Role of Pavlov-like Strategy on Emergence of Cooperation in Demographic Donor-Recipient Game*

Tsuneyuki Namekata
Information and Management Science
Otaru University of Commerce

Yoko Namekata

Abstract

We deal with Pavlov-like strategy as well as Tit for Tat-like strategy in Demographic Donor-Recipient (DR) game. We study the role of Pavlov-like strategy on the emergence of cooperation by Agent-Based Simulation.

We extend Tit for Tat (TFT) and Pavlov (Pav) up to three states from two and call them TFT-like and Pavlov-like strategy, respectively. Unlike TFT-like, Pav-like has the following feature: Pav-like changes to using C from using D or remains in using D if he is using D and experiences opponents' D's or C's, respectively. Thus we expect that some Pavlov-like strategies in the population may soften the tendency toward defection of the whole population and also the tendency toward full cooperation of the whole population. Although sole Pavlov-like strategy is not so effective to promote the cooperation, we found case where the cooperation emerges more frequently with both TFT-like and Pavlov-like strategy than with sole TFT-like (or Pav-like) strategy.

Keywords: Pavlov, Donor-Recipient game, emergence of cooperation, generalized reciprocity, Agent-Based Simulation

1 Introduction

This paper investigates the role of Pavlov-like strategy on the emergence of cooperation in Demographic DR game.

Epstein[1] introduces demographic model. He shows the emergence of cooperation where AllC and AllD are initially randomly distributed in a square lattice of cells. Here AllC always Cooperates and AllD always Defects. In each

*This article is the revised version of the original paper published in Vaclav Kratochvil and Jirina Vejnarova, eds. Proceedings of the 16th Czech-Japan Seminar on Data Analysis and Decision Making under Uncertainty (held in September 19-22, 2013 Marianske Lazne, Czech Republic). Faculty of Management University of Economics, Jindrichuv Hradec, Czech Republic, 2013.

period, players move locally and play Prisoner's Dilemma (PD) game against local player(s). If wealth (accumulated payoff) of a player becomes negative or his age becomes greater than his lifetime, he dies. If his wealth becomes greater than some amount and there is an unoccupied local cell, he has an offspring and gives the offspring some amount from his wealth.

Namekata and Namekata[2] extend Epstein's original model discussed above by introducing global move, global play, Reluctant players, who delay replying to changes and use extended forms of TFT, into demographic PD game and consider the effect of Reluctant players on the emergence of cooperation, and show cases where the reluctance promotes the emergence of cooperation. Here TFT Cooperates at first game and at later games uses the same move as the opponent did in the previous game. Namekata and Namekata[3] examine the effect of move-play pattern on the emergence of cooperation and the distribution of strategies. They restrict patters of move and play of a player to simple structure; local or global, where local or global means that with high probability the player moves (plays) locally or globally, respectively. For example, a player with global move and local play (abbreviated as *gl*) moves globally with high probability and plays DR games against (possibly different) local opponents with high probability at each period. They show that cooperative strategies evolutionarily tend to move and play locally, defective strategies do not, and AllC and AllD are abundant unless all strategies initially play locally.

Nowak and Sigmund[4] consider the emergence of cooperation in infinitely repeated PD game. Population consists of strategies that depend on one's own move as well as the opponent's at the last game, i.e., $(p_{\mathrm{CC}}, p_{\mathrm{CD}}, p_{\mathrm{DC}}, p_{\mathrm{DD}})$ where p_{XY} is the probability with which C is used at this game given that the outcome of the last game is XY. They do not use Demographic model. Players play *infinitely* repeated PD game at each period instead of one-shot PD game against randomly selected opponent. The frequency of each strategy in population at the next period is proportional to its payoff at this period. They show that not TFT but Pavlov $(0.999, 0.001, 0.007, 0.946)$ is most abundant strategy in the population in the long run. They argue that Pavlov's success is based on the following two advantageous features compared with TFT in *infinitely* repeated PD game: (1) Pavlovs can correct inadvertent defection and return to mutual cooperation. (2) Pavlov can exploit AllC.

We deal with Pavlov-like strategy as well as Tit for Tat-like strategy in *finitely* repeated Demographic Donor-Recipient (DR) game. Pavlov (Pav) is known to be one of the basic strategies in dilemma situations as well as Tit for Tat (TFT). TFT and Pav have two inner states whose label C(ooperate) or D(efect) indicates their current move. The state in the next game is determined based on the current opponent's move, differently between TFT and Pav. The next state of TFT is the immediate neighbor of the current state toward C or D (if possible) in case of the current opponent's C or D, respectively. On the other hand, that of Pav remains the same if the current opponent uses C or is changed from C to D (or from D to C) if the current opponent uses D, respectively. Alternatively Pav is described as follows: Win Stay, Lose Shift, that is, Pav remains in the same move if he feels comfortable, whereas Pav changes his move if he feels uncomfortable, because we configure the payoff matrix so that the payoff is positive if the opponent uses C or negative if the opponent uses D.

In this paper, we extend TFT and Pav up to three states and call them TFT-like and Pavlov-like strategy, respectively. Pavlov-like changes to using C from using D or remains in using D if he is using D and experiences opponents' D's or C's, respectively. Thus we expect that some Pavlov-like strategies in the population may soften the tendency toward defection of the whole population and also the tendency toward full cooperation of the whole population. We examine initial distribution of strategies that promote the emergence of cooperation and study the role of Pav-like strategy on the emergence of cooperation.

2 Model

We start with extending TFT and Pav as follows in order to introduce TFT-like and Pavlov-like (Pav-like) strategy. The idea is to introduce reluctance to immediate reply to its opponent's change: Let $m = 0, \ldots, n; t = 0, \ldots, m+1; s = 0, \ldots, m$. Strategy $(m, t; s)X$ is illustrated in Figure 1 where X is T for TFT-like or P for Pav-like. It has $m+1$ inner states. The inner states are numbered $0, \ldots, m$; thus m is the largest state number. State i is labeled D_i if $i < t$ or C_i if not. If current state is labeled C or D, then the strategy prescribes using C or D, respectively. In other words, the strategy prescribes using D if the current state $i < t$ and using C if not; thus the value t is the threshold which determines the move of a player. Initial state in period 0 is state s; its label is D_s if $s < t$ or C_s if not. If current state is i, then the next state of TFT-like is $\min\{i+1, m\}$ or $\max\{i-1, 0\}$ given that the opponent uses C or D, respectively, in this game. If current state is i and $i \geq t$, then the next state of Pav-like is $\min\{i+1, m\}$ or $\max\{i-1, 0\}$ given that the opponent uses C or D, respectively, in this game. If current state is i and $i < t$, then the next state of Pav-like is $\max\{i-1, 0\}$ or $\min\{i+1, m\}$ given that the opponent uses C or D, respectively, in this game. Thus TFT-like and Pav-like strategies act differently if their current state $i < t$; TFT-like strategy in De-

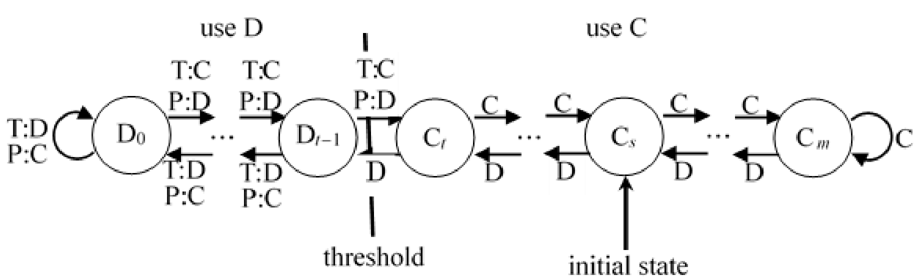

Figure 1: TFT-like and Pav-like strategies

fective state ($i < t$) tends to use the same move as the opponents, whereas Pav-like in Defective state ($i < t$) tends to use the opposite move as the opponents. If $m > 1$, then the strategy may delay replying to its opponent's change. Note that TFT or Pav is expressed as $(1,1;1)$T or $(1,1;1)$P, respectively, in this notation. Thus strategy $(m,t;s)X$ is an extended form of TFT or Pav. To sum up, our strategies are expressed as $(m,t;s)X$; m is the largest state number, t is the threshold, and s is the initial state number, X denotes TFT-like or Pav-like. We omit the initial state like $(m,t;*)X$ if it is determined randomly. We also omit the initial state like $(m,t)X$ if we have no need to specify it.

Table 1: Payoff Matrix of DR game ($b = 4.5$ and $c = 1$)

		Recipient
Donor	C	$-c-x, b-x$
	D	$-x, -x$

Table 2: Initial distribution of inheriting properties

property:	initial distribution
strategy:	We deal with 3 populations, $\mathrm{T}(x,m)$, $\mathrm{TP}(x,m)$, and $\mathrm{P}(x,m)$ for $x = 0.05$, 1/4 or 1/6 and $m = 4$ or ∞ as follows: $\mathrm{T}(x,m) := \{x\mathrm{AllD}(m), \frac{1-x}{2}(2,2;*)\mathrm{T}, \frac{1-x}{2}(2,1;*)\mathrm{T}, x\mathrm{AllC}(m)\}$, $\mathrm{TP}(x,m) := \{x\mathrm{AllD}(m), \frac{1-x}{4}(2,2;*)\mathrm{P}, \frac{1-x}{4}(2,2;*)\mathrm{T}, \frac{1-x}{4}(2,1;*)\mathrm{P}, \frac{1-x}{4}(2,1;*)\mathrm{T}, x\mathrm{AllC}(m)\}$, $\mathrm{P}(x,m) := \{x\mathrm{AllD}(m), \frac{1-x}{2}(2,2;*)\mathrm{P}, \frac{1-x}{2}(2,1;*)\mathrm{P}, x\mathrm{AllC}(m)\}$, where $\mathrm{AllC}(m) = (2,0)\mathrm{T}$ for $m = \infty$, $\mathrm{AllC}(m) = (4,1;4)\mathrm{T}$ for $m = 4$, and $\mathrm{AllD}(m) = (2,3)\mathrm{T}$ for $m = \infty$, $\mathrm{AllD}(m) = (4,4;0)\mathrm{T}$ for $m = 4$. The notation, for example, of $\mathrm{T}(x,m)$, means that with probability x strategy $\mathrm{AllC}(m)$ is selected, with probability $\frac{1-x}{2}$ strategy $(2,1;*)\mathrm{T}$ is selected, and so on, where $*$ indicates that initial state is selected randomly. Note that initially 50% of players use C on the average since both $\mathrm{AllC}(m)$ and $\mathrm{AllD}(m)$ are included with the same probability and so are both $(m,t;*)X$ and $(m,m-t+1;*)X$. As reference populations, we also deal with All:=$\{0.5\mathrm{AllD}(\infty), 0.5\mathrm{AllC}(\infty)\}$, All4:=$\{0.5\mathrm{AllD}(4), 0.5\mathrm{AllC}(4)\}$, and 2 inner states versions of $\mathrm{T}(0.05,4)$, $\mathrm{TP}(0.05,4)$, $\mathrm{P}(0.05,4)$.
(rGM, rGP)**:**	We deal with distribution $\{0.25ll, 0.25lg, 0.25gl, 0.25gg\}$. For example, gl means rGM is distributed in interval g and rGP in interval l, where $l := (0.05, 0.2)$ and $g := (0.8, 0.95)$. $\{0.25ll, 0.25lg, 0.25gl, 0.25gg\}$ means rGM and rGP are selected randomly among ll, lg, gl, and gg.

Note that AllC is denoted by $(m,0)$T and AllD by $(m,m+1)$T. If m is large, $(m,1;m)$T and $(m,m;0)$T are very close to AllC and AllD, respectively. We use these pseudo-AllC $(m,1;m)$T and pseudo-AllD $(m,m;0)$T for $m = 4$ later in this paper because we want to relax unrealistic fixed move strategy.

We deal with Donor-Recipient (DR) game as a stage game. DR game is a two-person game where one player is randomly selected as Donor and the other as Recipient. Donor has two moves, Cooperate (C) and Defect (D). C means Donor pays cost c in order for Recipient to receive benefit b ($b > c > 0$).

Table 3: Detailed Description of (1) Move and (2) Play

(1)	With probability rateOfGlobalMove (abbreviated as rGM), a player moves to random unoccupied cell in the whole lattice. If there is no such cell, he stays at the current cell. Or with probability $1 - rGM$, a player moves to random cell in von Neumann neighbors if it is unoccupied. If there is no such cell, he stays at the current cell.
(2)	With probability rateOfGlobalPlay (abbreviated as rGP), the opponent against whom a player plays dilemma game is selected at random from all players (except himself) in the whole lattice. Or with probability $1-rGP$, the opponent is selected at random from von Neumann neighbors (no interaction if none in the neighbors). This process is repeated 8 times. (Opponents are possibly different.)

Defect means Donor does nothing. Recipient has no move. Since it is common in demographic dilemma game that the sum of payoffs of a player, in two successive games once as Donor and once as Recipient, to be positive if the opponent uses C and negative if D and the worst sum of a player is equal to the best sum in absolute value, we transform the original payoffs to new ones by subtracting constant x. Constant x is given by $x = \frac{b-c}{4}$. We set $b = 4.5$ and $c = 1$ in this paper. Table 1 shows the transformed payoff matrix of DR game. We assume that each player plays 8 games against (possibly different) players at each period.

In period $0, N(= 100)$ players (agents) are randomly located in 30-by-30 lattice of cells. The left and right borders of the lattice are connected. If a player moves outside, for example, from the right border, then he comes inside from the left border. So are the upper and lower borders. Players use strategies of $(m, t; s)X$ form. Initial wealth of every player is 6. Their initial (integer valued) age is randomly distributed between 0 and deathAge (= 50). In each period, each player (1st) moves, and (2nd) plays DR games given by Table 1 against other players. Positive payoff needs opponent's C. (The detailed description of (1st) move and (2nd) play is given in Table 3.) The payoff of the game is added to his wealth. If the resultant wealth is greater than fissionWealth (= 10) and there is an unoccupied cell in von Neumann neighbors, the player has an offspring and gives the offspring 6 units from his wealth. His age is increased by one. If the resultant wealth becomes negative or his age is greater than deathAge (= 50), then he dies. Then next period starts.

In our simulation we use synchronous updating, that is, in each period, all players move, then all players play, then all players have an offspring if possible. We remark that the initial state of the offspring's strategy is set to the current state of the parent's strategy. There is a small mutationRate (= 0.05) with which inheriting properties are not inherited. Initial distributions of inheriting properties given in Table 2 are also used when mutation occurs. We assume that with errorRate (= 0.05) a player makes mistake when he makes his move. Thus AllC may defect sometime. If population consists of AllC and AllD, $rGM = 0$, and $rGP = 0$, then our model is similar to that of Epstein[1]. His model uses asynchronous updating while our model uses synchronous updating.

3 Simulation and Results

We use Ascape (http://sourceforge.net/projects/ascape/) to simulate our model. We execute 300 runs of simulations in each different setting. We judge that the cooperation emerges in a run if there are more than 100 players and the average C rate (average Cr) is greater than 0.2 at period 500, where the average Cr at a period is the average of the player's Cooperation rate (Cr) at the period over *all* players and the player's Cr at the period is defined as the number of move C used by the player divided by the number of games played as Donor at the period. (We interpret 0/0 as 0.) This average Cr is the rate at which we see cooperative move C as an outside observer. Since negative wealth of a player means his death in our model and he has a lifetime, it is necessary for many players to use C in order that the population is not extinct. We focus on emergence rate of cooperation that is the rate at which the cooperation emerges.

Table 4: Ce for pure AllC and AllD

$m = \infty$	All	T	TP	P
Ce(equal)	.473	.630	.557	.447
Ce($x = .05$)	.473	.657	.717	.420

Table 5: Ce, average actual Cr

$m = 4$	All4	T	TP	P
Ce($x = .05$)	.513	.487	.703	.590
aaCr$(4, 1)$T	.900	.931	.805	.804
aaCr$(4, 4)$T	.560	.636	.214	.222
2 states case	.513	.660	.490	.193

We are interested in cases where the cooperation emerges more frequently with both TFT- and Pavlov-like strategy than with sole Pavlov-like strategy and then than with sole TFT-like strategy. We examine how often the cooperation emerges in Demographic DR game with several different initial distributions of strategies. Pure AllC and AllD, and even with their low frequency 0.05 at period 0 prevent Pav-like strategy from promoting cooperation as shown in Table 4, e.g., $.557 < .630$ and $.420 < .657$. Table 4 shows emergence rate of cooperation Ce's for equal frequency at period 0 in the second row and for low 0.05 frequency in the third row, in pure AllC and AllD population. All column in Table 4 indicates pure AllC and AllD population defined in Table 2, T and P columns indicate the corresponding $x = 1/4$ and $x = 0.05$ population defined in Table 2 in case of $m = \infty$, and TP column indicates $x = 1/6$ and $x = 0.05$ population defined in Table 2 in case of $m = \infty$. In place of pure AllC and AllD, we use pseudo-AllC $(4, 1; 4)$T and pseudo-AllD $(4, 4; 0)$T with initial low frequency 0.05 at period 0. The emergence rate of cooperation Ce's and other related data are summarized in Table 5. Table 5 shows that Pav-like (.590) and TFT-like + Pav-like (.703) promote the cooperation in this order compared with T (.487). The third and fourth rows in Table 5, aaCr's are the *actual* average Cr of pseudo-AllC $(4, 1)$T and pseudo-AllD $(4, 4)$T. aaCr of a strategy is defined as the average of players' Cr over all player using the strategy and playing *at least one* game as Donor. The sum of aaCr's of pseudo-AllC and pseudo-AllD, for example, 1.46 for All4, is much larger than 1 for All4 and T, but is almost equal to 1 for TP and P. We conclude that introducing pseudo-AllC and pseudo-AllD in place of pure AllC and AllD is reasonable modelling if there exists Pav-like strategy in the population. The fifth row indicates the emergence rates of cooperation Ce of two states case instead of three states case in the second row. Unlike Pav-like and TFT-like + Pav-like, Pavlov (.193) and TFT + Pavlov (.490) do not promote the cooperation compared with T (.660).

Next we investigate the role of Pav-like strategy. We select 35 successful runs of T, TP and P populations, respectively, from data in Table 5. We trace average Cr from period 1 to period 500 at each successful run. We judge average Cr at a period is High(> 0.7) if it is greater than 0.7, or is Low(≤ 0.2) if it is less than or equal to 0.2, or is Middle otherwise. We see in Figure 2 that average Cr is almost High in T population, whereas it is mostly Middle in TP population and it is almost Middle in P population. Thus Pav-like makes average Cr Middle. We want to evaluate easily the change of average Cr over periods. We assign Low to 0 as a new vertical value different from the original value of average Cr, Middle to 1, and High to 2. Then we focus only on their local maximums and minimums. A transition of local optimums is classified into one of $\{-2, -1, 1, 2\}$. Suppose, for example, that local maximum is 2 at some period and the nearest local minimum is 1 at some later period, then the transition is evaluated as -1. We count all transitions of these local optimums over periods in each run. We show average of these number of transitions over 35 runs in Figure 3. We conclude that population T, TP, and P decreases the number of transitions in this order.

Next we concentrate on the av-

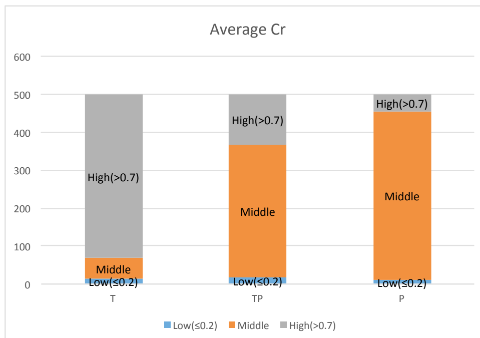

Figure 2: Average Cr

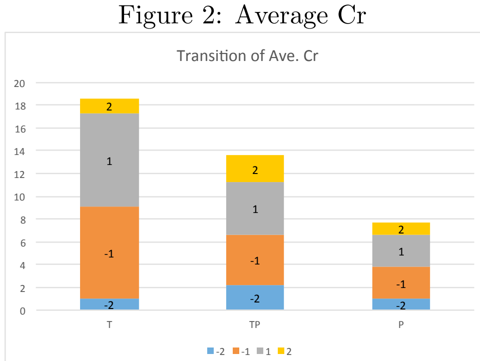

Figure 3: Transition of AveCr

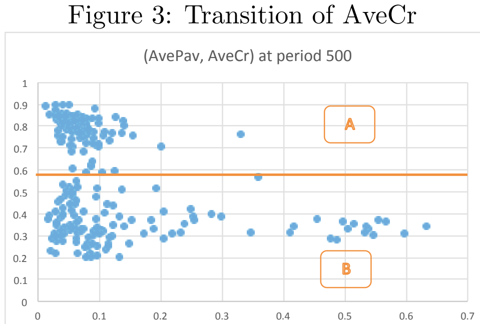

Figure 4: (AvePav,AveCr)

erage frequency of Pav-like strategy and the average Cr at period 500 in TP population. Figure 4 is scatter diagram of (average frequency of Pav-like strategy, average Cr) at period 500 of all successful runs in TP population. For convenience sake, let us divide all successful runs into two cases, A and B; A for average Cr >= 0.57, B for average Cr < 0.57. Figure 4 shows that the larger the average frequency of Pav-like strategy the smaller the average Cr at period 500 in TP population.

Figure 5 and 6 show the average distributions of strategies at period 500 for case A and B, respectively. We see that average frequency of Pav-like strategies, $(2,2)$P and $(2,1)$P, is not so large, around 0.15 even in case B.

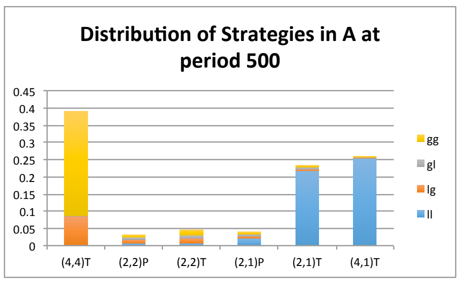

Figure 5: Distribution in A

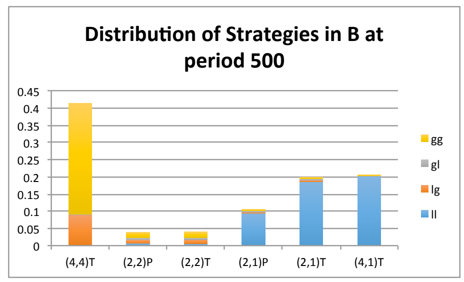

Figure 6: Distribution in B

4 Conclusion

We examine the role of Pav-like strategy on the emergence of cooperation in Demographic DR game by Agent-Based Simulation. We show that some Pav-like strategies promote cooperation and soften the tendency toward defection and toward full cooperation in whole population if there initially are low frequent pseudo-AllC and pseudo-AllD in stead of equal pure AllC and AllD.

References

[1] Epstein, J. M. (2006). "Zones of Cooperation in Demographic Prisoner's Dilemma." In: *Generative Social Science*. Princeton University Press, 199-221.

[2] Namekata, T. and Namekata, Y. 2011. "Effect of Reluctant Players in Demographic Prisoner's Dilemma Game." In: R. Bartak (ed.): *Proceedings of the 14th Czech-Japan Seminar on Data Analysis and Decision Making under Uncertainty* (held in September 18-21, 2011, Hejnice, Czech Republic), 102-109.

[3] Namekata, T. and Namekata, Y. 2012. "Emergence of cooperation and patterns of move-play in Demographic Donor-Recipient Game." In: Masahiro Inuiguchi, Yoshifumi Kusunoki and Hirosaki Seki (eds.): *Proceedings of the 15th Czech-Japan Seminar on Data Analysis and Decision Making under Uncertainty* (held in September 24-27, 2012 Osaka, Japan), 51-58.

[4] Nowak, M. A. and Sigmund, K. (1993). "A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game." *Nature*, No. 364, 56-58.