*J. Appl. Prob.* 27, 730–734 (1990)
*Printed in Israel*


# A PARADOX OF CONGESTION IN A QUEUING NETWORK

JOEL E. COHEN,* *Rockefeller University*
FRANK P. KELLY,** *University of Cambridge*

**Abstract**

In an uncongested transportation network, adding routes and capacity to an existing network must decrease, or at worst not change, the average time individuals require to travel through the network from a source to a destination. Braess (1968) discovered that the same is not true in congested networks. Here we give an example of a queuing network in which added capacity leads to an increase in the mean transit time for everyone. Self-seeking individuals are unable to refrain from using the additional capacity, even though using it leads to deterioration in the mean transit time. This example appears to be the first queuing network to demonstrate the general principle that in non-co-operative games with smooth payoff functions, user-determined equilibria generically deviate from system-optimal equilibria.

BRAESS'S PARADOX; NASH EQUILIBRIA; NON-CO-OPERATIVE GAME

## 1. Introduction

It seems intuitively obvious that adding routes and capacity to an existing transportation network should decrease, or at worst not change, the average time individuals require to travel through the network from a source to a destination. In an uncongested transportation network, the obvious is true. In a congested transportation network, the obvious need not be true.

Braess (1968) discovered a deterministic mathematical model of a congested network such that, paradoxically, when a link is added and each individual seeks his or her best possible route, at the new equilibrium the cost of travel for all individuals is higher than before. At equilibrium, independently self-seeking individuals are unable to ignore the added capacity that ends up increasing their travel cost. Braess's paradox is not a peculiarity of the parameter values or functional forms in Braess's example (Steinberg and Zangwill (1983); Dafermos and Nagurney (1984a,b)). The paradox may actually have occurred during 'development' in the center of Stuttgart (Knödel (1969)).

We report the first example of Braess's paradox in a mathematical model of a queuing network. Our example shows that the paradox is not a peculiarity of the mathematical formalism Braess used to describe a transportation network, but appears to be a more

Received 19 June 1989; revision received 20 July 1989.
* Postal address: Rockefeller University, 1230 York Avenue, Box 20, New York, NY 10021, USA.
** Postal address: Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, 16 Mill Lane, Cambridge CB2 1SB, UK.

general property of congested flows. Braess's paradox is not mentioned in recent reviews of optimal routing in queuing networks (Gallager (1977), Bertsekas (1982)), though it is known that decentralized routing based on the shortest expected delay may result in poor performance (Bertsekas (1982), p. 630).

## 2. Queuing network: behavioral assumptions

In an initial queuing network (Figure 1), individuals (customers, messages, manufacturing jobs, etc.) enter the network at node $A$ and exit at node $F$. Arrival streams are assumed to be Poisson. When a flow divides in two, e.g., at node $A$, the outgoing streams are assumed to be independent Poisson flows. Individuals know the mean delays in the queues in the network, but do not know the instantaneous queue lengths.

Each individual chooses a route to minimize his or her total mean time of transit from entry to exit, given the choices of other individuals. Equilibrium is defined to occur when no individual can lower his or her mean transit time by a change of route, if all other individuals retain their present routes. Thus the individuals in the network may be viewed as playing a non-co-operative game, each seeking to minimize the mean transit time from entrance to exit.

The network contains two kinds of servers, FCFS and IS. FCFS denotes a single-server queue with a first-come-first-served queue discipline. Service times are assumed to be independent exponential random variables with mean of $1/\varphi$ time units, where $\varphi > 0$ is fixed throughout. If individuals arrive at a stationary FCFS queue in a Poisson stream with a mean of $x$ individuals per unit of time, where $x < \varphi$, the mean sojourn time in the queue of an individual is $1/(\varphi - x)$.

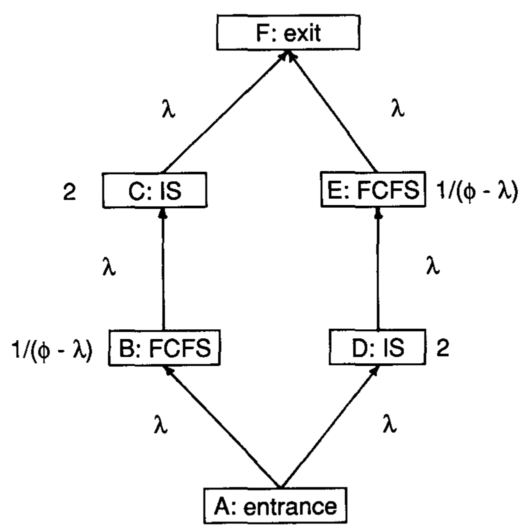


Figure 1. Initial queuing network. The entrance node is $A$. The exit node is $F$. Infinite-server queues are denoted IS; first-come-first-served single-server queues are denoted FCFS. The figure or formula adjacent to each node is the mean sojourn time in that node at equilibrium. The figure or formula adjacent to each arrow is the mean of the Poisson flow along that arrow at equilibrium

IS denotes an infinite-server queue at which each individual is delayed by some random amount of time, the average of which is independent of the number of individuals awaiting service. In the two IS queues in Figure 1, the average delay is 2 time units.

The augmented queuing network shown in Figure 2 differs from that in Figure 1 by the addition of an IS queue (node $G$) with a mean delay of 1 time unit and a route from node $B$ to node $G$ to node $E$.

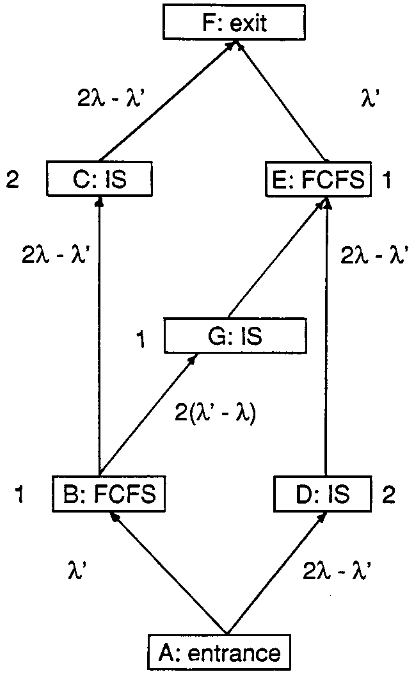


Figure 2. Augmented queuing network. An infinite-server queue (node $G$) has been added to the network of Figure 1. All symbols are as in Figure 1. The additional capacity in this queuing network *increases* the mean transit time of every individual at equilibrium

## 3. Analysis

For certain parameter values, if individuals choose a route from entry to exit so as to minimize their average transit time, given the choices of other individuals, then at equilibrium the mean transit time in the augmented network (Figure 2) is strictly larger than the mean transit time in the initial network (Figure 1).

*Theorem*. Let $2\lambda$ denote the total traffic departing from node $A$ and assume $2\lambda > \varphi - 1 > \lambda > 0$. Then the mean transit time in the initial network is strictly less than 3 time units, while the mean transit time in the augmented network equals 3 time units.

*Proof*. In the initial network (Figure 1), if $x$ is the Poisson flow from $A$ to $B$, then the total mean transit time along the route $ABCF$ is $1/(\varphi - x) + 2$; the first term is the mean wait at node $B$ and the second is the mean wait at node $C$. This is a convex function of $x$. By symmetry the mean transit time along $ADEF$ is $1/(\varphi - [2\lambda - x]) + 2$. It follows that at equilibrium individuals minimize their mean transit time by behaving so that $x = \lambda$, i.e., by distributing themselves equally between the two possible routes. Thus the mean transit time for all individuals is $1/(\varphi - \lambda) + 2$.

The assumption that $\varphi - 1 > \lambda > 0$ implies that $0 < 1/(\varphi - \lambda) < 1$, hence the mean transit time is less than 3 time units.

In the augmented network (Figure 2), there are three routes from entry to exit: $ABCF$, $ABGEF$ and $ADEF$. Whichever of these routes have positive flows of traffic at equilibrium must also have equal mean transit times, for if not individuals would shift from a route with greater mean transit time to a route with lesser mean transit time. We shall show that all three routes carry positive traffic flows and that the mean transit time for all is 3 time units.

Suppose the Poisson flow from node $A$ to node $B$ is $\lambda'$. Then the mean transit time along $ABCF$ is $1/(\varphi - \lambda') + 2$. Let $x$ be the Poisson flow from node $B$ to node $C$ and let $y$ be the Poisson flow from node $B$ to the newly added node $G$; $x + y = \lambda'$. The residual flow from node $A$ to node $D$ is $2\lambda - \lambda'$. Therefore the total Poisson stream arriving at node $E$ from nodes $D$ and $G$ is $2\lambda - \lambda' + y$.

An individual who departs from node $B$ and goes to node $C$ faces a mean transit time of 2 units before arriving at node $F$. An individual who travels from node $B$ to node $F$ via nodes $G$ and $E$ faces a mean transit time of $1 + 1/(\varphi - [2\lambda - \lambda' + y])$.

If $y > 0$, these transit times must be equal, whence $y = \varphi - 2\lambda + \lambda' - 1$. Therefore the Poisson stream arriving at node $E$ is just $\varphi - 1$ and the mean sojourn time in the queue at node $E$ is precisely 1 unit. Therefore the mean transit time along $ADEF$ is 3 units; hence the mean sojourn time at node $B$ must be 1 unit, i.e., $1 = 1/(\varphi - \lambda')$ or $\lambda' = \varphi - 1$ and $y = 2(\lambda' - \lambda)$ and $x = 2\lambda - \lambda'$. The mean transit time along each possible route is 3 time units.

Suppose $y = 0$. The argument given for the initial network implies that $\lambda' = \lambda$, so the flow from node $D$ to node $E$ is just $\lambda$ and the mean sojourn time in the queue at node $E$ is only $1/(\varphi - \lambda) < 1$. So an individual departing from node $B$ could arrive at node $F$ in only $1/(\varphi - \lambda) + 1$ time units via nodes $G$ and $E$, and this would be faster than the 2 time units required to travel by node $C$. Hence at equilibrium $y > 0$. By hypothesis, $2\lambda > \varphi - 1$, and we showed that $\varphi - 1 = \lambda'$. Since $x = 2\lambda - \lambda'$, it follows that $x > 0$ and the flows along all three routes are positive, with a mean transit time of 3 time units.

The additional node $G$ in the augmented network allows some individuals to pass through both FCFS servers and to avoid both of the original IS servers. The additional flow through the FCFS servers increases the mean delay there with no compensating reduction in the mean delay incurred at the IS servers. Thus the initial decrease in total waiting time of the customers that choose the new route is more than paid for by the increase in delay from the congestion at the merging of this new traffic.

This result illustrates a general method for constructing queuing networks that display paradoxes analogous to those of deterministic network models. After the first presentation of these results, Richard Stone (personal communication) used the same approach to construct a queuing analog of another novel paradox of deterministic congested network models.

## 4. Conclusion

In the simple queuing network reported here, added capacity leads to an increase in the mean transit time of everyone, because self-seeking individuals are unable to refrain from using the additional capacity even though using it leads to deterioration in the mean transit time. This example appears to be the first queuing network to demonstrate the general principle (Dubey (1986)) that in non-co-operative games with smooth payoff functions, user-determined equilibria generically deviate from system-optimal equilibria. In the language of game theory, Nash equilibria are generally Pareto-inefficient.

It would be valuable to learn how frequently this paradox arises in real queuing networks, and whether the paradox is sensitive to the assumption that individuals respond to mean sojourn times rather than to instantaneous queue lengths. It may also be rewarding to recognize analogous paradoxes in other technological and biological systems where individual actions have significant consequences for others.

## Acknowledgements

These results were first presented in an invited talk at the Third Advanced Research Institute on Discrete Applied Mathematics (ARIDAM III), Rutgers University, 23 May 1988. David Housman, Fred S. Roberts, Richard Steinberg, Richard E. Stone, D. de Werra, other participants of ARIDAM III, and the referee provided helpful suggestions and reprints. J.E.C. thanks Mr and Mrs William T. Golden for hospitality during this work. J.E.C. was supported in part by NSF grant BSR 87-05047.

## References

BERTSEKAS, D. P. (1982) Optimal routing and flow control methods for communication networks. In *Analysis and Optimization of Systems*, ed. A. Bensoussan and J. L. Lions. Lecture Notes in Control and Information Sciences **44**, Springer-Verlag, Berlin, 615–643.

BRAESS, D. (1968) Über ein Paradoxon aus der Verkehrsplanung. *Unternehmensforschung* **12**, 258–268.

DAFERMOS, S. AND NAGURNEY, A. (1984a) On some traffic equilibrium theory paradoxes. *Transportation Res. B* **18**, 101–110.

DAFERMOS, S. AND NAGURNEY, A. (1984b) Sensitivity analysis for the asymmetric network equilibrium problem. *Math. Programming* **28**, 174–184.

DUBEY, P. (1986) Inefficiency of Nash equilibria. *Math. Operat. Res.* **11**, 1–8.

GALLAGER, R. G. (1977) A minimum delay routing algorithm using distributed computation. *IEEE Trans. Comm.* **25**, 73–85.

KNÖDEL, W. (1969) *Graphentheoretische Methoden und ihre Anwendungen*, Springer-Verlag, Berlin, pp. 56–59.

STEINBERG, R. AND ZANGWILL, W. I. (1983) The prevalence of Braess' paradox. *Transportation Sci.* **17**, 301–318.