On Braess's Paradox

Henry Lin* Tim Roughgarden* Éva Tardos*

In 1968, Braess [1] demonstrated a remarkable and counterintuitive fact, now known as "Braess's Paradox": in a network in which users selfishly and independently choose minimum-latency paths and the latency of an edge increases with the amount of edge congestion, deleting edges from the network can *decrease* the latency encountered by traffic.

Our work here is primarily motivated by the paper of Roughgarden [3], which quantifies the maximum-possible severity of Braess's Paradox. Specifically, in [3] it was shown that there is a network with n vertices in which removing a set of edges decreases the latency encountered by the network traffic by a factor of $\lfloor n/2 \rfloor$, and that no greater decrease is possible in any such network. The construction for $n = 4$ essentially gives Braess's original example.

However, the "Braess graphs" of [3] differ from Braess's example in that $\lfloor n/2 \rfloor - 1$ edges need to be removed to effect an $\lfloor n/2 \rfloor$ factor of decrease in latency. It is thus natural to ask, for $n \geq 6$, can a similar decrease be caused by the removal of a single edge? In particular, when restricting attention to removing one edge, is there a network in which Braess's Paradox is more severe than in Braess's original example?

In this note, we introduce a simple but powerful combinatorial lemma that resolves these questions in the negative: for any integer $k \geq 1$, the only way to decrease the latency experienced by traffic by a factor strictly greater than k is by removing at least k edges from the network. Along the way, we give the first combinatorial proof of a useful monotonicity property of traffic equilibria, and give a new proof of the bound from [3] on the worst-case severity of Braess's Paradox.

Preliminaries

The Model. We work with a model of "selfish routing" defined by Wardrop and first studied in a theoretical computer science context by Roughgarden and Tardos (see [4] for detailed references). We consider a directed network $G = (V, E)$ with vertex set V, edge set E, source vertex s, and destination vertex t. We denote the set of s-t paths by $\mathcal{P}$. A *flow* is a function $f : \mathcal{P} \rightarrow \mathcal{R}^+$; for a fixed flow f we define the load $f_e = \sum_{P: e \in P} f_P$. With respect to a finite and positive *traffic rate* r, a flow f is said to be *feasible* if $\sum_{P \in \mathcal{P}} f_P = r$. Each edge $e \in E$ is given a load-dependent *latency* that we denote by $\ell_e(\cdot)$. We assume that each ℓ_e is a nonnegative, continuous, and nondecreasing function. The latency of a path P with respect to a flow f is then the sum of the latencies of the edges in the path, denoted by $\ell_P(f) = \sum_{e \in P} \ell_e(f_e)$. We call the triple (G, r, ℓ) an *instance*.

Flows at Nash Equilibrium. A flow f feasible for (G, r, ℓ) is said to be *at Nash equilibrium* (or is a *Nash flow*) if for every two s-t paths $P_1, P_2 \in \mathcal{P}$ with $f_{P_1} > 0$, $\ell_{P_1}(f) \leq \ell_{P_2}(f)$. In particular, if a flow f is at Nash equilibrium then all s-t flow paths have equal latency.

Finally, it is well known that Nash flows always exist, are (essentially) unique, and we can assume wlog that the graph of edges carrying flow is acyclic (see [4]). Thus, the following definition makes sense: for an instance (G, r, ℓ), let $L(G, r, \ell)$ denote the common latency of all flow paths in a flow at Nash equilibrium for (G, r, ℓ).

*Department of Computer Science, Cornell University, Ithaca NY 14853. Supported by ONR grant N00014-98-1-0589. Email: {hcl5@cornell.edu,timr@cs.cornell.edu,eva@cs.cornell.edu}.

Our Results

Alternating Paths. The crucial object in the proofs of this paper is an "alternating path", defined as follows.

Definition 1 *Let f be a Nash flow for (G, r, ℓ) and f' another flow in G. Edge e of G is* light *if $f_e \leq f'_e$ and $f'_e > 0$,* heavy *if $f_e > f'_e$, and* useless *if $f_e = f'_e = 0$. An undirected path P in G is* alternating *if it comprises only forward light edges and backward heavy edges.*

We now prove that alternating paths exist, when comparing a Nash flow to another flow in the network feasible for the same or an increased traffic rate.

Lemma 2 *Let f be at Nash equilibrium for (G, r, ℓ) and f' a flow feasible for (G, r', ℓ) with $r' \geq r$. Then there is an alternating s-t path. Moreover, there is such a path that begins and ends with a light edge.*

Proof: First, suppose there is no alternating s-t path and let S denote the set of nodes reachable from s via alternating paths. The set S contains s and, by assumption, does not contain t; it is therefore an s-t cut. By assumption, edges that exit S are heavy or useless; since f and f' are flows, at least one must be heavy. Edges that enter S are light or useless. It follows that there is strictly less net f'-flow than net f-flow crossing the cut S, which contradicts the fact that f' is a flow feasible for a rate no less than that of f. The first and last edge of any alternating path must be light, as in an acyclic flow edges entering s and leaving t do not carry flow. ■

A Monotonicity Theorem. We intuitively expect the latency encountered by traffic in a Nash flow to increase as we inject new traffic into the system. While true, this statement is not obvious (especially in light of Braess's Paradox) and requires proof. It was first proved by Hall [2], in a more general multicommodity flow network setting, using techniques from sensitivity analysis of convex programs. It was later applied and extended by several different authors (see [4]). Here, we give a direct, combinatorial proof for single-commodity flow networks. Also, the proof techniques employed will be crucial for our final theorem.

Theorem 3 ([2]) *If $r' \geq r$, then $L(G, r', \ell) \geq L(G, r, \ell)$.*

Proof: Let f and f' be Nash flows for (G, r, ℓ) and (G, r', ℓ), respectively, and $d(v)$ and $d'(v)$ the shortest-path distance from s to v w.r.t. edge lengths $\{\ell_e(f_e)\}$ and $\{\ell_e(f'_e)\}$, respectively. By the definition of a Nash flow, all flow paths of f and f' are shortest paths w.r.t. their respective edge lengths. Thus, the theorem asserts that $d(t) \leq d(t')$.

We aim to prove this inequality by induction on the edges of an arbitrary alternating s-t path P, which exists by Lemma 2. For the base case, $d(s) = d'(s) = 0$. Now suppose that $d(v) \leq d(v')$ holds for some vertex v on P, and let w be the subsequent vertex on P. There are now two cases.

First, suppose that edge $e = (v, w)$ is light. Then, $\ell_e(f_e) \leq \ell_e(f'_e)$ and $f'_e > 0$. By the triangle inequality for shortest-path distances, the inductive hypothesis, and the fact that f' routes only on shortest paths, we have $d(w) \leq d(v) + \ell_e(f_e) \leq d'(v) + \ell_e(f'_e) = d'(w)$.

Now suppose that edge $e = (w, v)$ is heavy. Then e is an f-flow-carrying edge and $d(v) = d(w) + \ell_e(f_e)$. As above, combining this fact with the inequalities $d'(v) \leq d'(w) + \ell_e(f'_e)$, $\ell_e(f'_e) \leq \ell_e(f_e)$, and $d(v) \leq d'(v)$ gives $d(w) \leq d'(w)$, which proves the inductive step and therefore the lemma. ■

In fact, the proofs of Lemma 2 and Theorem 3 can be extended to prove that $d(v) \leq d'(v)$ for all nodes $v \in V$, where d and d' denote the distances used in the proof above.

Bounding Braess's Paradox. We now show how a refinement of the preceding argument leads to the main theorem of this note: to improve the latency of a Nash flow by a factor of more than k, k edges must be removed from the network.

Theorem 4 *Let (G, r, ℓ) be an instance, and (H, r, ℓ) an instance obtained by removing at most k edges from G. Then, $L(G, r, \ell) \leq (k + 1) \cdot L(H, r, \ell)$.*

Proof: Let f and f' be Nash flows for (G, r, ℓ) and (H, r, ℓ), respectively. We view f' as a flow (no longer at Nash equilibrium) in the larger network G in the obvious way. Let P be an alternating path, which exists by Lemma 2. Let d and d' denote shortest path distances w.r.t. the edge latencies induced by f and f' in G and H, respectively. Our proof here must differ from that of Theorem 3 because f' is at Nash equilibrium in H but not in G.

In this proof, we will proceed by induction on the *segments* of P, where by a segment we mean a maximal subpath of P that contains only light or only heavy edges. We claim that if v is a vertex at the end of a segment of P, and c segments of P between s and v contain some edge in G but not H, then $d(v) \leq d'(v) + c \cdot L(H, r, \ell)$. This inequality trivially holds when $v = s$, so suppose we know that it holds for a vertex v that is last on a segment of P. We wish to prove the corresponding inequality for w, defined as the last vertex on the next segment. If all edges on the segment between v and w lie in H as well as G—in particular, if the segment is one of light edges—then the arguments from the proof of Theorem 3 apply directly here. It remains to prove the inductive hypothesis in the case where the segment between v and w is heavy backward edges, at least one of which is absent from H.

To finish the proof, note first that since there is a path of heavy edges (used by f), $d(w) \leq d(v)$. By the inductive hypothesis, $d(v) \leq d'(v) + c \cdot L(G, r, \ell)$. By the stronger form of Lemma 2, this segment has a predecessor segment in P, which must be a path of light edges (used by f') that terminates at v; hence, f' must route flow from v to t and $d'(v) \leq d'(t) = L(H, r, \ell)$. Putting the chain of inequalities together and noting that $d'(w) \geq 0$, we obtain $d(w) \leq d'(w) + (c + 1) \cdot L(H, r, \ell)$, which completes the inductive step and the proof of the theorem. ■

In the proof, the upper bound on d-values increases by at most $L(H, r, \ell)$ per segment of heavy edges that are not all in H, and distinct segments of heavy edges are vertex-disjoint. We have thus proved a stronger result.

Corollary 5 *Let (G, r, ℓ) be an instance, and (H, r, ℓ) an instance with a set S of edges from G removed, with the size of the maximum matching in S at most k. Then, $L(G, r, \ell) \leq (k + 1) \cdot L(H, r, \ell)$.*

Also, the stronger form of Lemma 2 implies that an alternating path can be chosen to avoid all edges incident to s or t that are in G but not H.

Corollary 6 *Let (G, r, ℓ) be an instance, and (H, r, ℓ) an instance with only edges incident to s or t removed. Then, $L(G, r, \ell) \leq L(H, r, \ell)$.*

Since the maximum matching size among nodes other that s and t is at most $\lfloor n/2 \rfloor - 1$, combining Corollaries 5 and 6 we obtain a new proof of one of the main results from [3].

Corollary 7 ([3]) *Let (G, r, ℓ) be an instance and H a subgraph of G. Then, $L(G, r, \ell) \leq \lfloor n/2 \rfloor \cdot L(H, r, \ell)$.*

References

[1] D. Braess. Uber ein paradoxon der verkehrsplanung. *Unternehmensforschung*, 12:258–268, 1968.

[2] M. A. Hall. Properties of the equilibrium state in transportation networks. *Transportation Science*, 12(3):208–216, 1978.

[3] T. Roughgarden. Designing networks for selfish users is hard. In *FOCS '01*, pages 472–481.

[4] T. Roughgarden. *Selfish Routing*. PhD thesis, Cornell University, 2002.