The Evolution of Conventions

H. Peyton Young

SFI WORKING PAPER: 1991-08-043

THE EVOLUTION OF CONVENTIONS

H. Peyton Young

School of Public Affairs/
Department of Economics
University of Maryland

Fourth Draft: June, 1991

Abstract

Consider an n-person game that is played repeatedly, but by different agents. In each period, n players are drawn at random from a large finite population. Each player chooses an optimal strategy based on a sample of information about what other players have done in the past. The sampling defines a stochastic process that, for a large class of games that includes coordination games and common interest games, converges almost surely to a pure strategy Nash equilibrium. Such an equilibrium can be interpreted as the "conventional" way of playing the game. If, in addition, the players sometimes experiment or make mistakes, then society occasionally switches from one convention to another. In this case some conventions (i.e., equilibria) are *a priori* more probable than others. Moreover, as the likelihood of mistakes goes to zero, only some of the equilibria have positive probability in the limit. We show how to compute these *stochastically stable equilibria* using the theory of perturbed Markov processes.

Acknowledgments

This research is part of the Economics Research Program at the Santa Fe Institute which is funded by grants from Citicorp/Citibank and the Russell Sage Foundation and by grants to SFI from the John D. and Catherine T. MacArthur Foundation, the National Science Foundation (PHY-8714918), the U.S. Department of Energy (ER-FG05-88ER25054), and the Alex C. Walker Educational and Charitable Foundation—Pittsburgh National Bank. I am indebted to Dean Foster, David Canning, Michael Cohen, David Lane, and George Mailath for stimulating conversations on this topic. The usual caveat applies.

THE EVOLUTION OF CONVENTIONS

H. Peyton Young

The individual is foolish but the species is wise.
Edmund Burke

A convention is a pattern of behavior that is customary, expected, and self-enforcing. Everyone conforms, everyone expects others to conform, and everyone wants to conform given that everyone else conforms [Lewis, 1967]. Familiar examples include driving on the right when others drive on the right, going to lunch at noon if others go at noon, accepting dollar bills in payment for goods if others accept them, and so forth. Conventions need not be symmetric. Men conventionally propose to women. Sailboats on the port tack yield the right-of-way to sailboats on the starboard tack. In some regions, tenant farmers customarily get one-third of the harvest and landlords get two-thirds, whereas in other regions the reverse convention holds [Bardhan, 1984]. For each role in such asymmetric interactions there is a customary and expected behavior, and everyone prefers to follow the behavior expected of him provided that others follow the behavior expected of them. Under these circumstances we say that people follow a convention. A convention is an equilibrium that everyone expects. But how do expectations become established when there is more than one equilibrium?

One explanation is that some equilibria are *a priori* more reasonable than others. A deductive theory of this type has been proposed by Harsanyi and Selten (1988). A second explanation is that agents focus their attention on one equilibrium because it is more prominent or conspicuous than the others (Schelling, 1960). Yet a third explanation is that, over time, expectations converge on one equilibrium through positive feedback effects. Suppose that a game is played repeatedly, either by the same or different agents. Past plays have a feedback effect on the expectations and behaviors of those playing the game now because people pay attention to precedent. Eventually, one equilibrium becomes entrenched as the conventional one, not because it is inherently prominent or focal, but because the dynamics of the process happen to select it.

This evolutionary explanation for the origin of conventions has been suggested in a variety of papers (Lewis, 1967; Sugden, 1986; Warneryd, 1991], but the precise dynamics of the process by which expectations and behaviors evolve has not been clearly spelled out. In particular it is not clear whether it works. Does the process converge to an equilibrium, and if so, are all equilibria equally likely to be selected? We shall show that the process does converge provided that the underlying game has an acyclic best reply structure, and provided there is sufficient stochastic variability in the players' respsones. In this case, society is at or close to a Nash equilibrium most of the time. Not all Nash equilibria are equally likely to be selected, however. In fact, typically only *one* Nash equilibrium will be observed with high probability in the long run. We shall show how to compute these equilibria using the theory of perturbed Markov processes.

2. Outline of the model

We consider a fixed n-person game that is played once each period. The players are drawn at random from a large, finite population of individuals. Each player chooses an optimal strategy based on his beliefs about his environment, which he takes to be stationary. He forms his beliefs by looking at what other agents have done in the recent past. Since gathering information is costly, however, each player knows only a small portion of the history, that is, he bases his current actions on a *sample* of plays from recent time periods. We shall also assume that the players occasionally experiment with different strategies, or simply make mistakes.

The strategies that the agents choose in the current period are recorded and the game is played again in the next period by another random draw of n agents from the fixed population. Each of these agents takes a random sample of previous plays and reacts accordingly. Actions in earlier periods therefore have a feedback effect on actions by agents in later periods. Given that the population is large, however, it is unlikely that the same agents will meet again, or that the action of any one individual will have a substantial effect on the evolution of the process. So the agents are more or less justified in ignoring the feedback effects of their own actions on future plays of the game.

The adaptive dynamics described above define a Markov chain whose states are the histories of play truncated to a finite number of periods. It is similar to fictitious play in that agents choose best replies to other agents' past actions. In fictitious play, however, agents base their decisions on the entire history of actions by other agents. Here we assume that agents base their decisions on *limited* information about actions of other agents in the *recent* past, and they do not always optimize. These assumptions seem less fictitious than fictitious play, hence we call this process *adaptive play*.

For general n-person games, adaptive play need not converge to a Nash equilibrium, either pure or mixed, as we shall show below by example. Nevertheless, there is an important class of games for which it does converge. These games have the property that, from any initial choice of strategies, there exists a sequence of best replies that leads to a strict, pure strategy Nash equilibrium. This class includes, but is substantially more general than, coordination games and common interest games. For these *weakly acyclic* games, adaptive play converges with probability one to a pure strategy Nash equilibrium provided that the samples are sufficiently incomplete and the players never make mistakes. Incompleteness is essential because it creates enough stochastic variability to prevent the process from becoming stuck in suboptimal cycles. Finite memory is essential because it allows past miscoordinations to be forgotten eventually. Once a given equilibrium has been played for as long as anyone can remember, then this equilibrium becomes entrenched as the "conventional" way of playing the game. It is an absorbing state of the process. One cannot say in advance, of course, which equilibrium will become the conventional one, since this depends on the vagaries of the process and on the initial state. What can be said is that some equilibrium will eventually be selected with probability one, and it will *not* be a mixed strategy equilibrium.

If the players occasionally experiment or make mistakes, however, then more can be said. In this case the process has no absorbing states; rather, it has a stationary distribution that describes the relative frequency with which different states are observed in the long run. We shall show that, if the probability of mistakes is small, then this stationary distribution is concentrated around a particular subset of pure strategy Nash equilibria. In fact, typically it puts almost all the weight on exactly one equilibrium. This *stochastically stable equilibrium* will be observed with probability close to one when the noise is very small (Foster and Young, 1990). This concept differs in an important respect from other notions of equilibrium stability (such as evolutionarily stable strategies), which are based on the idea that the equilibrium should be restored if it is subjected to a small, one-time shock.[1] In our model, stochastic flucuations due to sampling error and mistakes are not one-shot affairs, but form an integral part of the dynamics.

Several other recent papers deal with similar topics. Kandori, Mailath, and Rob (1991) consider an evolutionary learning process defined on 2x2 matrix games in which every player plays every other player in every period, and the strategy choices constitute the state in the next period. Successful strategies are adopted with higher probability than unsuccessful ones, and there is a small probability that players make mistakes. They show that this stochastic process selects the risk dominant Nash equilibrium when the mistake probability is small. Canning (1991) examines a more general class of learning models in which agents adapt their behavior to the current state and occasionally make mistakes. He shows that, under certain regularity conditions, the stationary distribution of the perturbed process converges to a stationary distribution of the unperturbed one. Canning's results are

[1] Models of equilibrium selection based on this concept include Axelrod (1984), Fudenberg and Maskin (1990), Samuelson and Zhang (1990), Crawford (1991), and Samuelson (1991). For other models of evolutionary dynamics in games see Samuelson (1988), Nachbar (1990), and Friedman (1991).

quite general, but they do not tell us what the limiting stationary distribution looks like.[2] In this paper we show specifically how to compute the limiting stationary distribution of such processes by solving a series of shortest path problems. We shall then apply this result to compute the stochastically stable equilibria of adaptive play. For 2x2 matrix games the risk dominant equilibrium is the unique stochastically stable equilibrium. In games with more strategies, the two concepts differ.

3. Adaptive play

Let G be an n-person game in normal form, and let S_i be the finite set of strategies available to player i. Let N be a finite population of individuals that is partitioned into n nonempty classes C_1, C_2, $\ldots$, C_n. Each member of C_i is a candidate to play role i in the game. For example, C_1 is the class of men, C_2 is the class of women, and the game is Battle of the Sexes. We shall assume that all individuals in class i have the same utility function $u_i(s)$ for strategy-tuples $s = (s_1, s_2, \ldots, s_n) \in \prod S_i$, which we shall identify with outcomes.

Let t = 1, 2, ... denote successive time periods. The game G is played once each period. In period t, one individual is drawn at random from each of the n classes and is assigned to play the appropriate role in the game. It will be convenient to refer to the individual playing role i as player "i" even though the identity of this individual may change from one period to the next. Player i chooses a pure strategy $s_i(t)$ from his strategy space according to a rule that will be defined below. The strategy-tuple $s(t) = (s_1(t), s_2(t), \ldots, s_n(t))$ is recorded and will be referred to as the *play at time* t. The *history* of plays up to time t is the sequence $h(t) = (s(1), s(2), \ldots, s(t))$. We assume that the histories are anonymous: it does not matter *who* played a given strategy in a given period, only that it was played by someone.

The agents decide how to choose their strategies as follows. Fix integers k and m such that $1 \leq k \leq m$. In period $t \geq m$, each player inspects k plays drawn without replacement from periods $t - 1, t - 2, \ldots, t - m$. The draws are independent for the various players. Concretely, we may think of each player "asking around" to find out how the game was played in recent periods. The agent stops when he has obtained information about k different plays within the last m periods. It is not necessary to assume that the agent samples every subset of k plays with equal probability; it is enough that he sample every such subset with positive probability. For notational simplicity, we shall

2 Other models of selection based on stochastic dynamical processes include Fudenberg and Harris (1991), Evans and Honkapohja (1991), and Kirman (1991).

assume that all agents have the same m and k, although this is not essential for the results. We shall also assume, for the sake of generality, that the first m plays are randomly selected. Thus we can think of the sampling process as beginning in period $t = m + 1$ from some arbitrary initial sequence of m plays $h(m) = (s(1), s(2), \ldots, s(m))$.

The above decision rule defines a finite Markov chain on the state space H consisting of all sequences of length m drawn from $\prod S_i$, beginning with some arbitrary "initial" state h(m).

SUCCESSOR. A *successor* of a state $h \in H$ is any state $h' \in H$ obtained by deleting the left-most element of h and adjoining a new right-most element.

The process moves from a state to a successor state in each period according to the following transition rule. Let h be the current state. For each $s \in S_i$, let $p_i(s|h)$ be the probability that agent i chooses s. We assume only that $p_i(s|h) > 0$ if and only if there exists a sample of size k to which s is i's best reply, and that $p_i(s|h)$ is independent of t. Then the probability of moving from h to h' is

$$P_{hh'} = \prod_{i=1,n} p_i(s_i|h) \quad \text{if } h' \text{ is a successor of } h \text{ and } s \text{ is the right-most element of } h' \qquad (1)$$
$$P_{hh'} = 0 \quad \text{otherwise.}$$

We call this process *adaptive play with memory* m *and sample size* k.

4. Convergence of adaptive play when there are no mistakes.

Let us begin by observing that, *if* adaptive play converges to an absorbing state, then that state is a strict pure strategy Nash equilibrium played m times in succession. Suppose, indeed, that $h = (s^1, \ldots, s^m)$ is an absorbing state. For each agent i let s_i be i's best reply to some subset of k plays drawn from h, and let $s = (s_1, \ldots, s_n)$. By assumption, there is a positive probability of moving from h to $h' = (s^2, \ldots, s^m, s)$ in one period. Since h is absorbing, $h = h'$ and hence $s^1 = s^2$. Continuing in this fashion we conclude that $s^1 = s^2 = \ldots = s^m = s$. Hence $h = (s, s, \ldots, s)$. By construction, s_i is a best reply to some sample of k elements from h. Hence s_i is a best reply to s_{-i} for each i. It must also be a unique best reply to s_{-i}, because otherwise the process could move to a successor that is different from h. So s is a strict, pure strategy Nash equilibrium. Conversely, any state h consisting of m repetitions of a strict, pure strategy Nash equilibrium is clearly an absorbing state. Such a state will be called a *convention*.

If adaptive play converges to an absorbing state, then clearly the game must have a strict Nash equilibrium in pure strategies. This is not a sufficient condition for convergence, however. Consider the following variation of an example due to Shapley :

EXAMPLE 1.

	a	b	c	d
a	2, 1	0, 0	1, 2	-1, -1
b	1, 2	2, 1	0, 0	-1, -1
c	0, 0	1, 2	2, 1	-1, -1
d	-1,-1	-1,-1	-1,-1	3, 3

Here d is a best response to itself, but it is not a best response to any mixture of a, b, and c. If the initial state does not involve d, then adaptive play (like fictitious play) cycles. Consider, for example, the case where m = 2 and k = 1. Let the first two plays be (a, a) and (a, c). In period 3, Column will sample one of Row's previous two choices (both a) and react by playing c. Row will sample one of Column's previous two choices (a or c) with equal probability and react by playing a or c. So the next play will be (a, c) or (c, c) with equal probability. The process therefore moves from state [(a, a), (a, c)] to state [(a, c), (a, c)] with probability one-half, or to state [(a, c), (c, c)] with probability one-half. The subsequent transitions form a cycle of length six imbedded within a cycle of length twelve, as shown in Figure 1. This cycle constitutes an irreducible, recurrent class of the Markov process defined by (1).

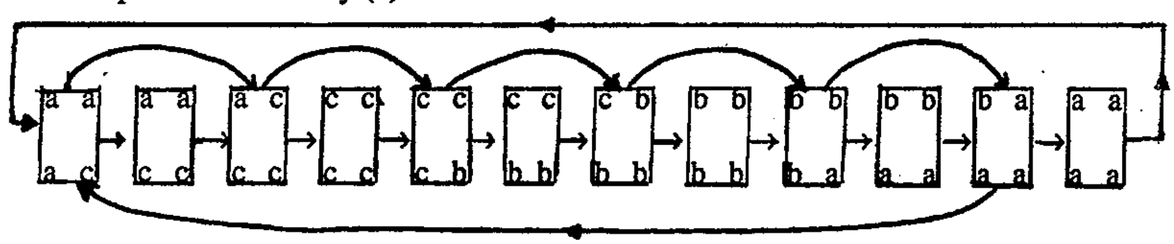

Figure 1. An irreducible class of adaptive play with m = 2, k = 1. The top line shows Row's choices, the bottom line shows Column's choices.

When cycling is built into the best reply structure of the game, as in the above example, we cannot expect adaptive play to converge. Nevertheless, there are many games that do not have a cyclic best-reply structure. Consider a two-person coordination game in which both agents have the same number of strategies, and each agent strictly prefers to play his jth strategy if and only if the other agent plays his jth strategy for every j. Clearly there is no cycling problem here: once one of them chooses a pure strategy and the other responds optimally, then they have achieved a coordination equilibrium.

To take another example, suppose that the agents have *common interests*: for every two strategy tuples s and s', either everyone strictly prefers s to s' or everyone strictly prefers s' to s. Given an arbitrary strategy-tuple s that is not a strict Nash equilibrium, there exists some agent i who can do better by playing s_i' instead of s_i. Let $s' = (s_i', s_{-i})$. If s' is not a strict Nash equilibrium, there is some agent j who can do better by playing s_j'' instead of s'_j. Let $s'' = (s_j'', s'_{-j})$, and so forth. At each step of this adjustment process everyone's utility increases, so the process cannot cycle and it must end at a strict, pure strategy Nash equilibrium.

This construction can be generalized as follows. Let G be an n-person game in normal form on a finite strategy space $\prod S_i$. Define a directed graph $\Gamma(G)$ such that each vertex of Γ is an n-tuple of strategies $s \in \prod S_i$, and for every two vertices s and s' there is a directed edge $s \rightarrow s'$ if and only if $s \neq s'$ and there exists exactly one agent i such that s_i' is a best reply to s_{-i} and $s'_{-i} = s_{-i}$. $\Gamma(G)$ is the *best reply graph* of G.

ACYCLIC GAME. A game G is *acyclic* if its best reply graph contains no directed cycles. It is *weakly acyclic* if, from any initial vertex s, there exists a directed path to some vertex s* from which there is no exiting edge (a *sink)*.

Every sink of $\Gamma(G)$ is clearly a strict Nash equilibrium in pure strategies. So a game is weakly acyclic if, and only if, from every strategy-tuple there exists a finite sequence of best replies by one agent at a time that ends in a strict, pure strategy Nash equilibrium. We shall show that, for this class of games, adaptive play converges with probability one provided that sampling is sufficiently incomplete and the players do not make mistakes.

Let G be a weakly acyclic n-person game. For each strategy-tuple s, let L(s) be the length of a shortest directed path in $\Gamma(G)$ from s to a strict Nash equilibrium, and let $L_G = \max_s L(s)$.

THEOREM 1. *Let* G *be a weakly acyclic n-person game. If* $m/k \geq L_G + 2$ *then adaptive play converges with probability one to a convention.*

Proof. Fix k and m, where $m/k \geq L_G + 2$. We shall show that there exists a positive integer M, and a positive probability p, such that from any state h, the probability is at least p that adaptive play converges within M periods to a convention. M and p are time-independent and state-independent. Hence the probability of *not* reaching a convention after at least rM periods is at most $(1 - p)^r$, which goes to zero as $r \rightarrow \infty$.

Let $h = (s(t - m + 1), \ldots, s(t))$ be the state in period $t \geq m$. In period $t + 1$ there is a positive probability that each of the n agents samples the last k plays in h, namely, $(s(t - k + 1), \ldots, s(t)) = \eta$. There is also a positive probability that, from periods $t + 1$ to $t + k$ inclusive, every agent draws the sample η every time. Finally, there is a positive probability that, if an agent has a choice of several best replies to η, then he will choose the same one k times in succession. Thus there is a positive probability of a run $(s, s, \ldots, s)$ from periods $t + 1$ to $t + k$ inclusive. Note that this argument depends on the agents' memory being at least $2k - 1$, since otherwise they could not choose the sample η in period $t + k$.

Suppose that s happens to be a strict Nash equilibrium. There is a positive probability that, from periods $t + k + 1$ through $t + m$, each agent will sample only the last k plays, in which case the unique best response of each agent i is s_i. So they play s for $m - k$ more periods. At this point an absorbing state has been reached, and they continue to play s forever.

Suppose instead that s is not a strict Nash equilibrium. Since G is weakly acyclic, there exists a directed path $s, s', \ldots, s^r$ in $\Gamma(G)$ such that s^r is a strict Nash equilibrium. The first edge on this path is $s \rightarrow s'$. Let i be the index such that $s'_{-i} = s_{-i}$ and s_i' is a best reply to s_{-i}. Consider the event in which agent i samples from the run of s established in periods $t + 1$ to $t + k$ and responds by playing s_i', while every agent $j \neq i$ draws the sample $\eta = (s(t - k + 1), \ldots, s(t))$. By assumption, the best response of every agent j to this sample is s_j. These events occur together with positive probability, and there is a positive probability that they occur in every period from $t + k + 1$ to $t + 2k$, assuming that $m \geq 3k - 1$. The result is a run of $s' = (s'_i, s_{-i})$ for k periods in succession.

Continuing in this fashion, we see that there is a positive probability of obtaining a run of s, followed by a run of s' . . . followed eventually by a run of s^r. Each run is of length k, and the run of s^r occurs from period $t + kr + 1$ to $t + kr + k$. To reach this point may require that some agent look back $kr + 2k - 1$ periods, namely, from period $t + kr + k$ to period $t - k + 1$. This is possible because of the assumption that $m/k \geq L_G + 2$.

After this, the process can converge to the absorbing state $(s^r, s^r, \ldots, s^r)$ by period $t + kr + m$ if each agent samples the previous k plays from periods $t + kr + k + 1$ to $t + kr + k + m$ inclusive.

Thus we have established that, given an initial state h, there is a probability $p_h > 0$ of converging to an absorbing state within $M = kr + k + m$ periods. Letting $p = \min_{h \in H} p_h > 0$, it follows that from any initial state the process converges with probability at least p to an absorbing state within at most M periods. This completes the proof.

We do not claim that the lower bound $L_G + 2$ is best possible for all weakly acyclic games, but incomplete sampling is a necessary condition for this result. Consider the following version of "Battle of the Sexes":

EXAMPLE 2.

	Yield	Not Yield
Yield	0, 0	1, $\sqrt{2}$
Not Yield	$\sqrt{2}$, 1	0, 0

Let k = m, so that both players sample the same m plays in each period. Consider any initial sequence of m plays in which the players have always miscoordinated, that is, they both yielded or they both failed to yield in each period. Let f be the frequency with which they yielded in this sequence. In the next period, Row yields if and only if $1 - f > f\sqrt{2}$, and Column does the same. So they miscoordinate again. (Since f is rational, the inequality is always strict so we never have to consider ties.) Thus, if they begin in a state of perfect miscoordination, then they miscoordinate forever. The same holds if memory is unbounded, as in fictitious play: if they miscoordinate on the first move, then they will continue to miscoordinate forever.

The virtue of incomplete sampling is that it introduces stochastic variation into the players' responses. They may coordinate by chance, and if they do so often enough the process eventually locks in to a pure strategy Nash equilibrium. This equilibrium then becomes the conventional way of playing the game, because for as long as anyone can remember, the game has always been played in this way. Therefore sampling does not matter any more, because no matter what samples the agents take, their optimal response will be to play the equilibrium that is already in place.

5. Adaptive play with mistakes

Theorem 1 relies on the assumption that, while agents base their decisions on limited information, they always choose a best response given their information. This assumption is clearly unrealistic. Agents sometimes make mistakes; they may also experiment with nonoptimal responses. In this case the stochastic process does not converge to an absorbing state, because it has no absorbing states. Mistakes constantly perturb the process away from equilibrium. If we assume, however, that all mistakes are possible and that the mistake probabilities are time-independent, then the process does have a unique stationary distribution. Hence we can study its asymptotic behavior. When the probability of mistakes is small, we shall show that this stationary distribution is concentrated around a particular convention (or, in the event of ties, a subset of conventions). These may be interpreted as the *stochastically stable* conventions, that is, the ones that will be observed with positive probability in the long run when the noise is small but nonvanishing.

Fix the sample size k and memory m. Suppose that, in each time period, there is some small probability $\varepsilon\lambda_i > 0$ that player i experiments by choosing a strategy randomly from S_i instead of optimizing based on a sample of size k. The ratio λ_i/λ_j is the relative probability with which a player of type i experiments as compared to a player of type j. The factor ε determines the probability with which players in general experiment. The event that i experiments is assumed to be independent of the event that j experiments for every $i \neq j$. For every i, let $q_i(s|h)$ be the conditional probability that i chooses $s \in S_i$ given that i experiments and the process is in state h, where $\sum_{s\in S_i} q_i(s|h) = 1$ for every i and h. We shall assume that $q_i(s|h)$ is independent of t, and that $q_i(s|h) > 0$ for all $s \in S_i$. The latter assumption is made for ease of exposition; as we shall see later, it suffices to assume that the $q_i(.)$ have enough positive support that every state is reachable from every other state in a finite number of periods by agents who experiment.

A priori we do not know the distributions $\mathbf{q} = (q_1(.), q_2(.), \ldots, q_n(.))$ or the relative probabilities of experimentation $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$. It turns out, however, that this does not matter. If the overall probability of experimentation ε is small, and if the agents experiment independently of one another, then the selected equilibria are independent of $\mathbf{q}$ and λ.

The perturbed process may be described as follows. Suppose that the process is in state h at time t. Let $J \subseteq N$ be a subset of j players, $1 \leq j \leq n$. The probability is $\varepsilon^j (\prod_{j\in J} \lambda_j) (\prod_{j\notin J} (1 - \varepsilon\lambda_j)$ that exactly the players in J experiment and the others do not. Conditional on this event, the transition probability of moving from h to h' is

$Q^J_{hh'} = \prod_{j\in J} q_i(s_i|h) \prod_{j\notin J} p_i(s_i|h)$ if h' is a successor of h and s is the right-most element of h'

$Q^J_{hh'} = 0$ otherwise.

If no agent experiments, then the transition probability of moving from h to h' in one period is $P_{hh'}$ as defined in (1). This event has probability $\prod_{i=1,n}(1 - \varepsilon\lambda_i)$. The perturbed Markov process therefore has the transition function:

$$P^*_{hh'} = (\prod_{1=1,n}(1 - \varepsilon\lambda_i))\, P_{hh'} + (\sum_{J\subseteq N,\, J\neq\emptyset} \varepsilon^{|J|} (\prod_{j\in J} \lambda_j)(\prod_{j\notin J} (1 - \varepsilon\lambda_j))\, Q^J_{hh'} \quad . \tag{2}$$

This process will be called *adaptive play with memory* m, *sample size* k, *and experiment probabilities* $\varepsilon\lambda_i q_i$.

6. Asymptotic behavior of adaptive play.

We shall now characterize the asymptotic behavior of adaptive play when the overall probability of experimenting ε is close to zero. Let us view $P^*(\varepsilon)$ as a function of ε, noting that $P^*(0) = P$. Let h and h' be two distinct states. If $P^*(\varepsilon)$ is in state h at time t, there is a positive probability that all players will experiment for m periods in succession and arrive at state h' at time t + m. Hence $P^*(\varepsilon)$ is irreducible. It is aperiodic because there is also a positive probability that the process could first arrive at state h' at period t + m + 1. Hence $P^*(\varepsilon)$ has a unique stationary distribution $\mu_h(\varepsilon)$, where $\mu_h(\varepsilon)$ is the limiting probability of being in state h after t periods as $t \to \infty$.

STOCHASTIC STABILITY.[3] A state $h\in H$ is *stochastically stable* relative to the process $P^*(\varepsilon)$ if $\lim_{\varepsilon\to 0} \mu_h(\varepsilon) > 0$.

Over the long run, states that are not stochastically stable will be observed infrequently compared to states that are, provided that the probability of mistakes ε is small. If there is a *unique* stochastically stable state, then it will be observed almost all of the time when ε is small.

For a general n-person game, we shall characterize the stochastically stable states of $P^*(\varepsilon)$ and show that they are independent of **q** and λ. If the game is weakly acyclic, then every stochastically stable state is a convention, and typically it is unique.

MISTAKE. Let h' be a successor of h and let s be the right-most element of h'. A *mistake* in the transition $h \to h'$ is a component s_i of s that is not an optimal response by agent i to any sample of size k from h.

A mistake can only arise if a player experiments, but an experimental choice need not be a mistake, since it could (by chance) be an optimal choice.

---

[3] This notion was introduced for general stochastic dynamical systems by Foster and Young (1990), and was applied to an adaptive model for playing 2x2 games by Kandori, Mailath, and Rob (1991).

RESISTANCE. For any two states h, h' the *resistance* r(h, h') is the total number of mistakes in the transition h $\to$ h' if h' is a successor of h; otherwise r(h, h') = ∞.

Let us now view the state space H as the vertices of a directed graph. For every pair of states h, h' insert a directed edge h $\to$ h' if r(h, h') is finite, and let r(h, h') be its "weight" or "resistance." The edges of zero resistance correspond to the transitions that can occur under P*(0). Let H_1, $H_2, \ldots, H_{v^*}$ be the irreducible, recurrent classes of P*(0). These classes have the property that, from every state h there exists a path of zero resistance to some class H_v. No state outside of a class H_v is reachable by a path of zero resistance from any state inside H_v, but within H_v there is a directed path of zero resistance from every state to every other. Given any two distinct classes H_u and H_v, and any two states $h \in H_u$ and $h' \in H_v$, consider all directed paths from h to h'. There is at least one such path, because the perturbed process P(ε) is irreducible. Among all such paths, choose one with *least total resistance*, and let this resistance be denoted by r_{uv}. Computing r_{uv} amounts to solving a shortest path problem in a directed graph. Note that r_{uv} is independent of the particular choice of h and h', because every two states within the same class are accessible from each other by paths of zero resistance.

Now define a directed graph $\mathcal{G}$ on the set of indices V = {1, 2, . . ., v*} as follows: for every u, $v \in V$ there is a directed edge (u, v) from u to v, and r_{uv} is its "weight" or "resistance." The following concept is due to Freidlin and Wentzell (1984).

v-TREE. For every $v \in V$, a *v-tree* on $\mathcal{G}$ is a subset τ of edges with the property that, for every vertex $u \neq v$, there exists exactly one directed path in τ from u to v.

For each $v \in V$, let $\mathcal{T}_v$ be the set of all v-trees on $\mathcal{G}$. The "resistance" of a v-tree τ is the sum of the resistances of its edges,

$$r(\tau) = \sum_{(u, u') \in \tau} r_{uu'} . \tag{3}$$

STOCHASTIC POTENTIAL. The *stochastic potential* of the irreducible class H_v is the least resistance among all v-trees:

$$\rho(v) = \min_{\tau \in \mathcal{T}_v} r(\tau) \tag{4}$$

Computing $\rho(v)$ for a given set of weights $r_{uu'}$ is a standard problem in combinatorial optimization known as the *arborescence problem*. There exist algorithms for solving it in the order of $(v^*)^2$ steps [Chu and Liu, 1965; Edmonds, 1967; Tarjan, 1977].

THEOREM 2. *Let* G *be an n-person game on a finite strategy space, and let* $P^*(\varepsilon) = P^*(k, m, q, \lambda, \varepsilon)$ *be adaptive play with parameters* $m \geq k \geq 1$, $q > 0$, $\lambda > 0$, and $\varepsilon > 0$. *The stochastically stable states of* P^* *are precisely the states contained in the irreducible recurrent classes of* $P^*(0)$ *that have minimum stochastic potential.*

COROLLARY. *If* G *is weakly acyclic and* $m/k \geq L_G + 2$, *then the stochastically stable states of adaptive play are the convention(s) of minimum stochastic potential.*

Theorem 2 follows from a general theorem on perturbed Markov processes that we prove in the Appendix. The corollary is a direct consequence of Theorem 2 and Theorem 1.

Note that the numbers r_{uv}, and hence the potential function ρ, depends only on the *number* of mistakes in making various transitions, not on the relative probability with which specific mistakes are made. Hence the stochastically stable states are independent of the probability distributions $\lambda_i q_i$. This is important, for in applications one would rarely know the relative probabilities of various mistakes, only that they are possible. What matters is that the probability of mistakes is small, and the agents make them independently of one another.

The potential function is computed in three steps. First we identify the irreducible classes of the process $P^*(0)$ without mistakes. In general, these classes can be quite complex, as the discussion in section 4 shows. If the game is weakly acyclic and the sampling is sufficiently incomplete, however, then Theorem 1 tells us that the irreducible classes correspond one-to-one with the strict pure strategy Nash equilibria. In this case the irreducible classes are easy to compute. The second step is to compute the path of least resistance from every irreducible class to every other. This involves solving a series of shortest path problems. The third and final step is to construct a complete directed graph with these resistances as weights, and to find the arborescence having least weight. This identifies the stochastically stable convention(s). In the remainder of the paper we shall illustrate the computations for 2x2 and 3x3 matrix games.

7. The 2x2 case.

Let G be a 2x2 matrix game with two strict Nash equilibria in pure strategies. It is clear that G is acyclic and $L_G = 1$. Without loss of generality we may write G in the form

	1	2
1	a_{11}, b_{11}	a_{12}, b_{12}
2	a_{21}, b_{21}	a_{22}, b_{22}

where $a_{11} > a_{21}$, $b_{11} > b_{12}$, $a_{22} > a_{12}$, and $b_{22} > b_{21}$. The strict, pure strategy Nash equilibria are (1, 1) and (2, 2). Theorem 1 implies that, if $m/k \geq 3$, then adaptive play without mistakes has two absorbing states: $h_1 = ((1,1), (1,1), \ldots, (1,1))$ and $h_2 = ((2,2), (2,2), \ldots, (2,2))$. To determine which of these states is stochastically stable, we must compute the path of least resistance from h_1 to h_2, and the path of least resistance from h_2 to h_1.

Let h_1 be the state at time $t = m$. To go from h_1 to h_2 requires that at least one player choose strategy 2 by mistake. Moreover, he must choose strategy 2 so often that the other's optimal reply is also strategy 2 for at least one sample of size k, for otherwise the process cannot lock in to the absorbing state h_2.

Suppose, for example, that Row chooses 2 by mistake from periods $t = m + 1$ to $t = m + k'$ inclusive, and from then on Row makes no further mistakes. These choices are marked 2* in Table 1.

Period	1	2 . . . m	m+1	m+2 . . .	m+k'	m+k'+1 . . .	m+k'+k	m+k'+k+1 . . .
Row	1	1 . . . 1	2*	2*	2*	1 . . .	1(2)	2
Column	1	1 . . . 1	1	1	1	2	2	2

Table 1. A succession of k' mistakes by Row causes the process to converge to h_2. 2* denotes a mistaken choice of 2, 1(2) an optimal choice of either 1 or 2.

If Column draws a sample that includes these k' choices of 2, as well as k - k' choices of 1, then Column's best reply is 2 provided that

$$(1 - k'/k)b_{12} + (k'/k)b_{22} \geq (1 - k'/k)\, b_{11} + (k'/k)\, b_{21}\,,$$

that is,

$$k' \geq \frac{b_{11} - b_{12}}{b_{11} - b_{12} - b_{21} + b_{22}} k. \qquad (5)$$

If equality holds in (5) then strategy 2 is among Column's best replies, so Column will play it with positive probability.

Suppose that (5) holds and that Column's sample happens to include Row's mistakes in every period from $m + k' + 1$ to $m + k' + k$ inclusive. This event has positive probability if $m \geq k + k' - 1$. Then Column's best reply is to play 2 from periods $m + k' + 1$ to $m + k' + k$ and none of these choices are mistakes. In period $m + k' + k + 1$ suppose that Row samples Column's choices of 2, while Column samples Row's choices of 2. This event has positive probability if $m \geq k' + k$. Then their best replies are to play 2. In the next period there is again a positive probability (if m is large enough) that both sample enough choices of 2 to want to play 2 again, and so forth. So with positive probability the process converges to the absorbing state h_2 with no further mistakes. In other words, k' mistakes is sufficient to move the process from h_1 to h_2 provided that k' satisfies (5) and m/k is large enough.

Similarly, the process converges with positive probability to h_2 if Column chooses 2 by mistake k" times, where

$$k'' \geq \frac{a_{11} - a_{21}}{a_{11} - a_{12} - a_{21} + a_{22}} k.$$

Let

$$R_1 = \min \left\{ \frac{a_{11} - a_{21}}{a_{11} - a_{12} - a_{21} + a_{22}}, \frac{b_{11} - b_{12}}{b_{11} - b_{12} - b_{21} + b_{22}} \right\}.$$

For every real number x, let [x] denote the least integer greater than or equal to x. We have just shown that the resistance in going from h_1 to h_2 is $[R_1 k]$. A similar argument shows that the resistance in going from h_2 to h_1 is $[R_2 k]$ where

$$R_2 = \min \left\{ \frac{a_{22} - a_{12}}{a_{11} - a_{12} - a_{21} + a_{22}}, \frac{b_{22} - b_{21}}{b_{11} - b_{12} - b_{21} + b_{22}} \right\}.$$

If $R_1 \geq R_2$, then (1, 1) *weakly risk dominates* (2, 2) [Harsanyi and Selten, 1988]. If $R_1 > R_2$, then the unique stochastically stable convention is h_1 for all sufficiently large values of k and m/k.

If $R_1 = R_2$, then both h_1 and h_2 are stochastically stable conventions for all sufficiently large values of k and m/k.

THEOREM 3. *Let G be a 2x2 matrix game with two strict Nash equilibria in pure strategies. For all sufficiently large k and m/k, the stochastically stable conventions correspond to the weakly risk dominant Nash equilibria.*

7. The 3x3 case

When the agents have three or more strategies, matters become more complicated. In this case there is no simple formula analogous to risk dominance that identifies the stochastically stable equilibria. First, the path of least resistance must be computed from every equilibrium to every other. Then a minimum arborescence problem must be solved for each equilibrium. We shall illustrate by solving an example.

EXAMPLE 3.

	1	2	3
1	6, 6	0, 5	0, 0
2	5, 0	7, 7	5, 5
3	0, 0	5, 5	8, 8

The pairs (i, i) are the strict, pure strategy Nash equilibria, i = 1, 2, 3. Let h_i denote the convention in which (i, i) is played m times in succession. Theorem 1 says that these are the absorbing states of the unperturbed process provided that $m/k \geq 3$. Let us compute the path of least resistance from every convention to every other for a fixed k and m, assuming always that $m/k \geq 3$.

Suppose that the perturbed process is in state h_1. To exit to h_2 or h_3, one agent must choose a sufficient number of 2's or 3's (or both) to cause the other agent to choose 2 or 3. Since the game is symmetric, it does not matter which player makes the mistakes and which player reacts. Assume that the Column player chooses 2 at least k" = [(1/8)k] times in succession. If Row samples these choices (plus k - k" choices of strategy 1), then Row's best reply is also strategy 2. At this point there is a positive probability that the process will converge to h_2 with no further mistakes. Thus the resistance in going from h_1 to h_2 is [(1/8)k]. Moreover, the least resistant path is *direct* in the sense that it only involves strategies 1 and 2.

Not all paths of least resistance are direct, however. For example, suppose that the process is in state h_3 and we want to exit to state h_2. The direct route is for one player to choose strategy 2 by mistake at least [(3/5)k] times, which causes the other player to reply with strategy 2. But if one player chooses strategy 1 by mistake at least [(3/8)k] times and at most (5/6)k times, then the best reply of the other player is strategy 2. Thus, if k is large enough, the indirect route has lower resistance.

The resistances between every pair of equilibria are shown in Figure 2. For each vertex h_i there are three h_i-trees, and the h_i-tree of least resistance determines the stochastic potential of h_i, as shown in Figure 3.

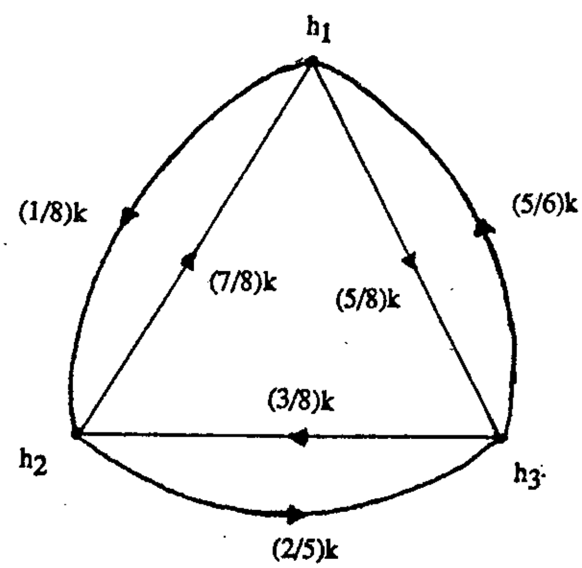

Figure 2. Pairwise resistances for the pure strategy equilibria of Example 3.

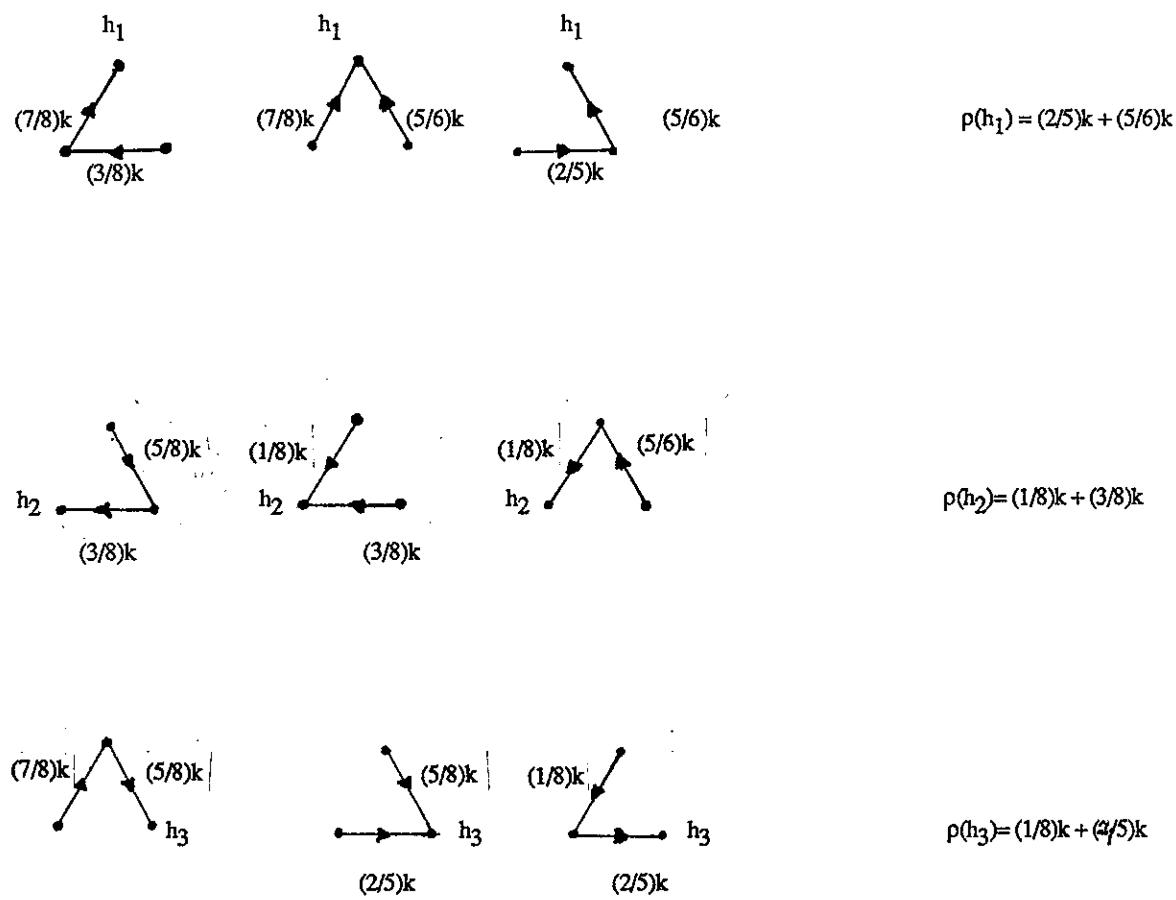

Figure 3. Computation of the potential for each of the three absorbing states h_i.

Thus the stochastically stable convention is uniquely h_2 whenever $k \geq 36$ and $m/k \geq 3$. But the risk dominant equilibrium is (3, 3). The distinction between the two concepts is roughly this: risk dominance selects the equilibrium that is easiest to flow into from every other equilibrium considered in isolation. Stochastic stability selects the equilibrium that is easiest to flow into from all other states combined, including both equilibrium and non-equilibrium states.

9. Conclusion

In this paper we have shown how an equilibrium can evolve in a game that is played repeatedly by different agents. The model is similar to fictitious play in that agents' expectations are shaped by precedent. It differs in that the agents base their choices on an incomplete knowledge of recent precedents and they occasionally make mistakes. These assumptions seem more natural than the deterministic dynamics of fictitious play, so we can justify them on the grounds of realism. They also play an important technical role: by introducing noise into the dynamic adjustment process, they select among pure strategy Nash equilibria for weakly acyclic games, and among more complex regimes for general n-person games. Unlike some other models of equilibrium selection, these perturbations are not one-shot affairs but are an integral part of the evolutionary process. This leads to a different criterion of equilibrium selection than the classical ones such as risk dominance (except in the 2x2 case).

Several questions remain to be explored. One is the sensitivity of the selection mechanism to memory and sample size. For 2x2 games we showed that the stochastically stable equilibria are independent of m and k so long as m/k and k are sufficiently large. It is not clear whether this result holds for weakly acyclic games in general, although we know of no examples in which it fails to hold. This issue will be examined elsewhere. Second, the model could be enriched by allowing the agents more decision-making scope. For example, they might learn as they play the game repeatedly, they might make inferences about the others' decision rules, and they might choose optimal sample sizes. These additions will complicate both the state space and the stochastic process, but if agents make mistakes infrequently and independently of each other, then the stochastically stable states can be analyzed using the techniques developed above. We have deliberately chosen to focus on the case where agents do *not* learn in order to make the point that the selection mechanism operates at the social rather than at the individual level. Individuals do not necessarily learn, but eventually society does.

Appendix

Here we shall prove a general result on finite Markov chains of which Theorem 2 is a special case. Let P be a stationary Markov chain defined on a finite state space X. Suppose that this process is subjected to a small perturbation or noise. By this we mean that with high probability the process follows the transition function P, but with small probability certain transitions occur that could *not* have occurred via P. We shall assume that the perturbed process can be modelled as a stationary Markov chain on X with transition function $P(\varepsilon)$, where ε is a scalar parameter that measures the overall level of noise, $P(0) = P$, and the following three conditions hold:

$$P(\varepsilon) \text{ is defined and continuous in some nontrivial interval } 0 \le \varepsilon \le a. \tag{6}$$

$$P(\varepsilon) \text{ is aperiodic and irreducible for all } 0 < \varepsilon \le a \tag{7}$$

$$\forall x, y \in X,\ \ P_{xy}(\varepsilon) \neq 0 \text{ for some } \varepsilon > 0 \text{ implies } \exists!\ r \ge 0 \text{ s.t. } 0 < \lim_{\varepsilon \to 0} \varepsilon^{-r} P_{xy}(\varepsilon) < \infty. \tag{8}$$

We shall call a finite Markov process $P(\varepsilon)$ satisfying (6) - (8) a *regular perturbation* of P(0). Condition (7) implies that the perturbed process $P(\varepsilon)$ has a unique stationary distribution $\mu(\varepsilon)$ for every $0 < \varepsilon \le a$. Condition (8) says that every transition $x \to y$ either has zero probability for all $\varepsilon > 0$, or its probability is on the order of ε^r (for some real $r \ge 0$) as ε becomes small. In the former case we set $r(x, y) = \infty$, and in the latter case we set $r(x, y) = r$. The number $r(x, y)$ will be called the *resistance* of the transition $x \to y$. Condition (6) implies that

$$\lim_{\varepsilon \to 0} \varepsilon^0 P_{xy}(\varepsilon) = \lim_{\varepsilon \to 0} P_{xy}(\varepsilon) = P_{xy}(0).$$

Hence $r(x, y) = 0$ if and only if $P_{xy}(0) > 0$, that is, feasible transitions under P(0) have zero resistance.

We remark that the process $P^*(k, m, \mathbf{q}, \lambda, \varepsilon)$ defined by (2) in the text is a regular perturbation with respect to ε, and the resistance of a one-period transition is the minimum number of mistakes required to make it.

The perturbed process $P(\varepsilon)$ has a unique stationary distribution $\mu(\varepsilon)$ for all $\varepsilon > 0$, while the unperturbed process may have many stationary distributions. We are going to show that $\lim_{\varepsilon \to 0}$

$\mu(\varepsilon) = \mu$, where μ is one of the stationary distributions of P(0). Thus the effect of the perturbations is to *select* among the stationary distributions of P(0).[4] We shall show further that the limiting distribution μ depends only on the resistances r(x, y). Hence it is enough to know the order of magnitude of various transition probabilities in order to compute the limiting distribution.

To characterize the limiting distribution μ, we shall view the states $x \in X$ as the vertices of a directed graph G, whose edges are the ordered pairs (x, y) such that r(x, y) is finite. Fix an arbitrary state z $\in$ X. A *z-tree* T is a subset of directed edges such that, for every vertex $x \neq z$, there exists a unique directed path in T from x to z. For every $z \in X$, let $\mathcal{T}_z$ be the set of all z-trees in G. Since $P(\varepsilon)$ is irreducible for each $\varepsilon > 0$, there exists a path of finite resistance from every state to every other state. Hence for every $z \in X$, $\mathcal{T}_z \neq \phi$. The *resistance* of a z-tree T is the sum total of the resistances of its edges, that is,

$$r(T) = \sum_{(x,y)\in T} r(x, y).$$

The *stochastic potential* of a state z is

$$\gamma(z) = \min_{T\in\mathcal{T}_z} r(T) . \tag{9}$$

LEMMA 1. Let $P(\varepsilon)$ be a regular perturbation of P(0) and let $\mu(\varepsilon)$ be its stationary distribution. Then $\lim_{\varepsilon\to 0} \mu(\varepsilon) = \mu$ exists, μ is a stationary distribution of P(0), and for every $x \in X$, $\mu_x > 0$ if and only if $\gamma(x)$ is a minimum.

The state x is *stochastically stable* if $\lim_{\varepsilon\to 0} \mu_x(\varepsilon) > 0$ [Foster and Young, 1990]. Lemma 1 says that the stochastically stable states are precisely the states with minimum stochastic potential.

PROOF. We employ a general method for computing stationary distributions due to Freidlin and Wentzell (1984). Let P' be the transition function of any aperiodic, irreducible, stationary Markov process defined on the finite space X. For each $z \in X$, define the number

---

[4] Canning (1991) proves that $\lim_{\varepsilon\to 0} \mu(\varepsilon)$ is a stationary distribution of $\lim_{\varepsilon\to 0} P(\varepsilon)$ under more general conditions, but he does not characterize the support of the limiting distribution.

$$p'_z = \sum_{T \in \mathcal{T}_z} \prod_{(x,y) \in T} P'_{xy} .$$

p'_z is positive because P' is irreducible. Let

$$\mu_z = p'_z / (\sum_{x \in X} p_x) > 0.$$

If may then be verified that $P'\mu = \mu$, from which it follows that μ is the unique stationary distribution of P' [Freidlin and Wentzell, 1984, Lemma 3.3].

Now let us apply this result to the process $P(\varepsilon)$ hypothesized in the Lemma. Let

$$p_z(\varepsilon) = \sum_{T \in \mathcal{T}_z} \prod_{(x,y) \in T} P_{xy}(\varepsilon) . \tag{10}$$

By the above result, the stationary distribution of $P(\varepsilon)$ is given by the formula

$$\forall z \in X, \qquad \mu_z(\varepsilon) = p_z(\varepsilon) / \sum_{x \in X} p_x(\varepsilon) . \tag{11}$$

Define $\gamma(z)$ as in (9) and let $\gamma = \min_z \gamma(z)$. Given $z \in X$, choose a z-tree T and consider the identity

$$\varepsilon^{-\gamma} \prod_{(x,y) \in T} P_{xy}(\varepsilon) = \varepsilon^{r(T)-\gamma} \prod_{(x,y) \in T} \varepsilon^{-r(x,y)} P_{xy}(\varepsilon). \tag{12}$$

Since r(T) is finite, r(x, y) is finite for every $(x, y) \in T$. Hence $0 < \lim_{\varepsilon \to 0} \varepsilon^{-r(x,y)} P_{xy}(\varepsilon) < \infty$. Taking the limit as $\varepsilon \to 0$ on both sides of (12) we obtain

$$\lim_{\varepsilon \to 0} \varepsilon^{-\gamma} \prod_{(x,y) \in T} P_{xy}(\varepsilon) = 0 \ \text{ if } \gamma < r(T),$$

and

$$\lim_{\varepsilon \to 0} \varepsilon^{-\gamma} \prod_{(x,y) \in T} P_{xy}(\varepsilon) > 0 \ \text{ if } \gamma = r(T),$$

With $p_z(\varepsilon)$ as defined in (10) it follows that

$$\lim_{\varepsilon \to 0} \varepsilon^{-\gamma} p_z(\varepsilon) = 0 \text{ if } \gamma < \gamma(z) \tag{13}$$

and

$$\lim_{\varepsilon \to 0} \varepsilon^{-\gamma} p_z(\varepsilon) > 0 \text{ if } \gamma = \gamma(z). \tag{14}$$

By (11),

$$\mu_z(\varepsilon) = \varepsilon^{-\gamma} p_z(\varepsilon) / \Sigma_{x \in X} \varepsilon^{-\gamma} p_x(\varepsilon). \tag{15}$$

From (13) - (15) it follows that, for every $z \in X$,

$$\lim_{\varepsilon \to 0} \mu_z(\varepsilon) = 0 \text{ if } \gamma < \gamma(z),$$

and

$$\lim_{\varepsilon \to 0} \mu_z(\varepsilon) > 0 \text{ if } \gamma = \gamma(z).$$

Hence $\mu = \lim_{\varepsilon \to 0} \mu(\varepsilon)$ exists, and its support is the set of states z that minimize $\gamma(z)$. Since $\mu(\varepsilon)$ satisfies $P(\varepsilon)\mu(\varepsilon) = \mu(\varepsilon)$ for every $\varepsilon > 0$, it follows from the continuity of $P(\varepsilon)$ that $P(0)\mu = \mu$. Hence μ is a stationary distribution of P(0). This completes the proof of Lemma 1.

Since μ is a stationary distribution of P(0), $\mu_z = 0$ for every state z that is not recurrent under P(0). To find the stochastically stable states, it therefore suffices to compute the potential function only on the recurrent states. Let the irreducible, recurrent communication classes of P(0) be denoted by $X_1, \ldots, X_{v^*}$. By definition, every two states within the same class X_v are accessible from each other by transitions in P(0), that is, by paths having zero resistance. From every state in X there exists a path of zero resistance to at least one of the classes X_v, but no state outside of a class X_v is accessible by such a path from any state inside it.

We shall show that the potential function is constant on every class X_v, and that it may be computed by finding a path of least resistance from every irreducible class to every other and then solving an arborescence problem on v* vertices rather than on the whole state space X.

Let the indices $\{1, 2, \ldots, v^*\} = V$ be the vertices of a graph. For each pair of indices (u, v), let r_{uv} be the least resistance among all directed paths that begin in X_u and end in X_v. (For this purpose it is sufficient to fix any two states $x \in X_u$ and $y \in X_v$ and find a least-resistant path from x to y.)

A *v-tree* on V is a subset τ of directed edges such that for every vertex $u \neq v$ there exists exactly one directed path from u to v. Let $\mathcal{T}_v$ be the set of all v-trees on V. For each $v \in V$ find the v-tree of least total resistance and let this resistance be denoted by $\rho(v)$:

$$\rho(v) = \min_{\tau \in \mathcal{T}_v} \sum_{(u, v) \in \tau} r_{uv} . \tag{16}$$

LEMMA 2. $\rho(v)$ is the stochastic potential of all states $x \in X_v$.

PROOF. Fix u such that $1 \leq u \leq v^*$ and fix an arbitrary element $x_u \in X_u$. Let ρ be defined as in (16). We shall show first that $\gamma(x_u) \leq \rho(u)$. Then we shall show the reverse inequality.

Fix a u-tree τ on the vertex set $\{1, \ldots, v^*\}$ such that $r(\tau) = \rho(u)$. For every $v \neq u$, there exists exactly one outgoing edge $(v, v') \in \tau$. In the graph G, choose a directed path $D_{vv'}$ from X_v to $X_{v'}$ having resistance $r_{vv'}$, and denote the initial vertex on this path by x_v. Since X_v is a communication class of P(0), we may also choose a directed tree T_v on the subset of vertices X_v such that from every vertex in X_v there is a unique directed path in T_v to v, and $r(T_v) = 0$.

Let E be the union of all of the edges in the trees T_v $(v \neq u)$ and all of the edges in the directed paths $D_{vv'}$ where $(v, v') \in \tau$. By construction, the set of directed edges E contains at least one directed path from every vertex in X to the fixed vertex x_u. Therefore it contains an x_u-tree T and

$$r(T) \leq r(E) = \sum_{1 \leq v \leq v^*} r(T_v) + \sum_{(v, v') \in \tau} r(D_{vv'}) = 0 + \sum_{(v, v') \in \tau} r_{vv'} = \rho(u).$$

By definition, $\gamma(x_u) \leq r(T)$, and hence $\gamma(x_u) \leq \rho(u)$ as claimed.

To show that $\gamma(x_u) \geq \rho(u)$, fix a state x_u in X_u and fix an x_u-tree T such that $\gamma(x_u) = r(T)$. Define a *junction* in the tree T to be any vertex y with at least two incoming T-edges. Label the junction "w" if there exists a path of zero resistance from y to the class X_w. There exists at least one such class X_w, because these are the absorbing classes of the process P(0). If there are several, then choose any one of them as the label.

For every $v \neq u$, fix a state x_v in X_v and label it "v". These vertices x_v, together with x_u, will be called *special vertices*. Every labelled vertex is either a special vertex or a junction (or both), and the label identifies a class to which there is a path of zero resistance. Define the *special predecessors* of a state $x \in X$ to be the special vertices x_v that strictly precede x in the tree T such that there is no other special vertex $x_{v'}$ strictly between x and x_v in T.

If a junction is labelled "w" and x_v *is a special predecessor of this junction, then the unique path in the tree from* x_v *to the junction has resistance at least* r_{vw}. (17)

This property clearly holds for the tree T because any path from x_v to a junction labelled "w" can be extended by a zero resistance path to the class X_w, and the total path must have resistance at least r_{vw}. We shall now perform certain topological operations on the tree T that preserve property (17).

Suppose that T contains a junction y that is not a special vertex, and let its label be "w". We distinguish two cases, depending on whether the special vertex x_w is or is not a predecessor of y in the tree.

Case 1. If x_w is not a predecessor of y in the tree (see Figure 4), cut off the subtree consisting of all edges and vertices that precede y and glue them onto the tree at the vertex x_w.

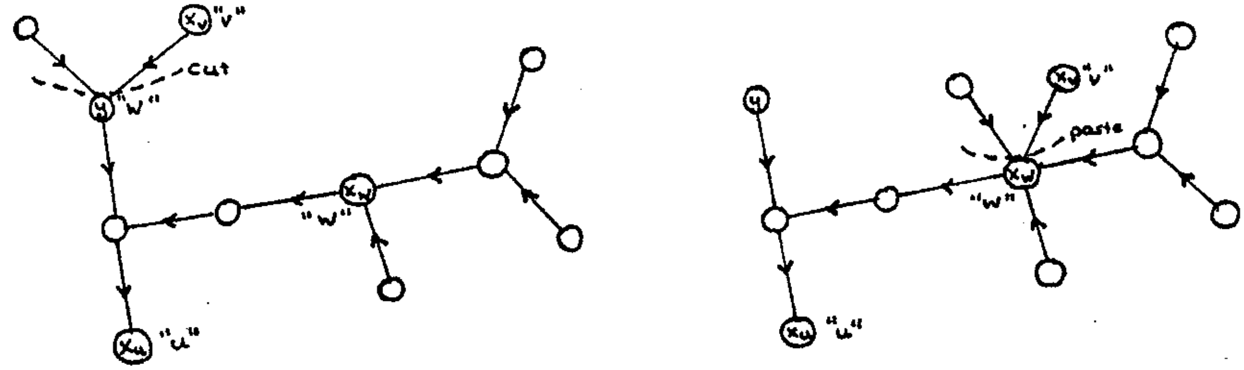

Figure 4. Case 1 surgery : before after.

Case 2. If x_w is a predecessor of y (see Figure 5), let (z, y) be the last edge on the unique path from x_w to y. Cut off the subtree consisting of all edges and vertices that precede y *except* for the subtree consisting of (z, y) and all the predecessors of z, and glue them onto x_w.

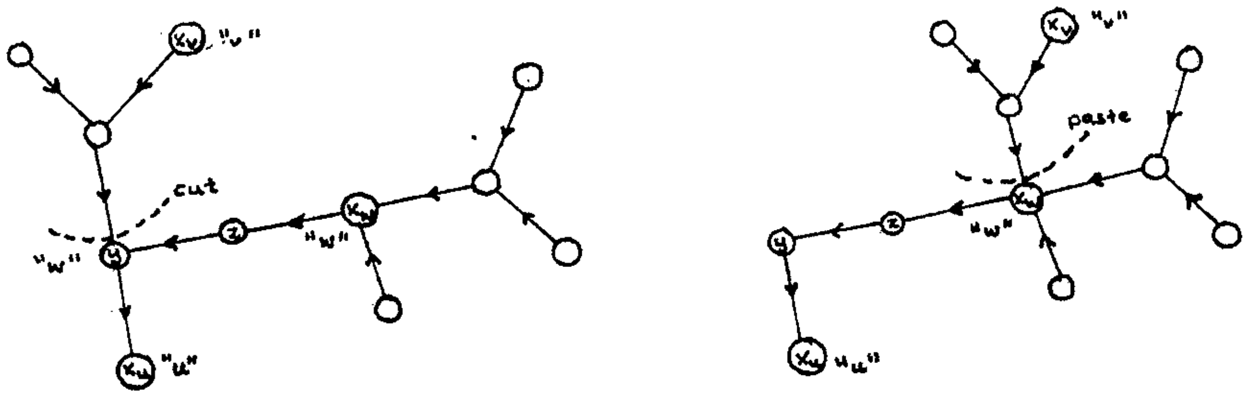

Figure 5. Case 2 surgery: before after.

Both of these operations preserve property (17) because x_w and y have the same label "w." Moreover, they do not change the total resistance of the tree. Each operation reduces by one the number of junctions that are not special vertices, so we eventually obtain an x_u-tree T* in which every junction is a special vertex. Moreover, $r(T^*) = r(T)$ and (17) is satisfied.

Now construct a u-tree τ on the vertex set V as follows. For every v and w in V put the directed edge (v, w) in τ if and only if x_v is a special predecessor of x_w in T*. By construction, τ forms a u-tree. Let D^*_{vw} be the unique path in T* from x_v to x_w. These paths are edge-disjoint because every junction is one of the special vertices. Hence

$$r(T) = r(T^*) \geq \Sigma_{(v,w)\in\tau}\, r(D^*_{vw}).$$

Property (17) says that $r(D^*_{vw}) \geq r_{vw}$ for all $(v,w) \in \tau$. Hence

$$\Sigma_{(v,w)\in\tau}\, r(D^*_{vw}) \geq \Sigma_{(v,w)\in\tau}\, r_{vw}\,.$$

Since τ is a u-tree on V, $\Sigma_{(v,w)\in\tau}\, r_{vw} \geq \rho(u)$. By choice of T, $\gamma(x_u) = r(T)$. Putting all of this together we obtain $\gamma(x_u) \geq \rho(u)$, as was to be shown. This completes the proof of Lemma 2.

Together, Lemmas 1 and 2 yield the following characterization of the stochastically stable states of the perturbed process $P(\varepsilon)$.

THEOREM 4. *Let* $P(\varepsilon)$ *be a regular perturbation of* $P(0)$ *on the state space* X, *and let* $X_1, \ldots, X_{v^*}$ *be the irreducible recurrent classes of* $P(0)$. *The stationary distribution of* $P(\varepsilon)$ *converges to a stationary distribution of* $P(0)$ *as* $\varepsilon \to 0$, *and the stochastically stable states are precisely the states contained in the irreducible classes* X_v *that minimize* $\rho(v)$.

References

Axelrod, Robert (1984): *The Evolution of Cooperation*. New York: Basic Books.

Bardhan, Pranab (1984): *Land , Labor, and Rural Poverty*. New York: Columbia University Press.

Canning, David (1990): "Social Equilibrium," Manuscript. Pembroke College, Cambridge, UK.

------------------(1991): "Average Behavior in Learning Models," Manuscript. Pembroke College, Cambridge, UK.

Chu, Y. J., and T. H. Liu(1965): " On the shortest arborescence of directed graphs," *Scientia Sinica*, 14, 1390-1400.

Crawford, Vincent (1991): "An Evolutionary Interpretation of Van Huyck, Battalio, and Beil's Experimental Results on Coordination, " *Games and Economic Behavior 3*, 25-59.

Edmonds, Jack(1967): "Optimum Branchings," *Journal of Research of the National Bureau of Standards*, 71B, 233-240.

Evans, George and S. Honkapohja (1991): "Convergence of Recursive Leaning Mechanisms to Steady States and Cycles in Stochastic Nonlinear Models." Manuscript. London School of Economics.

Foster, Dean, and H. Peyton Young (1990): "Stochastic Evolutionary Game Dynamics," *Theoretical Population Biology*, 38, 219-232.

Freidlin, Mark, and Alexander Wentzell (1984): *Random Perturbations of Dynamical Systems*. New York: Springer Verlag.

Friedman, Daniel (1991): "Evolutionary Games in Economics," *Econometrica*, 59, 637-666.

Fudenberg, Drew, and Eric Maskin (1990): "Evolution and Cooperation in Noisy Repeated Games," *American Economic Review*, 80, 274-279.

Fudenberg, Drew, and C. Harris (1991): "Evolutionary Dynamics in Games with Aggregate Shocks," Unpublished Manuscript. Massachusetts Institute of Technology.

Kandori, Michihiro, George Mailath, and Rafael Rob (1991): "Learning, Mutation, and Long Run Equilibria in Games," Manuscript. The University of Pennsylvania.

Kirman, Alan (1991): "Ants, Rationality, and Recruitment," Manuscript. European University Institute.

Harsanyi, John, and Reinhard Selten (1988): *A General Theory of Equilibrium Selection in Games*. Cambridge, Mass.: MIT Press.

Lewis, David (1967): *Convention:A Philosophical Study*. Cambridge, Mass: Harvard University Press.

Nachbar, John H. (1990): "'Evolutionary Selection Dynamics in Games: Convergence and Limit Properties," *International Journal of Game Theory*, 19, 59-89.

Samuelson, Larry (1988): "Evolutionary Foundations of Solution Concepts for Finite, 2-Player, Normal-Form Games," in Vardi, M. Y., ed., Theoretical Aspects of Reasoning About Knowledge. Los Altos: Morgan Kauffman.

------------------------(1991): "Limit Evolutionarily Stable Strategies in Two-Player, Normal Form Games, " *Games and Economic Behavior* 3, 110-128.

Samuelson, Larry, and J. Zhang (1990): "Evolutionary Stability in Asymmetric Games," Preprint. The Pennsylvania State University

Schelling, Thomas C. (1960): The Strategy of Conflict. Cambridge, Mass.:Harvard University Press.

Sugden, Robert (1986): *The Evolutions of Rights, Cooperation and Welfare*. New York: Basil Blackwell.

Tarjan, R.E. (1977): "Finding Optimum Branchings," *Networks*, 7, 25-35.

Warneryd, Karl (1990): "Conventions, " *Constitutional Political Economy* 1, 83-106.

Young, H. Peyton (1989): "The Evolution of Bargaining Conventions." Manuscript, University of Maryland.

Young, H. Peyton, and Dean Foster (1991): "Cooperation in the Short and in the Long Run," *Games and Economic Behavior,* 3, 145-156.