第 36 卷第 2 期
2019 年 2 月

控 制 理 论 与 应 用
Control Theory & Applications

Vol. 36 No. 2
Feb. 2019

局部信息约束下网络演化博弈的动力学与优化

王元华†, 刘希玉
(山东师范大学 商学院, 山东 济南 250014)

摘要: 网络演化博弈的优化问题是混合值逻辑网络的一个自然推广. 本文研究了一类网络演化博弈的优化控制问题, 其中每个控制个体在极大化自己的收益时只能获取到邻域信息. 首先, 利用矩阵的半张量积, 将局部信息约束下控制网络演化博弈的动力学转化为相应的代数形式. 然后得到了局部信息约束下确定型网络演化博弈的最优控制序列. 最后, 基于动态规划的解, 研究了局部信息约束下概率型网络演化博弈的优化控制问题, 得到了最优控制序列的简单计算公式. 两个数值例子验证了本文的理论结果.

关键词: 控制网络演化博弈; 优化; 局部信息; 半张量积
引用格式: 王元华, 刘希玉. 局部信息约束下网络演化博弈的动力学与优化. 控制理论与应用, 2019, 36(2): 279 – 285
DOI: 10.7641/CTA.2018.17070

Dynamics and optimization of control networked evolutionary games with local information

WANG Yuan-hua†, LIU Xi-yu
(Business School, Shandong Normal University, Jinan Shandong 250014, China)

Abstract: The optimization of networked evolutionary games (NEGs) is a natural extension of optimization for mix-valued logical networks. This paper studies the optimization problem for a class of control NEGs, where each controller can only use the information of its neighbors so as to maximize its payoff over a finite or infinite number of time steps. First, the dynamics of control NEGs with local information is converted into an algebraic form by using the semi-tensor product of matrices. Then the optimal control sequences for deterministic NEGs with local information are obtained. Finally, based on the dynamic programming solutions, some easily computable formulas are provided for stochastic NEGs with local information. Two examples are presented to illustrate the theoretical results.

Key words: control networked evolutionary games; optimization; local information; semi-tensor product
Citation: WANG Yuanhua, LIU Xiyu. Dynamics and optimization of control networked evolutionary games with local information. *Control Theory & Applications*, 2019, 36(2): 279 – 285

1 Introduction

Evolutionary game theory was first introduced by evolutionary biologists[1–2] for describing and modelling the evolution of lives. It has been widely applied to economics[3], social physics[4], engineering science[5], etc.

In a classical evolutionary game, it is assumed that each player is equally likely to interact with any other member of the population. However, sometimes players do not interact with all other players, but play with some of them. For example, the decision of an individual to whether buy or not a new product, attend a meeting, find a job is often influenced by the choices of its friends and acquaintances. In these situations we can use a network to describe the interactions between players, in which the nodes denote players and the interactions between players are described by edges[6]. Such a game that combines the evolutionary dynamics and a network is called a networked evolutionary game (NEG).

In recent years, the NEG has been proved to be a powerful tool for studying evolutionary dynamics[7], in which players learn by interacting with their neighbors to update their strategies in the next step. For instance, in biological communities, individual often communicates with its neighbors within limited ranges of seeing and hearing. These individuals are apt to collect the local information in their surrounding regions and to exchange the local information with their neighbors. In some cases, the strategies for a small portion of players can be assigned at each moment, who can be regarded

Received 3 November 2017; accepted 31 May 2018.
†Corresponding author. E-mail: wyh_1005@163.com; Tel.: +86 13583122799.
Recommended by Associate Editor CHEN Jie.
Supported by the National Natural Science Foundation of China (61333001, 61773371, 61733018).

as control players (or controllers), while other players can be called state players (or states), then such a NEG is named as a control NEG.

Recently, a new mathematical tool, called the semi-tensor product of matrices, was introduced[8–9]. It has been successfully applied to the analysis of Boolean networks[10–13], mix-valued logical networks[14–17], finite games[18–19], and evolutionary games[20–23]. In a NEG, the evolutionary process is actually a finitely logical dynamic process, then the NEG can be expressed as a k-valued logical network by using the semi-tensor product method. Under this framework, a control NEG can be expressed as a control k-valued logical network, then the known control results about k-valued logical networks are applicable to the control NEGs. In [22], the authors presented an algebraic framework for investigating the NEGs and studied some control problems. In [23], the strategy optimization was posed in term of maximizing the average payoff of the pseudo-player in the long run. However, these papers are based on global information, that is, the information of all states are available for the controllers. Such an assumption is unrealistic and difficult to hold in our real world situations. For instance, in a cancer treatment application, we may be able to track the status of only a limited number of genes and not necessarily all the ones for the reasons of cost, accessibility or other considerations. The control problems in this case, which are based on the presence of local information, arise in many real world problems. In this paper, we consider a class of control NEGs where each controller can only observe the information of its neighbourhoods, and the information of other states is uncertain. The main contributions consist of i) providing an algebraic expression of control NEGs with local information; ii) extending the existing results from k-valued logical networks to control NEGs with local information; iii) designing the optimal control sequences for both deterministic and stochastic NEGs with local information.

The remainder of this paper is organized as follows. Section 2 introduces some necessary preliminaries on the semi-tensor product of matrices. In Section 3, we give a problem formulation and an algebraic form for control NEGs with local information. Section 4 presents the optimization problems of control NEGs with local information. Section 5 is a conclusion.

2 Preliminaries

First, we give some necessary notations for ease:
$\mathcal{M}_{m\times n}$: the set of $m\times n$ real matrices.

$$\mathcal{D}_k := \{1,2,\cdots,k\},\ k\geqslant 2,$$
$$\Delta_n := \left\{\delta_n^i | i=1,\cdots,n\right\},$$

where δ_n^i is the i-th column of the identity matrix I_n.
$\mathrm{Col}(M)(\mathrm{Row}(M))$: the set of columns (rows) of M.
$\mathrm{Col}_i(M)(\mathrm{Row}_i(M))$: the i-th column (row) of M.

A matrix $L\in\mathcal{M}_{m\times n}$ is called a logical matrix if $\mathrm{Col}(L)\subset\Delta_m$. That is, $L=[\delta_m^{i_1}\ \delta_m^{i_2}\ \cdots\ \delta_m^{i_r}]$. It is briefly denoted as $L=\delta_m[i_1\ i_2\ \cdots\ i_r]$. The set of $m\times n$ logical matrices is denote by $\mathcal{L}_{m\times n}$.

$$\Upsilon_n := \{r\in\mathbb{R}^n \mid r_i\geqslant 0, \text{ and } \sum_{i=1}^n r_i = 1\}.$$

A matrix $T\in\mathcal{M}_{m\times n}$ is called a probabilistic matrix if $\mathrm{Col}(T)\subset\Upsilon_m$. The set of $m\times n$ probabilistic matrices is denoted by $\Upsilon_{m\times n}$.

Definition 1[9] Let $M\in\mathcal{M}_{m\times n}$, $N\in\mathcal{M}_{p\times q}$, $t=\mathrm{lcm}\{n,p\}$ be the least common multiple of n and p. The semi-tensor product (STP) of M and N is defined as

$$M\ltimes N := \left(M\otimes I_{t/n}\right)\left(N\otimes I_{t/p}\right), \tag{1}$$

where $\otimes$ is the Kronecker product.

The STP is a generalization of conventional matrix product, and we can omit the symbol "$\ltimes$" without confusion.

Proposition 1 Let $X\in\mathbb{R}^m$ be a column and M be any matrix. Then

$$X\ltimes M = (I_m\otimes M)\,X. \tag{2}$$

Next, we define the swap matrix[8–9]:

Definition 2 A matrix $W_{[m,n]}\in\mathcal{M}_{mn\times mn}$, defined by

$$W_{[m,n]}=\delta_{mn}\begin{bmatrix} 1 & m+1 & \cdots & (n-1)m+1 \\ 2 & m+2 & \cdots & (n-1)m+2 \\ \vdots & \vdots & \vdots & \vdots \\ m & 2m & \cdots & nm \end{bmatrix},$$

is called the (m,n)-th dimensional swap matrix.

Proposition 2 Let $X\in\mathbb{R}^m$ and $Y\in\mathbb{R}^n$ be two columns. Then

$$W_{[m,n]}\ltimes X\ltimes Y = Y\ltimes X. \tag{3}$$

Definition 3 Let $M\in\mathcal{M}_{p\times m}$, $N\in\mathcal{M}_{q\times m}$. Then the Khatri-Rao Product of M and N is defined as

$$M\times N = [\mathrm{Col}_1(M)\ltimes\mathrm{Col}_1(N)\ \cdots\ \mathrm{Col}_m(M)\ltimes\mathrm{Col}_m(N)].$$

To use vector expression of logical variables, we identify $i\sim\delta_k^i$, $i=1,2,\cdots,k$, then $\mathcal{D}_k\sim\Delta_k$. Using vector expression, a (pseudo) logical function can be expressed as an algebraic form.

Theorem 1[9] Let $f:\mathcal{D}_k^n\to\mathcal{D}_k$ be a k-valued logical function. Then there exists a unique $M_f\in\mathcal{L}_{k\times k^n}$, such that in vector form we have

$$f(x_1,\cdots,x_n) = M_f \ltimes_{i=1}^n x_i, \tag{4}$$

where $x_i\in\mathcal{D}_k$, $i=1,\cdots,n$, and M_f is called the structure matrix of f.

Corollary 1 Let $c:\mathcal{D}_k^n\to\mathbb{R}$ be a k-valued pseudo-logical function. Then there exists a unique row vector $V^c \in \mathbb{R}^{k^n}$, such that in vector form we have

$$c(x_1, \cdots, x_n) = V^c \ltimes_{i=1}^{n} x_i, \tag{5}$$

V^c is called the structure vector of c.

3 Dynamics of control NEGs with local information

3.1 Problem formulation

Definition 4[24] A normal finite game, denoted by $G = (N, S, C)$, consists of three factors:

1) $N = \{1, 2, \cdots, n\}$ is the set of players;

2) $S = \prod_{i=1}^{n} S_i$ is the strategy profile, where $S_i = \{1, 2, \cdots, k\}$ is the set of strategies for player i. The strategies of all players but the i-th one are denoted by $S_{-i} := \prod_{j \neq i} S_j$;

3) $C = (c_1, \cdots, c_n) \in \mathbb{R}^n$ with $c_i : S \to \mathbb{R}$ defined as

$$c_i := c_i(x_1, \cdots, x_n) = V_i^c \ltimes_{i=1}^{n} x_j, \quad x_j \in S_j,\ i = 1, \cdots, n \tag{6}$$

is called the payoff function of player i.

In this paper, a normal finite game is played repeatedly on a network and a small portion of players perform the role of active controllers, which is described as follows.

Definition 5 A control NEG, denoted by $G_N = (\mathcal{G}_c, G, \Pi)$, is composed by

1) A networked graph $\mathcal{G}_c = (N, E) = (X \cup U, E)$, where N is the set of nodes (or players) and $\{X, U\}$ is a partition of N, that is, $X \cup U = N$ and $X \cap U = \varnothing$. $X = \{x_1, x_2, \cdots, x_n\}$ is the states and $U = \{u_1, u_2, \cdots, u_m\}$ is the controllers, where x_i, $u_j \in \mathcal{D}_k$. $E \subseteq N \times N$ is the set of edges. If $(i, j) \in E$ implies $(j, i) \in E$, the graph is undirected. The set of all neighbors for i is denoted by $\mathcal{N}_i$ and $l_i = |\mathcal{N}_i|$. In this paper we assume that $i \in \mathcal{N}_i$, and $\mathcal{N}_p \cap \mathcal{N}_q = \varnothing$ for $u_p, u_q \in U$.

2) G is a normal finite game with two players, such that if $(i, j) \in E$, then i and j play G with strategies $x_i(t)$ and $x_j(t)$ at time t respectively. Particularly, if G is not symmetric, the corresponding network graph must be directed and the directed edge is used to distinguish different roles of two players. Assume $(i, j) \in E$, then there is an edge from i to j, and i is player one and j is player two.

3) Π is an updating rule which describes how a state player chooses a proper strategy for the next step. In this paper, we only consider a simple updating rule, called Myopic best response adjustment rule (MBRAR)[25], then the strategy dynamics is

$$x_i(t+1) = f_i(x_j(t), u_p(t);\ j, p \in \mathcal{N}_i), \tag{7}$$

where f_i is determined by MBRAR, which is described as

$$x_i(t+1) = \arg\max_{x_i \in S_i} c_i(x_i, x_{-i}(t)), \quad x_{-i}(t) \in S_{-i},\ i = 1, \cdots, n. \tag{8}$$

If $x_i(t) \in \arg\max_{x_i \in S_i} c_i(x_i, x_{-i}(t))$, then $x_i(t+1) = x_i(t)$. Else, we have the following two options:

1) Deterministic Model: choose the smallest j, such that $x_j \in \arg\max_{x_i \in S_i} c_i(x_i, x_{-i}(t))$, and set $x_i(t+1) = x_j$;

2) Stochastic Model: choose any j, such that $x_j \in \arg\max_{x_i \in S_i} c_i(x_i, x_{-i}(t))$, with equal probability.

Remark 1 The overall payoff of player i at time t is

$$c_i(x_i(t), x_{-i}(t)) = \sum_{j \in \mathcal{N}_i \setminus i} c_{ij}(x_i(t), x_j(t)),\ i \in \mathbb{N},$$

where $c_{ij}(x_i(t), x_j(t))$ is the payoff of player i from the game played with j.

In [26], the control problems of NEGs with global information have been studied. That is, the strategy profile $x(t) \in \mathcal{D}_k^n$ is available for the controllers, which is described as a state feedback,

$$u_p(t) = g_p(x_1(t), x_2(t), \cdots, x_n(t)) := G_p x(t), \quad p = 1, \cdots, m, \tag{9}$$

where G_p is the structure matrix of g_p.

In this paper, we only consider the local information case, that is, each controller only use the information of its neighbors,

$$u_p(t) = h_p(x_j(t);\ j \in \mathcal{N}_p),\ p = 1, \cdots, m. \tag{10}$$

3.2 Algebraic form

First, we give a useful projection result.

Lemma 1[22] Let $x_i \in \Upsilon_k$, $i = 1, \cdots, n$. Define a set of projection matrices $\pi_i = \mathbf{1}_{k^{i-1}}^{\mathrm{T}} \otimes I_k \otimes \mathbf{1}_{k^{n-i}}^{\mathrm{T}}$, then

$$\pi_i x = x_i,\ i = 1, \cdots, n,$$

where $x = \ltimes_{i=1}^{n} x_i$, and $\mathbf{1}_k := (\underbrace{1, 1, \cdots, 1}_{k})^{\mathrm{T}}$.

For (7), using Theorem 1, we can find the structure matrix $\tilde{M}_i \in \mathcal{L}_{k \times k^{l_i}}$ (or in stochastic model $\tilde{M}_i \in \Upsilon_{k \times k^{l_i}}$) for each logical function f_i. Then we have

$$x_i(t+1) = \tilde{M}_i \ltimes_{p \in \mathcal{N}_i} u_p(t) \ltimes_{j \in \mathcal{N}_i} x_j(t). \tag{11}$$

Similarly, for the controller's strategy dynamics (10), we have

$$u_p(t) = \tilde{H}_p \ltimes_{j \in \mathcal{N}_p} x_j(t),\ p = 1, 2, \cdots, m, \tag{12}$$

where $\tilde{H}_p \in \mathcal{L}_{k \times k^{l_p}}$ (or in stochastic model $\tilde{H}_p \in \Upsilon_{k \times k^{l_p}}$) is the structure matrix of h_p in (10).

Using Lemma 1 and Proposition 1, the dynamics (11) and (12) can be converted into

$$\begin{cases} x_i(t+1) = M_i u(t)x(t), & i = 1, 2, \cdots, n, \\ u_p(t) = H_p x(t), & p = 1, 2, \cdots, m, \end{cases} \tag{13}$$

where $M_i \in \mathcal{L}_{k\times k^{m+n}}$ (or $M_i \in \Upsilon_{k\times k^{m+n}}$) and $H_p \in \mathcal{L}_{k\times k^n}$ (or $H_p \in \Upsilon_{k\times k^n}$).

Multiply all equations in (13) together, then the algebraic form of control NEGs with local information is

$$\begin{cases} x(t+1) = Mu(t)x(t), \\ u(t) = Hx(t), \end{cases} \tag{14}$$

where $M = M_1 * M_2 * \cdots * M_n \in \mathcal{L}_{k^n\times k^{m+n}}$ (or $M \in \Upsilon_{k^n\times k^{m+n}}$) and $H = H_1 * H_2 * \cdots * H_m \in \mathcal{L}_{k^m\times k^n}$ (or $H \in \Upsilon_{k^m\times k^n}$).

4 Optimization problems of control NEGs with local information

In this section, we discuss the optimization problems of deterministic model and stochastic model respectively. First, we consider the deterministic case.

4.1 Optimal control of deterministic NEGs with local information

Let $x \in \Delta_k$. Define an order reducing matrix as

$$O_k^R := \delta_{k^2}[1\ \ k+2\ \ 2k+3\ \ \cdots\ \ k^2] \in \mathcal{L}_{k^2\times k}.$$

Then we have $x^2 = O_k^R x$.

The algebraic form of (14) can be expressed as

$$x(t+1) = Lx(t), \tag{15}$$

where $L = MHO_{k^n}^R \in \mathcal{L}_{k^n\times k^n}$.

It is obvious that (15) is a k-valued logical network without controllers, hence the properties of control NEGs with local information can be revealed from (15), equivalently from the transition matrix L.

Let $\Theta(t) = \{\theta_1(t)\ \ \theta_2(t)\ \ \cdots\ \ \theta_m(t)\}$ be the local information that is available for the controllers $U(t) = \{u_1(t)\ \ u_2(t)\ \ \cdots\ \ u_m(t)\}$ at time t. In this section, our purpose is to maximize the following value,

$$J(u) = \sum_{t=0}^{T} \lambda^t c(u_1(t), \cdots, u_m(t), \theta_1(t), \cdots, \theta_m(t)) = \sum_{t=0}^{T} \lambda^t c(U(t), \Theta(t)). \tag{16}$$

For each controller u_p, denote its neighbors' information by $\theta_p(t) = \ltimes_{i\in\mathcal{N}_p} x_i(t) \in \mathcal{D}_{k^{l_p}}$, and assume $\theta_p(t) \cap \theta_q(t) = \varnothing$ for $p \neq q$. Then we have $\theta(t) = \ltimes_{p=1}^{m} \theta_p(t) \in \mathcal{D}_{k^r}$, where $r = \sum_{i=1}^{m} |\mathcal{N}_i|$.

Define a payoff matrix as

$$\varPhi := (\varphi_{i,j}) \in \mathcal{M}_{k^m\times k^r}, \tag{17}$$

where $(\varphi_{i,j}) := c(\delta_{k^m}^i, \delta_{k^r}^j)$, $i = 1, 2, \cdots, k^m$, $j = 1, 2, \cdots, k^r$, is the i-th row, j-th column element of the matrix $\varPhi$, which is corresponding to the payoff of controllers $u(t)$ with respect to their neighbors' strategy profile $\theta(t)$. Then the controllers' payoff in (16) can be converted into

$$\lambda^t c(u(t), \theta(t)) = \lambda^t u^{\mathrm{T}}(t)\varPhi\theta(t). \tag{18}$$

Now the optimization problem can be expressed as

$$\max J_t(\theta(t)) = \max_{u(t)\in\Delta_{k^m}} \lambda^t c(u(t), \theta(t)) = \max_{u(t)\in\Delta_{k^m}} \lambda^t u^{\mathrm{T}}(t)\varPhi\theta(t). \tag{19}$$

Arranging $J_t(\theta(t))$ with different strategy profiles $\theta(t) \in \{\delta_{k^r}^j | j = 1, 2, \cdots, k^r\}$ into a vector form as

$$J_t(\theta(t)) = [J_t(\delta_{k^r}^1)\ \ J_t(\delta_{k^r}^2)\ \ \cdots\ \ J_t(\delta_{k^r}^{k^r})]^{\mathrm{T}},$$

then the equation (19) becomes

$$\max J_t(\theta(t)) = \lambda^t \begin{bmatrix} \max\limits_{u(t)\in\Delta_{k^m}} u^{\mathrm{T}}(t)\mathrm{Col}_1(\varPhi) \\ \max\limits_{u(t)\in\Delta_{k^m}} u^{\mathrm{T}}(t)\mathrm{Col}_2(\varPhi) \\ \vdots \\ \max\limits_{u(t)\in\Delta_{k^m}} u^{\mathrm{T}}(t)\mathrm{Col}_{k^r}(\varPhi) \end{bmatrix}. \tag{20}$$

Hence, to maximize the value $J_t(\theta(t))|_{\theta(t)=\delta_{k^r}^j}$ is equivalent to finding the maximum component of $\mathrm{Col}_j(\varPhi)$. That is, the optimal control for $J_t(\delta_{k^r}^j)$ is $u^*(t)|_{\theta(t)=\delta_{k^r}^j} = \delta_{k^m}^{i^*}$, where

$$i^* = \arg\max_i \mathrm{Col}_j^i(\varPhi),$$

and $\mathrm{Col}_j^i(\varPhi)$ is the i-th component of $\mathrm{Col}_j(\varPhi)$.

We give a simple example to illustrate the above results.

Example 1 Consider a control NEG $G_N = (\mathcal{G}_{\mathrm{c}}, G, \varPi)$, where i) $N = (X \cup U)$, $X = \{x_1, x_2, x_3\}$, $U = \{u\}$, the network graph is given by Fig. 1; ii) G is the Prisoner's Dilemma with the payoff bi-matrix shown in Table 1; iii) The strategy updating law is M-BRAR.

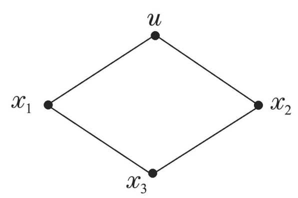

Fig. 1 Network graph

Table 1 Payoff bi-matrix of Prisoner's Dilemma

$P_1\backslash P_2$	1	2
1	(3, 3)	(0, 5)
2	(5, 0)	(1, 1)

First, for the controller u_p, we define the frequency vector $q_j(t)$ of player $j \notin \mathcal{N}_p$ as

$$q_j(t) = (q_j^1(t), q_j^2(t), \cdots, q_j^k(t))^{\mathrm{T}}, \tag{21}$$

where

$$q_j^i(t) := \frac{1}{t+1}\sum_{l=0}^{t} I\{x_j(l) = i\}$$

is the percentage of stage at which player j chosen the strategy $i \in \mathcal{D}_k$ from time 0 to time t, $x_j(l)$ denote the chosen strategy of player j at time $l = 0, 1, \cdots, t$, and $I\{\cdot\}$ is the indicator function.

From Fig. 1, the controller u can only observe the information of x_1 and x_2. Suppose x_3 selects strategy 1 and 2 with equal frequency at $t = 0$, then we have

$$x_3(0) = q_3(0) = (\frac{1}{2}, \frac{1}{2})^{\mathrm{T}}.$$

We calculate the expect payoffs for x_1 and x_2 at $t = 0$,

$$Ec_i = \sum_{x_{-i} \in \mathcal{D}_2^2} (c_i(x_i, x_{-i}) \prod_{x_j \in x_{-i}} q_j^{x_j}(0)).$$

Using MBRAR, the best strategies for players x_1 and x_2 at time $t = 1$ can be chosen, which is listed in Table 2 and Table 3.

Table 2 Best strategies when $u(0) = 1$

$x_1(0)x_2(0)$	11	12	21	22
Ec_1	4.5	4.5	8	8
Ec_2	4.5	8	4.5	8
$x_1(1)$	2	2	2	2
$x_2(1)$	2	2	2	2

Table 3 Best strategies when $u(0) = 2$

$x_1(0)x_2(0)$	11	12	21	22
Ec_1	1.5	1.5	4	4
Ec_2	1.5	4	1.5	4
$x_1(1)$	2	2	2	2
$x_2(1)$	2	2	2	2

Denote $\theta(t) = x_1(t)x_2(t)$, then we have

$$\theta(1) = x_1(1)x_2(1) = M^\theta u(0)\theta(0), \tag{22}$$

where $M^\theta = \delta_4[4\ 4\ 4\ 4\ 4\ 4\ 4\ 4]$.

From (22), given $\theta(0) \in \{\delta_4^1, \delta_4^2, \delta_4^4, \delta_4^4\}$, we have $\theta(1) = \delta_4^4$, then we calculate $c_3(\delta_2^1, \delta_4^4) = 0$ and $c_3(\delta_2^2, \delta_4^4) = 2$. It is clear that player x_3 selects strategy 2 at time $t = 1$,

$$x_3(1) = q_3(1) = (\frac{1}{3}, \frac{2}{3})^{\mathrm{T}}.$$

Then we do the above calculation process repetitively. After $t = N \geqslant 1$ steps, we can obtain

$$\theta(N+1) = M^\theta u(N)\theta(N), \tag{23}$$

and

$$x_3(N) = q_3(N) = (\frac{1}{N}, \frac{N-1}{N})^{\mathrm{T}}.$$

It follows that $\lim\limits_{N\to\infty} x_3(N) = (0, 1)^{\mathrm{T}} = \delta_2^2$. According to (14), when $t \to \infty$, let $x(t) = \ltimes_{i=1}^{3} x_i(t)$, we have

$$x(t+1) = Mu(t)x(t),$$

where $M = \delta_8[8\ 8\ 8\ 8\ 8\ 8\ 8\ 8\ 8\ 8\ 8\ 8\ 8\ 8\ 8\ 8]$.

From Table 1, the controller's payoff matrix can be easily calculated as

$$\varPhi = \begin{bmatrix} 6 & 3 & 3 & 0 \\ 10 & 6 & 6 & 2 \end{bmatrix}. \tag{24}$$

Assume $\lambda = 0.9$. From (24), we have

1) when $t = 0$, $u^*(0)|_{\theta(0)=\delta_4^1(_4^2,\ \delta_4^4,\ \delta_4^4)} = \delta_2^2$. And $J_0|_{\theta(0)=\delta_4^1} = 10$; $J_0|_{\theta(0)=\delta_4^2} = 6$; $J_0|_{\theta(0)=\delta_4^3} = 6$; $J_0|_{\theta(0)=\delta_4^4} = 2$.

2) when $t \geqslant 1$, according to (23), we have $\theta(t) = \delta_4^4$, then $u^*(t)|_{\theta(t)=\delta_4^4} = \delta_2^2$. And $J_t|_{\theta(t)=\delta_4^4} = \sum\limits_{t=0}^{\mathrm{T}} 0.9^t \times 2$.

It follows that for any $\theta(0)$, we select $u^*(t) = \delta_2^2$ for $t \geqslant 0$, such that the maximum of the controller's payoff will be maintained.

On the other hand, the controller is a time-invariant state feedback, that is,

$$u^*(t) = H^\theta x_1(t)x_2(t) = Hx(t),$$

where $H^\theta = \delta_2[2\ \ 2\ \ 2\ \ 2]$ and $H = H^\theta(I_4 \otimes \mathbf{1}_2^{\mathrm{T}}) = \delta_2[2\ \ 2\ \ 2\ \ 2\ \ 2\ \ 2\ \ 2\ \ 2]$.

According to (15), when $t \to \infty$, we have

$$x(t+1) = Lx(t),$$

where $L = MHO_{2^3}^R = \delta_8[8\ \ 8\ \ 8\ \ 8\ \ 8\ \ 8\ \ 8\ \ 8]$. It is clear that $x^* = \delta_8^8$ is an unique fixed point. Hence, the evolutionary dynamics can globally converge to x^* after finite steps, where $x^* = \delta_8^8 \sim (2, 2, 2)$ is the pure Nash equilibrium of this control NEG.

4.2 Optimal control of stochastic NEGs with local information

For the stochastic model, the optimization objective (16) becomes

$$J = \mathrm{E}[\sum_{t=0}^{T} \lambda^t c(u_1(t), \cdots, u_m(t), \theta_1(t), \cdots, \theta_m(t))], \tag{25}$$

where $\theta(t) = \ltimes_{p=1}^{m} \theta_p(t) \in \Upsilon_{k^r}$, $\theta_p(t) \in \Upsilon_{k^{l_p}}$, $r = \sum\limits_{p=1}^{m} |\mathcal{N}_p|$, and $M \in \Upsilon_{k^n \times k^{m+n}}$ is a probabilistic matrix in (14).

Proposition 3[27] Let $J^*(\theta(0))$ be the optimal value of (25), then

$$J^*(\theta(0)) = J_0(\theta(0)). \tag{26}$$

Where the function J_0 is given by the last step of a dynamic programming algorithm which proceeds backward in time from $t = N$ to $t = 0$:

$$J_N(\theta(N)) = \max_{u(N)} \lambda^N c(u(N), \theta(N)), \tag{27}$$

and

$$J_t(\theta(t)) = \max_{u(t)} \lambda^t E[c(u(t), \theta(t)) + J_{t+1}(\theta(t+1))]. \tag{28}$$

Note that the expectation on the right hand side of (28) is conditional on $\theta(t)$ and $u(t)$. Denote the transition probability of $\theta(t)$ from j to i under the controller $u(t)$ as

$$p_{i,j}(u) = P(\theta(t+1) = i|\theta(t) = j, u(t) = u).$$

In [16], the recursive solution for finite horizon optimization problem of mix-valued probabilistic logical networks with global information has been obtained. Similarly, we can provide the dynamic programming solution to the above optimization problem (27) and (28):

$$\begin{cases} J_N(\theta(N)) = \max\limits_{u(N)\in\Delta_{k^m}} \lambda^N u^{\mathrm{T}}(N)\Phi\theta(N); \\ J_t(\theta(t)) = \max\limits_{u(t)\in\Delta_{k^m}} [\lambda^t u^{\mathrm{T}}(t)\Phi\theta(t) + \\ \qquad \sum\limits_{i=1}^{k^r} p_{i,\theta(t)}(u(t))J_{t+1}(i)], \\ \qquad t = N-1, \cdots, 1, 0. \end{cases}$$

By using the formulas (25) to (31) in [16], we can obtain the complete solution for finite horizon optimization of stochastic NEGs with local information. Note that the difference between (29) and [16] for the solution to the finite horizon optimization is that $\theta(t)$ in this paper is the local information and $x(t)$ in [16] denotes the global information for the controllers.

Example 2 Recall Example 1. Assume G is the Matching Pennies game, where the payoff bi-matrix is given in Table 4. The network graph is shown in Fig. 2, where the edges are all directed.

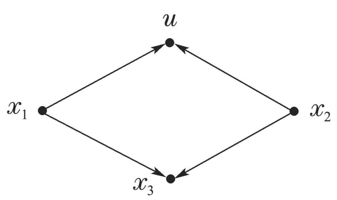

Fig. 2 Network graph

Table 4 Payoff bi-matrix of matching pennies

$P_1\backslash P_2$	1	2
1	(1, −1)	(−1, 1)
2	(−1, 1)	(1, −1)

Assume x_3 takes strategy 1 with probability 1 all the time and the strategy updating law is MBRAR. Then we can calculate the expected payoffs and the best responses of x_1 and x_2, which are shown in Table 5 and Table 6.

Table 5 Expect payoffs and best responses when $u = 1$

profile	11	12	21	22
Ec_1	2	2	−2	−2
Ec_2	2	−2	2	−2
$x_1(t+1)$	1	1	1	1
$x_2(t+1)$	1	1	1	1

Table 6 Expect payoffs and best responses when $u = 2$

profile	11	12	21	22
Ec_1	0	0	0	0
Ec_2	0	0	0	0
$x_1(t+1)$	$(\frac{1}{2},\frac{1}{2})$	$(\frac{1}{2},\frac{1}{2})$	$(\frac{1}{2},\frac{1}{2})$	$(\frac{1}{2},\frac{1}{2})$
$x_2(t+1)$	$(\frac{1}{2},\frac{1}{2})$	$(\frac{1}{2},\frac{1}{2})$	$(\frac{1}{2},\frac{1}{2})$	$(\frac{1}{2},\frac{1}{2})$

Denote $\theta(t) = x_1(t)x_2(t)$, we have

$$\theta(t+1) = M^\theta u(t)\theta(t), \tag{29}$$

where

$$M^\theta = \begin{bmatrix} 1 & 1 & 1 & 1 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}.$$

According to (14), we have

$$x(t+1) = Mu(t)x(t),$$

where $M \in \Upsilon_{8\times16}$.

Assume $\lambda = 0.9$ and from Table 4, we have

$$\Phi = \begin{bmatrix} -2 & 0 & 0 & 2 \\ 2 & 0 & 0 & -2 \end{bmatrix}. \tag{30}$$

Using (29) and the formulas in [16], the finite horizon ($N = 3$) optimal controls are as follows:

1) When $t = N = 3$, we have

$$u^*(3)|_{\theta(3)=\delta_4^1} = \delta_2^2,\ J_3|_{\theta(3)=\delta_4^1} = 0.9^3 \times 2 = 1.458;$$

$$u^*(3)|_{\theta(3)=\delta_4^2} = \delta_2^1(\text{or } \delta_2^2),\ J_3|_{\theta(3)=\delta_4^2} = 0.9^3 \times 0 = 0;$$

$$u^*(3)|_{\theta(3)=\delta_4^3} = \delta_2^1(\text{or } \delta_2^2),\ J_3|_{\theta(3)=\delta_4^3} = 0.9^3 \times 0 = 0;$$

$$u^*(3)|_{\theta(3)=\delta_4^4} = \delta_2^1,\ J_3|_{\theta(3)=\delta_4^4} = 0.9^3 \times 2 = 1.458.$$

Then $J_3 = [1.458\ \ 0\ \ 0\ \ 1.458\,]^{\mathrm{T}}$.

2) When $t = N-1 = 2$, we calculate that

$$u^*(2)|_{\theta(2)=\delta_4^1} = \delta_2^2;\ u^*(2)|_{\theta(2)=\delta_4^2} = \delta_2^1;$$

$$u^*(2)|_{\theta(2)=\delta_4^3} = \delta_2^1;\ u^*(2)|_{\theta(2)=\delta_4^4} = \delta_2^1.$$

Then $J_2 = [2.349\ \ 1.458\ \ 1.458\ \ 3.078\,]^{\mathrm{T}}$.

3) When $t = N-2 = 1$, we have

$$u^*(1)|_{\theta(1)=\delta_4^1} = \delta_2^2;\ u^*(1)|_{\theta(1)=\delta_4^2} = \delta_2^1;$$

$$u^*(1)|_{\theta(1)=\delta_4^3} = \delta_2^1; u^*(1)|_{\theta(1)=\delta_4^4} = \delta_2^1.$$

And then $J_1 = [3.88575\ \ 2.349\ \ 2.349\ \ 4.149\,]^{\mathrm{T}}$.

4) When $t = N-3 = 0$, we have

$$u^*(0)|_{\theta(0)=\delta_4^1} = \delta_2^2; u^*(0)|_{\theta(0)=\delta_4^2} = \delta_2^1;$$

$$u^*(0)|_{\theta(0)=\delta_4^3} = \delta_2^1; u^*(0)|_{\theta(0)=\delta_4^4} = \delta_2^1.$$

And then

$$J_0 = [5.1831875\ \ 3.88575\ \ 3.88575\ \ 5.88575\ ]^{\mathrm{T}}.$$

Finally, according to Proposition 5.2 in [16], we have the infinite horizon ($0 \leqslant t \leqslant \infty$) optimal control

$$u^*(t) = H^{\theta}\theta(t),$$

where $H^{\theta} = \delta_2[\ 2\ \ 1\ \ 1\ \ 1\ ]$.

According to (15), the algebraic form of control NEGs with local information is $x(t+1) = Lx(t)$, where $L = MH^{\theta}(I_4 \otimes \mathbf{1}_2^{\mathrm{T}})O_{2^3}^R$. When $k \geqslant 36$, we calculate that L converges to L^*, where

$$\mathrm{Col}_i(L^*) = [0.571\ \ 0\ \ 0.143\ \ 0\ \ 0.143\ \ 0\ \ 0.143\ \ 0]^{\mathrm{T}},\quad i = 1, 2, \cdots, 8.$$

Using Lemma 1, we have

$$x_1(\infty) = x_2(\infty) = [0.7143\ \ 0.2858\ ]^{\mathrm{T}}.$$

That is, the control NEGs with local information can globally converge to a stationary distribution $x(\infty) = x_1(\infty) \ltimes x_2(\infty) \ltimes x_3(\infty)$, where $x_3(\infty) = [\ 1\ \ 0\ ]^{\mathrm{T}}$.

5 Conclusions

The optimization problem of control NEGs with local information was considered. Using the STP method, the dynamics of control NEGs with local information has been converted into an algebraic form. Then the optimal control sequences for the deterministic and stochastic NEGs were presented respectively. Accordingly, some easily computable formulas were provided by generalizing the corresponding results in [16] for mix-valued logical networks with global information.

References:

[1] HAMILTON W D. The genetical evolution of social behaviour. *Journal of Economic Theory*, 1964, 7(11): 1 – 16.

[2] SMITH J M, PRICE G R. The logic of animal conflict. *Nature*, 1973, 246: 15 – 18.

[3] SUGDEN R. *The Economics of Rights, Cooperation and Welfare*. Oxford, England: Blackwell, 1986.

[4] OHTSUKI H, HAUERT C, LIEBERMAN E, et al. A simple rule for the evolution of cooperation on graphs and social networks. *Nature*, 2006, 441: 502 – 505.

[5] MOJICA-NAVA E, MACANA C A, QUIJANO N. Dynamic population games for optimal dispatch on hierarchical microgrid control. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2014, 44(3): 306 – 317.

[6] NOWAK M A, MAY R M. Evolutionary games and spatial chaos. *Nature*, 1992, 359: 826 – 829.

[7] TANG C, WU B, WANG J, et al. Evolutionary origin of asymptotically stable consensus. *Scientific Reports*, 2014, Article number: 4590.

[8] CHENG D, QI H, LI Z. *Analysis and Control of Boolean Networks: A Semi-tensor Product Approach*. London: Springer, 2011.

[9] CHENG D, QI H, ZHAO Y. *An Introduction to Semi-tensor Product of Matrices and Its Applications*. Singapore: World Scientific, 2012.

[10] LI F, SUN J. Stability and stabilization of Boolean networks with impulsive effects. *Systems & Control Letters*, 2012, 61(1): 1 – 5.

[11] LI H, XIE L, WANG Y. On robust control invariance of Boolean control networks. *Automatica*, 2016, 68(C): 392 – 396.

[12] MENG M, LIU L, FENG G. Stability and l_1 gain analysis of Boolean networks with Markovian jump parameters. *IEEE Transactions on Automatic Control*, 2017, 62(8): 4222 – 4228.

[13] LU J, ZHONG J, HUANG C, et al. On pinning controllability of Boolean control networks. *IEEE Transactions on Automatic Control*, 2016, 61(6): 1658 – 1663.

[14] LIU Z, WANG Y, LI H. Two kinds of optimal controls for probabilistic mix-valued logical dynamic networks. *Science China Information Sciences*, 2014, 57(5): 1 – 10.

[15] ZHAO G, FU S. Matrix approach to trajectory control of higher-order k-valued logical control networks. *IET Control Theory and Applications*, 2017, 11(13): 2110 – 2115.

[16] CHENG D, ZHAO Y, XU T. Receding horizon based feedback optimization for mix-valued logical networks. *IEEE Transactions on Automatic Control*, 2015, 60(12): 3362 – 3366.

[17] LI H, ZHAO G, MENG M, et al. A survey on applications of semi-tensor product method in engineering. *Science China Information Sciences*, 2018, 61(1): 010202.

[18] WANG Y, LIU T, CHENG D. From weighted potential game to weighted harmonic game. *IET Control Theory and Applications*, 2017, 11(13): 2161 – 2169.

[19] WANG Y, CHENG D, LIU X. Matrix expression of Shapley value and its application to distributed resource allocation. *Science China Information Sciences*, 2018, DOI: 10.1007/s11432-018-9414-5.

[20] GUO P, ZHANG H, ALSAADI F E, et al. Semi-tensor product method to a class of event-triggered control for finite evolutionary networked games. *IET Control Theory and Applications*, 2017, 11(13): 2140 – 2145.

[21] CHENG D, QI H, HE F, et al. Semi-tensor product approach to networked evolutionary games. *Control Theory and Technology*, 2014, 12(2): 198 – 214.

[22] CHENG D, HE F, QI H, et al. Modeling, analysis and control of networked evolutionary games. *IEEE Transactions on Automatic Control*, 2015, 60(9): 2402 – 2415.

[23] GUO P, WANG Y, LI H. Algebraic formulation and strategy optimization for a class of evolutionary networked games via semi-tensor method. *Automatica*, 2013, 49(11): 3384 – 3389.

[24] RASMUSEN E. *Games and Information: An Introduction to Game Theory*. Oxford, England: Blackwell, 2007.

[25] YOUNG H P. The evolution of conventions. *Econometrica*, 1993, 61(1): 57 – 84.

[26] GUO P, WANG Y. Matrix expression and vaccination control for epidemic dynamics over dynamic networks. *Control Theory and Technology*, 2016, 14(1): 39 – 48.

[27] BERTSEKAS D P. *Dynamic Programming and Stochastic Control*. New York: Academic Press, 1976.

作者简介:

王元华 博士, 讲师, 目前研究方向为博弈论及逻辑动态系统, E-mail: wyh_1005@163.com;

刘希玉 教授, 博士生导师, 目前研究方向为信息管理与电子商务及计算智能等, E-mail: sdxyliu@163.com.