
Access from the University of Nottingham repository:
http://eprints.nottingham.ac.uk/30423/1/PID3941653.pdf

Copyright and reuse:

The Nottingham ePrints service makes this work by researchers of the University of Nottingham available open access under the following conditions.

This article is made available under the University of Nottingham End User licence and may be reused according to the conditions of the licence. For more details see:
http://eprints.nottingham.ac.uk/end_user_agreement.pdf

A note on versions:

The version presented here may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the repository url above for details on accessing the published version and note that access may require a subscription.

For more information, please contact eprints@nottingham.ac.uk
Computing Nash Equilibria and Evolutionarily Stable States of Evolutionary Games

Jiawei Li, Member, IEEE, Graham Kendall, Senior Member, IEEE, and Robert John, Senior Member, IEEE

Abstract—Stability analysis is an important research direction in evolutionary game theory. Evolutionarily stable states have a close relationship with Nash equilibria of repeated games, which are characterized by the folk theorem. When applying the folk theorem, one needs to compute the minimax profile of the game in order to find Nash equilibria. Computing the minimax profile is an NP-hard problem. In this paper we investigate a new methodology to compute evolutionary stable states based on the level-k equilibrium, a new refinement of Nash equilibrium in repeated games. A level-k equilibrium is implemented by a group of players who adopt reactive strategies and who have no incentive to deviate from their strategies simultaneously. Computing the level-k equilibrium is tractable because the minimax payoffs and strategies are not needed. As an application, this paper develops a tractable algorithm to compute the evolutionarily stable states and the Pareto front of n-player symmetric games. Three games, including the iterated prisoner’s dilemma, are analyzed by means of the proposed methodology.

Index Terms—Evolutionary game theory, evolutionary stability, folk theorem, iterated prisoner’s dilemma, Nash equilibrium.

I. INTRODUCTION

Evolutionary game theory has been successful in helping to explain many complex and challenging aspects of biological and social phenomena in recent decades [1, 2]. Based on the idea that biological organisms that are more fit in a given environment will tend to produce more offspring, evolutionary game theory provides us with the methodology to study strategic interactions among individuals in evolving populations.

Evolutionary stability analysis is one of the major research directions in evolutionary game theory. It considers strategic interactions in the situations when mutant strategies invade an infinite or finite population of homogeneous or heterogeneous strategies, without concerning the parameters of evolutionary dynamics such as the selection scheme [3, 28].

Classical evolutionary stability analysis is based on the concept of evolutionarily stable strategy (ESS). An ESS is a strategy such that, if all the members of a population adopt it, then no mutant strategy can invade the population under the influence of natural selection [3]. According to [4], the condition for a strategy \( x \) to be ESS is that for any strategy \( y \neq x \),

\[
\begin{align*}
&u(x, x) \geq u(y, x) \\
&u(x, y) > u(y, y)
\end{align*}
\]  

where \( u(x, y) \) is the payoff of strategy \( x \) when interacting with another strategy \( y \).

This condition guarantees that an ESS always outperforms mutant strategies so that a homogeneous population can be maintained in evolutionary dynamics. However, the definition of ESS is so strict that there is frequently no ESS in an infinite length or indefinite length two-player [5, 6] or \( n \)-player [7] repeated game. Except in specific situations, the condition of ESS cannot be used to analyze the evolutionary stability of strategies in evolutionary games.

A population is considered to be in an evolutionarily stable state if its genetic composition is restored by selection after a disturbance [8]. In evolutionary game theory, an evolutionarily stable state has a close relationship with Nash equilibrium (NE) and the folk theorem for infinitely repeated games [9, 10]. The folk theorem, which is the fundamental theory of non-cooperative repeated games, states that any feasible payoff profile that strictly dominates the minimax profile is a NE profile in an infinitely repeated game [11, 12]. The evolutionary stable states are a subset of those Nash equilibria in the corresponding evolutionary game.

Computing a NE is generally hard in either a one-shot game or a repeated game. It has been proved in complexity theory that computing whether a two-player \( n \times n \) game has any NE in which both players get non-negative payoffs is NP-hard [13]. Recent results have shown that the problem of finding a NE is PPAD-complete, even for a two-player game, and even if all payoffs are \( \pm 1 \), which suggests that these problems are as hard as finding Brouwer fixed-point and thus are computationally intractable [14-16]. Computing an approximate NE, such as \( \varepsilon \)-NE, is also PPAD-complete [17]. In infinitely repeated games, finding a NE or \( \varepsilon \)-NE of \((k+1)\)-player games is as hard as finding NE of \( k \)-player one-shot games [18], unless some changes are made to simplify the model [19].
The main difficulty of this problem lies in computing the minimax profile because the strategy space of a repeated game is much more complicated than that of a one-shot game. In [20], a polynomial algorithm for finding a NE in a two-player average-payoff repeated game is designed by restricting the strategies of the players. Some learning algorithms, fictitious play for example [47, 49], assume that each player adopts a stationary strategy at each round. These methods are flawed if reactive strategies are taken into account. The computational complexity of finding a NE for a repeated game is still an open problem.

Some specific subsets of NE are tractable. Classical game theorists have developed a number of refinements of NE, such as subgame perfect NE [21], strong NE [22], and coalition proof NE [23, 24], because many games, especially repeated games, have multiple equilibria and those equilibria can be significantly different in terms of simplicity, stability, and commitment. One can refine equilibria to distinguish them in which implicit commitments are credible due to incentives [25]. Some refinements of NE require less computational complexity than computing the Brouwer fixed-point or the minimax payoff profile. Unlike a NE, however, none of the refinements of NE so far guarantees its existence in a game.

In this paper, we propose a new refinement of NE in repeated games, namely the level-k equilibrium, in order to develop tractable and general algorithms to compute NE of repeated games. Based on this concept, we show that specific subsets of NE are tractable. Classical game theories of the players. Some learning algorithms, fictitious average-payoff repeated game is designed by restricting the sources of disturbance.

II. EVOLUTIONARILY STABLE STATE AND NASH EQUILIBRIUM

An evolutionarily stable state is a state such that a disturbance cannot change the genetic compositions of the population, if that disturbance is not too large. Besides intrinsic disturbances such as stochastic decisions and noise in game dynamics, crossover and mutation in selection process are also sources of disturbance.

The dynamic of an evolutionary game is generally expressed by the replicator equation,

$$\dot{x}_i = x_i(u_i(x) - \bar{u}(x))$$

(2)

where $x = (x_1, \ldots, x_n)$ is the vector of the distribution of types $i = 1, \ldots, n$ in the population, $u_i(x)$ is the fitness (payoff) of type $i$, and $\bar{u}(x) = \sum_{j=1}^n x_j u_j(x)$ is the average population fitness. It is generally difficult to acquire accurate solutions of the replicator equation.

A state $\hat{x}$ is said to be evolutionary stable if for all $x \neq \hat{x}$ in some neighborhood of $\hat{x},$

$$u_i(x) \geq \bar{u}(x) \quad \text{if} \quad x_i \leq \hat{x}_i$$

$$u_i(x) < \bar{u}(x) \quad \text{if} \quad x_i > \hat{x}_i$$

(3)

When the population deviates from the state $\hat{x}$, due to a disturbance, it will be restored by a selection process.

NE is a stable state in strategic games such that no player can be better off by unilaterally deviating from it. Nash has proven that any strategic game has at least one NE if mixed strategies are taken into consideration [26]. However, there is no upper bound of the number of NE. A strategic game may have multiple, sometimes unlimited, NE [48].

Consider an infinitely repeated $n$-player game $G = \{I, S, U\}$ where $I = \{1, \ldots, n\}$ is the player set and $S = \{s_1, \ldots, s_n\}$ and $U = \{U_1, \ldots, U_n\}$ are the strategy set and the payoff set respectively. Let $s_{-i} = (s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_n)$ denote the strategy profile excluding the strategy of player $i$. The iteration of the game is counted by $t$, starting from $t = 0$. Each player has a pure action space $A_i$ in the $t$th stage game.

A strategy profile $(s_1, \ldots, s_n)$ is a NE if, for any $i \in n$ and $s'_i \in S_i$ there are

$$u_i(s_i', s_{-i}) \leq u_i(s_i, s_{-i})$$

(4)

In repeated games, the set of NE includes any feasible payoff profile that dominates the minimax profile, which is characterized by the folk theorem.

The concept of an evolutionarily stable state is equivalent to the concept of strong NE [9]. A strong NE is an equilibrium in which no coalition of a group of players can be better off by deviating from their current strategies simultaneously. In repeated games, a strong NE is both a Pareto optimum and a NE of the stage game.

A strategy profile $(s_1, \ldots, s_n)$ and the corresponding payoff profile is a strong NE in a $n$-player repeated game if, for any $s'_i \neq s_i$ and $s'_{-i} \neq s_{-i}$, there are
These inequalities are necessary and sufficient conditions for (3) to hold in a selection process. (5.a) is the condition of NE while (5.b) is the condition of Pareto optimum. For a two-player symmetric game, (5.a) and (5.b) can be simplified to

\[ u_i(s_i, \bar{s}_{-i}) \geq u_i(s'_i, \bar{s}_{-i}) \] (5.a)
\[ u_i(s_i, \bar{s}_{-i}) \geq u_i(s'_i, s'_{-i}) \] (5.b)

The main difficulty in computing a NE of a repeated game is that it is difficult to find a suitable expression of the strategies of players. There are not only pure strategies and mixed strategies, but also reactive strategies that the choice of a player at time \( t \) is a function of past choices of all players. In iterated prisoner’s dilemma, for example, tit for tat [1], grim trigger, Pavlov [27], group strategies [28], and newly appeared zero-determinant strategies [29] are all reactive strategies, as well as those learning and evolving strategies [30-36]. We will show in the next section that a group of players in a game can coordinate their repeated choices by means of reactive strategies, which may lead to equilibria that are not characterized by the folk theorem.

III. REACTIVE STRATEGIES AND LEVEL-K EQUILIBRIUM

One assumption in classical game theory is that the players believe that a deviation in their own strategy will not cause any other player’s deviation from their strategies. This is not a reasonable assumption in repeated or evolutionary games because of the existence of reactive strategies. A player’s strategy is reactive if it is a function of some other players’ past actions. Here we give a formal definition of a reactive strategy.

Let \( h^t_i = (a^t_i, a^{t-1}_i, \ldots, a^{0}_i) \) be the sequence of actions chosen by player \( i \in I \) within \( t-1 \) periods, and \( h^{t-1}_i = (h^t_i, h^{t-1}_i, h^{t-2}_i, \ldots, h^0_i) \) be the past choices made by all players other than \( i \). Player \( i \)'s strategy, \( s_i, \) is a reactive strategy when there is

\[ s^t_i = \begin{cases} s^0_i & t = 0 \\ f(h^t_i, h^{t-1}_i) & t > 0 \end{cases} \] (7)

The strategy in the first stage game, \( s^0_i \), is either a pure strategy or a mixed strategy. Obviously, reactive strategies do not exist in one-shot games since there always are \( h^t_i = h^{t-1}_i = \emptyset \) for any \( i \).

Reactive strategies provide a way of coordination among a group of players in repeated games. In a repeated game with multiple Nash equilibria, for example, convergence to a designated Nash equilibrium is not guaranteed unless the players adopt specific reactive strategies.

Reactive strategies also provide a way of maintaining coordination among a group of players. Grim trigger, for example, is a reactive strategy for the players in iterated prisoner’s dilemma to maintain mutual cooperation. There exists a set of trigger strategies in a repeated game, by which the coordination among a group of players can be enforced. Once a group of players have coordinated their actions, they switch to the trigger strategy that one player will choose the minimax strategy if any other player in the group deviates from their coordination strategy.

Coordination among a group of players can be achieved when they adopt specific reactive strategies, which may form equilibrium in repeated games.

**Definition 1:** In a repeated \( n \)-player game, a level-\( k \) coordination \((2 \leq k \leq n)\) denotes that a group of \( k \) players coordinate their actions by adopting some trigger strategies such that they will change their strategies simultaneously once any player in the group deviates from the assigned action.

The necessary condition of a level-\( k \) coordination is that \( k \) players can be better off by coordinating their actions. Let \( v_i \) be the minimax payoff of player \( i \in I \) and \( s'_i \) the minimax strategy. Let \( K = \{j, \ldots, j + k - 1\} \) denote a group of \( k \) players where \( j = 1, \ldots, n - k + 1 \) and \( K \subseteq I \). The necessary condition for a level-\( k \) coordination is that there exists a strategy profile \( \bar{s} = (s_1, s_2, \ldots, s_n) \) such that, for any \( s_i (i \in K) \), there are

\[ v_j < u_j(s_1, \ldots, s_j, \ldots, s_{j+k-1}, \ldots, s_n) \]

\[ v_{j+k} < u_{j+k-1}(s_1, \ldots, s_j, \ldots, s_{j+k-1}, \ldots, s_n) \] (8)

A level-\( k \) coordination can be maintained if all players involved adopt a trigger strategy such as: keep playing the coordination strategy if all other players play their coordination strategies; otherwise, play the minimax strategy. The best responses of the players who do not belong to \( K \) are determined given the strategies of \( k \) players who coordinate their actions.

If the \( k \) players cannot further improve their payoffs by deviating from \( \bar{s} \) simultaneously, the strategy profile \( \bar{s} \) is a stable state (equilibrium) in the repeated game. This equilibrium is different from the concept of NE in that \( k \) players coordinate their actions.

**Definition 2:** In an infinitely repeated \( n \)-player game, we call it a level-\( k \) equilibrium \((2 \leq k \leq n)\) if a group of \( k \) players coordinate their actions and they have no incentive to deviate from their strategies simultaneously.

A strategy profile \( \bar{s} \) is a level-\( k \) equilibrium if, for any \( s'_i \) \((i = 1, \ldots, j-1, j+k, \ldots, n)\) and \( s'_j \neq s_j, \ldots, s'_{j+k-1} \neq s_{j+k-1} \), there are

\[ u_j(s'_1, \ldots, s'_j, \ldots, s'_{j+k-1}, \ldots, s'_n) \leq u_j(s'_1, \ldots, s'_j, \ldots, \bar{s}_{j+k-1}, \ldots, s'_n) \]

\[ u_{j+k-1}(s'_1, \ldots, s'_j, \ldots, s'_{j+k-1}, \ldots, s'_n) \leq u_{j+k-1}(s_1, \ldots, s_j, \ldots, \bar{s}_{j+k-1}, \ldots, s_n) \] (9)
We prove that any level-$k$ equilibrium is also a NE in infinitely repeated games.

**Theorem 1:** In an $n$-player infinitely repeated game, any level-$k$ equilibrium $(2 \leq k \leq n)$ is a NE.

**Proof:** Let $\bar{s}$ denote a level-$k$ equilibrium. We first consider the case of $k = n$. We have $v_i < u_i(\bar{s}_i, \ldots, \bar{s}_n)$ for all $i \in I$. According to the folk theorem, $\bar{s}$ is a NE strategy profile.

In the case of $k < n$, if $v_i < u_i(\bar{s}_i, \ldots, \bar{s}_n)$ are satisfied for all $i \in I$, $\bar{s}$ must be a NE strategy profile according to the folk theorem. It is impossible that, for any player $i$, there is $v_i > u_i(\bar{s}_i, \ldots, \bar{s}_n)$ because that player could deviate from $\bar{s}_i$ to the minimax strategy so that the payoff is guaranteed to be $v_i$. This conflict with the fact that $\bar{s}_i$ is player $i$’s best response. We simply need to consider $v_i = u_i(\bar{s}_i, \ldots, \bar{s}_n)$ for some players $i \notin K$.

Let $M$ be the group of players who receive their minimax payoffs. Any player $i \in M$ cannot improve his/her payoff by deviating from $\bar{s}_i$ since $\bar{s}_i$ is the best response to $\bar{s}_{-i}$.

Any player $i \in K$ cannot improve their payoff by deviating from $\bar{s}_i$. If player $i$ did deviate from $\bar{s}_i$ in order to gain a higher payoff in the current round, all other members of $K$ would play their minimax strategies in the future rounds. Player $i$ will have to play the minimax strategy and will receive $v_i$ in the future rounds. Knowing this, player $i$ has no incentive to deviate from $\bar{s}_i$.

Since any player has no incentive to deviate from $\bar{s}_i$, it is a NE.

Every level-$k$ equilibrium is a NE and a NE is not necessarily a level-$k$ equilibrium. Thus, the level-$k$ equilibria are refinements of NE in repeated games. The payoff profiles of the level-$k$ equilibrium form the Parteto front in an infinitely repeated game. We prove the existence of level-$k$ equilibrium in general repeated games in proposition 2.

**Theorem 2:** In an infinitely repeated $n$-player game, there must be at least one level-$k$ equilibrium.

**Proof:** A game must have at least one NE. We first consider the case that a repeated game has only one NE. Let $(u_1, \ldots, u_n)$ denote the payoff profile of the NE and $(v_1, \ldots, v_n)$ the minimax payoff profile. There must be $u_i = v_i$ for any $i$ because, if they are not, the game should have more NE according to the folk theorem. Obviously, $(u_1, \ldots, u_n)$ is a level-$n$ equilibrium payoff profile.

When there exists two or more NE, there must be at least one NE that is different from the minimax profile. Let $(s_1, \ldots, s_n)$ denote the strategy profile of such a NE. We first prove that there must be $v_i < u_i(s_1, \ldots, s_n)$ for at least two players. Assume that there is $v_a < u_a(s_1, \ldots, s_n)$ for the player $a$ and $v_i = u_i(s_1, \ldots, s_n)$ for any $i \neq a$. Since all players except $a$ play their minimax strategies and they have no incentive to deviate unilaterally (because $(s_1, \ldots, s_n)$ is a NE), $s_a$ is the minimax strategy for $a$. This conflicts with the premise that $(s_1, \ldots, s_n)$ is different from the minimax profile. Thus, there must be $v_i < u_i(s_1, \ldots, s_n)$ for at least two players.

Suppose that there are $v_i < u_i(s_1, \ldots, s_n)$ for $k (k \geq 2)$ players in the NE. If those $k$ players cannot improve their payoffs by changing their strategies simultaneously, this NE is a level-$k$ equilibrium. Otherwise, there must be a strategy profile $(s'_1)$ such that those $k$ players cannot further improve their payoffs by changing their strategies simultaneously and $(s'_1)$ is a level-$k$ equilibrium.

A level-$k$ equilibrium is not stable if it is not a NE in the stage game because once a player within the coalition changes his/her strategy in a level-$k$ equilibrium, all other $k-1$ players will be triggered to change their strategies.

Some level-$k$ equilibrium can be evolutionary stable. In the following section, we give the condition of evolutionarily stable states and propose an algorithm to compute evolutionarily stable states in $n$-player symmetric games.

**IV. EVOLUTIONARILY STABLE PAYOFF**

The space of strategy profiles of a repeated game is very complicated even for two-player $2 \times 2$ games. On the other hand, the space of payoff profiles is simple. For example, the space of payoff profiles of an $n$-player $n \times n$ game can be precisely defined as an $m$-dimension polyhedron. In this section, we study evolutionarily stable state in payoff space rather than in strategy space.

Consider a symmetric two-player game. Let $u(s, s')$ denote the payoff for playing strategy $s$ against strategy $s'$.

**Definition 3:** The payoff $u(s, s)$ is called *evolutionarily stable payoff* if, for any $s' \neq s$, there are $u(s, s) \geq u(s', s)$ and $u(s, s) \geq u(s', s')$.

The evolutionarily stable payoff (ESP) is different from the concept of ESS in that ESP is defined in payoff space rather than in strategy space. The ESP denotes the population payoff in the evolutionary stable state. It is easy to verify that there is at most one ESP in a two-player symmetric game.

An ESP may correspond to a group of strategy profiles, rather than a unique strategy profile. In an evolutionarily stable state, it is not necessary that all players adopt the same strategy even if they receive the same payoff. There might be an unlimited set of strategies that the players could adopt in order to receive the ESP.

The ESP is easy to compute because the payoff space of a game can be precisely defined. Given the set of NE payoff profiles characterized by means of the folk theorem, we simply need to check whether or not one point on the Pareto front of the set of NE payoff profiles is the ESP, as illustrated in Fig.1. The Pareto front intersects with the diagonal at point $P$. Point $P$ represents the subset of NE payoff profiles that satisfy (7b). Point $P$ is the ESP profile if it weakly dominates any other payoff profiles within its neighborhood as in the case of Fig. 1(b). Otherwise, there is no ESP in the game. Based on this idea, we propose an algorithm to compute the ESP and/or the Pareto front in $n$-player symmetric games.

The definition of ESP can be extended to $n$-player games. Let $i = 1, \ldots, n$ denote the $i$th player and $(u_1, \ldots, u_n)$ be a payoff profile. Since the game is symmetric, there exists a set of payoff profiles that the payoffs of all players are identical, $u_1 = u_2 = \cdots = u_n$ (those profiles located on the main
diagonal of the payoff matrix). Let \( u = (\bar{u}, \cdots, \bar{u}) \) denote the payoff profile that every player receives the maximum identical payoff \( \bar{u} \). We simply need to make clear whether or not \( u \) is the ESP profile. A simple algorithm can compute the ESP and the Pareto front.

The profile \( u \) dominates another profile \( u' = (u'_1, \cdots, u'_n) \) if \( \bar{u} > u'_i \) for any \( i \). If \( u \) dominates every other payoff profile, it is the ESP. Otherwise, there is no ESP in the game. Note that we simply need to compare \( u \) with other pure-strategy payoff profiles.

![Figure 1](image1.png)

**Figure 1** In a two-player symmetric game, the set of NE payoff profiles characterized by the folk theorem is illustrated by the grid area in the payoff space. X-Y coordinates are the average payoff per round of two players respectively. The Pareto front intersects with the diagonal at point P. (a) P is not an ESP profile because there exists \( P' \) such that (7.a) is not satisfied. (b) P is the ESP profile.

If \( u \) is not ESP, it is not necessarily a Pareto optimum. When there is \( \bar{u} < \frac{1}{n} \sum_{i=1}^{n} u'_i \), \( u \) must be dominated by some NEs in

which the players receive the identical payoff \( \frac{1}{n} \sum_{i=1}^{n} u'_i \). By comparing all payoff profiles, we could find the NE payoff profile that is also the Pareto optimum for \( n \) players. Given that each player has \( m \) pure strategies, it needs at most \( \frac{1}{2} nm^n \) comparisons. This algorithm can be illustrated by the flowchart in Fig. 2.

![Figure 2](image2.png)

**Figure 2** An algorithm to compute ESP and a Pareto optimum of a repeated \( n \)-player symmetric game.

Given the Pareto optimum \( u = (\bar{u}, \cdots, \bar{u}) \), the Pareto front can be computed by comparing all pure-strategy payoff profile with \( u \). A payoff profile \( u' = (u'_1, \cdots, u'_n) \) belongs to the Pareto front if \( \bar{u} a < u'_i \) for some \( i \) and \( u'_i < u'_i \) for any \( u'' \neq u' \).

V. Evolutionary Stability Analysis of Three Games

In this section, we compute the evolutionarily stable states of three repeated games, the iterated prisoner’s dilemma, a coordination game and a three-player symmetric repeated game, by means of the proposed methodology in previous sections. The iterated prisoner’s dilemma and the coordination game are typical repeated games in that the former does not have an ESP and the latter has one. The third game acts as a computational example to show the efficiency of the proposed algorithm.

A. Iterated Prisoner’s Dilemma

The prisoner’s dilemma (PD) is a two-player non-cooperative game in which two players try to maximize their payoffs by cooperating with, or betraying the other player. Also an \( n \)-player version PD is introduced in [37]. A PD can be represented as the following matrix (see Fig. 3):

![Figure 3](image3.png)

**Figure 3** The payoffs of two players in the prisoner’s dilemma is expressed as a matrix. Two players choose between two options, \( C \) and \( D \). The values of \( R, S, T \) and \( P \) denote Reward for mutual cooperation, Sucker's payoff, Temptation to defect, and Punishment for mutual defection respectively.

There are two constraints, \( T > R > P > S \) and \( R > (T + S) / 2 \), which motivate each player to play non-cooperatively and prevent any incentive to alternate between cooperation and defection.

The "dilemma" faced by the prisoners is that, whatever the other does, each is better off defecting than cooperating. However, the payoff when both defect is worse for each player than the outcome they would have received if they had both cooperated.

Under the assumption of rationality, game theory predicts that both players choose to defect and their payoff profile is \( (P, P) \), which is the unique NE of this game.

In the Iterated Prisoner’s Dilemma (IPD) game, two players have to choose their mutual strategies repeatedly, and they also have a memory of their previous choices and the choices of the opponents. IPD has been heavily studied as an ideal model to study how cooperation emerges and persists in a non-cooperative environment [38-40]. According to the folk theorem, any payoff profile that strictly dominates the minimax
payoff profile \((P, P)\) is a NE in an infinite length IPD. The grid area within ADBC in Fig. 4 illustrates the set of payoff profiles of all NE in an IPD can be illustrated by the grid area within ADBC, which is characterized by the folk theorem. X-Y coordinates are the average payoff per round of two players respectively. Point A denotes the minimax payoff profile and B denotes the mutual cooperation payoff profile. The shadow area within AFBE denotes the set of feasible payoff profiles.

Any payoff profile on CB and BD is a Pareto optimum so that any player cannot improve their payoff without reducing the other player’s payoff.

The Pareto front intersects with the diagonal at B that is the mutual cooperation payoff profile. B is not an ESP profile because it does not dominate all other payoff profiles, \(B'\) for example, in the neighborhood. Thus, no ESP exists in IPD.

An evolutionary game does not necessarily converge to a stable state when there is no ESP. We ran a series of simulations to study how evolving players interact with each other in evolutionary IPD.

The evolutionary IPD model reflects the payoffs received by players in one generation in terms of copies of themselves represented in the next generation. Stochastic universal sampling is used to ensure that players produce offspring in proportion to payoffs received so that those with higher payoffs reproduce at a proportionately higher rate than those with lower payoffs. We set \(R = 3\), \(S = 0\), \(T = 5\), and \(P = 1\) in our simulations. Every player plays 50 rounds of IPD against all other players in the population in each generation. The fitness of an individual player is expressed by average payoff per round. The parents simply copy their strategies to produce offspring and neither mutation nor crossover is carried out.

An evolving player adjusts their strategy in order to adapt to the evolutionary dynamics. Let \(\rho_R, \rho_T, \rho_S, \rho_P\) denote an IPD strategy where \(\rho_x (x = R, T, S, P)\) is the probability of choosing C given that the payoff in previous round is \(x\). The evolving players in our simulations update \(\rho_x\) every round according to

\[
\rho_x(t) = \rho_x(t - 1) + \Delta \\
\Delta = -\Delta \text{ if } u < \bar{u} \tag{10}
\]

where \(\Delta\) is a small constant value, \(\bar{u}\) is the average payoff and \(u\) is the average payoff in the past five rounds. The idea is that, if a change of \(\rho_x\) leads to a higher payoff, keep changing it in that direction. Otherwise, change \(\rho_x\) in the opposite direction. Note that \(\rho_x(t) \in [0, 1]\) and it was bounded if the computed value exceeded the limits.

In each evolutionary IPD game, we choose two evolving players randomly (by randomly setting the initial values of \(\rho_x\)) and generate a population of 20 individuals. Each strategy has 10 copies. The population evolves for 2,000 generations.

Some typical results are shown in Fig. 5 (a)-(d). In some situations, the evolution converges to stable states where two players receive equal payoffs, such as Fig. 5 (a) and (b). In other situations, one player is dominated by another player, such as Fig. 5 (c) and (d).

As the outcome of any single game is affected by chance, we repeat the evolutionary IPD for 1,000 times. The payoff profiles of two evolving players in 1,000 games are shown in Fig. 6. They distribute in the whole area of feasible payoff profiles, which shows the diversity and instability of evolutionary IPD games. The evolutionary dynamics is significantly influenced by the initial strategy combination of the population.
B. A Coordination Game

Consider a coordination game with the payoff matrix as shown in Fig. 7. There are two pure-strategy NEs in this game when one player chooses $L$ and the other chooses $R$. The players do not have any a priori knowledge about which NE strategy profile to choose. The probability that any NE is achieved is $0.5$ no matter what pure or mixed strategies are adopted.

\[
\begin{array}{c|cc}
\text{X Player} & L & R \\
L & 0, 0 & 1, 1 \\
R & 1, 1 & 0, 0 \\
\end{array}
\]

If this game is played repeatedly, it may converge to a NE given that some reactive strategies are adopted. The coordination between $X$ and $Y$ can be achieved with probability $\rho \to 1$ in an infinite repeated coordination game if two players adopt the below strategies:

\[
s^X_t = \begin{cases} 
L & \text{if } X \text{ chose } L \text{ and } Y \text{ chose } R \text{ at time } t - 1 \\
R & \text{if } Y \text{ chose } L \text{ and } X \text{ chose } R \text{ at time } t - 1 \\
\text{rand}(L, R) & \text{otherwise}
\end{cases}
\]

\[
s^Y_t = \begin{cases} 
R & \text{if } X \text{ chose } L \text{ and } Y \text{ chose } R \text{ at time } t - 1 \\
L & \text{if } Y \text{ chose } L \text{ and } X \text{ chose } R \text{ at time } t - 1 \\
\text{rand}(L, R) & \text{otherwise}
\end{cases}
\]

The set of NE payoff profiles is represented by a line segment in $X$-$Y$ coordinates as shown in Fig. 8. The endpoint $A$ denotes the ESP profile, which corresponds to two pure-strategy NEs. The two NEs are evolutionarily stable states.

C. A Three-player Symmetric Game

In this game, three players, $X$, $Y$, and $Z$, play a strategic game. Each player chooses a number in the range of $(0, 1]$ independently. Let $s_X$, $s_Y$, and $s_Z$ denote three players’ choices. Their payoffs are determined by

\[
u_i = \begin{cases} 
s_i & \text{if } s_X^i + s_Y^i + s_Z^i \leq 1 \\
0 & \text{otherwise}
\end{cases} \quad (i = X, Y, Z)
\]

This is a quadratic version of Divide-the-Dollar Game \[46\]. Assume that the players can only choose between the multiples
of 0.01. Thus, each player has 100 pure strategies and there are totally $10^5$ pure-strategy payoff profiles.

The algorithm to compute the ESP and the Pareto front was coded in Visual C++ and run on a PC with dual 2.66GHz Intel CPU and 3.25GB RAM. It took approximately 0.023 second to compute the ESP and 17 seconds to compute the Pareto front. The intersection of the axis of symmetry and the hull of feasible payoff profiles is computed to be $(0.55, 0.61, 0.57)$, which is close to the theoretically solution $(\sqrt[3]{3}/3, \sqrt[3]{3}/3, \sqrt[3]{3}/3)$. This payoff profile is not an ESP and thus this game has no ESP. The computed Pareto front contains 7,535 payoff profiles, as shown in Fig. 9.

Many games do not have an ESP. The evolutionary dynamic of these games may be complex and it is not likely to converge to a stable state unless some restrictions are imposed. Generally, a game without ESP contains multiple NE and none of those NE is dominant. In an IPD game, for example, the NE payoff profiles form a convex polygon as shown in Fig. 4. The Pareto front (points on CB and BD) denotes a set of level-$k$ equilibrium, which dominates those NE that are not Pareto optimum. How the players choose between multiple NE, which is probably the main reason for chaos and unpredictability in evolutionary game dynamics, is still an open question.

VI. CONCLUSION

Stability analysis based on the concept of ESS is flawed in that it is difficult to find an expression of possible strategies in a repeated game and that ESS is too strict to exist in most repeated games. We propose a new stability analysis of repeated games and evolutionary games based on a subset of NE, which we call them the level-$k$ equilibrium. Evolutionarily stable states can be computed in the payoff space rather than in the strategy space so that there is no need to search and compare strategies in a complicated space.

Computing a NE of repeated games is difficult because computing the minimax profile is NP-hard. The possible payoff profiles in an $n$-player repeated game form an $n$-D convex polyhedron in the payoff space. We prove that some Pareto optimums of the hull of the polyhedron, the level-$k$ equilibria, are NE payoff profiles. Computing the minimax profile is not necessary in computing the level-$k$ equilibria, which means that finding specific subset of NE is tractable.

A strategy profile determines a unique payoff profile, and not vice versa. A payoff profile probably corresponds to a set of strategy profiles. For example, the players in IPD could implement mutual cooperation by adopting the strategies such as always cooperate, tit-for-tat, and numerous other strategies. Given a payoff profile, there must be at least one strategy profile for the players to adopt in order to receive the given payoffs.

The concept of ESP is defined in the payoff space and thus it can be easily computed. Generally, the strategy space of a repeated game is much more complicated than the payoff space when reactive strategies are taken into consideration. Given an ESP, there must be at least one corresponding pure strategy profile.

The level-$k$ equilibria can be considered as refinement of NE of repeated games. In a level-$k$ equilibrium not only does any individual player not have incentive to unilaterally change their strategies but also a group of $k$ players has no incentive to deviate from it collectively. A level-$k$ equilibrium is not necessarily a NE of the stage game because its corresponding strategies are reactive strategies which may not exist in one-shot games.

The existence of level-$k$ equilibria may help to explain some phenomena in evolution that has not been well explained. Experiments have shown that cooperation can emerge and persist in multiple levels in the population [41, 42], and intermediate choices lead to less mutual cooperation [43].
level-k equilibrium suggests that cooperation in a local group can also be evolutionarily stable. With more intermediate choices, the set of level-k equilibrium increases and there are more possible evolutionarily stable states, with which the rate of cooperation is very likely to decrease because mutual cooperation in the whole population is only one of those stable states.

We also propose an algorithm to compute the ESP and the Pareto front of n-player symmetric games. The problem of computing the ESP can be simplified significantly because the convex hulls of feasible payoff profiles are symmetric for symmetric games. We simply need to compute the intersection point of the Pareto front and the main diagonal, which point denotes the possible ESP of the game.

This research helps to explain why the processes of evolution are diverse in some games. In those games where there is no ESP, IPD for example, the evolutionary dynamic could be complex and the results of evolution are hard to predict. We have shown that any feasible payoff profile is possible for evolving players in IPD simulations. The existence of reactive strategies suggests that individual players could coordinate their actions even if there is no communication among them, which provides another explanation for collective behaviors in evolution.

Choosing between multiple NE is still an unsolved problem in game theory. It is essentially a decision making under uncertainty because the subjective preferences of players are involved. Fuzzy logic seems to be a promising tool to solve this problem by using linguistic preference and fuzzy rules to prioritize payoff profiles [44, 45]. This will be the focus of our future work.

ACKNOWLEDGEMENT

We would like to thank Prof. Xin Yao for his insightful comments. The research of this paper was supported by the Engineering and Physical Sciences Research Council UK (EPSRC) project EP/H000968/1.

REFERENCES


Robert John has a BSc Mathematics, a MSc in Statistics and a PhD in Fuzzy Logic. He worked in industry for 10 years as a mathematician and knowledge engineer developing knowledge based systems for British Gas and the financial services industry. Bob spent 24 years at De Montfort University in various roles including Head of Department, Head of School and Deputy Dean. He led the Centre for Computational Intelligence research group from 2001 until 2012. He has over 150 research publications of which about 50 are in international journals. Bob joined the University of Nottingham in 2013 where he heads up the research group ASAP in the School of Computer Science. The Automated Scheduling, Optimisation and Planning (ASAP) research group carries out multi-disciplinary research into mathematical models and algorithms for a variety of real world optimisation problems.

Jaiew Li (M12) received the B.Sc. and Ph.D. degree in engineering from the Harbin Engineering University, Harbin, China in 1992 and 1998 respectively. He became a lecturer in Robotics Institute at Harbin Institute of Technology, China in 1999. Currently, he is a research fellow in the School of Computer Science at University of Nottingham, Nottingham, U.K. His research interests include a number of decision making and optimization methodologies including evolutionary game theory, fuzzy logic, interval probability theory, hyper-heuristics and economic risk benefit analysis. He is especially interested in solving complex real world problems by integrating multiple decision making methods.

Graham Kendall (M03-SM10) received the Bachelor’s degree in computation (1st class, honors) from the University of Manchester Institute of Science and Technology, Manchester, U.K. in 1997 and the Ph.D. degree in computer science from the University of Nottingham, Nottingham, U.K., in 2001. His previous experience includes almost 20 years in the Information Technology industry where he held both technical and managerial positions. He is a Professor of Computer Science at University of Nottingham and is currently based at their Malaysia Campus where he holds the position of Vice-Provost (Research and Knowledge Transfer). He is a Director of two companies (EventMAP Ltd., Nottingham, U.K.; Aptia Solutions Ltd., Nottingham, U.K.) and CEO of two companies (MyRIAD Solutions Sdn Bhd, Malaysia and MyResearch Sdn Bhd, Malaysia). He is a Fellow of the Operational Research Society. He is the editor-in-chief of the IEEE TRANSACTIONS ON COMPUTATIONAL INTELLIGENCE AND AI IN GAMES and an Associate Editor of eight international journals, including the IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION. He chaired the Multidisciplinary International Conference on Scheduling: Theory and Applications in 2003, 2005, 2007, 2009, and 2011, and has chaired several other international conferences, which has included establishing the IEEE Symposium on Computational Intelligence and Games. He has been awarded externally funded grants worth over £6 million from a variety of sources including Engineering and Physical Sciences Research Council (EPSRC) and commercial organizations.