equilibrium, and it has cost k.

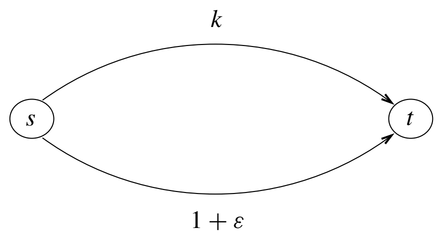

Figure 3. Multiple Nash equilibria in Shapley network design games (Example 4.2).

Example 4.2 shows that Shapley network design games are more ill-behaved than selfish routing networks and resource allocation games in a fundamental respect: there can be multiple equilibria, and different equilibria can have very different objective function values. (Cf., Theorem 2.3 and Proposition 3.8.) The definition of the price of anarchy is ambiguous in games with multiple equilibria – we would like to quantify the inefficiency of an equilibrium, but of which one?

The price of anarchy is historically defined as the ratio between the objective function value of the *worst* equilibrium and that of an optimal solution [23], [26]. This definition is natural from the perspective of worst-case analysis. In Example 4.2, the price of anarchy is (arbitrarily close to) k. It is also easy to show that the price of anarchy in every Shapley network design game is at most k.

In this section, we instead focus on the ratio between the cost of the *best* Nash equilibrium of a Shapley network design game and that of an optimal solution. This measure is called the *price of stability* [1]. Our motivation is twofold. First, as Example 4.2 shows, the price of anarchy is large and trivial to determine. Second, the price of stability has a reasonably natural interpretation in network design games – if we envision the network as being designed by a central authority for subsequent use by selfish players, then the best Nash equilibrium is an obvious solution to propose. In this sense, the price of stability measures the necessary degradation in solution quality caused by imposing the game-theoretic constraint of stability. See [1], [2], [6], [7] for further discussion and examples of the price of stability.

The price of stability in Example 4.2 is 1. We conclude this subsection with an example showing that this is not always the case.

Example 4.3 ([1]). Consider the network shown in Figure 4. There are k players, all with the same sink t, and $\varepsilon > 0$ is arbitrarily small. For each $i \in \{1, 2, \dots, k\}$, the edge (s_i, t) has cost $1/i$. In the minimum-cost outcome, each player i chooses the path $s_i \to v \to t$ and the cost is $1+\varepsilon$. This is not a Nash equilibrium, as player k can decrease its cost from $(1+\varepsilon)/k$ to $1/k$ by switching to the direct path $s_k \to t$. More generally, this direct path is a *dominant strategy* for the kth player – it is the minimum-cost strategy, independent of the paths chosen by the other players. It follows that in every Nash equilibrium, the kth player selects its direct path. Arguing inductively about the players $k-1, k-2, \ldots, 1$, we find that the unique Nash equilibrium is the outcome in which each player i chooses its direct path $s_i \to t$ to the sink. The cost of this outcome is exactly the kth harmonic number $\mathcal{H}_k = \sum_{i=1}^{k}(1/i)$, which is roughly $\ln k$. The price of stability can therefore be (arbitrarily close to) $\mathcal{H}_k$ in Shapley network design games.

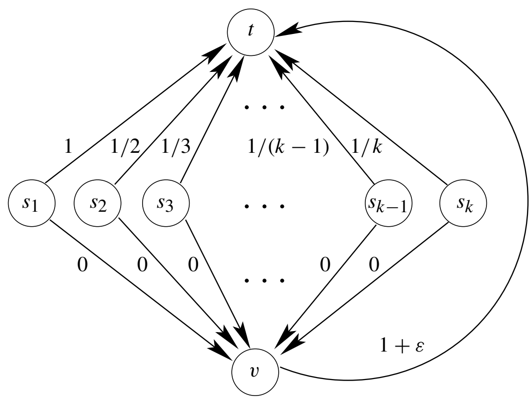

Figure 4. The price of stability in Shapley network design games can be at least $\mathcal{H}_k$ (Example 4.3).

4.2. A potential function for Nash equilibria. In this subsection we use a potential function to prove the existence of pure-strategy Nash equilibria and upper bound the price of stability in Shapley network design games. Recall that for both selfish routing networks and resource allocation games, we designed potential functions using a characterization of optimal solutions as a guide (see Propositions 2.6 and 3.6). In Shapley network design games, computing an optimal solution is an NP-hard network design problem [15], and we cannot expect to find an analogous characterization.

There are two ways that Shapley network design games differ from selfish routing networks that prevent the characterization of optimal solutions (Proposition 2.6) from carrying over. First, there are a finite number of players in the former model, and a continuum of players in the latter model. Second, cost functions in selfish routing networks are nondecreasing, whereas Shapley network design games effectively have cost functions that are *decreasing* in the "congestion" – if $x \geq 1$ players use an edge e with fixed cost c_e, then the per-player cost on that edge is c_e/x.

On the bright side, the potential function (2) for selfish routing networks is easily modified to account for these two differences. First, note that this function Φ remains well-defined for decreasing cost functions. Second, passing from an infinite player set to a finite one merely involves changing the integrals in (2) to sums. This motivates the following proposal for a potential function for a Shapley network design game:

$$\Phi_{ND}(P_1, \ldots, P_k) = \sum_{e \in E} \sum_{i=1}^{f_e} \frac{c_e}{i}, \tag{12}$$

where f_e denotes the number of paths P_i that include edge e. While equilibria in selfish routing networks and resource allocation games can be characterized as the *global* optima of their respective potential functions (2) and (7), we will see that the Nash equilibria of a Shapley network design game are characterized as the *local* optima of the potential function (12). This idea is originally due to Rosenthal [29], [30], who also considered the broader context of "atomic congestion games".

The next lemma, which is crucial for the rest of this section, states that the potential function "tracks" the change in cost experienced by a deviating player.

Lemma 4.4 ([1], [30]). *Let (G, c) denote a Shapley network design game with k players and Φ_{ND} the corresponding potential function* (12). *Let $i \in \{1, 2, \ldots, k\}$ be a player, and let (P_i, P_{-i}) and $(\tilde{P}_i, P_{-i})$ denote two outcomes that differ only in the strategy chosen by the ith player. Then*

$$c_i(\tilde{P}_i, P_{-i}) - c_i(P_i, P_{-i}) = \Phi_{ND}(\tilde{P}_i, P_{-i}) - \Phi_{ND}(P_i, P_{-i}). \tag{13}$$

Proof. Let f_e denote the number of players that choose a path containing the edge e in the outcome (P_i, P_{-i}). Then both sides of (13) are equal to

$$\sum_{e \in \tilde{P}_i \setminus P_i} \frac{c_e}{f_e + 1} - \sum_{e \in P_i \setminus \tilde{P}_i} \frac{c_e}{f_e}.$$

□

In the game theory literature, equation (13) is often taken as the definition of a potential function in the context of finite games. See Monderer and Shapley [24] for a fairly general treatment of potential functions for finite games.

While simple, Lemma 4.4 has a number of non-trivial consequences. First, Nash equilibria of a Shapley network design game are the local minima of the corresponding potential function. Formally, two outcomes of a Shapley network design game are *neighbors* if they differ in at most one component, and an outcome is a *local minimum* of Φ_{ND} if it has no neighbor with strictly smaller potential function value.

Corollary 4.5 ([1], [30]). *An outcome of a Shapley network design game is a Nash equilibrium if and only if it is a local minimum of the corresponding potential function Φ_{ND}.*

Proof. Immediate from the definitions and Lemma 4.4. □

Since every Shapley network design game has a finite number of outcomes, its corresponding potential function has a global (and hence local) minimum.

Corollary 4.6 ([1], [30]). *In every Shapley network design game, there is at least one (pure-strategy) Nash equilibrium.*

We note in passing that several related classes of network games do not always have pure-strategy Nash equilibria [2], [6], [13], [30].

A stronger version of Corollary 4.6 also holds. In a finite game, *better-response dynamics* refers to the following process: start with an arbitrary initial outcome; if the current outcome is not a Nash equilibrium, pick an arbitrary player that can decrease its cost by switching strategies, update its strategy to an arbitrary superior one, and repeat. Better-response dynamics terminate if and only if a Nash equilibrium is reached. Even in extremely simple two-player games, better-response dynamics need not terminate (e.g., in "rock-paper-scissors"). On the other hand, the potential function (12) ensures that such dynamics always converge in Shapley network design games.

Corollary 4.7 ([1], [30]). *In every Shapley network design game, better-response dynamics always converges to a Nash equilibrium in a finite number of iterations.*

Proof. By Lemma 4.4, every iteration of better-response dynamics strictly decreases the value of the potential function Φ_{ND}. Better-response dynamics therefore cannot visit an outcome more than once and eventually terminates, necessarily at a Nash equilibrium. □

Corollary 4.7 does not address the number of iterations required to reach a Nash equilibrium; see [1], [12] for further study of this issue.

Finally, the potential function (12) has direct consequences for the price of stability in Shapley network design games. Comparing the definitions of cost (11) and potential function value (12) of such a game, we have

$$c(P_1, \ldots, P_k) \leq \Phi_{ND}(P_1, \ldots, P_k) \leq \mathcal{H}_k \cdot c(P_1, \ldots, P_k) \tag{14}$$

for every outcome $(P_1, \ldots, P_k)$. As a result, a global minimum of the potential function Φ_{ND} of a Shapley network design game is both a Nash equilibrium (by Corollary 4.6) and has cost at most $\mathcal{H}_k$ times that of optimal (by the argument in the proof of Theorem 2.10). This gives the following theorem

Theorem 4.8 ([1]). *In every k-player Shapley network design game, the price of stability is at most $\mathcal{H}_k$.*

A similar argument shows that the bound of $\mathcal{H}_k$ in Theorem 4.8 applies to every Nash equilibrium reachable from an optimal solution via better-response dynamics. The bound also carries over to numerous extensions of Shapley network design games; see [1] for details.

Example 4.3 shows that the bound in Theorem 4.8 is tight for every $k \geq 1$. Thus, unlike for selfish routing networks and resource allocation games, a direct application of a potential function argument yields an optimal upper bound on the inefficiency of equilibria.

The upper bound in Theorem 4.8 is not optimal for some important special cases of Shapley network design games, however. For example, suppose we insist that the underlying network G is undirected. There is no known analogue of Example 4.3 for undirected Shapley network design games – the best lower bound known on the price of stability in such games is 2. On the other hand, it is not clear how to significantly improve the $\mathcal{H}_k$ bound in Theorem 4.8 for undirected networks.

Open Question 2. Determine the largest-possible price of stability in undirected Shapley network design games.

5. Conclusion

This survey has discussed three natural types of games: selfish routing networks, resource allocation games, and Shapley network design games. These classes of games differ from each other, both conceptually and technically, in a number of ways. Despite this, the worst-case inefficiency of selfish behavior is fairly well understood in all of these models, and in each case can be determined using a potential function characterization of equilibria.

While the entire field of algorithmic game theory is still in a relatively nascent stage, several broad research agendas are emerging. For the problem of quantifying the inefficiency of noncooperative equilibria, a central research issue is to understand characteristics of games that guarantee approximately optimal equilibria, and to develop flexible mathematical techniques for proving such guarantees. While many research accomplishments from the past few years have improved our understanding of these intertwined goals, there is clearly much left to be done. Perhaps the current state of the art in bounding the inefficiency of equilibria can be compared to the field of approximation algorithms circa twenty-five years ago, when the most fundamental problems and the most powerful algorithmic techniques (such as linear programming) were only beginning to crystallize. Motivated by this analogy, we conclude with the following question: will potential functions be as ubiquitous in bounds on the inefficiency of equilibria as linear programming is in bounds on the performance of approximation algorithms?

Open Question 3. We have seen that a potential function characterization of equilibria leads a bound on the inefficiency of equilibria. Under what conditions and to what extent does a converse hold? When does a bound on the inefficiency of the equilibria of a game imply the existence of some form of a potential function for the game?

References

[1] Anshelevich, E., Dasgupta, A., Kleinberg, J., Tardos, É., Wexler, T., and Roughgarden, T., The price of stability for network design with fair cost allocation. In *Proceedings of the 45th Annual Symposium on Foundations of Computer Science*, IEEE Computer Society Press, Los Alamitos, CA, 2004, 295–304.

[2] Anshelevich, E., Dasgupta, A., Tardos, É., and Wexler, T., Near-optimal network design with selfish agents. In *Proceedings of the 35th Annual ACM Symposium on Theory of Computing*, ACM Press, New York 2003, 511–520.

[3] Awerbuch, B., Azar, Y., and Epstein, E., The price of routing unsplittable flow. In *Proceedings of the 37th Annual ACM Symposium on Theory of Computing*, ACM Press, New York 2005, 57–66.

[4] Beckmann, M. J., McGuire, C. B., and Winsten, C. B. *Studies in the Economics of Transportation*. Yale University Press, 1956.

[5] Bertsekas, D. P., Nedic, A., and Ozdaglar, A. E., *Convex Analysis and Optimization*. Athena Scientific, 2003.

[6] Chen, H., and Roughgarden, T., Network design with weighted players. Submitted, 2005.

[7] Correa, J. R., Schulz, A. S., and Stier-Moses, N. E., Selfish routing in capacitated networks. *Math. Oper. Res.* **29** (4) (2004), 961–976.

[8] Correa, J. R., Schulz, A. S., and Stier-Moses, N. E., On the inefficiency of equilibria in congestion games. In *Integer Programming and Combinatorial Optimization*, Lecture Notes in Comput. Sci. 3509, Springer-Verlag, Berlin 2005, 167–181.

[9] Cramton, P., Shoham, Y., and Steinberg, R., *Combinatorial Auctions*. MIT Press, 2006.

[10] Doyle, P. G., and Snell, J. L., *Random Walks and Electrical Networks*. Mathematical Association of America, 1984.

[11] Dubey, P., Inefficiency of Nash equilibria. *Math. Oper. Res.* **11** (1) (1986), 1–8.

[12] Fabrikant, A., Papadimitriou, C. H., and Talwar, K., The complexity of pure Nash equilibria. In *Proceedings of the 36th Annual ACM Symposium on Theory of Computing*, ACM Press, New York 2004, 604–612.

[13] Fotakis, D., Kontogiannis, S. C., and Spirakis, P. G., Selfish unsplittable flows. *Theoret. Comput. Sci.* **348** (2–3) (2005), 226–239.

[14] Friedman, E. J., and Shenker, S., Learning and implementation on the Internet. Working paper, 1997.

[15] Garey, M. R., and Johnson, D. S., *Computers and Intractability: A Guide to the Theory of NP-Completeness*. Freeman, 1979.

[16] Hajek, B., and Gopalakrishnan, G., Do greedy autonomous systems make for a sensible Internet? Presentation at the Conference on Stochastic Networks, Stanford University, June 2002 (Cited in [20]).

[17] Hartline, J. D., Optimization in the Private Value Model: Competitive Analysis Applied to Auction Design. PhD thesis, University of Washington, 2003.

[18] Jain, K., A polynomial time algorithm for computing the Arrow-Debreu market equilibrium for linear utilities. In *Proceedings of the 45th Annual Symposium on Foundations of Computer Science*, IEEE Computer Society Press, Los Alamitos, CA, 2004, 286–294.

[19] Johari, R., Efficiency Loss in Market Mechanisms for Resource Allocation. PhD thesis, MIT, 2004.

[20] Johari, R., and Tsitsiklis, J. N., Efficiency loss in a network resource allocation game. *Math. Oper. Res.* **29** (3) (2004), 407–435.

[21] Kelly, F. P., Network routing. *Philos. Trans. Roy. Soc. London Ser. A* **337** (3) (1991), 343–367.

[22] Kelly, F. P., Charging and rate control for elastic traffic. *European Transactions on Telecommunications* **8** (1) (1997), 33–37.

[23] Koutsoupias, E., and Papadimitriou, C. H., Worst-case equilibria. In *STACS 99*, Lecture Notes in Comput. Sci. 1563, Springer-Verlag, Berlin 1999, 404–413.

[24] Monderer, D., and Shapley, L. S., Potential games. *Games Econom. Behav.* **14** (1) (1996), 124–143.

[25] Nash, J. F., Equilibrium points in N-person games. *Proc. National Academy of Science* **36** (1) (1959), 48–49.

[26] Papadimitriou, C. H., Algorithms, games, and the Internet. In *Proceedings of the 33rd Annual ACM Symposium on Theory of Computing*, ACM Press, New York 2001, 749–753.

[27] Pigou, A. C., *The Economics of Welfare*. Macmillan, 1920.

[28] Rapoport, A., and Chammah, A. M., *Prisoner's Dilemma*. University of Michigan Press, 1965.

[29] Rosenthal, R. W., A class of games possessing pure-strategy Nash equilibria. *Internat. J. Game Theory* **2** (1) (1973), 65–67.

[30] Rosenthal, R. W., The network equilibrium problem in integers. *Networks* **3** (1) (1973), 53–59.

[31] Roughgarden, T., The price of anarchy is independent of the network topology. *J. Comput. System Sci.* **67** (2) (2003), 341–364.

[32] Roughgarden, T., *Selfish Routing and the Price of Anarchy*. MIT Press, 2005.

[33] Roughgarden, T., Selfish routing and the price of anarchy. *OPTIMA* **71** (2006), to appear.

[34] Roughgarden, T., and Tardos, É., How bad is selfish routing? *J. ACM* **49** (2) (2002), 236–259.

[35] Schmeidler, D., Equilibrium points of nonatomic games. *J. Statist. Phys.* **7** (4) (1973), 295–300.

[36] Smith, M. J., The existence, uniqueness and stability of traffic equilibria. *Transportation Res. Part B* **13** (4) (1979), 295–304.

[37] Tardos, É., CS684 course notes. Cornell University, 2004.

[38] Wardrop, J. G., Some theoretical aspects of road traffic research. In *Proceedings of the Institute of Civil Engineers, Pt. II*, volume 1, 1952, 325–378.

Stanford University, Department of Computer Science, 462 Gates Building, 353 Serra Mall, Stanford, CA 94305, U.S.A.

E-mail: tim@cs.stanford.edu