Physica A 430 (2015) 81–92

Contents lists available at ScienceDirect

Physica A

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

Evolutionary behavior of generalized zero-determinant strategies in iterated prisoner's dilemma

Jie Liu [a], Y. Li [b], C. Xu [b,*], P.M. Hui [a,*]

[a] *Department of Physics, The Chinese University of Hong Kong, Shatin, Hong Kong*

[b] *College of Physics, Optoelectronics and Energy, Soochow University, Suzhou 215006, China*

HIGHLIGHTS

- Studied how generalized ZDS competes in well-mixed and structured populations.
- Presented expected payoffs for ZDS against AllC, AllD, TFT, and WSLS strategies.
- For finite and well-mixed populations, explained phase diagrams analytically.
- For square lattice, explained results by payoff band theory qualitatively.

ARTICLE INFO

Article history:
Received 15 December 2014
Received in revised form 7 February 2015
Available online 26 February 2015

Keywords:
Zero-determinant strategies
Iterated prisoner's dilemma
Evolutionary games
Structured populations

ABSTRACT

We study the competition and strategy selections between a class of generalized zero-determinant (ZD) strategies and the classic strategies of always cooperate (AllC), always defect (AllD), tit-for-tat (TFT), and win-stay-lose-shift (WSLS) strategies in an iterated prisoner's dilemma comprehensively. Using the generalized ZD strategy, a player could get a payoff that is χ ($\chi > 1$) times that of his opponent's, when the payoff is measured with respect to a referencing baseline parameterized by $0 \leq \sigma \leq 1$. Varying σ gives ZD strategies of tunable generosity from the extortionate-like ZD strategy for $\sigma \ll 1$ to the compliance-like strategy at $\sigma \approx 1$. Expected payoffs when ZD strategy competes with each one of the classic strategies are presented. Strategy evolution based on adopting the strategy of a better performing neighbor is studied in a well-mixed population of finite size and a population on a square lattice. Depending on the parameters, extortion-like strategies may not be evolutionarily stable despite a positive surplus over cooperative strategies, while extortion-like strategies may dominate or coexist with other strategies that tend to defect despite a negative surplus. The dependence of the equilibrium fraction of ZD strategy players on the model parameters in a well-mixed population can be understood analytically by comparing the average payoffs to the competing strategies. On a square lattice, the success of the ZD strategy can be qualitatively understood by focusing on the relative alignments of the finite number of payoff values that the two competing strategies could attain when the spatial structure is imposed. ZD strategies with properly chosen generosity could be more successful in evolutionary competing systems.

1. Introduction

The emergence of cooperation or altruism among egoists has long been a fascinating and important research topic [1–5]. The Prisoner's Dilemma (PD) provides a paradigm for studying possible cooperative behavior among selfish players in

* Corresponding authors.
E-mail addresses: cxu@suda.edu.cn (C. Xu), pmhui@phy.cuhk.edu.hk (P.M. Hui).

http://dx.doi.org/10.1016/j.physa.2015.02.080

evolutionary games. The PD is set up as a two-player game in which each player has two possible actions: to cooperate (C) or to defect (D) (not to cooperate). The payoff to each player depends on the actions of the two players. If they cooperate, i.e. situation CC indicating the actions of the first and second players, each gets a reward payoff R. If both players use D, i.e., situation DD, each gets a punishment payoff P. When one player chooses C and the other chooses D, i.e., situations CD and DC, the cooperator gets a payoff S and the defector gets T. In PD, the payoffs are ordered as $T > R > P > S$ [6], with $2R > T + S$. For a single-round PD, using D is the rational choice as $T > R$ and $P > S$. Referring to the four situations in the order of (CC, CD, DC, DD), the payoffs can be represented by a vector $\mathbf{S}_1 = (R, S, T, P)$ for the first player and another vector $\mathbf{S}_2 = (R, T, S, P)$ for the second player.

If the players are to play the game repeatedly, however, cooperation may be possible. In 1980s, Axelrod [1,2] established the Iterated PD (IPD) in which players compete in infinite rounds of PD. Players in IPD need to consider the strategy for deciding on C or D action in the forthcoming round based on the competing outcomes in the past rounds. It is common to assume that the players only remember the outcomes in the last round, i.e., a memory-one game [7–13]. As the possible outcomes in a memory-one IPD are the four situations (CC, CD, DC, DD), and the corresponding reaction strategies can be described by an array of four probabilities ($p_{CC}, p_{CD}, p_{DC}, p_{DD}$), each giving the probability of taking the C action if the outcome of the most recent round is that labeled by the subscripts. Therefore, the players have access to a large strategy pool described by the set of ($p_{CC}, p_{CD}, p_{DC}, p_{DD}$). Early studies showed that the Tit-for-Tat (TFT) strategy described by (1, 0, 1, 0), which amounts to imitate the opponent's latest action and thus retaliate immediately to any harsh action D of the opponent, is a highly successful strategy [1]. However, later studies showed that TFT is not evolutionarily stable [14]. It cannot dominate the AllC strategy of always playing C as described by (1, 1, 1, 1) [15] and it may coexist with the AllD strategy of always playing D as described by (0, 0, 0, 0) when players compete on a lattice [16,17]. A more generous TFT strategy in which there is a certain probability of forgiving the opponent's unfriendly D action was shown to dominate in the heterogeneous population [18] or in the spatial lattice [19]. Furthermore, Nowak and Sigmund showed that the WSLS strategy of win-stay-lose-shift as described by (1, 0, 0, 1), where a player switches action after encountering a D-action opponent, outperforms TFT in the IPD [20]. However, WSLS performs poorly against AllD as it will switch strategy every round and thus will be continually exploited by the AllD opponent.

There has been much work in the past two years related to a class of zero-determinant (ZD) strategies in the strategy pool of memory-one game after their proposal by Press and Dyson [10]. The ZD strategy has the property of enforcing a linear relation between the payoffs of a player and his opponent in IPD. Thus, the player using the ZD strategy can unilaterally set his opponent's payoffs. More interestingly, there exist extortionate ZD strategies which could let the extortioner get a higher surplus than that of his opponent, when the payoff is measured relative to P as the referencing baseline payoff. Within the context of PD, an extortioner's payoff is never below the payoff of its opponent [10]. When two players using extortionate ZD meet, they would both get the payoff P. Although the extortionate ZD strategy given in Ref. [10] would catalyze cooperation [12], it is evolutionarily unstable [12,13]. Stewart and Plotkin [11] showed that a more generous ZD strategy gets the highest scores among the important strategies such as AllC, AllD, TFT, WSLS and other strategies in a tournament similar to that of Axelrod [21,22], and the strategy can outperform classic IPD strategies and dominate in a large population [23]. It indicates that there may exist robust ZD strategies beyond the one given in Ref. [10] that can have an edge in the two-player IPD. More recently, studies on the effectiveness of ZD strategies have been extended to multi-player games [24–28] as well as to IPD with co-evolution of strategies and payoffs [29]. Furthermore, when there exist multiple time scales in a co-evolving system, the interplay between the time scales becomes an important factor [30] and it has been shown that the different time scales could help sustain the extortioners in the system [12]. In the present work, we study a class of ZD strategies that extend the previous extortionate ZD strategies [10,23] and investigate how the ZD strategies compete with the other important strategies, i.e., AllC, AllD, TFT, and WSLS, in populations that are well-mixed and with a spatial structure of a two-dimensional square lattice.

The plan of the paper is as follows. In Section 2, we define a class of generalized ZD strategies in which the excessive payoff to the players is measured with respect to a tunable baseline value. In Section 3, the expected payoffs when the generalized ZD strategy competes against the classic IPD strategies of AllC, AllD, TFT and WSLS in a two-person IPD are evaluated and discussed. The calculated expected payoffs are then used in an evolutionary game in which the players may imitate the strategy of their better performing neighbors. The case of generalized ZD strategy competing with each of the classic strategies in a well-mixed population is discussed in Section 4. Both simulation and analytic results are presented. In corresponding case of a population on a square lattice is discussed in Section 5. A qualitative explanation of the strategy selection dynamics is presented within the physical picture of alignments of payoff levels. A summary is given in Section 6.

2. A class of generalized ZD strategies

The class of ZD strategies ($p_{CC}, p_{CD}, p_{DC}, p_{DD}$) among the memory-one strategies can be described by [10]

$$
\begin{aligned}
p_{CC} &= 1 + \alpha R + \beta R + \gamma, \\
p_{CD} &= 1 + \alpha S + \beta T + \gamma, \\
p_{DC} &= \alpha T + \beta S + \gamma, \\
p_{DD} &= \alpha P + \beta P + \gamma,
\end{aligned} \tag{1}
$$