Tadeusz Sozański

PRISONER'S DILEMMA, ITERATED PD, AND EVOLUTIONARY PROCESSES IN POPULATIONS PLAYING IPD

Lecture note evolving into a research note

May 2005

(na końcu komentarz po polsku dopisany w maju 2006)

One-shot Prisoner's Dilemma (PD)

	c	*d*
c	*R* , *R*	*S* , *T*
d	*T* , *S*	***P*** , ***P***

c = cooperate *d* = defect

T – Temptation payoff *S* – Sucker's payoff
R – Reward payoff *P* – Punishment payoff

Assumptions on payoffs defining the PD game

$$T>R>P>S, \quad R>½(T+S)$$

PD game with payoffs used in Axelrod's tournament

	c	*d*
c	3 , 3	0 , 5
d	5 , 0	1 , 1

d –- strongly dominating strategy ($T>R$, $P>S$)
(d,d) – the only pair of strategies in Nash equilibrium

PD repeated twice

Strategies

xy = Play x in the 1st round; play y in the 2nd round (x=c or d; y=c or d)

I \ II	*cc*	*cd*	*dc*	*dd*
cc	2R \ 2R	R+S \ R+T	S+R \ T+R	2S \ 2T
cd	R+T \ R+S	R+P \ R+P	S+T \ T+S	S+P \ T+P
dc	T+R \ S+R	T+S \ S+T	P+R \ P+R	P+S \ P+T
dd	2T \ 2T	T+P \ S+P	P+T \ P+S	**2P** \ **2P**

dd – dominating strategy

A representation in which the other party's action in the first round is taken into account in the second round

x/yz = play x in the 1st round (x=c or x=d) ; play y in the 2nd round if the other party played c in the 1st round; play z if the other party played d in the 1st round

a_1: c/cc – always cooperate
a_2: c/*cd* – tit for tat
a_4: *d*/*dd* – always defect
a_6: *d*/*cd* – suspicious tit for tat

Payoffs gained from best replies to one's partner' actions are printed in boldface (it is assumed that T+S>2P, that is, point (P,P) lies below the straight line passing through points (S,T) and (T,S).

Every pair of dynamical strategies generates a history of the game played twice. Different strategy pairs may generate the same history.

If $T+S>2P$, then (a_8,b_8) is the only strategy pair in Nash equilibrium

If $T+S\leq 2P$, then (a_8,b_6) and (a_6,b_8) are in Nash equilibrium as well.

All equilibrium pairs generate the same history (both actors defect twice)

2 1	b1:c/cc	b2:c/cd	b3:c/dc	b4:c/dd	b5:d/cc	b6:d/cd	b7:d/dc	b8:d/dd
a1:c/cc	2R 2R	2R 2R	R+T R+S	R+T R+S	T+R S+R	T+R S+R	**2T** 2S	**2T** 2S
a2:c/cd	2R 2R	2R 2R	**R+T** R+S	**R+T** R+S	T+S S+T	T+S **S+T**	T+P S+P	T+P S+P
a3:c/dc	R+S R+T	R+S **R+T**	R+P R+P	R+P R+P	T+R S+R	T+R S+R	**2T** 2S	**2T** 2S
a4:c/dd	R+S R+T	R+S **R+T**	R+P R+P	R+P R+P	T+S S+T	T+S **S+T**	**T+P** S+P	**T+P** S+P
a5:d/cc	S+R T+R	S+T T+S	S+R T+R	S+T T+S	P+R P+R	**P+T** P+S	P+R P+R	**P+T** P+S
a6:d/cd	S+R T+R	**S+T** T+S	S+R T+R	**S+T** T+S	P+S **P+T**	2P 2P	P+S **P+T**	2P **2P**
a7:d/dc	2S **2T**	S+P T+P	2S **2T**	S+P **T+P**	P+R P+R	**P+T** P+S	P+R P+R	**P+T** P+S
a8:d/dd	2S **2T**	S+P T+P	2S **2T**	S+P **T+P**	P+S **P+T**	**2P** 2P	P+S **P+T**	**2P** **2P**

Conclusion: if the PD game is repeated a fixed finite number of times, rational players cannot optimize their choices in round t by taking into account what happened in rounds $1,\ldots,t-1$.

Repeating a game with a constant probability that a round will be followed by another round

Z_t – a zero-one random variable

Z_t = 1 / 0 – the game will/will not be played at time t, t=1.,2,…

$P\{Z_{t+1} = 1 / Z_t=0\} = 0$ – if the game has not been played at time t, it will not be played at time t+1, either.

$P\{Z_{t+1} = 1 / Z_t=1\} = \delta$ – if the game has been played at time t, it will be played at time t+1 with probability δ, $0<\delta<1$

$P\{Z_1 = 1\} = 1$ – the game will be played at least once

$P\{Z_t = 1\} = P\{Z_t = 1, Z_{t-1} = 0\} + P\{Z_t = 1, Z_{t-1} = 1\} =$

$P\{Z_t = 1 / Z_{t-1} = 0\} \cdot P\{Z_{t-1} = 0\} + P\{Z_t = 1 / Z_{t-1} = 1\} \cdot P\{Z_{t-1} = 1\} =$

$0 \cdot P\{Z_{t-1} = 0\} + \delta \cdot P\{Z_{t-1} = 1\} = \delta \cdot P\{Z_{t-1} = 1\} = \delta \cdot \delta \cdot P\{Z_{t-2} = 1\} = \dots$

$\delta^{t-1} \cdot P\{Z_1 = 1\} = \delta^{t-1}$

T – the random variable whose value gives the number of times the game is played

$P\{T = t\} = P\{Z_t = 1, Z_{t+1} = 0\} = P\{Z_t = 1\} \cdot P\{Z_{t+1} = 0 / Z_t = 1\} = \delta^{t-1}(1-\delta)$

$$\sum_{t=1}^{\infty} P\{T=t\} = \sum_{t=1}^{\infty} (1-\delta)\delta^{t-1} = (1-\delta)\sum_{t=0}^{\infty} \delta^t = \frac{1-\delta}{1-\delta} = 1$$

$$E(T) = \sum_{t=1}^{\infty} t P\{T=t\} = \sum_{t=1}^{\infty} t(1-\delta)\delta^{t-1} = (1-\delta)(\sum_{t=1}^{\infty} \delta^t)' = (1-\delta)(\frac{1}{1-\delta} - 1)'$$
$$= \frac{1-\delta}{(1-\delta)^2} = \frac{1}{1-\delta}$$

The probability that the game ends after a finite number of steps equals 1. E(T) is the expected duration of the game.

$\delta = .5$	$E(T) = 2$
$\delta = .9$	$E(T) = 10$
$\delta = .99$	$E(T) = 100$
$\delta = .995$	$E(T) = 200$

Histories and dynamical strategies

$A = \{a_1, \dots, a_m\}$ – pure strategies of player 1

$B = \{b_1, \dots, b_n\}$ – pure strategies of player 2

Mixed strategies (probability distributions)

$$P_m = \{p = (p_1, \dots, p_m):\ p_i \geq 0, \sum_{i=1}^{m} p_i = 1\} \qquad Q_n = \{q = (q_1, \dots, q_n):\ q_j \geq 0, \sum_{j=1}^{n} q_j = 1\}$$

H_t – the set of all possible histories up to time t, t = 1,2,…

$h = (a_{i1},b_{j1}),\dots,(a_{it},b_{jt})$ – history as a sequence of co-actions
$h = (a_{i1},\dots,a_{it}), (b_{j1},\dots,b_{jt})$ – history as two parallel action sequences

$a_{i1},\dots,a_{it}$
$b_{j1},\dots,b_{jt}$

$H_o = \{\varnothing\}$ – empty history (preceding the first move)

$H^* = \bigcup_{t=0}^{\infty} H_t$ – the set of all histories of the game

$s_1 : H^* \to P_m$ – a dynamical strategy of player 1

$s_2 : H^* \to Q_n$ – a dynamical strategy of player 2

$s_1(\varnothing)$ – strategy used in the 1st round by player 1

For any $h \in H_t$, $t \geq 1$, $s_1(h)$ is the strategy used by player 1 in round t+1

If player 1 uses dynamical strategy s_1 and player 2 uses dynamical strategy s_2, then the payoff of player 1 is defined by the formula

$$u(s_1,s_2) = \sum_{t=1}^{\infty} P\{T=t\} u_t(s_1,s_2) = \sum_{t=1}^{\infty} (1-\delta)\delta^{t-1} u_t(s_1,s_2) \text{ where } u_t(s_1,s_2) = \sum_{h\in H_t} P\{h\}\, u(h)$$

is the mean total payoff of player 1 if the game is repeated t times

The total payoff $u(h)$ earned by player 1 as a result of history h in which player 1 used strategies $(a_{i1},\dots,a_{it})$ while player 2 used strategies $(b_{j1},\dots,b_{jt})$ is given by the formula

$$u(h) = \sum_{g=1}^{t} u_{i_g j_g} \quad \text{if } h = (a_{i1},b_{j1}),\dots,(a_{it},b_{jt})$$

P{h} – probability of a history h under dynamical strategies s_1 and s_2

$h \in H_1$, $h = (a_{i1},b_{j1})$, P{h}=$p_{i1}q_{j1}$, where $p = s_1(\varnothing)$ and $q = s_2(\varnothing)$ are mixed strategies used by the players in the first round

$h \in H_t$, t>1

$h = h'(a_{it},b_{jt})$, $h' \in H_{t-1}$, P{h}=P{h'}$p_{it}q_{jt}$, where $p = s_1(h')$ and $q = s_2(h')$ are mixed strategies used by the players in round t (dependent on the history h' of the game in rounds 1 through $t-1$.

s_1 is a deterministic strategy if $s_1(h)$ is a pure strategy for any $h \in H^*$, that is, if $h \in H_t$, $s_1(h) = a_{i_{t+1}}$

Any two deterministic strategies s_1 and s_2 determine an infinite history

$$a_{i_1} a_{i_2} \cdots$$
$$b_{j_1} b_{j_2} \cdots$$

Game played infinitely many times with a discount parameter

If s_1 and s_2 are two deterministic strategies, then

$$u(s_1,s_2) = \sum_{t=1}^{\infty} \delta^{t-1}(1-p)\sum_{g=1}^{t} u_{i_g j_g} = \sum_{k=1}^{\infty} u_{i_k j_k}(1-\delta)\sum_{t=k}^{\infty} \delta^{t-1} = \sum_{k=1}^{\infty} u_{i_k j_k}\frac{1-\delta}{\delta}\delta^k \sum_{t=k}^{\infty} \delta^{t-k} =$$
$$\sum_{k=1}^{\infty} u_{i_k j_k}\frac{1-\delta}{\delta}\delta^k \sum_{t=0}^{\infty} \delta^{t} = \sum_{k=1}^{\infty} u_{i_k j_k}\frac{1-\delta}{\delta}\delta^k \frac{1}{1-\delta} = \sum_{k=1}^{\infty} \delta^{k-1} u_{i_k j_k}$$

which allows for an alternative interpretation of δ as a discount parameter

$u_{i_1 j_1}$	–	payoff in the 1st round
$\delta u_{i_2 j_2}$	–	payoff in the 2nd round
$\delta^{k-1} u_{i_k j_k}$	–	payoff in the *k*th round

The game is (in theory) always played infinitely many times, but the payoff shrinks from round to round by a constant factor and converges to 0 as t goes to infinity.

Symmetric games

Let $A = \{a_1,\ldots,a_m\}$ and $B = \{b_1,\ldots,b_m\}$ denote strategy sets of player 1 and player 2. The game with payoff functions u (player 1) and v (player 2) is called *symmetric* if $u(a_i,b_j)=v(b_i,a_j)$, for $i,j=1,\ldots m$ ($u_{ij}=v_{ji}$ in matrix notation).

Example: PD

$a_1 = b_1 = c$, $a_2 = b_2 = d$

$u_{11} = v_{11} = R$, $u_{22} = v_{22} = P$
$u_{12} = v_{21} = S$, $u_{21} = v_{12} = T$

Dynamical strategies in Iterated PD

Deterministic strategies: examples

ALL D – always defect

$s(h) = d$

ALL C – always cooperate

$s(h) = c$

TFT – Tit for Tat

$s(\varnothing) = c$

$h = (i_1,j_1),\ldots,(i_t,j_t)\in H_t$, $t>0$

$s_1(h) = j_t$ (player 1), $s_2(h) = i_t$ (player 2)

In each round, play the strategy your partner played in the previous round

STFT – Suspicious Tit for Tat

Defect in round 1, play TFT from round 2 onwards

TF2T – Tit for 2 Tats

Defect in round t only if your partner defected in rounds $t-2$ and $t-1$

PR – Permanent Retaliation (Never Forgive)

Cooperate until your partner defects. If he defects for the first time in round t, defect from round $t+1$ onwards

ALT – Alternate defection with cooperation (*dcdc*....)

Probabilistic strategies: examples

RND ALL π: Cooperate in every round with constant probability π
RND 1 π/TFT: Play c with probability π in round 1, play TFT since round 2 on.

Probabilistic strategies should not be confused with finitely mixed dynamical strategies of the form $\alpha_1: s_1, \ldots,\alpha_k: s_k$ where $\alpha_i \geq 0$, $\alpha_1+\cdots+\alpha_k=1$, and $s_1,\ldots,s_k$ are dynamical strategies. Some probabilistic strategies can be represented as mixed strategies with deterministic components, e.g. RND 1 π/TFT works equivalently as π: TFT, $1-\pi$: STFT.

Computing payoffs in the IPD

$$u(ALLD,\ ALLD) = \sum_{t=1}^{\infty} \delta^{t-1}P = P\sum_{t=1}^{\infty} \delta^{t-1} = \frac{P}{1-\delta}$$

$$u(TFT,\ ALLD) = S+\sum_{t=2}^{\infty} \delta^{t-1}P = S+P\delta\sum_{t=2}^{\infty} \delta^{t-2} = S+\frac{\delta}{1-\delta}P$$

$$u(ALLD,\ TFT) = T+\sum_{t=2}^{\infty} \delta^{t-1}P = T+\frac{\delta}{1-\delta}P$$

$$u(TFT,\ TFT) = \sum_{t=1}^{\infty} \delta^{t-1}R = \frac{R}{1-\delta}$$

Collective stability

Let *s* be a dynamical strategy of player 1. The respective strategy of player 2 will be noted with the same symbol s by virtue of the symmetry of the infinitely iterated game resulting from the symmetry of the underlying one-shot game.

A dynamical strategy *s* of player 1 is said to be *collectively stable* if $u(s,s)\geq u(s',s)$ for any dynamical strategy *s*' of player 2, that is, *s* is the best response of player 1 to the same strategy *s* used by player 2. Then $v(s,s)=u(s,s)\geq u(s',s)=v(s,s')$, so that *s* is the best response of player 2 to *s* used by player 1. Thus, *s* is collectively stable iff (s,s) is in Nash equilibrium in the game with infinitely many strategies.

A *dominating* dynamical strategy *s** of player 1 is defined by the condition: for any dynamical strategy *t* of player 2. $u(s^*,t)\geq u(s,t)$ for any dynamical strategy *s* of player 1, that is, *s** is the best response to any strategy of the other player. In an infinitely iterated PD, there is no *s** with this property if δ is sufficiently high. Indeed, for *t*=ALL D, we have

$$u(ALLD,\ ALLD) = \frac{P}{1-\delta} > \frac{S}{1-\delta} = u(ALLC,\ ALLD)$$

that is, ALL D is better than ALL C against ALL D. However, ALL C is better than ALL D against t=TFT.

$$u(ALLC,\ TFT) = \frac{R}{1-\delta} > T+\frac{\delta}{1-\delta}P = u(ALLD,\ TFT) \text{ if } \delta > \frac{T-R}{T-P}$$

Theorem 1: ALL D is collectively stable, i.e., $u(\text{ALL D, ALL D})\geq u(s, \text{ALL D})$ for any *s*

Proof

$$u(s,\ ALLD) = \sum_{t=1}^{\infty}(1-\delta)\delta^{t-1}\sum_{h\in H_t}P\{h\}u(h) \le \sum_{t=1}^{\infty}(1-\delta)\delta^{t-1}\sum_{h\in H_t}P\{h\}tP = u(ALLD,\ ALLD)$$

Theorem 2: TFT is collectively stable iff $\delta \ge \frac{T-R}{T-P},\ \frac{T-R}{R-S}$

Proof of Theorem 2:

Necessity

(1) The inequality $\delta \ge \frac{T-R}{T-P}$ is equivalent to $u(TFT,\ TFT) \ge u(ALLD,\ TFT)$, which means that TFT is a better reply to TFT than ALL D.

(2) The inequality $\delta \ge \frac{T-R}{R-S}$ is equivalent to $u(TFT,TFT) \ge u(ALT,TFT) = \frac{T+\delta S}{1-\delta^2}$

(we omit the computation of u(ALT,TFT)).

(3) The collective stability of TFT implies that $u(TFT,TFT) \ge u(ALLD,TFT)$ and $u(TFT,TFT) \ge u(ALT,TFT)$. Hence $\delta \ge \frac{T-R}{T-P}$ and $\delta \ge \frac{T-R}{R-S}$ by (1) and (2) and the proof of necessity is completed.

Sufficiency

(4) Assume that $\delta \ge \frac{T-R}{T-P},\ \frac{T-R}{R-S}$. We show that there is no deterministic strategy s such that $u(s,\text{TFT})>u(\text{TFT, TFT})$. Suppose that there exists such an s. We construct a strategy s_1 such that $u(s_1,\text{TFT})\ge u(s,\text{TFT})$ and player 1 chooses c in round 1.

Let $a_{i1},a_{i2},....$ be the infinite sequence of actions of player 1 using s against TFT. If $a_{i1}=c$, then $s_1=s$. If $a_{i1}=d$ then s_1 is obtained from s by changing from d to c in round 1, and modifying the prescription on what player 1 should to in rounds t=2,3,... so that the history generated by s_1 and TFT coincide with the history generated by s and TFT from round 3 on.

Assume that $a_{i2}=d$. Then player 1 using s plays dd in round 1 and 2, while his partner using TFT plays cd, which implies that $u(s,\text{TFT})=T+\delta P+\cdots$. If player 1 uses s_1, then his actions are cd, while his partner actions are cc, so that $u(s_1,\text{TFT})=R+\delta T+\cdots$. We have, therefore, $u(s_1,\text{TFT})-u(s,\text{TFT})=(R+\delta T)-(T+\delta P)=\delta(T-P)-(T-R)\ge 0$ since $\delta\ge(T-R)/(T-P)$.

Assume in turn that $a_{i2}=c$. Then, the histories (1:dc)(2:cd), (1:cc)(2:cc), generated by s and s_1 and TFT yield the formulas $u(s,\text{TFT})=T+\delta S+\cdots$, and $u(s_1,\text{TFT})=R+\delta R+\cdots$. We have $u(s_1,\text{TFT})-u(s,\text{TFT})=(R+\delta R)-(T+\delta S)=\delta(R-S)-(T-R)\ge 0$.

Similarly, given s_1, we construct s_2 which prompts player 1 to cooperate in rounds 1 and it is no worse reply to TFT than s_1. Repeating the procedure, we obtain a sequence s_n of strategies, each cooperating in rounds 1 through n.

Notice that $u(s_n, TFT) = (\sum_{k=1}^{n} \delta^{k-1}R) + x_n$ where $x_n \leq \sum_{k=n+1}^{\infty} \delta^{k-1}T$ converges to 0 so that $\lim u(s_n,\mathrm{TFT}) = R/(1-\delta) = u(\mathrm{TFT},\mathrm{TFT})$ and $u(\mathrm{TFT},\mathrm{TFT}) > u(\mathrm{TFT},\mathrm{TFT})$, which is a contradiction.

(5) Once it is known that $u(\mathrm{TFT},\mathrm{TFT}) \geq u(s,\mathrm{TFT})$ for any deterministic s, it remains to be shown that TFT does no worse than any s (how to prove this?)

General dynamical systems with discrete time

Let X be a set endowed with a distance ρ and let $F: X \to X$ be a continuous mapping of X into X. (X,F) is called a *dynamical system* with *space of states* X and transformation F. If x is the system's state in time t (t=0,1,…), then $F(x)$ is the system's state in time t+1.

If x^0 is an initial state (t=0), then x^0, $x^1=F(x^0)$, $x^2=F^2(x^0)=F(F(x^0))$ are successive states (t=1,2,…). The sequence $(x^t=F^t(x^0))_{t=0,1,\ldots}$ is called a *trajectory* going out from x^0.

A point x^* is called an *equilibrium* if it is a *fixed point* of F, that is,

$$F(x^*)=x^*.$$

A point x^* is an equilibrium if and only if $x^*=\lim F^t(x^0)$ for some x^0.

An equilibrium x^* is said to be *stable* if for any point x sufficiently close to x^* the trajectory going out of x converges to x^*, that is, if the system has for any reason left the equilibrium state x^*, it will be returning to it provided that it has not moved too far away from x^*. Points which are “sufficiently close” to x^* are those lying in a *neighborhood* $U(x^*,r)$ of x^* for some $r>0$ where $U(x^*,r)=\{x\in X: \rho(x^*,x)<r\}$ (the set of points whose distance from x^* is smaller than r).

Symmetric games played by a population of players

Consider a symmetric game with a finite set of strategies $\{s_1,\ldots,s_m\}$ and payoffs of the player 1 given by the matrix (u_{ij}) with nonnegative entries.

Let $X = \{x = (x_1,\ldots,x_m): x_i \geq 0, \sum_{i=1}^{m} x_i = 1\}$ denote the set of m-dimensional probability distributions. For any vector x in X, we interpret its ith coordinate x_i as the relative frequency of players who use strategy s_i in all games they play at a given period t with other members of a closed population. A measure of ith strategy's *fitness* in state x is defined as the mean payoff of a player who uses strategy s_i in games played with other members of the population, symbolically.

$$V_i(x) = x_1 u_{i1} + x_2 u_{i2} + \cdots + x_m u_{im} = \sum_{j=1}^{m} x_j u_{ij}$$

For example, if all players use strategy s_1, then the current system state is given as the probability vector (1,0,…,0). We have then $V_i(1,0,\ldots 0)=u_{i1}$

The average fitness in a population in state x is defined by the formula

$$\overline{V}(x) = x_1 V_1(x) + \cdots + x_m V_m(x) = \sum_{i=1}^{m} x_i V_i(x)$$

Evolution

To construct a dynamical system with such a space of states X, we need to define a transformation F of X into X. Assume that an F is defined in such a way that the following conditions are satisfied for any x in X:

If $x_i=0$ then $F(x)_i=0$. If $x_i>0$ and $x_j>0$, then

$$\frac{F(x)_i}{x_i} < \frac{F(x)_j}{x_j} \quad \textit{iff} \quad V_i(x) < V_j(x)$$
$$\frac{F(x)_i}{x_i} = \frac{F(x)_j}{x_j} \quad \textit{iff} \quad V_i(x) = V_j(x)$$

or, equivalently,

$$\frac{\Delta x_i}{x_i} < \frac{\Delta x_j}{x_j} \quad \textit{iff} \quad V_i(x) < V_j(x)$$
$$\frac{\Delta x_i}{x_i} = \frac{\Delta x_j}{x_j} \quad \textit{iff} \quad V_i(x) = V_j(x)$$

where $\Delta x_i = F(x)_i - x_i$.

Every F with these properties is called an *evolutionary process*. The quantity $F(x)_i/x_i$ shows the direction (growth or decline) and intensity of change in ith strategy frequency between the current and next stage of the process. Under evolutionary process, the frequency of a strategy which is currently fitter than another strategy grows relatively faster.

If F is an evolutionary process, then

$F(x)_i=x_i$ for all i iff $V_i(x)=V_j(x)$ for any i,j such that $x_i>0$ and $x_j>0$,

that is, x is an equilibrium iff any two strategies which have survived are equally fit. The sufficiency of this condition results from adding up the two sides of equation $F(x)_i x_j=F(x)_j x_i$ over all j

Any probability vector z^i such that $z^i_i=1$ and $z^i_j=0$ for any $j\neq i$ is automatically an equilibrium. If z^i stable, then strategy s_i is said to be *evolutionarily stable under F*.

The concept of equilibrium stability has a definite meaning only if we define a distance in the space X of m-dimensional probability vectors. Let the *distance* between x and y in X be given by the formula $\rho(x,y)=\max_j |x_j - y_j|$. Then $\rho(z^i,x) = Max_j |z^i_j - x_j| = 1 - x_i = \sum_{j \neq i} x_j$. The neighborhood of z^i with radius r consists of all probability vectors x such that $x_i > 1-r$.

Strategy s_i is evolutionarily stable under F if there is an $r>0$ such that $\lim F^t(x) = z^i$ for any x such that $x_i > 1-r$. The infinite sequence $F^t(x)_{t=0,1,\ldots}$ converges to z^i, by definition, iff for any $\varepsilon>0$ there is a t_ε such that for every $t \geq t_\varepsilon$ $\rho(F^t(x),z^i)<\varepsilon$, that is, $F^t(x)_i > 1-\varepsilon$, As a consequence, $\lim F^t(x) = z^i$ iff $\lim F^t(x)_i = 1$ (hence $\lim F^t(x)_j = 0$ for all $j \neq i$).

Proportional fitness rule (PFR)

Let us define F by the following formula known as "proportional fitness rule"

$$F(x)_i = \frac{V_i(x)}{\overline{V}(x)} x_i$$

Thus, $F(x)_I \geq x_i$ iff $V_i(x) \geq \overline{V}(x)$. Any strategy occurs more frequently in time t+1 ("in the next generation") if and only if its fitness in time t is above the current average in the population.

Example: IPD with the strategy set restricted to s_1=ALL D, s_2=TFT

1 \ 2	ALL D	TFT
ALL D	$\frac{P}{1-\delta}$; $\frac{P}{1-\delta}$	$T+\frac{\delta}{1-\delta}P$; $S+\frac{\delta}{1-\delta}P$
TFT	$S+\frac{\delta}{1-\delta}P$; $T+\frac{\delta}{1-\delta}P$	$\frac{R}{1-\delta}$; $\frac{R}{1-\delta}$

Let us represent any $x=(x_1,x_2)$ as $(\varepsilon,1-\varepsilon)$. Since $V_1(x)$ and $V_2(x)$ are functionally related to ε, we write $V_i(\varepsilon)$ instead of $V_i(\varepsilon,1-\varepsilon)$

$$V_1(\varepsilon) = \varepsilon u(ALL\ D,ALL\ D) + (1-\varepsilon) u(ALL\ D,TFT) = \frac{(\varepsilon + (1-\varepsilon)\delta)P + (1-\varepsilon)(1-\delta)T}{1-\delta}$$

$$V_2(\varepsilon) = \varepsilon u(TFT, ALL\ D) + (1-\varepsilon)u(TFT, TFT) = \frac{\varepsilon((1-\delta)S + \delta P) + (1-\varepsilon)R}{1-\delta}$$

$$\overline{V}(\varepsilon) = \varepsilon V_1(\varepsilon) + (1-\varepsilon)V_2(\varepsilon)$$

A state $(\varepsilon, 1-\varepsilon)$ is an equilibrium if $\varepsilon=1$ (all players play ALL D) or $\varepsilon=0$ (all play TFT) or $0<\varepsilon<1$ and $V_1(\varepsilon)=V_2(\varepsilon)$. The only solution of the latter equation is given by the formula

$$\varepsilon = \frac{\delta(T-P)-(T-R)}{\delta(T-P)-(T-R)+(R-S)-\delta(P-S)}$$

For a fixed value of δ, the state of equilibrium $(\varepsilon, 1-\varepsilon)$ where ε is given by the above equation represents balanced co-habitation of ALL D and TFT in a population.

Let ε_0 represent the initial frequency of players using ALL D, and let $\varepsilon_t = F(\varepsilon_{t-1}) = F^t(\varepsilon_0)$ stand for the respective frequency at time t. If $\varepsilon_0=\varepsilon(\delta)$, then the frequency ε_t does not change over time as shown below where T, R, P, S are given Axelrod values 5, 3, 1, 0 and $\delta=⅔$, $\varepsilon_0=\varepsilon(\delta)=⅔$.

t	$V_1(\varepsilon_{t-1})$	$V_2(\varepsilon_{t-1})$	$V(\varepsilon_{t-1})$	$\varepsilon_t=F(\varepsilon_{t-1})$	$1-\varepsilon_t$
1	4.33	4.33	4.33	0.6667	0.3333
2	4.33	4.33	4.33	0.6667	0.3333
3	4.33	4.33	4.33	0.6667	0.3333

The equilibrium is not stable: a slightest change of ε results in a process that converges to (1,0) or (0,1). If $\varepsilon>\varepsilon(\delta)$ then ALL D supersedes TFT; if $\varepsilon<\varepsilon(\delta)$, then TFT wins the contest. The processes of approaching these states are shown below.

ALL D beats TFT

$\varepsilon_0 = 0.8 > ⅔ = \delta$

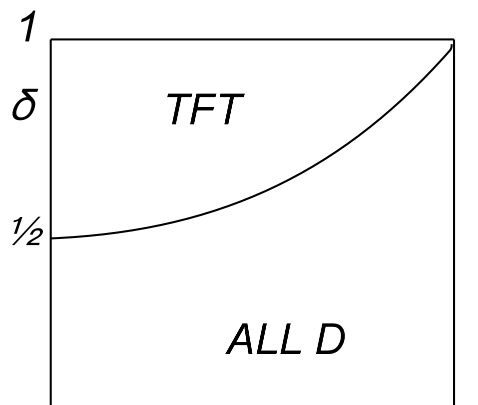
1	3.80	3.40	3.72	0.8172	0.1828
2	3.73	3.28	3.65	0.8357	0.1643
3	3.66	3.15	3.57	0.8552	0.1448
4	3.58	3.01	3.50	0.8752	0.1248
5	3.50	2.87	3.42	0.8952	0.1048
6	3.42	2.73	3.35	0.9144	0.0856
7	3.34	2.60	3.28	0.9321	0.0679
8	3.27	2.48	3.22	0.9478	0.0522
9	3.21	2.37	3.16	0.9610	0.0390
10	3.16	2.27	3.12	0.9716	0.0284

TFT beats ALL D

$\varepsilon_0 = 0.5 < ⅔ = \delta$

1	5.00	5.50	5.25	0.4762	0.5238
2	5.10	5.67	5.39	0.4498	0.5502
3	5.20	5.85	5.56	0.4208	0.5792
4	5.32	6.05	5.74	0.3895	0.6105
5	5.44	6.27	5.95	0.3563	0.6437
6	5.57	6.51	6.17	0.3217	0.6783
7	5.71	6.75	6.42	0.2865	0.7135
8	5.85	6.99	6.67	0.2515	0.7485
9	5.99	7.24	6.93	0.2177	0.7823
10	6.13	7.48	7.18	0.1857	0.8143

If $\delta < \frac{T-R}{T-P}$ (0.5 for Axelrod payoffs), then $\varepsilon(\delta)<0$ which is impossible. Thus, for such a δ only ALL D is evolutionarily stable. If $\delta \geq \frac{T-R}{T-P}$ so are both ALL D and TFT. Note that for any $0<\delta<1$ and $0\leq\varepsilon\leq1$ the trajectory going out of $(\varepsilon,1-\varepsilon)$ converges to an equilibrium.



A necessary and sufficient condition for a strategy to be evolutionarily stable under PFR

Suppose *m* strategies $s_1,\ldots,s_m$ are available to 100 players, but only s_1 is actually used in bilateral contests by all of them. What happens if one player switches to strategy s_2 while 99 players go on playing s_1? If s_1 is evolutionarily stable under PFR, then the deviant (“mutant”) will leave the population (“die”) or will play s_1 back again.

The problem of characterizing strategies which are evolutionarily stable under PFR was dealt with by Bendor and Świstak (1992, unpublished paper). According to their theorem given below every evolutionarily stable strategy is the best reply to itself, but collective stability alone does not entail evolutionary stability.

Theorem 3. In a population playing a symmetric game with strategies $s_1,\dots,s_m$ and payoff matrix (u_{ij}), a strategy s_i is evolutionarily stable under PFR if and only if for all $h\neq i$, $u_{ii} \geq u_{hi}$, if $u_{ii} = u_{hi}$, then $u_{ih} > u_{hh}$.

Proof of necessity

To simplify notation, let i=1. If s_1 is evolutionarily stable under PFR in the game with m strategies, then so it is in the game with two strategies: s_1 and s_h. Let h=2, again to simplify notation. Assume for an indirect proof that (a) $u_{11}<u_{21}$ or (b) $u_{11}=u_{21}$ and $u_{12}\leq u_{22}$. Put $d_1=u_{21}-u_{11}$, $d_2=u_{22}-u_{12}$ and rewrite $V_2(x)-V_1(x)$ as $x_1d_1+x_2d_2$.

Consider case (b) first. Then, for any x, $V_2(x)-V_1(x)=x_2d_2\geq 0$, so that $V_2(x)\geq V_1(x)$. Hence, $V_2(x)\geq \overline{V}(x)$ which implies in turn that $F(x)_2=\dfrac{V_2(x)}{\overline{V}(x)}x_2\geq x_2$ We conclude that $F^t(x)_2\geq x_2$ for any t , so that *lim* $F^t(x)_2 \geq x_2$ for any $x_2>0$, which contradicts the assumption that (1,0) is a stable equilibrium. Indeed, stability implies that *lim* $F^t(x)_2=0$ if x is sufficiently close to (1,0), that is, if $x_2<r$ for some $r>0$.

If (a) is the case, then $V_2(x)-V_1(x)=(1-x_2)d_1+x_2d_2=d_1-x_2(d_1-d_2)$ and $d_1>0$. There exists an $r'>0$ such that $d_1-x_2(d_1-d_2)>0$ if $x_2<r'$. Therefore, $V_2(x)>V_1(x)$, $V_2(x)>\overline{V}(x)$ $F(x)_2=\dfrac{V_2(x)}{\overline{V}(x)}x_2>x_2$ if $x_2<r'$. Let us fix an x such that $0<x_2<r''=\mathrm{Min}\{r,r'\}$. Then *lim* $F^t(x)_2=0$, which means that there is an integer $t'\geq 0$ such that $F^t(x)_2<r'$ for all $t\geq t'$. The sequence $(F^t(x)_2)$, $t=t',t'+1,\dots$ whose terms are bounded from above by 1 and increase with t must converge to a number greater than 0 (because $F^{t'+1}(x)_2>F^{t'}(x)_2\geq 0$), which is a contradiction because the sequence beginning from t' has the same limit 0 as the whole sequence $(F^t(x)_2)$, $t=0,1,\dots$

Proof of sufficiency (complete only for m=2))

To simplify notation, let i=1. Assume that for any $h\neq 1$, $u_{11}\geq u_{h1}$; if $u_{11}=u_{h1}$, then $u_{1h}>u_{hh}$.

(1) The first step will be to show that for $h=2,\dots,m$, $V_1(x)>V_h(x)$ for any $x\neq z^1$ in some neighborhood $U(z^1,r_h)$ of $z^1=(1,0,\dots,0)$. For h=2, $V_1(x)=x_1u_{11}+x_2u_{12}+\dots+x_mu_{1m}$, $V_2(x)=x_1u_{21}+x_2u_{22}+\dots+x_mu_{2m}$, so that $V_1(x)-V_2(x)= x_1e_1+x_2e_2+\dots+x_me_m$ where $e_j=u_{1j}-u_{2j}$, $j=1,\dots,m$. We distinguish two cases: (a) $u_{11}>u_{21}$, or $e_1>0$; (b) $u_{11}=u_{21}$ and $u_{12}>u_{22}$, or $e_1=0$ and $e_2>0$.

(a) Since the linear function $f(x)=x_1e_1+x_2e_2+\dots+x_me_m$ is continuous and $f(z^1)=e_1>0$, there must exist a neighborhood U of z^1 such that $f(x)>0$ for any $x\in U$.

(b) If m=2, then $V_1(x)-V_2(x)=x_2e_2>0$ for any $x_2>0$. **How to prove this for m>2 ?**

(2) We conclude that $V_1(x)>V_h(x)$, $h=2,\dots,m$, $x\in U=U(z^1,r)$, $x\neq z^1$, where $r=\mathrm{Min}\ r_h$. Hence, for any $x\in U$, $V_1(x)>V\text{–}(x)$, which implies that $F(x)_1>x_1$. Notice that $F(x)\in U$ for $x\in U$. The terms of the increasing sequence $F^t(x)_1$ are bounded from above by 1, so the sequence has a limit which must be equal to 1.

Problems for further study

How to prove (or possibly disprove) Theorem 3 for symmetric games with more than 2 strategies? Is the sequence $F^t(x)$ convergent for any x? (does evolution always lead to an equilibrium?) Can other equilibria (those in which two or more strategies have non zero frequencies) be stable?

Temat „Gry w populacjach. Procesy ewolucyjne” włączyłem do programu mojego kursu „Teoria gier i decyzji dla socjologów i psychologów” po raz pierwszy w roku akademickim 2004/05. W maju 2005 przygotowałem dla siebie notatkę do wykładu (tekst zredagowałem po angielsku). Za główne źródło (poza klasyczną książką Axelroda *The Evolution of Cooperation*) posłużył mi artykuł J. Bendora i P. Świstaka. “The Evolutionary Stability of Cooperation.” (*American Political Science Review* 91 (1997): 290–307; tłum. „Ewolucyjna stabilność kooperacji”. *Studia Socjologiczne* 1998 nr 3: 127–171). Czytając tę pracę, starałem się zrozumieć *sens matematyczny* przedstawionych tam wywodów (w tym sens kluczowego pojęcia „ewolucyjna stabilność”), a gdy pojawiły się trudności, uznałem, że łatwiejsze będzie samodzielne poruszanie się (tzn. korygowanie niejasności w definicjach i dowodzenie twierdzeń na własną rękę). Stąd określenie mojej notatki jako ”Lecture note evolving into a research note”. Utknąłem, próbując dowieść Twierdzenia 3 dla populacji z większą od 2 liczbą strategii (wszakże nie mam pewności czy twierdzenie to jest adekwatną rekonstrukcją oryginalnego twierdzenia Bendora i Świstaka) i dalsze badania odłożyłem na później.

Powtarzając kurs w roku akademickim 2005/2006, zdecydowałem się udostępnić słuchaczom (i wszystkim innym zainteresowanym) swoją prywatną notatkę, umieszczając ją na stronie domowej. Może ktoś pomoże mi doprowadzić badania do końca?

Do notatki dołączyłem dwa programy (pliki *tournmnt.exe* i *ipd.exe* spakowane razem z plikiem *ipd.pdf* w pliku *ipd.zip*): napisany w 1998 roku program ilustrujący turniej Axelroda, oraz nowy program (maj 2005) pokazujący przebieg procesu ewolucji (PFR) w populacji, w której dostępne są trzy strategie w Powtarzanym Dylemacie Więźnia: ALL D, ALL C i TFT (wypłaty w grze pojedynczej mają wartości jak u Axelroda). Użytkownik proszony jest najpierw o podanie wartości parametru *delta*, a następnie początkowego rozkładu strategii ALL D i ALL C – dwu liczb, których suma nie przekracza 1; częstość TFT program oblicza odejmując tę sumę od 1.

http://www.cyf-kr.edu.pl/~ussozans/