

© Drew Fudenberg 2004. These lecture notes are intended to accompany a classroom presentation, and may be incomplete without it. Do not copy or post without explicit permission from the author.

In a Nash equilibrium, each player's strategy is a best response to the strategies of the other players.

Nash equilibrium and its refinements are the basis for most applications of game theory to applied theoretical questions, case studies, and data from the field and the laboratory.

Questions: When and why does play resemble that of a Nash equilibrium? When there are multiple Nash equilibria, which ones are more likely to be observed?

Background:

- The assumption that the players are rational, or even that their rationality is common knowledge, is neither necessary nor sufficient for standard equilibrium concepts.
- In my view, the best way to think about equilibrium in games is as the long-run outcome of a non-equilibrium dynamic process of learning, imitation, or evolution.
- There are many such processes, with varying degrees of sophistication on the part of the agents. These lectures will focus on “evolutionary” models in a continuum population, which make assumptions about the aggregate behavior of the entire population.

Evolutionary Game Theory

Two key concepts:

a) *Evolutionary stable strategy* (“ESS”).

This is an equilibrium concept that is intended to capture the force of evolution. It strengthens the concept of Nash equilibrium by also asking that the strategy be “resistant to invasion by mutants.”

b) The *replicator dynamic*.

This is a non-equilibrium dynamic process that supposes the mass of agents using a given strategy grows at rate proportional to the strategy’s current payoff.

The classic result in evolutionary game theory is that ESS implies asymptotic stability under the replicator dynamic. Some recent work has studied alternative evolutionary processes.

Reasons this may be of interest to economists:

- 1) Although the replicator dynamics was originally motivated by a simplified version of biological evolution, the process can also describe the result of some types of “emulation” by economic agents.
- 2) Some of the properties of the replicator dynamic extend to various classes of more general processes that may correspond to other sorts of learning or emulation.
- 3) The study of evolutionary models has proved helpful (if controversial) in understanding animal behavior, and has been applied by some economists and anthropologists to the study of human behavior.

Defining the Replicator Dynamic

Specialize to the case of a single population playing a symmetric two-player stage game.

Suppose there is a continuum of players, each of whom uses a pure strategy.

Let $\phi_t(s)$ be the measure of players using pure strategy s at date t , and let

$$\theta_t(s) = \frac{\phi_t(s)}{\sum_{s'} \phi_t(s')}$$

be the share of the population playing strategy s .

Then the expected payoff to using pure strategy s against a randomly chosen opponent at date t

is $u_t(s) \equiv \sum_{s'} \theta_t(s') u(s, s')$,

the average expected payoff in the population is

$$\bar{u}_t = \sum_s \theta_t(s) u_t(s).$$

The *replicator dynamic* is the system

$$\dot{\phi}_t(s) = \phi_t(s)u_t(s), \quad \text{or equivalently}$$

$$\dot{\theta}_t(s) = \theta_t(s)[u_t(s) - \bar{u}_t].$$

This is equivalent from the quotient rule:

$$\begin{aligned} \dot{\theta}_t(s) &= \frac{d}{dt} \left(\frac{\phi_t(s)}{\left(\sum_{s'} \phi_t(s') \right)} \right) = \\ &= \frac{\dot{\phi}_t(s) \sum_{s'} \phi_t(s') - \phi_t(s) \sum_{s'} \dot{\phi}_t(s')}{\left(\sum_{s'} \phi_t(s') \right)^2} = \\ &= \frac{\phi_t(s)u_t(s) \sum_{s'} \phi_t(s') - \phi_t(s) \sum_{s'} u_t(s') \phi_t(s')}{\left(\sum_{s'} \phi_t(s') \right)^2} = \\ &= \frac{\phi_t(s)u_t(s) - \phi_t(s)\bar{u}_t}{\left(\sum_{s'} \phi_t(s') \right)} = \theta_t(s)[u_t(s) - \bar{u}_t]. \end{aligned}$$

The standard motivation for this system is as follows:

Each individual is genetically programmed to play some pure strategy, and this programming is inherited. Each individual is randomly matched with another individual from the population, and the reproduction rate of each individual is proportional to its score in the stage game. In the spirit of the law of large numbers, the total number of offspring of s -strategists is then $\phi_t(s)u_t(s)$.

Even in biological applications this is a highly simplified and stylized model, but it has proven to be successful there. If there is time, I will later discuss some motivations for the replicator dynamic based on imitation instead of reproduction.

Since $\dot{\phi}_t(s) = \phi_t(s)u_t(s)$,

$sign(\dot{\phi}_t(s)) = sign(u_t(s))$.

In particular, if all payoffs are negative, the entire population is shrinking. There is no problem with this on the biological interpretation. In most applications to economics the number of agents playing the game is held constant, while the share playing each strategy changes due imitation of successful play.

Note that even if payoffs are negative, the sum of the population shares is always 1. Note also that if the initial share of strategy s is positive, then its share remains positive: the share can shrink toward 0, but 0 is not reached in finite time.

The equation

$$\dot{\theta}_t(s) = \theta_t(s)[u_t(s) - \bar{u}_t]$$

says that the share of a strategy is increasing whenever that strategy does better than the population average, even if the strategy is not the best response to the current state.

Moreover, rate of increase is proportional to the amount that the strategy is better than average.

Despite this ability of sub-optimal strategies to increase their share, there is still a close connection between steady states of the replicator dynamic and Nash equilibria.

Steady States of the Replicator compared to Nash Equilibria:

1) Every Nash equilibrium is a steady state:

In the state corresponding to a Nash equilibrium, all strategies being played have the same average payoff, so the population shares are constant.

2) Interior steady states, where all actions have positive probability, must be Nash equilibria, but steady states on the boundary need not be Nash.

For example, any state where all agents use the same strategy is a steady state, since the dynamic does not allow the “entry” of strategies that are “extinct”.

3) Non-Nash steady states cannot be asymptotically stable: If the state is perturbed by introducing a small weight on an improving deviation, the share playing that deviation it will grow.

This last fact is true for the more general class of continuous-time deterministic processes where the growth rates are a strictly increasing function of the payoff differences.

Proof that non-Nash steady states are not stable:

Suppose that θ^* is a steady state, but the corresponding strategy profile σ^* is not a Nash equilibrium. Then, since payoffs are continuous, there exists a pure strategy $s \in \text{support}(\sigma^*)$, a pure strategy s' and a $\delta, \varepsilon > 0$ such that $u(s', \sigma') > u(s, \sigma') + \varepsilon$ for all σ' within δ of σ^* . Hence if there is a path that remains in a δ -neighborhood of σ^* , the growth rate of strategy s' exceeds that of strategy s by an amount that is bounded away from zero. Thus the share of strategy s must converge to 0, which is a contradiction.

What are the other implications of asymptotic stability?

Proposition [Bomze 1986]: An asymptotically stable steady state in the homogenous-population replicator dynamic corresponds to a Nash equilibrium that is trembling-hand perfect and isolated.

This result shows that asymptotic stability will be hard to satisfy in games with a non-trivial extensive form, for such games typically have connected sets of equilibria that differ only in their off-path play.

As with most dynamical systems, there is no guarantee that the replicator dynamics converge, and indeed, there are examples of games with no asymptotically stable steady states. In particular, even a totally mixed equilibrium need not be asymptotically stable. A simple example in which there is no asymptotically stable steady state is the game “rock-scissors-paper.”

Example Consider the “rock-scissors-paper” game

| | R | S | P |
|-----|------|------|------|
| R | 0,0 | 1,-1 | -1,1 |
| S | -1,1 | 0,0 | 1,-1 |
| P | 1,-1 | -1,1 | 0,0 |

The unique Nash equilibrium is the symmetric profile $(1/3, 1/3, 1/3)$. This steady state is surrounded by closed orbits, so that it is a “center;” hence is stable but not asymptotically stable. (One way to show this is to prove that $\theta_1\theta_2\theta_3$ is constant along any trajectory.)

From the viewpoint of general dynamical systems, centers are a knife-edge case, in the sense that small changes to the system (that is, small changes to the flow or vector field) can make the steady state a sink or a source. In this particular case such a change can be made simply by changing the payoffs slightly so that each strategy gets a small $\varepsilon > 0$ when matched against itself.

(In general, changing the payoffs need not lead to all possible small changes in the flow.)

When the diagonal payoffs are perturbed by adding $\varepsilon > 0$, the unique Nash equilibrium is still $(1/3, 1/3, 1/3)$, but now the Nash equilibrium is an unstable steady state, and the trajectories of the replicator dynamics spiral outward towards the boundary of the simplex without reaching it. Conversely, the Nash equilibrium is an asymptotically stable steady state for small $\varepsilon < 0$.)

(To show this, substitute $\theta_3 = 1 - \theta_1 - \theta_2$, linearize the two-dimensional system $\dot{\theta} = F(\theta)$ at the steady state, and compute the eigenvalues.)

Evolutionarily Stable Strategies

In applications, instead of working with explicit evolutionary dynamics, analysts often use the static concept of an *evolutionary stable strategy* or ESS (Maynard Smith, J. Theor. Biology 1974.) The idea of ESS is to require that the equilibrium be able to “repel invaders.”

Suppose that the population is originally at some profile σ , and then a small ε of “mutants” start playing σ' . ESS asks that the existing population gets a higher payoff against the resulting mixture than the mutants do. Specifically we ask that

$$(*) \quad u(\sigma, (1-\varepsilon)\sigma + \varepsilon\sigma') > u(\sigma', (1-\varepsilon)\sigma + \varepsilon\sigma')$$

for all sufficiently small positive ε , or equivalently

(**)

$$(1-\varepsilon)u(\sigma, \sigma) + \varepsilon u(\sigma, \sigma') > (1-\varepsilon)u(\sigma', \sigma) + \varepsilon u(\sigma', \sigma')$$

(**)

$$(1 - \varepsilon)u(\sigma, \sigma) + \varepsilon u(\sigma, \sigma') > (1 - \varepsilon)u(\sigma', \sigma) + \varepsilon u(\sigma', \sigma')$$

Since (**) need only hold for ε close to 0, it is equivalent to requiring that for all $\sigma' \neq \sigma$, either

(a) $u(\sigma, \sigma) > u(\sigma', \sigma)$ or

(b) $u(\sigma, \sigma) = u(\sigma', \sigma)$ and $u(\sigma, \sigma') > u(\sigma', \sigma')$

Note that (**) requires that σ is a Nash equilibrium. Otherwise the first term on the left hand side of the inequality is less than the corresponding term on its right. More formally, ESS is equivalent to requiring that

(a') σ is a NE: there is no other strategy σ' with $u(\sigma', \sigma) > u(\sigma, \sigma)$

and

(b') If $u(\sigma', \sigma) = u(\sigma, \sigma)$
(so σ' is an *alternate best reply* to σ)
then $u(\sigma', \sigma') < u(\sigma, \sigma')$.

Any strict Nash equilibrium (where the equilibrium strategy yields a strictly higher strategy than any alternative) is an ESS.

But many games do not have any strict equilibria, because, for example, mixed strategy equilibria can never be strict. And ESS need not exist.

Although mixed strategy equilibria can never be strict, they can, however, be ESS. Consider a “hawk-dove” game like

$$\begin{array}{cc} 0,0 & 1,1 \\ 1,1 & 0,0 \end{array}$$

(A hawk-dove game has a with a symmetric mixed-strategy equilibrium and two pure-strategy equilibria that are asymmetric.)

The unique mixed equilibrium is $(1/2, 1/2)$ and if either player is playing this strategy, both players get an expected utility of $1/2$.

Suppose most players play $\sigma = (1/2, 1/2)$. Then the payoff to an “invading strategy” σ' is

$$\begin{aligned} u(\sigma', (1-\varepsilon)\sigma + \varepsilon\sigma') &= \\ (1-\varepsilon)u(\sigma', \sigma) + \varepsilon u(\sigma', \sigma') &= \\ (1-\varepsilon)(1/2) + \varepsilon u(\sigma', \sigma'). \end{aligned}$$

Since $u(\sigma', \sigma') < 1/2$ for all $\sigma' \neq \sigma$, the invading strategy does strictly worse, so the definition of an ESS is satisfied.

Theorem: (Hofbauer and Sigmund 1988?) If σ is an ESS, then for some $\delta > 0$, $u(\sigma, \sigma') > u(\sigma', \sigma')$ for all σ' within δ of σ .

In words- not only is σ a Nash equilibrium, there are no other Nash equilibria nearby, and σ is a profitable deviation against any nearby σ' .

Theorem (Taylor and Jonker [Math BioSciences 1978]; Hofbauer et al [J. Theor. Biology 1979]; Zeeman [1980]): Every ESS is an asymptotically stable steady state of the replicator dynamics.

Proof: To see that ESS implies asymptotic stability, suppose that σ is an ESS, and let $\sigma(s)$ denote the weight that σ assigns to the pure strategy s . We will show that the “entropy” function $P_\sigma(\theta) = \prod_s \theta(s)^{\sigma(s)}$ is a strict local Lyapunov function at σ , that is, it has a local (actually global here) maximum at σ and that it is strictly increasing over time along trajectories in some neighborhood of σ .

To see this, note that

$$\begin{aligned}\frac{\dot{P}_\sigma}{P_\sigma} &= \frac{d}{dt}(\log P_\sigma(\theta_t)) = \sum_s \sigma(s) \frac{d \log(\theta_t(s))}{dt} \\ &= \sum_s \sigma(s) \frac{\dot{\theta}_t(s)}{\theta_t(s)} = \sum_s \sigma(s) [u(s, \theta_t) - u(\theta_t, \theta_t)] \\ &= u(\sigma, \theta_t) - u(\theta_t, \theta_t)\end{aligned}$$

where we have used $\bar{u}_t = u(\theta_t, \theta_t)$.

Since σ is an ESS, the final expression is positive in a neighborhood of σ , so P is an increasing function of time in this neighborhood as well.

Since P is maximized at σ it is a strict local Lyapunov function at σ , and hence σ is asymptotically stable.

The converse is false: van Damme [1987] shows by example that not every asymptotically stable steady state is an ESS.

Selected Applications and Extensions of ESS and Evolutionary Game Dynamics

1) The selection of equilibria in repeated games, particularly the prisoner's dilemma. Repeated play never eliminates static equilibria, so "All D" is a Nash equilibrium of the repeated prisoner's dilemma, but we feel that a fair bit of cooperation is likely.

Does ESS explain this cooperation? And which repeated game strategies are selected?

Consider the prisoner's dilemma with payoff matrix

| | | |
|----------|----------|----------|
| | <i>C</i> | <i>D</i> |
| <i>C</i> | 2,2 | 3,-1 |
| <i>D</i> | -1,3 | 0,0 |

Simplest version of the idea is due to Axelrod and Hamilton [*Science 1981*]: suppose the only two repeated game strategies are "Always Defect" and "Tit for Tat" and that players are completely patient.

("Tit for Tat" is the strategy "Play C in the first period, and from then on play the action the opponent played yesterday.")

Then the infinitely repeated game payoff matrix is

| | | |
|-------------|------------|-------------|
| | <i>TfT</i> | <i>AllD</i> |
| <i>TfT</i> | 2,2 | 0,0 |
| <i>AllD</i> | 0,0 | 0,0 |

Now imagine applying an evolutionary process to this repeated game. Both pure strategies are Nash equilibria, but *AllD* is not an ESS, since it can be “invaded” by a small group of *TfT*.

Drawbacks of this argument:

- a) If players discount, or if they play a finite number of times, then AIID is a strict equilibrium and hence an ESS. Thus we need an additional argument (e.g. on size of basins, or by adding noise) to select an equilibrium. (See e.g. Nowak, Fudenberg, Taylor 2004)

- b) It is artificial to consider only two repeated game strategies, why not more complex ones? With time averaging, the strategy “Alternate between C and D as long as this has happened in the past otherwise always play D” is a strict equilibrium, and hence an ESS in the game with the three strategies “Alternate,” “TfT”, and “AIID.”

Fudenberg and Maskin *AER* [1990] and Binmore and Samuelson [JET 1992] select the efficient outcome in the repeated prisoner's dilemma when allowing for general finitely-complex strategies by considering more complicated settings. (noise in FM, complexity costs in BS). FM select the strategy "perfect tit for tat," which says to play C if yesterday's outcome was (C,C) or (D,D), and otherwise play D.

(Both players using this strategy is a subgame-perfect equilibrium. It is not a subgame-perfect equilibrium for both to play TfT, since under this strategy a single deviation leads to the two-cycle (C,D), (D,C), which is not a Nash equilibrium in that subgame.)

Nowak and Sigmund [Nature 1994] analyze a two-parameter family of strategies by simulation, and conclude that the selected strategy is "win-stay, lose-shift," which is the same as perfect TfT.

- c) Binmore-Samuelson and Fudenberg-Maskin consider general two-player games, but two players is a special case.

Boyd and Richerson [J. Theor. Biology 1998] analyze N-player public good games.

N player symmetric game.

Two actions, C and D .

Payoff functions $V(C|i)$ and $V(D|i)$,

where i is the total number (including self) who play C .

They focus on the case where payoffs are linear in number of cooperators, as in standard model of voluntary provision of public good.

In these games $V(C|i) = \beta i - c$ and $V(D|i) = \beta i$, where c is the cost of cooperation/provision, and β is a scale factor that may depend on the size N of the group. Here we need $c < N\beta$ for cooperation to be efficient.

Consider the discounted infinitely repeated version of this game, where the strategy space is restricted to the two strategies:

U (Uncooperative): Always D

T_a (tft-a): C on the first period, and then C iff at least a other people played C yesterday. (In the two-player case, T_1 is the strategy “Tit for Tat.”)

“All U ” is a strict NE for any discount factor.

“All T_a ” is a steady state of the replicator dynamic, but for $a < N - 1$ it isn’t a Nash equilibrium, so it can’t be an ESS.

“All T_{N-1} ” is a strict NE if discount factor is close enough to 1, so it is an ESS.

“Exactly $a + 1$ play T_a ” is a NE for large enough discount factors, , but if $a + 1 < N$ it isn’t a symmetric NE (it requires a coordinating device) so it can’t arise in a one-population model.

Consider $a = N - 1$, and take discount factor large enough that “all T_{N-1} ” is a strict NE. Then the two pure-strategy equilibria are strict and stable, and there is a unique mixed equilibrium strategy equilibrium, so it is unstable. What do the basins of attraction look like?

- Their main point is that holding the discount factor fixed, the basin of all T_{N-1} shrinks to a point as N increases, since we need a higher fraction using T_{N-1} to have the same probability of a group of all T_{N-1} . This effect is geometric- we need $p^N = q$ or $p = q^{1/N}$.
- Holding the group size fixed, the basin of all T_{N-1} increases to cover the interval as the discount factor goes to 1: even a very small chance of permanent cooperation outweighs a likely one period gain.

Using “representative” numbers for the payoff matrix and the discount factor, their model predicts no cooperation, so they conclude that the prevalence of cooperation in large human societies is a puzzle—although they claim the conclusion of the model is right for animals since large groups of animals don’t cooperate. .

Missing: social sanctions/punishments that directly target deviators.

Also: what about populations with non negligible amounts of a number of strategies?

And more sophisticated strategies: why just unconditional triggers? Does this matter?

2) The selection of equilibria in “cheap talk” games

(e.g. Robson [*J. Theor. Biol.* 1990], Blume, Kim, and Sobel [*GEB* 1993].)

Idea: Suppose that players will play the coordination game

| | | |
|----------|----------|----------|
| | <i>L</i> | <i>R</i> |
| <i>L</i> | 2,2 | -100,0 |
| <i>R</i> | 0,-100 | 1,1 |

Before actually playing, they simultaneously announce their “intended actions;” once the announcements are made they simultaneously choose actions. There is a Nash (subgame-perfect) equilibrium where both players announce R and play R, but this equilibrium is not an ESS, since it can be invaded by the strategy “announce L, and play whatever the opponent announces.” More generally, if there is an unsent message, the mutants can invade a Pareto-inefficient outcome by using a “secret handshake.”

Once again, the most straightforward application of ESS doesn't deliver the conclusion we want: the unique ESS is for players to randomize over all signals so that mutants can never "recognize each other." (Schlag [1993].) To predict that meaningful communication will occur, we either need that there is always the possibility of creating a previously unsent signal, or a solution concept that incorporates drift, as in Matsui [*JET* 1990] .

- 3) Evolution of preferences-e.g. Samuelson [*Econometrica* 2003].
- 4) Mutations, flexibility, and genetic diversity in a changing environment. E.g. Ben Porath, Dekel, and Rustichini [*GEB* 1993], Heller [*GEB* 2004].
- 5) Evolutionary models of language formation
- 6) Other models of social learning, with local structure, peer effects, etc.
- 7) Evolutionary dynamics in finite populations: is the continuum population a good approximation?

Generalizations of the Replicator Dynamic

Samuelson and Zhang [*JET* 1992] say that an adjustment process (that is, a flow on the state space $\Theta^1 \times \Theta^2 = \Sigma^1 \times \Sigma^2$) is *regular* if

- (i) it is Lipschitz continuous,
- (ii) the sum of the flows in each population equals 0, and
- (iii) strategies with 0 shares have non-negative growth rates.

This implies that when it starts in the interior of the simplex, the process cannot hit the boundary in finite time.

(SZ consider two-population models; in a symmetric game, the one-population model can be viewed as a special case corresponding to a symmetric initial position.)

A process is *payoff monotone* if at all interior points,

$$u_t^i(s^i) > (=) u_t^i(s^{i'}) \Rightarrow \frac{\dot{\theta}_t^i(s^i)}{\theta_t^i(s^i)} > (=) \frac{\dot{\theta}_t^i(s^{i'})}{\theta_t^i(s^{i'})}.$$

Recall that strategy σ^i is *strictly dominated* if there is some other (possibly mixed) strategy $\hat{\sigma}^i$ such that

$$u^i(\hat{\sigma}^i, \sigma^{-i}) > u^i(\sigma^i, \sigma^{-i})$$

for all profiles of opponents' strategies σ^{-i} .

Iterated strict dominance is the process of first removing all strictly dominated strategies for each player, then removing all strategies that become strictly dominated once the dominated strategies are deleted, and so on until no further deletions are possible. (It can be shown that the order of deletions does not matter.)

Following Samuelson and Zhang, define the process of *iterated pure-strategy strict dominance* to be the analogous iterative process when only dominance by pure strategies is considered.

This process may delete fewer strategies, since a strictly dominated strategy may not be dominated by any pure strategy.

Theorem (Samuelson and Zhang): Under any regular, monotone dynamics, if strategy s is eliminated by the process of iterated pure-strategy strict dominance, then starting from any interior point the share of strategy s converges to 0 asymptotically, irrespective of whether the state itself converges.

Sketch of Proof: To begin, suppose that s is strictly dominated by some other pure strategy \hat{s} . Then the growth rate of s is always some fixed amount less than the growth rate of \hat{s} , and so the share of s in the population must go to 0.

Intuitively, we expect the adjustment process to run through iterative deletion: Once the dominated strategies have shares close to 0, then strategies that are removed at the second round of iterated pure-strategy dominance must have lower payoffs than those of other strategies with non-negligible shares, so their share starts to shrink to 0, and so on.

To make this intuition more precise, suppose that s is strictly dominated by pure strategy s' . Then

$$\frac{\dot{\theta}_t^i(s^i)}{\theta_t^i(s^i)} < \frac{\dot{\theta}_t^i(s^{i'})}{\theta_t^i(s^{i'})} - \varepsilon \Rightarrow$$

$$\frac{d}{dt} \left(\ln \left[\frac{\theta_t^i(s^i)}{\theta_t^i(s^{i'})} \right] \right) = \frac{\dot{\theta}_t^i(s^i)}{\theta_t^i(s^i)} - \frac{\dot{\theta}_t^i(s^{i'})}{\theta_t^i(s^{i'})} < -\varepsilon$$

so

$$\theta_t^i(s^i) < \exp(-\varepsilon t) \left(\frac{\theta_0^i(s^i)}{\theta_0^i(s^{i'})} \right) \theta_t^i(s^{i'}) \leq \exp(-\varepsilon t) \left(\frac{\theta_0^i(s^i)}{\theta_0^i(s^{i'})} \right)$$

Thus the share of s^i is bounded above by a function that converges to 0 at an exponential rate. Since there are only a finite number of such dominated strategies, there is, for any positive ε , a finite time T such that at all $t > T$ every one of them has share less than ε .

Let s'^i be a strategy for player i that is not strictly dominated by a pure strategy but is strictly dominated by some \hat{s}^i once the first round of deletions is performed. Since payoff functions are continuous functions of the mixed strategies, \hat{s}^i has a strictly higher payoff than s'^i once the shares of all of the “pure-strategy-strictly dominated” strategies are less than some sufficiently small ε , and by taking ε small enough we can ensure that this is true uniformly over all of the strategies removed at the second round of the iteration. Thus after some finite time T' , the shares of all of the strategies that are removed at the second round of iteration are bounded by a function that converges to 0 at an exponential rate, so the shares of these strategies become negligible at some finite time T'' . Now iterate. Since the process of iterated deletion ends in a finite number of rounds in finite games, only a finite number of iterations are required.

In order to eliminate strategies that are strictly dominated by a mixed strategy, Samuelson and Zhang introduce the condition of “aggregate monotonicity;” this was relaxed by Hofbauer and Weibull [*JET* 1996] to *convex monotonicity*.

Interpretation of the Replicator Equation

Two related types of learning stories have been proposed to explain the replicator dynamic.

- Social learning, in which players learn from other players in the population.
- “Aspiration models,” in which players “satisfice” rather than “optimize.”

There are many ways of formulating such models so that they generate a *payoff monotone* dynamic, that is, a system in which the growth rates of strategies are ordered by their expected payoff against the current population, so that strategies that are “doing better” grow faster.

The replicator is the particular form of a payoff monotone dynamic in which the rates of growth are proportional to payoff differences with the mean; it corresponds to particular functional form assumptions in the learning models. This makes it important to understand which conclusions based on the replicator dynamic or ESS extend to “similar” processes.

Social Learning

Here is an example of a model of social learning. Suppose that each period, some fraction α of the agents leaves the system, and are replaced by new agents, each of whom learns observes the strategy and payoff of one exiting agent and of one other agent drawn randomly from the same population, where the probability of sampling an agent using a given strategy equals the population fraction using that strategy.

New agents make a once-and-for all choice of strategy, which they do by adopting the strategy with the higher observed payoff, or in case of a tie, the strategy they “inherited” from the exiting agent. If the agent they sample is using the same strategy that they “inherited,” the agents do not switch, even if that strategy is performing poorly.

(The literature has extensions to sample N agents, other decision rules, etc.)

Since the rule “switch if the other strategy’s payoff is higher” depends only on the ordinal rankings of the payoffs, it will not lead to the replicator dynamic, since in the replicator the speed of adjustment depends on the size of the payoff differences.

Moreover, if agents observe the realized payoff of the agent they sample, as opposed to its average payoff against the current population, the resulting process need not select the best response.

To see this, consider the following decision problem:

Each period player 1 picks U or D , and Nature picks L or R , with the probability of L fixed and equal to p ; player 1’s payoff is

| | | |
|-----|-----|-----|
| | L | R |
| U | 9 | 0 |
| D | 2 | 2 |

New agents who sample someone using the same strategy as their “parent” do not switch; at date t , fraction $\theta_1^t(U)$ is using U , so a fraction $\theta_1^t(U)^2$ of the active agents is composed of agents with U parents who sample another agent using U .

Agents with parents using U who sample someone using D stick with U if and only if their own current payoff is 9; the fraction of such agents is $p\theta_1^t(U)\theta_1^t(D)$.

Agents whose parents used D switch to U if they meet a U -user whose payoff was 9; this corresponds to fraction $p\theta_1^t(U)\theta_1^t(D)$.

Combining terms yields the difference equation $\theta_1^{t+1}(U) = (1 - \alpha)\theta_1^t(U) + \alpha(\theta_1^t(U)^2 + 2p\theta_1^t(U)\theta_1^t(D))$; substituting $\theta_1^t(D) = 1 - \theta_1^t(U)$ and simplifying, yields $\theta_1^{t+1}(U) - \theta_1^t(U) = \alpha\theta_1^t(U)(1 - \theta_1^t(U))(2p - 1)$.

Thus the system converges to all player 1's using D whenever $\alpha < 1/2$ even though U is a best response whenever $\alpha > 2/9$.

Schlag [*JET* 1998] supposes that if the agent's parent's payoff is u , and the agent samples an agent with payoff u' , the agent switches with probability $\max\{0, b(u' - u)\}$ where b is a constant that depends on the range of the utility function. He proves that there is a sense in which this rule "performs well," and that it produces a system that converges to the replicator dynamic in the limit of shorter and shorter time periods.

There are other justifications for the assumption that aggregate play looks like the replicator dynamic. One is that agents sample as above, and that there is a distribution of switching costs in the population, and agents only switch if the gain exceeds the switching cost.

Another, due to Binmore and Samuelson [JET 1997], is that there is a distribution of “aspiration levels,” and that agents only switch if they see an alternative that is better and their current payoff is below their aspiration level.

Instead of supposing that agents observe the realized payoffs of other players, Bjornerstedt and Weibull [1995] assume that agents receive noisy statistical information directly about the current expected payoff of the strategy they sample. They show that this assumption, together with the assumption that the support of the noise is sufficiently large, leads to a resulting process that is monotone.

To see this, suppose that the distribution of noise is such that the difference between any two noise terms has c.d.f. Φ_i . Then the probability that a player i who is currently using s^i and who samples a player using \tilde{s}^i will switch is the probability that the noise term is less than the payoff difference, namely $\Phi_i(u_t^i(\tilde{s}^i) - u_t^i(s^i))$.

Under proportional sampling, the fraction that uses s^i and samples \tilde{s}^i equals the fraction that uses \tilde{s}^i and samples s^i . The net growth rate of the share using s^i is the inflow minus the outgo, so the population evolves according to the dynamic

$$\dot{\theta}_t^i(s^i) = 2\theta_t^i(s^i) \left[\sum_{\tilde{s}^i} \theta_t^i(\tilde{s}^i) \left(\Phi_i(u_t^i(s^i) - u_t^i(\tilde{s}^i)) - \Phi_i(u_t^i(\tilde{s}^i) - u_t^i(s^i)) \right) \right]$$

This is payoff monotone whenever the Φ_i are strictly increasing over the range of all payoff differences, which will be the case whenever the support of the noise is big enough.

If, moreover, the noise has a uniform distribution over a sufficiently large interval, and the distribution is the same for the various players, then this simplifies to the replicator dynamics.

Bjornerstedt [1995] develops an alternative derivation of the replicator, based on the idea that only “dissatisfied” agents change their strategy, with the probability of dissatisfaction depending on the agent’s own payoff and on some function of the current state, such as the current average payoff in the population, or the current lowest payoff.

Agents who are dissatisfied choose another agent at random (under proportional sampling) and copy that agent’s choice regardless of its current payoff.

- If agents with lower payoffs are more likely to be dissatisfied, the resulting dynamic is payoff monotone.
- The system is exactly the replicator dynamics in the special case where the probability of dissatisfaction is a suitably scaled linear function of the payoffs. (The scaling must ensure that the revision probabilities stay between 0 and 1, and so depends on the payoff function of the particular game.)