Journal of Theoretical Biology 430 (2017) 215–220

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/jtbi

Duality between cooperation and defection in the presence of tit-for-tat in replicator dynamics

Seung Ki Baek[a,*], Su Do Yi[b], Hyeong-Chai Jeong[c,d]

[a] Department of Physics, Pukyong National University, Busan 48513, Republic of Korea
[b] CCSS, Department of Physics and Astronomy, Seoul National University, Seoul 08826, Republic of Korea
[c] Department of Physics and Astronomy, Sejong University, Seoul 05006, Republic of Korea
[d] Quantum Universe Center, Korea Institute for Advanced Study, Seoul 02455, Republic of Korea

A R T I C L E I N F O

Article history:
Received 7 July 2017
Accepted 25 July 2017
Available online 26 July 2017

Keywords:
Iterated prisoner's dilemma
Evolution of cooperation
Mutation

A B S T R A C T

The prisoner's dilemma describes a conflict between a pair of players, in which defection is a dominant strategy whereas cooperation is collectively optimal. The iterated version of the dilemma has been extensively studied to understand the emergence of cooperation. In the evolutionary context, the iterated prisoner's dilemma is often combined with population dynamics, in which a more successful strategy replicates itself with a higher growth rate. Here, we investigate the replicator dynamics of three representative strategies, i.e., unconditional cooperation, unconditional defection, and tit-for-tat, which prescribes reciprocal cooperation by mimicking the opponent's previous move. Our finding is that the dynamics is self-dual in the sense that it remains invariant when we apply time reversal and exchange the fractions of unconditional cooperators and defectors in the population. The duality implies that the fractions can be equalized by tit-for-tat players, although unconditional cooperation is still dominated by defection. Furthermore, we find that mutation among the strategies breaks the exact duality in such a way that cooperation is more favored than defection, as long as the cost-to-benefit ratio of cooperation is small.

1. Introduction

Although a society consists of individuals, the collective interest is not an aggregate of individual ones. The prisoner's dilemma (PD) game is a toy model to illustrate such a social dilemma. The PD game can be formulated as follows: Suppose that we have two players, say, Alice and Bob. When Alice cooperates, it benefits Bob by a certain amount of b at her own cost c. If she defects, on the other hand, it does not incur any cost and Bob gains nothing. If c exceeds b, defection obviously drives out cooperation, so we restrict ourselves to $0 < c < b$. The cost-to-benefit ratio, c/b, is thus limited to an open interval (0, 1). The resulting payoff matrix between cooperation (C) and defection (D) is expressed as

$$\begin{array}{c} \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cc} C & D \end{array} \\ \begin{pmatrix} b-c & -c \\ b & 0 \end{pmatrix} \end{array}, \tag{1}$$

from the row-player Alice's point of view, and the game is symmetric to both players. The collective interest is maximized when both choose C, but D is the rational choice for each individual, hence a dilemma.

By construction of the PD game, unconditional defection (AllD) always constitutes a Nash equilibrium. However, it has been widely known by folk theorems that a cooperative strategy can also be rational if the PD game is repeated *indefinitely* with high enough probability because one's cooperation can be reciprocated by the other's in future. This is called direct reciprocity and has been popularized by Axelrod's tournament of the iterated prisoner's dilemma (IPD) (Axelrod, 1984). We assume that the repetition probability approaches one. An archetypal strategy of direct reciprocity is Tit-for-tat (TFT). It begins with C at the first encounter and then replicates the co-player's last move. Except the first round, therefore, it cooperates only if the co-player cooperated last time. We may call it a conditional cooperator, opposed to an unconditional cooperator (AllC). We will explain that the interactions between the aforementioned strategies, i.e., AllD, TFT, and AllC, are rather subtle, indicating the complexity in evolution of cooperation. Earlier studies have already focused on the dynamics of these three representative strategies (Brandt and Sigmund, 2006; Imhof, et al., 2005; Toupo et al., 2014).

All these fall into a class of reactive strategies (Baek et al., 2016) represented by a two-component array $\alpha = (P_C, P_D)$, where P_C (P_D)

* Corresponding author.
E-mail addresses: seungki@pknu.ac.kr (S.K. Baek), esudoz@snu.ac.kr (S.D. Yi), hcj@sejong.edu (H.-C. Jeong).

http://dx.doi.org/10.1016/j.jtbi.2017.07.026
0022-5193/

means the probability to cooperate when the co-player cooperated (defected) last time. In this notation, we have AllC = (1, 1), AllD = (0, 0), and TFT = (1, 0). If error occurs with probability e at each time step, the effective behavior is described as $\alpha' = ((1-e)P_C + e(1-P_C), (1-e)P_D + e(1-P_D)) = (P'_C, P'_D)$. The error rate e is assumed to be small, and this statement will be made quantitative later. Suppose that two strategies $\alpha = (P_C, P_D)$ and $\beta = (Q_C, Q_D)$ meet in the IPD. They effectively behave as α' and β', respectively, and stochastically visit four states, *CC, CD, DC*, and *DD*, where the former (latter) symbol means the move of the player adopting α (β). The transition probabilities between the states can be arranged in the following matrix (Nowak and Sigmund, 1989, 1990):

$$\tilde{M} = \begin{pmatrix} P'_C Q'_C & P'_D Q'_C & P'_C Q'_D & P'_D Q'_D \\ P'_C(1-Q'_C) & P'_D(1-Q'_C) & P'_C(1-Q'_D) & P'_D(1-Q'_D) \\ (1-P'_C)Q'_C & (1-P'_D)Q'_C & (1-P'_C)Q'_D & (1-P'_D)Q'_D \\ (1-P'_C)(1-Q'_C) & (1-P'_D)(1-Q'_C) & (1-P'_C)(1-Q'_D) & (1-P'_D)(1-Q'_D) \end{pmatrix}. \tag{2}$$

This stochastic matrix is irreducible and positive definite, so the Perron–Frobenius theorem guarantees the existence of a unique right eigenvector $\vec{v} = (v_{CC}, v_{CD}, v_{DC}, v_{DD})$ with the largest eigenvalue $\Lambda = 1$. If we normalize $\vec{v}$ in such a way that $v_{CC} + v_{CD} + v_{DC} + v_{DD} = 1$, it is the stationary probability distribution over the four states when the strategies α and β are adopted in the IPD. The long-term payoff of α against β per round is obtained by calculating an inner product $p_{\alpha\beta} = \vec{v} \cdot \vec{h}_1$, where $\vec{h}_1 = (b-c, -c, b, 0)$. Likewise, we obtain $p_{\beta\alpha} = \vec{v} \cdot \vec{h}_2$ with $\vec{h}_2 = (b-c, b, -c, 0)$. If we list the three strategies in the order of AllC, AllD, and TFT, the matrix $\tilde{p} = \{p_{\alpha\beta}\}$ can be written as follows:

$$\tilde{p} = \begin{pmatrix} (b-c)(1-e) & be - c(1-e) & b(1-2e+2e^2) - c(1-e) \\ b(1-e) - ce & (b-c)e & 2b(1-e)e - ce \\ b(1-e) - c(1-2e+2e^2) & be - 2c(1-e)e & (b-c)/2 \end{pmatrix}. \tag{3}$$

Note that the limit of $e \to 0$ does not coincide with the case of $e = 0$: If e was strictly zero between two TFT players, each of them would earn $b - c$ at each round. For any $e > 0$, however, the average payoff per round reduces to $(b-c)/2$ as written in Eq. (3). All these results are fully consistent with existing ones such as in Refs. Molander (1985) and Imhof et al. (2007).

In an evolutionary framework, we consider dynamics of a well-mixed population in which random pairs of individuals play the IPD game. Let us assume that the population is so large that stochastic fluctuations can be ignored. If a certain strategy earns a higher payoff than the population average, we can expect that its fraction will grow at a rate proportional to the payoff difference from the population average. Likewise, a strategy with a lower payoff than the population average will decrease in its fraction. Replicator dynamics (RD) expresses this idea by using a set of deterministic equations for the time evolution of the fractions. Let N_s be the total number of strategies in the population. We have $N_s = 3$ in a set of the three strategies, i.e., {AllC, AllD, TFT}. We are interested in the fraction x_α of strategy α, with a normalization condition that $\sum_\alpha x_\alpha = 1$. The long-term payoff of strategy α from the whole population is denoted as

$$p_\alpha = \sum_\beta p_{\alpha\beta} x_\beta. \tag{4}$$

RD describes the time evolution of x_α as follows:

$$\frac{dx_\alpha}{dt} = \sum_\beta q_{\alpha\beta} p_\beta x_\beta - \langle p \rangle x_\alpha, \tag{5}$$

where $q_{\alpha\beta}$'s are elements of a transition matrix between strategies. The average payoff of the population is denoted as $\langle p \rangle \equiv \sum_\alpha p_\alpha x_\alpha = \sum_{\alpha\beta} p_{\alpha\beta} x_\alpha x_\beta$. If we choose the transition matrix as

$$q_{\alpha\beta} = \begin{cases} 1-\mu & \text{for} \quad \alpha = \beta \\ \mu/(N_s - 1) & \text{for} \quad \alpha \neq \beta, \end{cases} \tag{6}$$

RD takes the following form:

$$\frac{dx_\alpha}{dt} = (1-\mu) p_\alpha x_\alpha - \langle p \rangle x_\alpha + \frac{\mu}{N_s - 1} \sum_{\beta \neq \alpha} p_\beta x_\beta, \tag{7}$$

where μ is a mutation rate, assumed to satisfy $\mu \ll e$. The first term on the right-hand side means growth with a rate proportional to the payoff, the second term normalizes the total sum of x_α's, and the last term describes mutation. Note that the fitness of strategy α is identified with its payoff $p_\alpha(t)$, so that it produces offspring in proportion to $p_\alpha(t) x_\alpha(t)$ between time t and $t + dt$. The mutation structure in Eq. (6) means that some of these offspring are randomly picked up and change the strategy to one of the others.

In this work, we will show the following: If μ vanishes, the time evolution of x_{AllC} in RD is the same as that of x_{AllD} under time reversal, $t \to -t$, and vice versa. The duality does not exactly hold for $\mu > 0$, and we will discuss its consequences by analyzing the system perturbatively.

2. Fixed-point structure

For the sake of notational convenience, we define $x_1 \equiv x_{\text{AllC}}$, $x_2 \equiv x_{\text{AllD}}$, and $x_3 \equiv x_{\text{TFT}}$ henceforth. Due to the normalization condition, we have only two independent variables, which we choose as x_1 and x_2. Plugging Eq. (4) into Eq. (7), we find a set of equations, which can be formally written as follows:

$$\frac{dx_1}{dt} = f_1(x_1, x_2; e, \mu) \tag{8}$$

$$\frac{dx_2}{dt} = f_2(x_1, x_2; e, \mu). \tag{9}$$

After a little algebra, one can show that

$$f_1(x_1, x_2; e, \mu) + f_2(x_2, x_1; e, \mu) = \frac{1}{2}\mu(b-c)(1-3x_1), \tag{10}$$

which becomes zero as μ vanishes. Note that x_1 and x_2 exchange their positions when they are arguments of f_2 in Eq. (10). If we set $\mu = 0$ and define $\tau \equiv -t$, therefore,

$$\frac{dx_1}{d\tau} = -\frac{dx_1}{dt} = -f_1(x_1, x_2; e, 0) = f_2(x_2, x_1; e, 0) \tag{11}$$

$$\frac{dx_2}{d\tau} = -\frac{dx_2}{dt} = -f_2(x_1, x_2; e, 0) = f_1(x_2, x_1; e, 0) \tag{12}$$