How Braess' paradox solves Newcomb's problem: not!

LOUIS MARINOFF
Department of Philosophy, The City College of the City University of New York, USA

'For what is the Heart, *but a* Spring; *and the* Nerves, *but so many* Strings..."
(Hobbes, 1651, Introduction to the Leviathan*)*

Abstract *In an engaging and ingenious paper, Irvine (1993) purports to show how the resolution of Braess' paradox can be applied to Newcomb's problem. To accomplish this end, Irvine forges three links. First, he couples Braess' paradox to the Cohen-Kelly queuing paradox. Second, he couples the Cohen-Kelly queuing paradox to the Prisoner's Dilemma (PD). Third, in accord with received literature, he couples the PD to Newcomb's problem itself. Claiming that the linked models are "structurally identical", he argues that Braess solves Newcomb's problem. This paper shows that Irvine's linkage depends on structural similarities—rather than identities—between and among the models. The elucidation of functional disanalogies illuminates structural dissimilarities which sever that linkage. I claim that the Cohen-Kelly queuing paradox cloaks a fine structure that decouples it from both Braess' paradox and the PD (Marinoff, 1996a). I further assert that the putative reduction of the PD to a Newcomb problem (e.g. Brams, 1975; Lewis, 1979) is seriously flawed. It follows that Braess' paradox does not solve Newcomb's problem via the foregoing and herein sundered chain. I conclude by substantiating a stronger claim, namely that Braess' paradox cannot solve Newcomb's problem at all.*

1. Structural similarities among four models

Braess' paradox

Braess' original paradox (1968), set in the context of general transportation networks, illustrates that additional carrying capacity can lead to more costly travel for all. Cohen & Horowitz (1991) introduce both mechanical and electronic models of Braess' problem. Irvine (1993) selects the mechanical model for his linkage. Figure 1a depicts a network of identical, ideally massless and perfectly elastic springs ($S1$, $S2$) connected by a massless and perfectly inelastic string (L). A suspended mass (M) extends the network to an equilibrium length ($H1$). Note the additional pair of identical strings ($L1$, $L2$) of sufficient length so as to remain slack when attached as shown.

The question is: if L were cut, would the resultant equilibrium length ($H2$) raise or lower mass M with respect to its previous position? The answer is: it depends. Although one's intuition might proclaim that the new position of M would surely be lower than

0269-8595/96/030217-20

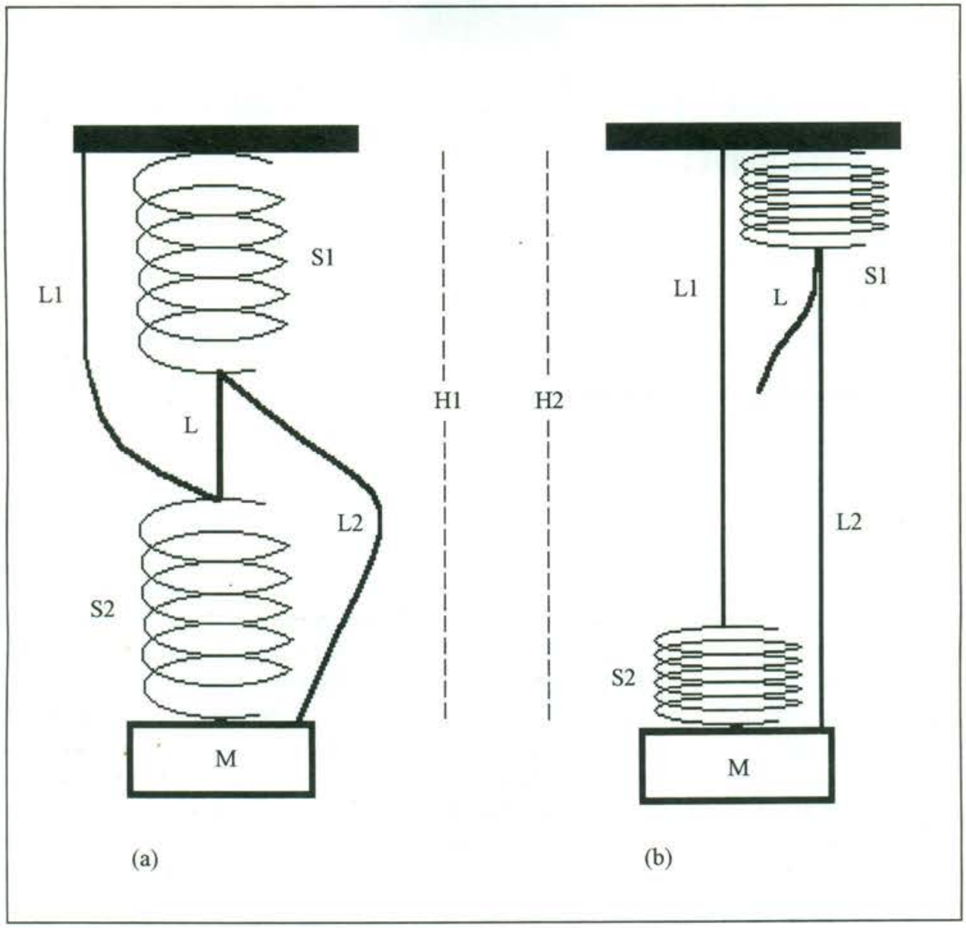

Figure 1. *Braess' paradox.*

the old, in fact the new equilibrium length of the network can be greater than, less than, or equal to the previous length. In both configurations, equilibrium extension depends upon the relative values of systemic constants—the unstretched spring-length (S), the spring-constant or modulus of elasticity (K), the string-lengths (L, $L1$, $L2$), the local gravitational constant (g), and the upon the variable mass. Explicitly,

$$H1 = L + 2S + 2Mg/K$$
$$H2 = L1 + S + Mg/2K$$

Thus

$$H2 > H1 \text{ iff } (L1 - L) > (S + 3Mg/2K)$$
$$H2 < H1 \text{ iff } (L1 - L) < (S + 3Mg/2K)$$
$$H2 = H1 \text{ iff } (L1 - L) = (S + 3Mg/2K)$$

If the unstretched spring is sufficiently long, the spring-constant sufficiently large, and the difference in string-lengths sufficiently small, then the suspended mass will ascend, not descend, when the central string is cut. As Irvine rightly points out, this result is not at all paradoxical; merely counter-intuitive. If your intuition misinformed you that the mass would necessarily descend, it did so by maintaining consistency with an implicit but unsound assumption: namely, that the suspended mass produces the same extension of the springs in both configurations. In fact, it does not. In Figure 1a, each spring bears the full weight (Mg) and therefore sustains the full resultant extension (Mg/K); whereas in Figure 1b, each spring bears only half the weight ($Mg/2$) and therefore sustains only half the resultant extension ($Mg/2K$). As Irvine says, once we appropriately modify our background assumption(s), we accordingly dispel the counter-intuition.

Many so-called "paradoxes" (e.g. Zeno's, Bertrand Russell's, Allais') have been similarly resolved (e.g. Marinoff, 1994). These problems do not entail contradictions; they only appear to do so in light of unsound or incompatible background assumptions which require, and which eventually receive, appropriate modification. It remains an open question whether every paradox can be similarly resolved. If so, then "paradox" is only an apotheosized synonym for an inconsistent set of premises (e.g. Sorensen, 1988). If not, then our faculties of speech and reason both humble and torment our cherished illusion of understanding.

The Cohen-Kelly queuing paradox

The Cohen-Kelly queuing paradox appears as a particular variation on Braess' general theme. Cohen & Kelly (1990) (and Cohen and Horowitz (1991), and Kelly (1991) represent Braess's (1968) paradox as "a more general property of congested flows" in a transportation network. The paradox arises because the addition of an alternative route through a congested network appears to increase, rather than decrease, mean transit time through the network. In this context, "congested" and "non-congested" are technical terms. A non-congested network is one "in which the cost for traversing an arc is independent of the number of users on that arc" (Steinberg & Zangwill, 1983); similarly, a congested network is one in which the cost of traversing an arc is dependent upon the number of users on that arc.

In the initial queuing network, depicted in Figure 2a, clients (e.g. travellers, customers, messages, jobs, etc.) enter at node A. The single incoming stream is a Poisson flow, and the two outgoing streams are independent Poisson flows. The network contains two kinds of servers: FCFS (first-come-first-served) and IS (infinite-server). The delay incurred at an FCFS server is a (convex) function of traffic flow to that server; i.e. clients are queued and processed in sequence. Explicitly, the mean FCFS delay time is $1/(b-x)$ units, where b is a systemic constant and x is the rate of flow ($b > x > 0$). The delay incurred at an IS server is independent of traffic flow to that server; i.e. clients are not queued but instead are processed in parallel. Explicitly, the mean IS delay time is a units, where a is another systemic constant. To reach the exit at F, clients must travel via one of two routes: either $ABCF$, or $ADEF$.

In the augmented queuing network, depicted in Figure 2b, an additional IS server, with an associated mean delay time of $a/2$ units, is placed between and linked to the two FCFS servers. This affords clients an alternative route through the network, namely $ABGEF$. Since each client seeks to minimize individual transit time, whether in light or in spite of collective traffic flow, the augmented network can be viewed as an N-player non-cooperative game.

Cohen & Kelly (1990) show that, under certain conditions, mean transit time through the augmented network is strictly larger than mean transit time through the

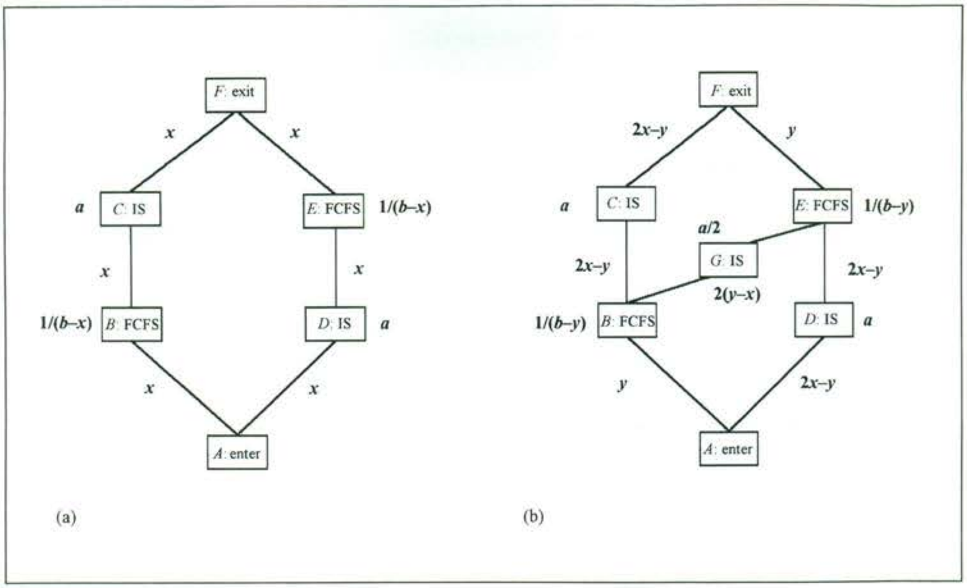

Figure 2. *Cohen-Kelly queuing paradox.*

initial network. In this section I offer a more succinct and also more general proof of their theorem. Following Cohen & Kelly, I denote the total traffic arriving at node A as $2x$ units of flow, and assume that $2x > b - 1 > x > 0$.

In the initial network, let the Poisson flow through $ABCF$ be y. Thus the flow along $ADEF$ is $2x - y$. The mean transit times through these branches are, respectively, $1/(b - y) + a$ and $1/[b - (2x - y)] + a$. At equilibrium, clients distribute themselves such that their mean transit times are equal, from which it follows algebraically that $y = x$. This result, which is also intuitively obvious, gives the mean transit time through the initial network as $1/(b - x) + a$.

There are three possible routes through the augmented network: $ABCF$, $ADEF$, and $ABGEF$. Again, assuming an equilibrium condition, mean transit times along all routes are equal (although in this case the flows are not). As before, let the Poisson flow along AB be y. Then the mean transit time along $ABCF$ is $1/(b - y) + a$. From the equilibrium assumption, the mean transit times along routes $ABCF$ and $ADEF$ are equal. So the mean transit time along $ADEF$ is also $1/(b - y) + a$. Since the flow along ADE is $2x - y$, it follows that the flow along BGE must be $2(y - x)$, which yields the necessary algebraic sum y for the flow through server E.

To prove the theorem, it is sufficient to observe that since $2(y - x)$ (the traffic flow between servers B and E) is by definition greater than zero, and since $x > 0$, therefore $y > x$. It follows that

$$1/(b - y) + a > 1/(b - x) + a$$

or, in other words, that the mean transit time through the augmented network is strictly greater than that through the initial network.

Irvine links this and the foregoing model by graphical representation; that is, by direct mapping of components: he allows that IS gates correspond to strings, FCFS

(a)

		Column Player	
		c	*d*
Row Player	*C*	*R*,*R*	*S*,*T*
	D	*T*,*S*	*P*,*P*

C and *c* denote cooperation; *D* and *d*, defection

$T > R > P > S$

(b)

		Column Player, or *N* Other Players	
		c	*d*
Row Player	*C*	*R*	0
	D	$R+P$	*P*

$S = 0; T = R+P$

$R > P$

Figure 3. *Prisoner's dilemma.*

gates to springs, mean travel times to equilibrium extensions, and items of traffic to units of downward force. The addition of the alternative route in the augmented network, which leads to increased mean travel time, similarly corresponds to the addition of the central string, which under appropriate conditions increases the equilibrium extension of the mechanical apparatus. Thus the first link is forged.

The Prisoner's Dilemma

The PD is perhaps the paradigmatic non-cooperative game, characterized by a payoff matrix whose strong transitive ordering has spawned a gargantuan literature that spans game theory, evolutionary biology, social psychology, sociology, economics, political science, ethics, computer modeling and rational choice theory. Figure 3a depicts the generic two-player PD. The fundamental tension in the original two-player model arises because the dominance strategy dictates that row player is better off defecting regardless of column player's choice. Owing to symmetric payoffs, the converse prescription also obtains. Hence both players defect, to their mutual detriment. Mutual defection results in the attainment of a so-called "Nash equilibrium", a state such that neither player can realize a better payoff by making a different choice, given that the other player does not make a different choice. Then again, assuming a sufficient degree of probabilistic dependence, the strategy of maximizing expected utilities prescribes that each player cooperate. Mutual cooperation results in the attainment of a so-called "Pareto efficient" (or "Pareto optimal") outcome, a state which confers the best aggregate payoff.

The PD can readily accommodate multiple pairs of players (a possibility that has engendered strategic competition in computer tournaments, e.g. Axelrod, 1980a, 1980b, Marinoff, 1992, 1996b). It can also accommodate an undifferentiated horde of players. Hobbes's (1651) state of nature, Hume's (1729) meadow-draining experiment, Hardin's (1973) tragedy of the commons, Pettit's (1986) free-riding scenario, and Glance's & Huberman's (1994) office luncheon are exemplary many-player PDs.

From an egoistic disposition, each player in a many-player PD can view himself as arrayed against a collective opponent, a Leviathanesque corpus compounded of all the other players. In such a game, entries in the payoff matrix are not constants; rather, are partial functions of the proportion of cooperators or defectors in the player population. Prompted by Irvine (1993), I have developed one plausible family of such functions, which consistently extends the maximization principle from the two-player to the many-player case (Marinoff, 1996c). It does so on the assumption that maximizing expected utilities prescribes that the lone player cooperate just in case sufficient numbers of other players also cooperate. One determines the minimum sufficient number, or threshold frequency of cooperation, as follows.

Availing oneself of the arbitrarily-fixed zero-point of the ordinal scale on which the payoffs are expressed, one chooses $S = 0$, and $T = R + P$ (see Figure 3b). In the two-player case, one posits a probability of conditional cooperation: call it x. Maximizing expected utilities on this schema yields $\text{EUC} = xR$, $\text{EUD} = (1 - x)(R + P) + xP$. Letting q stand for the quotient P/R, it follows that MEU prescribes cooperation if $x > (1 + q)/2$. In the many-player case, one assumes that each of n other players cooperates with some probability x (either uniform or uniformly averaged over n). Now the lone player evaluates probabilities of cooperative proportions in the opposing population, such that MEU prescribes cooperation only if the proportion of probable cooperators exceeds the required threshold $(1 + q)/2$.

Given that one knows the number of other players (n) and their average probability of cooperation (x), one can readily determine whether the threshold frequency is exceeded. The probability that k of n players cooperate is the product of (1) the number of possible states of k cooperative players (i.e. combinatorially distinct states), and (2) the probability that each state obtains. Thus:

$$p(f_k) = \binom{n}{k} p_k$$

Explicitly:

$$p(f_k) = \frac{n!}{k!(n-k)!} x^k (1-x)^{n-k}$$

Similarly, the probability that $k + 1$ other players cooperate is:

$$p(f_{k+1}) = \frac{n!}{(k+1)!(n-k-1)!} x^{k+1} (1-x)^{n-k-1}$$

Now, if k is the smallest integer such that $f_k > (1 + q)/2$; that is, the smallest integer such that $k/n > (1 + q)/2$, then:

$$p(f) = p(f_k) + p(f_{k+1}) + \ldots + p(f_n)$$

that is:

$$p(f) = \sum_{j=k}^{n} \frac{n!}{j!(n-j)!} x^j (1-x)^{n-j}$$

Irvine's linkage between the N-player PD and the Cohen-Kelly queuing paradox is straightforward. Obviously, they both belong to the class of many-player, non-cooperative, non-zero-sum games of imperfect information. The shared taxonomy entails a functional analogy. In the Cohen-Kelly problem, the best possible payoff obtains for a given individual just in case all the others choose the initial routes, while that individual alone chooses the augmented route [on the assumption that $a > 2/(b - x)$]. In the PD, the best possible payoff obtains for a given individual just in case all the others cooperate while that individual defects. Choosing the initial route is thus analogous to cooperation; the augmented route, to defection. In both models, individual defection is strongly

		Demon	
		predicts opaque box	predicts both boxes
Player	chooses one box	\$M	0
	chooses both boxes	\$M + \$T	\$T

\$M denotes one million dollars
\$T denotes one thousand dollars

Figure 4. *Newcomb's problem.*

dominant, yet in both models, mass defection leads away from Pareto-efficiency and toward Nash equilibrium. Thus the second link is forged.

Newcomb's problem

Since its original articulation by Nozick (1969), Newcomb's problem has engendered a rich, philosophical literature. In its generic formulation, an arbitrarily lengthy but finite succession of humans play one at a time against an arbitrarily wealthy and uncannily prescient demon. Each human player is confronted by both a transparent box which contains \$1000, and an opaque box that will subsequently contain either 1 million dollars or else nothing. The moves are made in strict sequence. First, the demon predicts whether the human player will choose the contents either of the opaque box alone, or else of both boxes. If the former prediction is made, the demon places 1 million dollars in the opaque box; if the latter, nothing. The human then chooses either the contents of the opaque box alone, or else of both boxes. A caveat: the human is aware that arbitrarily many humans have already played this game, and that the demon's frequency of correct prediction is arbitrarily close to unity. The payoff matrix is depicted in Figure 4.

As Figure 4 illustrates, the payoff structures of Newcomb's problem and the PD share the same strong transitive ordering. Choosing both boxes is analogous to defecting and, once again, defection is strongly dominant. Similarly, choosing the opaque box alone is analogous to cooperating and, given the demon's outstanding predictive performance, maximizing expected utilities prescribes cooperation (but not unequivocally). Albeit with varying justification, contributors to the debate on this problem are known, in a predictable vernacular, as either "one-boxers" or "two-boxers".

The standard, simplistic but plausible principle that prescribes choosing both boxes is dominance, here reinforced by causality. Choosing two boxes strongly dominates choosing one box only. Moreover, given the rigid temporal succession of moves, choosing both boxes can neither cause the demon to refrain from placing the million in the opaque box, nor cause the demon to remove the million from the opaque box, in the event that it wrongly predicted the player's choice and placed the million therein. At the same time, choosing one box alone cannot cause the demon to place 1 million in it, just in case it predicted wrongly and refrained from doing so.

The standard, simplistic but plausible principle that prescribes choosing the opaque box alone is maximizing expected utilities, here reinforced by evidentiality. Given the demon's successful prediction rate, the margin of prescription is convincingly wide. Let the demon's relative frequency of correct prediction be x. Assume that the utility of money is linear in its amount. Then the expected utilities are:

$$EU1 = xM;\ EU2 = (1 - x)(M + T) + xT$$

In order that choosing one box be prescribed, EU1 must exceed EU2. Thus:

$$xM > (1 - x)(M + T) + xT$$
$$x > (1 + T/M)/2$$
$$x > 0.5005$$

But given that x is arbitrarily close to unity, x far exceeds this threshold. Thus, choosing one box is forcefully prescribed. Note that the demon's high frequency of correct prediction lends apparent evidentiary support to the subjunctive conditional proposition that, if a player were to choose only the opaque box, then it would contain 1 million dollars.

Then again, it has been shown by Lewis (1976) that conditional probabilities are not in general equal to probabilities of conditionals, and by Levi (1974) that conditional probabilities are not in general equal to converse conditional probabilities. It transpires that alternative formulations of the expected utility calculus—for instance by Gibbard & Harper (1978), Eells (1982), and Jeffrey (1983)—can lead to prescriptions which converge with that of dominance; i.e. which suggest taking both boxes. The relevance of such formulations to Newcomb's problem has been ably challenged by Horwich (1987) and Price (1986, 1991), who save the catholic one-box evidentiary picture from two-box reformation. But these developments do not resolve the theoretical debate; they rather shift its ground to issues involving the underdetermination of decision theory by counterexample. Such issues are important, but tangential to this treatment.

Empirically, Nozick (1969) reports that people's preferences seem to divide about evenly with respect to these conflicting principles of choice. One each occasion that I have presented this dilemma to a class of students, I have found a similar distribution of choice. It also turns out, as Nozick remarks, that each group usually deems the other mistaken or even foolish for adhering to its respective principle.

The linkage between the N-player PD and Newcomb's problem is well-rehearsed in the literature. Notably, both Brams (1975) and Lewis (1979) have argued that the PD is actually two Newcomb problems, running simultaneously and in parallel. Each prisoner, in effect, is engaged in a separate Newcomb problem, because each prisoner can (and possibly should) suppose that the other has already made his choice. This is not to assert the temporal impossibility that both prisoners choose first; rather, to assert the logical possibility that each supposes the other to have chosen first. In brief, the claim that the PD is equivalent to two Newcomb problems is supported by their common payoff structures, and by similarly conflicting prescriptions of the relevant principles of choice. Thus the third link is forged.

Having coupled the four models, Irvine's argument approaches its conclusion:

...why is it that in the case of Newcomb's problem the argument from expected utility *appears* persuasive?...why is it that it is regularly those players who *avoid* dominance who obtain the largest payoffs?...The answer is a simple one....it is to deny that, over the long run, it will continue to be those who avoid dominance, or who fail to defect, who will continue to receive the largest payoffs...the solution to Newcomb's problem...involves an important modification to our original background assumptions...we simply abandon the (false) assumption that past observed frequency is an infallible guide to probability and, with it, the claim that Newcomb's problem is in any sense a paradox of rationality...The solution to the problem is simply to deny that selecting box B [the opaque box] alone in any way affects or is indicative of the probability of its containing $100,000 [$M]. Similarly in the case of the prisoner's dilemma, the solution is to deny that a single prisoner's failure to defect in any way affects or is indicative of the likelihood that other prisoners will also defect. Similarly in the case of the Cohen-Kelly queuing network, the solution is to deny that a single traveller's failure to select route ABGEF in any way affects or is indicative of the likelihood that other travellers will do the same. Such outcomes are no more likely in the case of Newcomb's problem, prisoner's dilemma, or the Cohen-Kelly queuing paradox than they are in the case of Braess' paradox. Which is to say, they are not likely at all.

2. Functional disanalogies among four models

Braess' paradox revisited

We are dealing here with an idealized deterministic system, whose state of extension is always unambiguously specifiable in terms of physical constants (i.e. string lengths, unstretched spring lengths, moduli of elasticity, a gravitational constant) and a single variable, suspended mass. Note also that the network is subject to some fixed carrying capacity, which if exceeded by the suspended mass would result in a broken string or a deformed spring. Thus, in order to maintain the network's integrity, we demand that the suspended mass not exceed a certain threshold, whose value is determined not by explicit systemic constants, rather by intrinsic physical properties of systemic components.

Adopting Irvine's analogy, we begin to introduce units of mass into the system. Imagine that the strings and springs consist of hollow tubes, through which one drops a succession of marbles. They enter the system at the top, and roll down to the bottom, where they accumulate as a collective suspended mass. The "path" taken by each marble is represented precisely by the increase in tension that it effects in a given branch of the network. All possible tensions in the network—and hence all possible extensions produced by them—are sufficiently described by a set of three independent equilibrium equations, one pertaining to each branch (see Figures 5a and 5b):

$$\{(T + T1 = K\Delta S2),\ (T + T2 = K\Delta S2),\ (T + T1 + T2 = Mg)\}$$

Note that for each unit of mass so introduced, the system reacts by attaining a static equilibrium extension such that these equations are simultaneously satisfied. Although the conformation of the network is admittedly unusual, the fundamental laws of statics are neither compromised nor violated by it.

A closer examination of these systemic equations reveals that it is physically

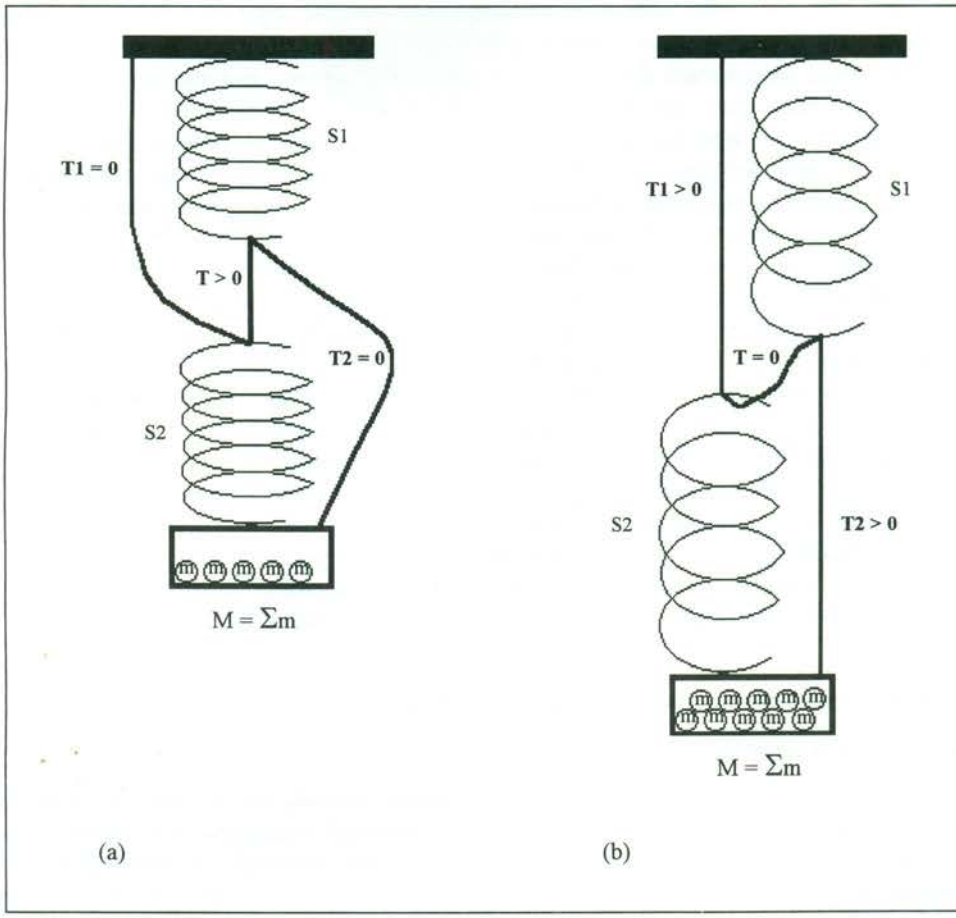

Figure 5. *Braess' problem revisited.*

impossible for all three branches to be simultaneously loaded. Specifically, given that $L1$, $L2 > L$, an initial accumulation of suspended mass creates tension only in the central branch through L. As long as L is taut, $L1$ and $L2$ necessarily remain slack. Eventually, a sufficient accumulation of suspended mass stretches the springs enough to necessitate the transfer of tension to $L1$ and $L2$, whereupon they become taut while L necessarily reverts to slackness. There is no balance point such that tension is distributed throughout. In consequence, while every aggregate mass produces a corresponding static equilibrium extension, no aggregate mass can produce an extension such that all branches are simultaneously under tension.

Thus we already have a *prima facie* rebuttal of the thesis that Braess' paradox "solves" Newcomb's problem. For every unit of mass introduced into Braess' network, the branch along which it produces tension is strictly determined by the aggregate accumulated mass (in light of systemic constants). Physical determination repudiates Irvine's implied denial that a single unit of mass's failure to produce tension in a given branch "in any way affects or is more indicative of the likelihood" that other units of mass will do the same. In fact, if a single unit of mass fails to produce tension in the central branch

Table 1. *Cohen-Kelly queuing paradox revisited*

State of Network:	Specification:	Minimum Mean Transit Time:	Traffic Pattern:
Underpopulated	$x \ll b-2/a$	Lesser of: $1/b + a$ or $2/b + a/2$	Free-flowing
Sparsely-populated	$x < b-2/a$	$1/(b-y) + a = 3a/2$ $(y = b-2/a)$	Nash equilibrated
Modulated	$x = b-2/a$	$1/(b-x) + a = 3a/2$ $(y = x)$	Pareto-efficient
Densely-populated	$x > b-2/a$	Possibly: $2/(b-x) + a/2$ or $1/(b-2x) + a$	Indeterminate
Overpopulated	$x \gg b-2/a$	None (all routes clogged)	Arrested flow

of Braess' network, then it is a matter not merely of likelihood, but rather of certainty, that every subsequent unit of mass must similarly fail. But while Irvine (like Hume) denies that observed frequency is an infallible indicator of future likelihood, he (unlike Hume) does not assert that physical necessity is a fallible indicator of future certainty. Yet this is implied by his ultimate linkage between Braess' and Newcomb's problems. Leaving that on one side, we proceed to refute the putative structural identity of Braess' model and the Cohen-Kelly queuing paradox.

The Cohen-Kelly queuing paradox revisited

If the two models are, as Irvine claims, "structurally identical", then they ought to be functionally analogous. But while Braess' network cannot sustain simultaneous non-zero tensions in all its branches, the augmented Cohen-Kelly network can obviously tolerate simultaneous non-zero transit times along all its routes. Thus they are not functional analogues. Since tensions and transit times (effected respectively by items of mass and numbers of travellers) turn out to be functionally disanalogous, the models cannot be structurally identical. They are merely graphically—and perhaps topographically—similar. This observation suffices to sever the first link.

The Cohen-Kelly model itself embodies a peculiar functionality that bears further discussion (see Marinoff, 1996a). We can specify five different densities of traffic flow in terms of given systemic constants. Each of these five densities is indicative of a distinct state of the augmented network, with which we can associate a minimum mean transit time and a characteristic traffic pattern (see Table 1).

We begin with an unpopulated system, and gradually introduce traffic at node A. We assume that all travellers entering node A know both the instantaneous rate of flow at that node (call it $2x$) and the current state of the network itself. In the first case, $x \ll b - 2/a$. I term this state "under-populated", because traffic flow along all branches is non-zero but negligible. A traveller can minimize mean transit time simply in accordance with systemic constants. In the second case, $x < b - 2/a$. I term this state "sparsely-populated", because the flow is non-negligible but strictly less than a certain critical value. There are compelling reasons for asserting that traffic in this state would divide along all three routes such that the mean individual transit times will be equal but not minimal (Cohen & Kelly, 1990; Marinoff, 1996a). This state thus gives rise to the Nash equilibrium; Cohen & Kelly subsume their treatments completely under it. (For re-emphasis, I repeat that a corresponding equilibrium cannot be attained in the spring-and-string analog of Braess' model.)

But if we continue to increase traffic flow, a surprising thing happens: the very condition that sustains the Nash-equilibrated state suddenly transforms it into a Pareto-efficient state. Note that the rate of flow through node G is necessarily $2(y-x)$; the traffic must divide in this way to equalize mean individual transit times in the sparsely-populated state. The value of y remains constant; it is a function of systemic constants ($y = b - 2/a$). Now let us increase the traffic flow until $x = b - 2/a$. But now x equals y, and consequently there is zero traffic flow through node G. I term this state "modulated", because x is tuned to y. The significance of the modulation should be clear: traffic divides evenly between the two external routes, and Pareto-efficiency is therefore attained. The mean individual transit time in the modulated state is in fact the same as that in the sparsely-populated state, although traffic flow in the latter state is by definition greater than that in the former. The sudden but algebraically necessary emergence of Pareto-efficiency from Nash equilibrium is a counter-intuitive result indeed.

If we now increase traffic flow such that $x > b - 2/a$, Pareto-efficiency itself gives way to decision-theoretic indeterminacy, and does so with similar suddenness and force of algebraic necessity. I term this state "densely-populated", because the flow is now greater than the previous critical value. The densely-populated state conceals a counter-intuition of its own: while Pareto-efficiency is theoretically attainable (just in case all traffic divides evenly between the two external routes), Nash equilibrium is demonstrably unattainable. Recall that a necessary condition of Nash equilibrium is that flow through node G equal $2(y-x)$. But in the densely-populated state, x is greater than y. Hence, for a Nash equilibrium to occur in this state, traffic flow through node G would have to be negative. If we attempt to reinterpret negative traffic flow as backward traffic flow (*à la* Feynman), we will be hoist with a Newtonian petard: backward traffic flow would necessarily elapse in negative time, and the associated equations would assert that travellers can reverse time's arrow merely by driving in reverse along particular routes. While backward time-travel on select closed curves remains a Gödelian possibility (e.g. see Savitt, 1994), the foregoing methodology—namely driving in reverse gear—will not reliably allow you to arrive home from the office prior to your departure from it: on the contrary. It follows that there is no Nash equilibrium in the densely-populated state and, because of this, there is no characteristic traffic pattern either. For travellers in this state, decision theory is prescriptively mute.

If we continue to increase traffic flow until $x \gg b - 2/a$, we enter the "over-populated" state, in which decision theory is once again prescriptive but now useless. Since all routes are clogged, traffic flow is arrested. There is no mean minimum transit time at all.

The N*-player PD revisited*

In the N-player PD—in stark contrast to the Cohen-Kelly model—we find no constraint on the possibility of mass defection as the number of players increases. For any N players in a PD, it is always possible that all, or very nearly all, will defect. The Nash equilibrium is therefore always attainable in a PD (whether it is attained or not). Moreover, unlike the Cohen-Kelly model, the PD entails no modulated state, such that Pareto-efficiency necessarily emerges from the Nash equilibrium, and is itself necessarily superseded by indeterminate patterns of choice, as the number of players continues to increase. In sum, the specification of traffic patterns according to graduated flows in the Cohen-Kelly queuing paradox has no analogue in the N-player PD. This functional disanalogy arises from structural non-identity; thus the second link is severed.

I proceed directly to sever the third and final link, by offering two independent arguments which refute Brams's and Lewis's claims that the PD is a Newcomb problem, or rather is two Newcomb problems side-by-side. Lewis's version has permeated the philosophical literature without much serious opposition. A noteworthy dissenter is Sobel (1985), who argues that not every PD is a Newcomb problem, but in so doing assents that some are. My claim is stronger than Sobel's, and cedes no ground: no PD is a Newcomb problem.

I begin with two uncontroversial if not incontrovertible premises: first, that all rational players in the PD share similar interests—at least with respect to obtaining the better payoffs and avoiding the worse; and second, that the interests of the players and the demon in Newcomb's problem are distinctly dissimilar. In Newcomb's problem, all rational players still wish to obtain the better payoffs and avoid the worse, whereas the poor but far from impecunious demon reaps no payoffs whatsoever, except in the normally pejorative sense of paying out. Yet the demon remains indifferent to losing money; it is both fabulously wealthy and supremely disinterested in how often it may be compelled by the rules to shell out thousands and perforce millions of dollars. However, the demon does maintain one abiding interest, and that lies in sustaining a high frequency of correct predictions.

The first argument is meta-game-theoretic. It seems not generally well-known, or is perhaps largely ignored, that the two-player PD actually admits of a "solution", if by solution we understand an analysis that impels players away from the Nash equilibrium and toward the Pareto-efficient outcome. Here is a synopsis of Rapoport's (1967) presentation of Howard's meta-game solution, which I transpose from a two-player to an N-player mode.

Let the lone player have four strategies: A_1, cooperate unconditionally; A_2, choose expected majority choice; A_3, choose complement of expected majority choice; and A_4, defect unconditionally. These give rise to the second-order game depicted in Figure 6. As the matrix shows, unconditional defection is weakly dominant for the lone player. Defection is therefore meta-dominant for the n other players. Thus the lone player defects, and a Nash meta-equilibrium persists at $[A_4,d]$.

In the third-order game, each of the n other players generates 16 meta-strategies, which represent all possible combinations of responses to the concatenated string of the

		Lone Player			
		A_1	A_2	A_3	A_4
n Other Players	c	R,R	R,R	S,T	S,T
	d	T,S	P,P	T,S	P,P

Figure 6. *PD revisited: first-order metagame.*

	A_1	A_2**	A_3	A_4*
CCCCx	R,R	R,R	S,T	S,T
CCCDx	R,R	R,R	S,T	P,P
CCDCx	R,R	R,R	T,S	S,T
CDCCx	R,R	P,P	S,T	S,T
DCCCx	T,S	R,R	S,T	S,T
CCDDz	R,R	R,R	T,S	P,P
CDCDx*	R,R	P,P	S,T	P,P
DCCDx	T,S	R,R	S,T	P,P
CDDCx	R,R	P,P	T,S	S,T
DCDCx	T,S	R,R	T,S	S,T
DDCCx	T,S	P,P	S,T	S,T
DDDCx	T,S	P,P	T,S	S,T
DDCDx	T,S	P,P	S,T	P,P
DCDDz**	T,S	R,R	T,S	P,P
CDDDy	R,R	P,P	T,S	P,P
DDDDy	T,S	P,P	T,S	P,P

Figure 7. *PD revisited: second-order metagame.*

lone player's four strategies. This third-order game is depicted in Figure 7 (in which *C* means cooperate; *D*, defect). As the matrix shows, the lone player's better strategies are A_2 and A_4, because both avoid the worst payoff (*S*) and enable the better payoffs (*R* and *T* respectively). At first blush, A_4 is preferable.

Similarly, the *n* other players first eliminate all 12 meta-strategies tagged "*x*", which might confer the worst payoff (*S*) on them. Then, knowing that the lone player's better strategies are A_2 and A_4, they eliminate the two meta-strategies tagged "*y*", which both confer the second-worst payoff (*P*). By default, there remain the two meta-strategies tagged "*z*", of which the second is weakly dominant. Thus the *n* other players (or a majority of them) choose meta-strategy *DCDD*.

Knowing this, the lone player chooses strategy A_2, which now dominates A_4. That is, the lone player chooses as he expects the majority to choose, and cooperates. The Pareto-optimal outcome obtains at [*DCDD*, A_4]. Moreover, Howard allegedly proves that there is no re-emergence of the Nash equilibrium in any higher-order decision space.

I applied Howard's meta-game methodology to Newcomb's problem and, initially under the spell of Brams and Lewis, was predisposed to suspect that Howard's results would be replicated. The first-order meta-game for Newcomb's problem confirms this suspicion, as it leads to a similar impasse. Let the player have four corresponding strategies: A_1, choose one box unconditionally; A_2, choose as demon predicts; A_3, choose not as demon predicts; and A_4, choose both boxes unconditionally. Figure 8 depicts a situation of strong meta-dominance, in which the player prefers that the demon predict that the player choose one box. But on the evidence of successful prediction, the player must intend to choose one box in order to exercise this meta-preference. At the same time, strategy A_4 (choosing both boxes unconditionally) weakly

		Player			
		A_1	A_2	A_3	A_4
	predicts one box	\$M	\$M	\$M+\$T	\$M+\$T
Demon					
	predicts both boxes	\$0	\$T	\$0	\$T

Figure 8. *Newcomb's problem revisited: first-order metagame.*

dominates the other strategies, and causalists (e.g. Sorensen, 1988) claim that an intention to take one box entails a decision to take both. Thus, as in the second-order PD, the Nash equilibrium persists.

Foolishly disregarding the ghost of Hume, I immediately succumbed to the inductive fallacy and became convinced that the third-order game would similarly yield nothing new. But the conviction proved utterly unfounded. Although the expanded matrix of the third-order Newcomb problem resembles that of the PD so closely that one might be seduced into pronouncing them "structurally identical" on the basis of inspection alone, meta-strategic analysis of the expanded Newcomb matrix follows an unequivocal but functionally different route.

Consider Figure 9 (in which *O* means the opaque box; *B*, both). As in the PD, the player might decide that A_2 or A_4 is best. However, the player might reason further that some meta-strategies are not in the demon's interest. Specifically, the eight meta-strategies tagged "x" reward the player just in case the demon will have predicted incorrectly, so they are eliminable. (The demon's interest in correct prediction arguably entails a reluctance to reward players who falsify its predictions.) Of the remaining options, the six meta-strategies tagged "*y*" result in correct meta-prediction frequencies of only 1/4 or 2/4, and so are likewise eliminable. The two remaining meta-strategies tagged "*z*", namely *OOBB* and *OBBB*, are such that both they do not reward the player just in case the demon will have predicted incorrectly, and they confer the highest possible correct meta-prediction frequency, namely 3/4.

Given the demon's two viable meta-strategies by default, the player observes that strategy A_1 weakly dominates the other strategies, resulting in a payoff of \$M. So the player will choose one box unconditionally. As the demon need not strain its predictive power to arrive at this conclusion, it places \$M in the opaque box. This outcome is Pareto-optimal, in that it gives the player a desirable payoff and affords the demon a correct prediction.

But the two equivalent choices that generate this outcome, [*OOBB*, A_1] and [*OBBB*, A_1], are functionally distinct from the sole choice that generated the Pareto-optimal outcome in the corresponding PD. So the PD and Newcomb's problem are functionally disanalogous, at least in third-order decision space. Yet functional disanalogy implies structural non-identity, wherever it may be found.

	A_1*	A_2	A_3	A_4
OOOOx	$M	$M	$M+$T	$M+$T
OOOBx	$M	$M	$M+$T	$T
OOBOy	$M	$M	$0	$M+$T
OBOOx	$M	$T	$M+$T	$M+$T
BOOOx	$0	$M	$M+$T	$M+$T
OOBBz*	$M	$M	$0	$T
OBOBx	$M	$T	$M+$T	$T
BOOBx	$0	$M	$M+$T	$T
OBBOy	$M	$T	$0	$M+$T
BOBOy	$0	$M	$0	$M+$T
BBOOx	$0	$T	$M+$T	$M+$T
BBBOy	$0	$T	$0	$M+$T
BBOBx	$0	$T	$M+$T	$T
BOBBy	$0	$M	$0	$T
OBBBz*	$M	$T	$0	$T
BBBBy	$0	$T	$0	$T

Figure 9. *Newcomb's problem revisited: second-order metagame.*

The second argument speaks directly to first-order structural dissimilarities, instead of merely implying them. We re-iterate that each player in the PD seeks to attain one of the two better outcomes, and to avoid both of the worse ones. The same can be said of each (human) player in Newcomb's problem. This much follows from the identical transitive ordering of the respective payoff structures. Identically-ordered payoff structures, however, do not imply identical games. Von Neumann & Morgenstern (1944, p. 49) define a game as the totality of its rules, and one can here instantiate the sagacity of that definition, which smacks of elegance rather than oversimplification.

Recall the demon's sole abiding interest: making correct predictions. Now I ask whether any player in either game maintains a similar interest. The player in Newcomb's problem apparently cannot do so, for the rules stipulate that this player is not a predictor; rather, a predictee. If we allow the player to have an interest in making a correct prediction about what the demon will predict concerning that player's future choice, then the demon's prediction becomes a retrodiction concerning the player's prediction about the demon's prediction about the player's choice, which in turn obliges the player's prediction to concern itself with the demon's future retrodiction about the player's prediction about the demon's prediction about the player's choice. To avoid an infinite regress, we must deny that the player in Newcomb's problem has any interest in making correct predictions—or else he will never be able to make a prediction at all.

Do the players in a PD have any such interest? Here it turns out that they just might, but if they do, it is not the same kind of interest as the demon's. *Contra* Irvine's denial that past observed frequency is an infallible guide to probability stands the frequentist position, championed by von Mises (1928), who asserts that probability has no meaning other than observed limiting frequency. (Ineluctable fallibilities of von Mises' redux are somewhat mollified by margins of statistical confidence). But we need not be drawn by the hoary debate between Bayesians and frequentists, at least not in this context. The classic PD treated herein rules out the frequentist interpretation altogether: it is a non-repeated (or "one-shot") affair, in which the players have no recourse to any record of prior play. Their game has no history, and so has no observed relative frequency of outcomes, and thus has no limiting frequency at all. For that matter, neither do we require an observed frequency to run Newcomb's problem, at least in its early stages. The first player in any Newcomb sequence has no statistics on the demon's performance—and indeed the first few players have very poor ones—but still they may harbor subjective degrees of belief, arbitrarily close to certainty, about the demon's predictive power.

A rational player in a PD seems committed either to the dominance principle, or else to some personalist assessment of expected utilities. The latter entails holding a subjective degree of belief—in short, making a prediction—about what the other player will do. However, we can show that both disjuncts lead to a denial that the PD is a Newcomb problem.

If either player is committed to the dominance principle, then by definition that player defects regardless of the other player's choice. If both players are so committed, then neither player has any interest in predicting what the other will do, in which case the PD is not a Newcomb problem.

If either player is committed to maximizing expected utilities, then that player must posit a likelihood that the other will cooperate conditional on one's own cooperation. With respect to the functional calculus, this amounts to selecting some real number from the continuous closed interval [0,1], which stands for the conditional probability that the other will cooperate if one does. Note that this kind of prediction is quite different from the demon's. The demon does not select some real number from that continuum to stand for the probability that the player will choose one box conditional on the demon's placement of 1 million dollars in it; rather, the demon assigns a truth-value, a Boolean probability of zero or unity, to the proposition that the player will choose one box (and assigns its complement to the proposition that the player will choose both boxes). The rules stipulate clearly that the demon's placement (or non-placement) of the 1 million is contingent on the demon's prediction; thus the demon's prediction cannot be conditional on its placement (or non-placement) of the 1 million. Since the two kinds of prediction are not equivalent, the PD is not a Newcomb problem.

If one endeavors to finesse this argument by reducing conditional probabilization in the PD to pure prediction, then one is trumped either by dominance or else by absence of compulsion. Suppose that either player in the PD maintains an interest in predicting unconditionally what the other will choose. If you predict that the other will defect, then you must defect yourself to avoid incurring the worst payoff. But if you predict that the other will cooperate, then you must choose whether to cooperate or defect yourself. You incur the best payoff by defecting; but if you defect, then your interest in predicting the other player's choice was entirely vacuous, since you defected regardless of his choice. But if you cooperate, you must account for not having defected, because you are voluntarily relinquishing a larger payoff for a smaller. While such behavior in a one-shot game is apparently irrational, its rationale can be salvaged by the introduction of some auxiliary mitigating principle of choice. For instance, if convinced that the other player will cooperate in a PD, you could readily invoke Hobbesian, Kantian, or utilitarian arguments for cooperating yourself. But there is no rule in the PD which compels your cooperation just in case you predict that the other player will cooperate. However, there is a rule in Newcomb's problem which compels the demon's placement of 1 million dollars in the opaque box, just in case the demon predicts that the player will choose that box alone. Since the functional consequences of a player's prediction in a PD are not equivalent to those of the demon's prediction in a Newcomb problem, the two problems are not structurally identical.

In *fine*, whether or not either player in a PD has interests in predicting the other player's choice, the PD is not a Newcomb problem.

Arguing that the PD is really two Newcomb problems side-by-side is analogous to arguing that the human female chromosome is really two male chromosomes side-by-side. Since the males are both *XY*, if we remove *X* from one and *X* from the other and recombine them, we obtain an *XX*. But it does not follow that the female is structurally or functionally identical to the male. Similarly, if we remove the human player from one Newcomb problem and the human player from another Newcomb problem, and recombine them in a new game with the same transitive payoff ordering, but with different rules, we obtain a PD. But it does not follow that the PD is structurally or functionally identical to a Newcomb problem. That they possess similar payoff "structures" yields an equivocation, not an identity.

(And what of the two residual demons? Like a *YY* chromosome, which presumably does not constitute a viable biological entity, two demons could not engage in a playable game. Neither demon would be able to move, since each must first predict what the other will do.)

Thus, the third link is severed and the chain is sundered. I have demonstrated that Braess' paradox is not structurally identical to the Cohen-Kelly queuing problem, which in turn is not structurally identical to the *N*-player PD, which in turn is not structurally identical to Newcomb's problem. Thus Braess' paradox solves Newcomb's problem: Not!

3. A stronger claim, and a neo-Augustinian entreaty

In this final section, I should like to show something stronger: not merely *how* Braess' paradox fails to solve Newcomb's problem (via some linkage), but *that* Braess' paradox fails to solve Newcomb's problem (via any linkage).

The mechanical analogue of Braess' paradox has a deterministic albeit infinite solution set of static equilibria; that is, an infinite set of steady-state solutions. For any (real) number of units of mass suspended from the system within its carrying capacity, the resultant tensions and extensions are predictable from systemic equations and constants. If Newcomb's problem does not admit of a deterministic steady-state solution set, such that for any number of players the demon's predictions and the players' subsequent choices are similarly predictable from systemic equations and constants, then Braess' paradox cannot in any sense solve Newcomb's problem. Since there is no means extant by which we can predict the demon's predictions, let alone the players' subsequent choices, it appears that this stronger claim is substantiated.

Even if we modify Newcomb's problem, with a view to contriving a steady-state solution, we cannot guarantee its prolongation. Given the demon's interest in making correct predictions, further suppose that each player knows not only the demon's relative frequency of correctness, but also the precise distribution of prior outcomes. In that case, it appears that the demon could impose a steady-state solution by repeatedly placing nothing in the opaque box. After a negligible number of transient outcomes, during which a few one-boxers would be sorely disappointed, all subsequent players would surely choose two boxes, not because they are convinced that the demon has predicted their choices, but because they are convinced that the demon has placed

		Demon	
		places $M in opaque box	does not place $M in opaque box
Players	choose opaque box	0	10
	choose both boxes	0	990

Figure 10. *Newcomb's problem: steady-state outcome matrix.*

nothing in the opaque box. The situation is depicted in Figure 10. Given the demon's high "predictive" success rate, maximizing expected utilities misleadingly prescribes choosing one box only. Dominance (as it were, on the half-shell) prescribes choosing both. But common sense dictates that the demon is predicting nothing; rather, that it is pseudo-deterministically compelling a sequence of identical outcomes. You or I, presumably lacking predictive powers, could play the demon's role in such a game.

This pseudo-determinism, however, can be decisively undermined. For while the demon must sustain an interest in making correct predictions, the players need not abet the demon's cause. If significant numbers of players were predisposed to martyrdom, and began to choose one box instead of both, then the demon's "predictive" success rate would be compromised, and it would necessarily abandon its attempt to compel a steady-state solution.

Moreover, the payoff structure of Newcomb's problem precludes a pseudo-deterministic compulsion of the complementary pure sequence; that is, precludes a steady-state solution in which the demon repeatedly places 1 million dollars in box one, and the players choose only that box. For were the demon to do so, and were each player to inspect the associated matrix of previous outcomes, it seems plausible that significant numbers of players would choose both boxes. But this in turn would compromise the demon's "predictive" success rate, and thus the pseudo-deterministic steady-state would collapse yet again.

In conclusion, I have not claimed that Newcomb's problem cannot be solved; only that Braess' paradox does not and cannot solve it. My modest but avowed task is accomplished. Now let the more ambitious, who would ponder how Newcomb's problem can be solved, consider this caveat: an unequivocal solution has not been found in the exclusive disjunction of one box or two. A rationally satisfactory but morally unsavory solution may perhaps be found in a neo-Augustinian entreaty; not to the demon, but to the uncognized source of those unknown laws from which the demon's formidable foresight flows. I offer you Augustine's putative game-theoretic supplication: "Make me a two-boxer, but not yet!"

Acknowledgements

I am grateful to Joel Cohen, Hartry Field, Dan Greenberger and Andrew Irvine for their valuable comments and insights. Any technical errors are mine alone. This paper was first read in November 1995, at a colloquium in the Philosophy Department at the Graduate Center of the City University of New York.

References

AXELROD, R. (1980a) Effective choice in the prisoner's dilemma, *Journal of Conflict Resolution*, 24, pp. 3–25.

AXELROD, R. (1980b) More effective choice in the prisoner's dilemma, *Journal of Conflict Resolution*, 24, pp. 379–403.

BRAESS, D. (1968) Über ein paradoxon aus der verkehrsplanung, *Unternehmensforchung*, 12, pp. 258–268.

BRAMS, S. (1975) Newcomb's problem and prisoner's dilemma, *Journal of Conflict Resolution*, 19, pp. 596–612.

COHEN, J. & KELLY, F. (1990) A paradox of congestion in a queuing network, *Journal of Applied Probability*, 27, pp. 730–734.

COHEN, J. & HOROWITZ, P. (1991) Paradoxical behaviour of mechanical and electrical networks, *Nature*, 352, pp. 699–701.

EELLS, E. (1982) *Rational Decision and Causality* (New York, Cambridge University Press).

GIBBARD, A. & HARPER, W. (1978) Counterfactuals and two kinds of expected utility, in: C. HOOKER, J. LEACH & E. MCCLENNAN (Eds) *Foundations and Applications of Decision Theory, Volume 1* (Dordrecht, Reidel).

GLANCE, N. & HUBERMAN, B. (1994) The dynamics of social dilemmas, *Scientific American*, March, pp. 76–81.

HARDIN, G. (1973) The tragedy of the commons, in: A. BAER (Ed.) *Heredity and Society* (New York, The MacMillan Company).

HOBBES, T. (1651) *Leviathan* (Oxford, Basil Blackwell, 1957).

HORWICH, P. (1987) *Asymmetries in Time* (Cambridge, The MIT Press).

HUME, D. (1729) *A Treatise of Human Nature* (Oxford, Clarendon Press, 1958).

IRVINE, A. (1993) How Braess' paradox solves Newcomb's problem, *International Studies in the Philosophy of Science*, 7, pp. 141–160.

JEFFREY, R. (1983) *The Logic of Decision*, 2nd edition (Chicago, The University of Chicago Press).

KELLY, F. (1991) Network routing, *Philosophical Transactions of the Royal Society*, Series A, 337, pp. 343–367.

LEWIS, D. (1976) Probabilities of conditionals and conditional probabilities, *Philosophical Review*, 85, pp. 297–315.

LEWIS, D. (1979) Prisoner's dilemma is a Newcomb problem, *Philosophy and Public Affairs*, 8, pp. 235–240.

LEVI, I. (1974) Newcomb's many problems, *Theory and Decision*, 6, pp. 161–175.

MARINOFF, L. (1992) Maximizing expected utilities in the prisoner's dilemma, *Journal of Conflict Resolution*, 36, pp. 183–216.

MARINOFF, L. (1994) A resolution of Bertrand's paradox, *Philosophy of Science*, 61, pp. 1–24.

MARINOFF, L. (1996a) The Cohen-Kelly queuing paradox revisited, under review by *Transportation Science*.

MARINOFF, L. (1996b) The failure of success: intrafamilial exploitation in the prisoner's dilemma, in: P. DANIELSON (Ed.) *Modeling Rational and Moral Agents, Vancouver Cognitive Science Series* (Oxford, Oxford University Press), forthcoming.

MARINOFF, L. (1996c) Effects of uncertainty in the repeated prisoner's dilemma, under review by the *Journal of Mathematical Sociology*.

NOZICK, R. (1969) Newcomb's problem and two principles of choice, in N. RESCHER, (Ed.) *Essays in Honor of Carl Hempel* (Dordrecht, Reidel).

PETTIT, P. (1986) Free riding and foul dealing, *Journal of Philosophy*, 83, pp. 361–379.

PRICE, H. (1986) Against causal decision theory, *Synthese*, 67, pp. 195–212.

PRICE, H. (1991) Agency and probabilistic causality, *The British Journal for the Philosophy of Science*, 42, pp. 157–176.

RAPOPORT, A. (1967) Escape from paradox, *Scientific American*, July, pp. 50–56.

SAVITT, S. (1994) The replacement of time, *The Australasian Journal of Philosophy*, 72, pp. 463–474.

SOBEL, J. (1985) Not every prisoner's dilemma is a Newcomb problem, in: R. CAMPBELL, L. SOWDEN (Eds) *Paradoxes of Rationality and Cooperation* (Vancouver, The University of British Columbia Press).

SORENSEN, R. (1988) *Blindspots* (Oxford, Clarendon Press).

STEINBERG, R. & ZANGWILL, W. (1983) The prevalence of Braess' paradox, *Transportation Science*, 17, pp. 301–318.

VON MISES, R. (1928) *Probability, Statistics and Truth* (New York, Dover Publications Inc., 1981).

VON NEUMANN, J. & MORGENSTERN, O. (1944) *Theory of Games and Economic Behavior* (New York, John Wiley & Sons Inc.).

Note on contributor

After taking a first degree in Mathematical Physics in his home-town of Montreal, Louis Marinoff earned a PhD in the Philosophy of Science at University College London. Following a post-doctoral year at the Hebrew University of Jerusalem, he spent three years at the University of British Columbia, where he was Executive Moderator of the Canadian Applied Ethics Research Networks and a Lecturer in Philosophy. He has been Assistant Professor of Philosophy at the City College of the City University of New York since 1994. Louis Marinoff's research interests include decision theory, computer modeling of rational and moral agents, and philosophical practice. *Correspondence:* Louis Marinoff, Department of Philosophy, The City College of New York, 137th St. at Convent Ave., New York, NY10031. Fax: (212) 650-7649. Email: marinoff.@cnct.com