The Evolution of Cooperation Through Imitation

David K. Levine and Wolfgang Pesendorfer

Overview

- The problem of multiple equilibria and the folk theorem
- One answer: an evolutionary approach to see what equilibrium results

- Propogation through imitation rather than innovation
- Strategies that depend on the opponent's type in a random matching game
- Example: cooperate if opponent is same type, punish if he is different – usual folk theorem result, this is an equilibrium, but so is never cooperate
- No errors: long-run equilibrium is efficient, but bursts of conflict in which players minmax the difference in their payoffs
- Errors in an additively separable environment: maximize a weighted sum of own and opponents payoff – preferences for altruism and spite

Literature

- Kandori, Mailath and Rob [1993] and Young [1993]
- Morris, Rob and Shin [1993] one-half dominance
- Ellison [2000] co-radius
- Bergin and Lipman [1994] relative probabilities of noise matters
- Johnson, Pesendorfer and Levine [2000] emergence of cooperation in a trading game
- Kandori and Rob [1993] winning pairwise contests sufficient because implies 1/2-dominance

The Model and Basic Characterization of Long Run Outcomes

- symmetric normal form game
- single population of players
- finitely many pure strategies $s \in S$
- mixed strategies $\sigma \in \Sigma$
- utility on his own pure strategy and mixed strategy of population $u(s, \sigma)$
- $u(s, \sigma)$continuous in σ
- two player matching games: $u(s, \sigma)$ linear in σ

Evolution

- m player population
- each plays a pure strategy
- distribution of strategies at t denoted by $\sigma_t \in \Sigma$
- initial distribution σ_0

The Imitative Process

- σ_t is determined from σ_{t-1} according to the “imitative” process

1) One player i chosen at random; only this player changes strategy

2) probability $C\varepsilon$ *imitation* - strategies chosen in proportion to previous period frequency - player i chooses from S randomly using the probabilities σ_{t-1}.

3) probability ε^n *innovation* - strategies are entirely at random -player i chooses each strategy from S with equal probability

4) probability $1 - C\varepsilon - \varepsilon^n$ *relative best response* - best response among those strategies that are actually used - player i randomizes with equal probability among the strategies that solve

$$\max_{s \in \mathrm{supp}(\sigma_{t-1})} u(s, \sigma_{t-1})$$

The Markov Process

- The imitative process is a Markov process M on state space $\Sigma^m \subset \Sigma$ of all mixed strategies consistent with the grid induced by each player playing a pure strategy
- process M is positively recurrent because of innovation
- so M has unique invariant distribution μ^ε
- goal: characterize $\mu \equiv \lim_{\varepsilon \to 0} \mu^\varepsilon$

Assumption: imitation is much more likely than innovation.

Unlikely Innovation: $n > m$

as $\varepsilon \to 0$ probability of *every* player changing strategy by imitation much greater than probability a single player innovates

Basic Results

Pure Strategies

- mixed strategies less stable than pure strategies
- mixed strategy can evolve to pure only using imitation
- pure cannot evolve at all without at least one innovation

Theorem 1: $\mu = \lim \mu^{\varepsilon}$ *exists and* $\mu(\sigma) > 0$ *implies that* σ *is a pure strategy*

Pairwise Contests

what it means to win pairwise contests:

$0 \leq \alpha \leq 1$mixed strategy that plays s with probability α and $\tilde{s}$ with probability $1-\alpha$ denoted by $\alpha s + (1-\alpha)\tilde{s}$.

Definition 1: *The strategy* s **beats** $\tilde{s}$ *iff*

$$u(s, \alpha s + (1-\alpha)\tilde{s}) - u(\tilde{s}, \alpha s + (1-\alpha)\tilde{s}) > 0$$

for all $1/2 \leq \alpha < 1$

Definition 2: *The strategy* s **weakly beats** $\tilde{s}$ *iff*

$$u(s, \alpha s + (1-\alpha)\tilde{s}) - u(\tilde{s}, \alpha s + (1-\alpha)\tilde{s}) > 0$$

for all $1/2 < \alpha < 1$ *and* $u(s, \tfrac{1}{2}s + \tfrac{1}{2}\tilde{s}) - u(\tilde{s}, \tfrac{1}{2}s + \tfrac{1}{2}\tilde{s}) = 0$

Definition 2': *The strategy* s *is* **tied** *with* $\tilde{s}$ *iff*

$$u(s, \alpha s + (1-\alpha)\tilde{s}) - u(\tilde{s}, \alpha s + (1-\alpha)\tilde{s}) = 0$$

for all $1 \geq \alpha \geq 0$

Definition 3: *If* s *beats all* $\tilde{s} \neq s$ *we say that* s **beats the field**. *If for all* $\tilde{s} \neq s$ *either* s *weakly beats* $\tilde{s}$ *or is tied with* $\tilde{s}$ *we say that* s **weakly beats the field**

A Sufficient Condition for Long Run Equilibrium

Theorem 2: *If m is sufficiently large and s beats the field then $\mu(s) = 1$. If m is sufficiently large and s weakly beats the field then $\mu(s) > 0$. Moreover, if $\mu(\tilde{s}) > 0$ then $u(s, \tfrac{1}{2} s + \tfrac{1}{2} \tilde{s}) - u(\tilde{s}, \tfrac{1}{2} s + \tfrac{1}{2} \tilde{s}) = 0$*

Uses standard Kandori-Mailath-Rob-Young tree-trimming arguments

This condition is similar to ½ dominance, but much weaker – it isn't necessary to beat combinations of other strategies, just one at a time

We'll see how that gets used later

Matching Games with Behavioral Types

- each period players matched into pairs to play a symmetric normal form game
- prior choosing an action, each player receives a “signal” about how his opponent will behave in the game
- how does the long-run outcome depends upon the amount of information contained in the signals?

Formalities

- action space A
- payoff of a player who takes action a and whose opponent takes action a' is $U(a, a')$
- actual strategy space is a finite abstract space S

strategies serve two roles

1. influence the information that is generated about the player and his opponent
2. govern the behavior as a function of the generated information

Informational Function of Strategies

- each player receives a signal $y \in Y$, a finite set
- probability of signal given by $\pi(y \mid s, s')$ if the player uses s and the opponent s'
- signals are private information and reflects what the opponent can learn about his opponent prior to the interaction

Behavioral Function of Strategies

- each strategy s gives rise to a map $s : Y \rightarrow A$
- several strategies may induce the same map, yet differ in the probability with which they send signals

for every map from signals to actions there is some strategy that induces that map

Assumption 0: *If* $f : Y \to A$ *there is a strategy* $s \in S$ *with* $s(y) = f(y)$.

the space of signals is necessarily smaller than the set of strategies

the cardinality of the space of strategies is at least A^Y, which is greater than that of Y provided that there are at least two signals

Motivation

- People have some ability to detect lying
- Frank [1987] some people blush whenever they tell a lie
- A good lie is costly to construct – Ben Gurion airport
- The signal is a combination of a self-report of intentions together with an involuntary signal about lying
- “I’ll still respect you in the morning [blush]”

example

$Y = \{0,1\}$

$\pi(y = 0 \mid s', s) = 1$ if $s' = s$

$\pi(y = 1 \mid s', s) = 1$ if $s' \neq s$

two players meet who use the same strategy then both receive the signal 0 whereas when two players meet who use different strategies then both receive the signal 1

players recognize if the opponent uses the same strategy or a different strategy prior to play

important, because it turns out that strategies that recognize themselves are likely to emerge in the long-run equilibrium.

signals affect payoffs only indirectly by affecting behavior

player i uses strategy s and opponent uses strategy s' expected payoff of player i

$$\sum_{y' \in Y} \sum_{y \in Y} U(s(y), s'(y'))\pi(y \mid s, s')\pi(y' \mid s', s)\sigma(s')$$

the function $u(s,\sigma)$

$$u(s,\sigma) = \sum_{s' \in S} \sum_{y' \in Y} \sum_{y \in Y} U(s(y), s'(y'))\pi(y \mid s, s')\pi(y' \mid s', s)\sigma(s')$$

Perfect Identification

each player can identify with certainty whether the opponent is using the same strategy

assume for notational simplicity one signal, denoted by y_0 that perfectly identifies that the opponent is using the same strategy.

Assumption 1: *There is a* $y_0 \in Y$ *such that for every* $s \in S$ $\pi(y_0 \mid s,s) = 1$ *and* $\pi(y_0 \mid s,s') = 0$ *for* $s \neq s'$.

two assumptions on U

Assumption 2 requires that there is no asymmetric profile that makes both players strictly better off than any symmetric profile.

Assumption 2: $U(a,a') > \max_a U(a,a) \Rightarrow U(a',a) \leq \max_a U(a,a)$

with public randomization device Assumption 2 always satisfied if include actions that may depend on the outcome of the public randomization: use coin flip to decide which player is the row player and which player is the column player o

once roles are assigned, players choose the Pareto optimal actions

Assumption 3 requires an action $\underline{a} \in A$ that ensures that the player gets a payoff that is at least as large as the payoff of his opponent.

Assumption 3: *There is an* $\underline{a} \in A$ *such that* $U(\underline{a}, a) - U(a, \underline{a}) \geq 0$ *for all* $a \in A$.

the payoff difference $U(a', a) - U(a, a')$ defines symmetric zero-sum game hence has a (possibly mixed) minmax strategy

Assumption 3 says the minmax strategy is an element of A: the game defined by the payoff differences has a pure minmax strategy

Assumption 3 always satisfied if we include the possibly mixed minmax action as one of the elements of A

Let $\overline{a}$ be a Pareto optimal symmetric outcome

$$\overline{a} \in \arg\max_{a \in A} U(a,a)$$

By assumption 0, there is a strategy s_0 for which.

$$s_0(y) = \begin{cases} \overline{a} & \text{if } y = y_0 \\ \underline{a} & \text{if } y \neq y_0 \end{cases}$$

strategy s_0 takesPareto efficient action when the opponent uses the same strategy and punishes the opponent by taking action $\underline{a}$ when the opponent chooses a different strategy

note that punishment action maximizes minimum difference between the player's payoff and his opponent's payoff

Theorem 3 shows that long run outcome of the evolutionary dynamics puts positive probability on the strategy s_0

every other strategy s that used with positive probability is similar to s_0: when s meets s both players receive the payoff $U(\overline{a},\overline{a})$; when s meets s_0 both players receive the same payoff.

Theorem 3: Under A0-A3, $\mu(s_0) > 0$. *If* $\mu(s) > 0$ *then* $u(s,s) = u(s_0,s_0) = U(\overline{a},\overline{a})$ *and* $u(s,s_0) = u(s_0,s)$.

if $\overline{a}$ is the unique symmetric Pareto optimal outcome and if $U(\underline{a}, a) - U(a, \underline{a}) > 0$ for all $a \neq \underline{a}$ then s_0 is the unique outcome in the long-run limit

suppose some redundancy description of the game

there may be other strategies that induce the same map as s_0

there may be an action $\widehat{a}$ with $U(\widehat{a}, a) = U(\overline{a}, a)$, for all $a \in A$

in either case two or more strategies that satisfy the requirement of s_0

strategies differ in extraneous detail only but will not recognize each other as the "same strategy"

long-run distribution places positive weight on every such strategy

during this brief period of transition from one to another players using different versions will punish one another by choosing $\underline{a}$, the action that maximizes the difference between the two player's payoff proof of Theorem 3 shows that the strategy s_0 weakly beats the field

that is; weakly beats every other strategy in a pairwise contest

s_0 need not be ½ dominant in the ordinary sense

underlying game is Prisoner's dilemma

$\tilde{s}$ be a constant strategy that always plays "defect"

signals that enable a strategy $\underline{s}$ to play "defect" against s_0 and "cooperate" against $\tilde{s}$

s_0 plays "cooperate" against s_0 and "defect" otherwise against $\frac{1}{2}s + \frac{1}{2}\underline{s}$ the strategy $\tilde{s}$ does better than s and therefore s is not ½ dominant

such a strategy seems to serve no useful purpose except to make $\overline{s}$ look good against s_0

our theory of infrequent innovation provides a rigorous account of why we should not expect such strategies to play a role in determining the long-run equilibrium

they do not themselves do well against s_0 and will not remain around long enough for players to discover that they should play $\overline{s}$

Gift Exchange and Imperfect Identification

simple additively separable structure

each action a has a cost $c(a)$ and yields a benefit $b(a)$ for the opposing player

payoff of a player who takes action a and whose opponent chooses action a' is

$$U(a, a') = b(a') - c(a)$$

example: two players meet and have an opportunity to exchange goods. c denotes the cost of the good and b describes the benefit of the good for the opposing player

resemble a prisoner's dilemma in that the cost minimizing action is dominant normalize so that $c(a) \geq 0$, with $c(a) = 0$ for some action $\hat{a} \in A$

note that this utility function satisfies Assumptions 2 and 3

Simple Strategies

consider strategies that may not be able to identify with certainty when the opponent is using the same strategy

To keep things simple, we first assume that all strategies use the same symmetric information structure

Assumption 4:

$$\pi(y \mid s, s') = \begin{cases} p(y) & \text{if } s = s' \\ q(y) & \text{if } s \neq s' \end{cases}$$

$p(y)$ probability that a player has type y if he and his opponent use the same strategy

$q(y)$ probability when the two players use different strategies

suppose prior probability a player uses s is α and that he receives the signal y. posterior probability that the opponent will also play according to s is

$$\frac{\alpha p(y)}{\alpha p(y) + (1 - \alpha) q(y)}.$$

This posterior is greater than α when $p(y) > q(y)$ and less than α when $q(y) > p(y)$

strategy s_0

for every signal y the action $s_0(y)$ solves

$$\max_{a \in A}(p(y) - q(y))b(a) - (p(y) + q(y))c(a)$$

assume the maximization problem has unique solution for every y

s_0 rewards opponent when $p(y) > q(y)$

punishes opponent when $q(y) > p(y)$

in limiting case where the type allows no inference about play ($p(y) = q(y)$) strategy s_0 minimizes cost c

Theorem 4 shows that in the long run only s_0 is played

Theorem 4: $\mu(s) > 0$ if and only if $s(y) = s_0(y)$ for all $y \in Y$

signal is y then s_0 maximizes

$$\frac{p(y)-q(y)}{p(y)+q(y)}b(a)-c(a)$$

Let s denote the opponent's strategy choice, and suppose that there a prior of ½ that the opponent is using s_0. In that case,

$$\Pr(s=s_0 \mid y)=\frac{p(y)}{p(y)+q(y)};\Pr(s\neq s_0 \mid y)=\frac{q(y)}{p(y)+q(y)}$$

and therefore

$$\frac{p(y)-q(y)}{p(y)+q(y)}=\Pr(s=s_0 \mid y)-\Pr(s\neq s_0 \mid y)$$

objective function puts a larger weight on the opponent's benefit when he is more likely to use strategy s_0

in long-run stable outcome all players are choosing the same strategy

probability that opponent using s_0 is one irrespective of signal

Nevertheless players punish each other for appearing to be different

so equilibrium is typically inefficient

Complex Strategies

Allow general strategies, but continue to keep strategy s_0

consider only strategies s_0 and s being played

strategy s_0 is *informationally superior* to strategy s if the signal generated by s_0 provides better information about the opponent's strategy than the signal generated by s

signal generated by s_0 provides better information (in the sense of Blackwell (1954)) than the signals generated by s if there is a non-negative matrix

$$\left(\lambda_{yz}\right)_{y\in Y, z\in Y}$$

such that

$$\sum_{y\in Y}\lambda_{yz} = 1, \forall z$$
$$\pi(y \mid s, s) = \sum_{z\in Y}\lambda_{yz}p(z),$$
$$\pi(y \mid s, s_0) = \sum_{z\in Y}\lambda_{yz}q(z);$$

In other words, the signals generated by $\pi(\cdot \mid s, \cdot)$ are a *garbling* of the signals generated by s_0

strategy s_0 is *informationally dominant*, if it is informationally superior to every other strategy s

informational dominance only better information in situations where s_0 and *one* other competing strategy are played

s_0 may be informationally dominant strategy even though strategy s does better at identifying a third strategy $\overline{s}$

trivial example of an informationally dominant strategy is a strategy that cannot be distinguished from any other strategy

$\pi(y \mid s, s_0) = \pi(y \mid s, s)$ for all s and hence strategy s_0 is informationally dominant even if strategy s_0 does not generate any information, that is, $p(y) = q(y)$ for all y

strategy s_0 is informationally dominant because it successfully masquerades as other strategies when strategy s_0 is informationally dominant, it emerges as an outcome of the long-run stable distribution

every strategy that is a long-run stable outcome is similar to strategy s_0: if $\mu(s) > 0$ then the payoff when s meets s is the same as the payoff when s_0 meets s_0.

Theorem 5: *If s_0 informationally dominant then,* $\mu(s_0) > 0$*. Moreover, for every strategies s with $\mu(s) > 0$ we have $u(s, s) = u(s_0, s_0)$ and* $u(s, s_0) = u(s_0, s)$.

we have restricted the informationally dominant strategy to generate symmetric information, that is, generate the same information for every opponent

this allows us to identify a behavior (a map from signals to actions) that is successful against every opponent

if we give up the symmetry assumption we must replace it with a requirement that preserves this ability to do well against all opponents

for example, can assume that there is a reference strategy s such that any signal realization generated by s_0 against an arbitrary opponent is at least as informative as it is against strategy s

Informational dominance would then require that the signal generated against s is informationally superior to the signal generated by any opponent.

Example 1: Perfect identification and additive separability

Against self

$$\max_{a \in A} b(a) - c(a)$$

against difference

$$\max_{a \in A} - b(a) - c(a) = - \min_{a \in A} b(a) + c(a)$$

punishment action minimizes the sum of the player's cost and of the opponent's benefit, so player willing to incur a cost if it leads to a negative payoff to a different opponent

Example 2.

add the strategy $\bar{s}$, which can masquerade as any other strategy and which takes the least cost action $\hat{a}$ for every signal realization

a player using strategy s cannot determine whether the opponent uses $\bar{s}$ or s, hence $\pi(y \mid s, \bar{s}) = \pi(y \mid s, s)$ for all signals $y \in Y$.

also players who use $\bar{s}$ do not receive informative signals about their opponents

such a strategy is informationally dominant

Theorem 5 implies that every strategy that is a long-run stable outcome must play the least cost action

introduction of a cost minimizing strategy that successfully masquerades as other strategies eliminates cooperation between players

Example 3. s_0 need not be ½ dominant in the ordinary sense

environment of Theorem 5

assume that s_0, the informationally dominant long-run outcome is not constant

$\tilde{s}$ be a constant strategy that always plays $\hat{a}$

signals that enable a strategy $\underline{s}$ to identify $\tilde{s}$ with certainty and choose an action that maximizes b. Otherwise, $\underline{s}$ chooses $\hat{a}$

For an appropriate choice of b, the strategy $\tilde{s}$ does better than s against $\frac{1}{2} s + \frac{1}{2} \underline{s}$ and therefore s_0 is not ½ dominant.

Example 4: symmetric two-signal scenario, $Y = \{0,1\}$ and $p(0) = q(1) = p, p \geq 1/2$

signal is $y = 0$ then this is an indication that the two players are using the same strategy

signal is $y = 1$ it is an indication that the strategies are different

three actions $a \in \{-1,0,1\}$, $b(a) = \beta a, c(a) = 2a^2 + a$

trading game with a cooperative action ($a = 1$), a no-trade action ($a = 0$), and a hostile action ($a = -1$)

hostile and the cooperative action are costly for players, whereas the no-trade action is costless apply Theorem 4 and distinguish the following cases

$$\frac{1}{1-2p} > \beta,$$

then in the unique long run outcome all players to take the no-trade action

$$\beta > \frac{3}{1-2p},$$

then in the unique long run outcome players choose the cooperative action when the signal is 0 and the hostile action when the signal is 1

$$\frac{3}{1-2p} > \beta > \frac{1}{1-2p},$$

then in the unique long run outcome players take the no trade action when the signal is 0 and the hostile action when the signal is 1

in this case, the long run outcome is worse than unique equilibrium of the normal form game

Players choose no trade and hostility and do not realize any of the gains from trade