A GAME-CHAIN-BASED APPROACH FOR DECISION MAKING

An Tingyu, Watanabe tsunami
Kochi University of Technology

***ABSTRACT*:** Nowadays, with the rapid development of information society, decision-making problems become more and more complicated especially in large scale systems such as infrastructure, environmental and industrial fields, which are usually accompanied by psychological competition between involved parties in a complicated, uncertain and dynamic situation. From a holistic perspective of system, a specific decision-making method which is described as game-chain-based decision making has been proposed in this paper in order to seek a solution for these problems. It takes other involved parties' thinking into consideration and explores consequences of holistic instead of reductionist in order to pursue scientific proof and optimal choice for decision makers.

Based on systematic and game ideas, this paper attempts to find a solution by mathematical modeling. First, characteristics and descriptions of this sort of decision-making problems are summarized which can be abstracted as game chain. Through giving proper definitions of decision point, state, action, state transition probability and immediate utility, a mathematical model is set up which translates the process of game chain decision making into Markov decision process with a list of 5 objects. Second, with support of Game Theory and Markov Decision Process, the corresponding equilibrium (system equilibrium) in a holistic view of system is developed under the principle function of expected total utility and then proved to be existed under some certain conditions. Then a pathway to find the system equilibrium is given. Finally the proposed method is demonstrated through a tank virtual game.

KEYWORDS: game chain, system equilibrium, Markov decision process

1. INTRODUCTION

Currently, management for large-scale, highly interconnected systems such as infrastructure, environmental and industrial systems seems to become more complicated partly because it is usually required to balance the involved parties' requirements and benefits in complex, uncertain and dynamic situations from a holistic perspective. Take infrastructure for example, which herein refers to the technical structures that support a society, such as roads, water supply, sewers, telecommunications, power grids, and so forth. One typical attribute of infrastructure is that the system or network tends to evolve over time as it is continuously modified, improved, enlarged, and as various components are rebuilt, decommissioned or adapted to other uses. It usually brings a complex, uncertain and dynamic environment for infrastructure management. Another attribute is that system components are commonly interdependent, not usually capable of subdivision, and moreover, the management process is composed of several stages or steps involving multiple parties. Consequently, management from holistic and game perspectives seems quite necessary. As such, an organization's success will largely depend on its ability to manage the problems induced by those attributes. One important involvement during the management process is decision making, which is especially made more difficult by those attributes mentioned above, and consequently accompanied by psychological competition between involved parties in a complicated, uncertain and dynamic situation. This paper is an illustration of the application of system process modeling and game idea to the resolution of complex decision problems.

A holistic illustration is needed to understand the overall situation by showing the interconnectivity between involved parties, which is the foundation of the system processes modeling. The system thinking movement (Checkland, 1999; Senge, 2006) is an exploration of consequences of holistic instead of reductionist ways of thinking. System approaches are a way of grasping and managing situations of complexity and uncertainty in which there are no simple answers and people and their attitudes are an integral part of the problem. That is to say, such a decision-making method from a holistic perspective of system should take involved parties' thinking into account and allow for complexity to be managed. Based on above requirements, it seems reasonable to combine systematic idea with game idea in order to pursue optimal choice for decision makers.

Game theory is considered as a branch of applied mathematics which can be used in the social sciences, biology, engineering, political science, international relations, computer science (mainly for artificial intelligence) etc. Game theory attempts to capture behavior in strategic situations mathematically, in which an individual's success in making choices depends on the choices of others. It plays an important role in solving practical decision-making problems existed in many fields. Nowadays, "game theory is a sort of umbrella or 'unified field' theory for the rational side of social science, where social is interpreted broadly, to include human as well as non-human players (computers, animals, plants)" (Aumann 1987). Traditional applications of game theory attempt to find equilibrium in these games, in which each player of the game has adopted a strategy that they are unlikely to change. Many equilibrium concepts have been developed (most famously the Nash equilibrium) in an attempt to capture this idea. However, as it is generally focused on one game, it seems necessary to add something to game theory in order to find a solution for problems with a sequence of games which can be depicted as a game chain (Figure 1). This paper is just an attempt with structure showed in Figure 2.

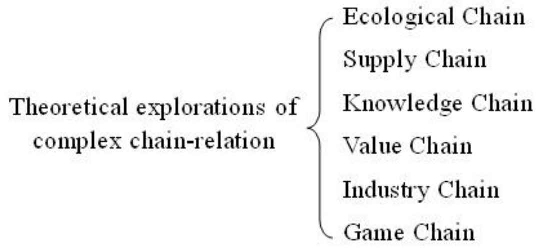

Figure1 Chain-relation exploration

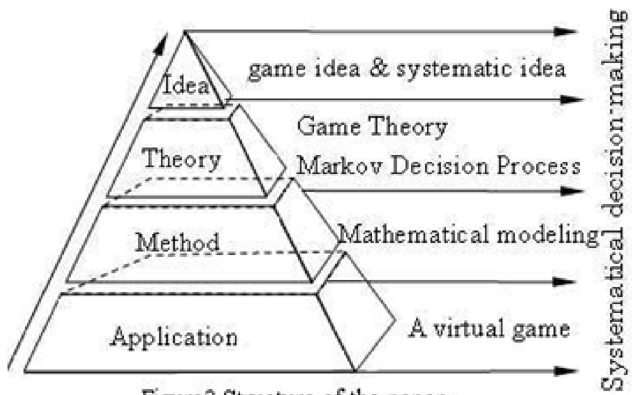

Figure2 Structure of the paper

2. CHARACTERISTICS OF GAME CHAIN DECISION PROBLEMS

The game chain decision-making problems have characteristics as follows:

(1) It can be abstracted as a chain made up of a sequence of game units in which the players take the control of decision variables by themselves, respectively.

(2) Some game unit will exert an influence on the following game' condition by taking action, which

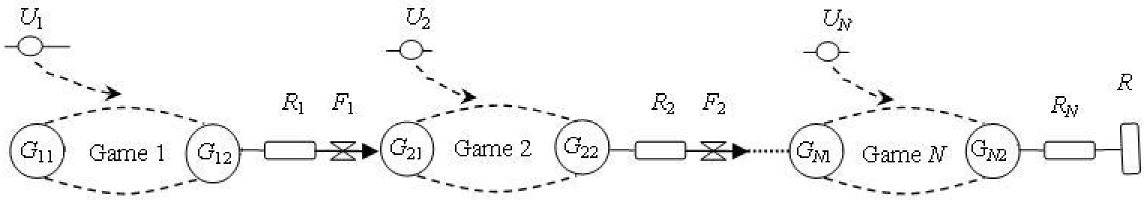

Figure 3 Description Graph of the game chain

requires that, any decision-maker should not only consider the immediate reward but also take care of long-term benefit. In other words, today's decision will influence tomorrow, and tomorrow's decision will influence future too. If you disregard impact on future and only take the interests of current stage into account, you do not make a wise decision from the long-term perspective. In a word, decision makers are facing a decision process which is frequently made up of more than one related phases.

(3) The decision-making problem is displayed with a hierarchically discrete structure, thus a system strategy for the game chain should be formed from a sequence of strategies for each stage.

(4) The decision-making process is an orderly process which means that the players make decisions follow the sequence of game units.

(5) The final system strategy should be acceptable for each participant, containing two layers of meaning: first, decision-makers seek their own "optimal" strategy on the premise of taking others' possible strategies into account; second, they balance the immediate reward with possible future reward in various circumstances. As a result, all parties are unlikely to change or else they will get less benefit.

In conclusion, game chain decision problem is firstly a game problem, which means that we should take others' strategies into consideration. Secondly, it is a sequential decision-making process usually under a complicated, uncertain and dynamic environment. As decision makers are frequently faced with the complexity of not only taking care of immediate reward but also attempting to create opportunities for future, it seems a challenge to find a solution for this problem.

3. MODELING FOR A CERTAIN KIND OF GAME CHAIN

3.1 Decision process description

The certain game chain to be analyzed is illustrated as $GC=(G_1,G_2,\cdots,G_N)$ (see Figure 3). $G_{ij}\,(j=1,2)$ represent the players of game unit $i\,(i=1,2,\cdots,N)$; U_i and R_i $(i=1,2,\cdots,N)$ represent the payoff and outcome vectors of game unit *i*, respectively; F_i is the relation function which reflects the influence of previous game on the one behind.

The decision-making process based on game chain is showed in Figure 4. Our concerning is that how to make a proper decision in such a process in a holistic view which means taking all stages into consideration.

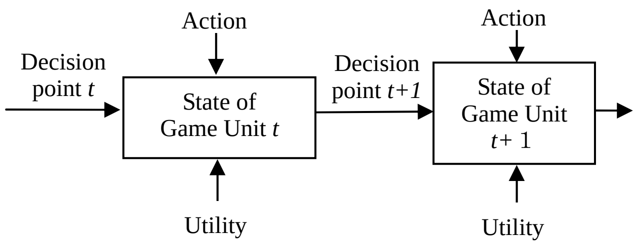

Figure 4 The game-chain decision process

3.2 Elements of the game chain decision process

(1) Decision-point and stage

Each decision should be made at a certain moment called decision-point in this paper. The corresponding set of decision-points in the game chain $GC=(G_1,G_2,\cdots,G_N)$ is then denoted as $T=\{1,2,\cdots,N\}$, while the process between two adjacent decision-making points is known as Stage.

(2) State and action

Condition of the moment when game unit t starts is defined as the state of decision-point t and all possible states form the set S_t. Given $i_t \in S_t$, denote $A_t(i_t)$ as the action set of game unit t which gives all possible action profiles of players in game unit t. Notice that under a certain state, a certain action may occur or an uncertainty may also be available with a number of possible actions to be selected randomly. Thus for decision-point t, we define $Dis(A_t(i_t))$ as all probability distributions on $A_t(i_t)$. Then selecting action or making decision under the state $i_t \in S_t$ at the decision-point t is equivalent to selecting a probability distribution $\pi_t(\cdot) \in Dis(A_t(i_t))$ which displays the probability of action $a_t \in A_t(i_t)$ as $\pi_t(a_t)$, satisfying

$$\sum_{a_t \in A_t(i_t)} \pi_t(a_t) = 1$$

If this distribution is degenerated, it means that action is selected in determinacy.

Assumption 1

$\forall t \in T$, S_t is a finite set.

Assumption 2

$\forall t \in T$, $i_t \in S_t$, strategy sets of both players in game unit t are all finite sets which guarantees that $A_t(i_t)$ be a finite set too.

(3) State transition probability and immediate utility

We notice that the action $a_t \in A_t(S_t)$ selected at the decision-point t would influence on state of the next decision-point t+1. Naturally, we interpret what will be brought by taking action $a_t \in A_t(i_t)$ under the state of $i_t \in S_t$ at the decision-point t as follows:

1) A utility profile at the tth stage in the form of $r_t(i_t, a_t) = (r_t^{(G_{t1})}(i_t, a_t), r_t^{(G_{t2})}(i_t, a_t))$, which is expressed as immediate utility in this paper. In each game unit, either player's immediate utility is dependent on strategies of both players at the corresponding stage. Therefore, once the action a_t is known, we can determine the immediate utility of the tth stage.

Note: As a matter of convenience, we suppose that immediate utility of each game unit has the same dimension (otherwise we can achieve it by method of normalization).

2) Influence on state of the next decision-point t+1, which is represented by the probability distribution $p^{(t)}(\cdot | i_t, a_t)$ described as follows:

If the system takes action $a_t \in A_t(i)$ under the state of $i \in S_t$ at the decision-point t, then it will transfer to state $j \in S_{t+1}$ at the decision-point t+1 with probability $p_{ij}^{(t)}(a_t)$, satisfying

$$\sum_{j \in S_{t+1}} p_{ij}^{(t)}(a_t) = 1 \qquad (1)$$

Denote $p = \{p^{(t)}(j | i, a_t), i \in S_t, j \in S_{t+1}, a_t \in A_t(i), t \in T\}$ as the state transition probability distribution family, which is generally related to decision-point. The state transition probability distribution displays us the relation of two adjacent game units, which is expressed as the relation function before. For example, if the state of game unit t+1 can be only proved to be j_{a_t}, then the corresponding transition probability is

$$p^{(t)}(j | i, a_t) = \begin{cases} 1, j = j_{a_t} \\ 0, others \end{cases} \quad i \in S_t, a_t \in A_t(i), t \in T \qquad (2)$$

Thus we have the relation function as

$$F_t(i, a_t, j) = p_{ij}^{(t)}(a_t) \qquad (3)$$

for $i \in S_t, a_t \in A_t(i), j \in S_{t+1}$. In the game chain decision problems, the relationship between action at decision-point t and state of decision-point t+1 can be determined using game theory or obtained from experience, thus given the state of decision-point t, transition probability to the state of decision-point t+1 is independent of previous states or actions, which indicates that game-chain-based decision process possesses the Markov property. Therefore, a mathematical model for game chain based on

Markov decision process is set up with a list of 5 objects as follows:

$$\{T, S = \bigcup_{t \in T} S_t, A = \bigcup_{t \in T} A_t(i_t), p^{(t)}(\cdot|i_t, a_t), r_t(i_t, a_t)\} \qquad (4)$$

(4) Decision policy and strategy

From the description above, we notice the game chain decision process is composed of successive states and actions. Label one trace of decision process as h_t, which denotes a series of states and actions from decision-point 1 to decision-point *t* with state of the decision-point *t* as i_t:

$$h_t \overset{def}{=} (i_1, a_1, i_2, a_2, \ldots, i_t), \qquad t \in T \qquad (5)$$

in which $i_k \in S_k, a_k \in A_k(i_k)\left(k = 1, \cdots, t-1\right)$, $i_t \in S_t$. The universal set of all that kind of traces is denoted as H_t, which actually displays all possible decision pathways to reach the state of i_t at the decision-point *t*.

A decision policy should give the principle to follow when making decisions or selecting actions under different states at the decision-points.

Definition 1: If function f_t satisfies: for any $i_t \in S_t$, there is $f_t(i_t) \in A_t(i_t)$, then f_t is defined as a deterministic decision policy or decision function for short. The corresponding decision function sequence $\pi = (f_1, f_2, \ldots, f_N)$ is called a deterministic strategy.

Definition 2: If a family of probability distribution π_t on the state space satisfies: given any pathway $h_t \in H_t$ to approach the decision-point *t* with the state $i_t \in S_t$, $\pi_t(\cdot|h_t) \in Dis(A_t(i_t))$ is a probability distribution on $A_t(i_t)$, such that

$$\begin{cases} \pi_t(a_t|h_t) \geq 0 \\ \sum_{a_t \in A_t(i_t)} \pi_t(a_t|h_t) = 1 \end{cases} \qquad (6)$$

then π_t is defined as a generic decision policy. The corresponding decision policy sequence

$$\pi = (\pi_1, \pi_2, \ldots, \pi_N)$$

is called a generic strategy. The universal set of generic strategies is denoted as Π.

Definition 3: If the decision policy π_t at the decision-point *t* is independent on previous states and actions, which means

$$\pi_t(\cdot|h_t) = \pi_t(\cdot|i_t), h_t \in H_t, t \in T \qquad (7)$$

then π_t is a Markov decision policy and the corresponding decision policy sequence

$$\pi = (\pi_1, \pi_2, \ldots, \pi_N)$$

is called random Markov strategy. The universal set of random Markov strategies is denoted as Π_m. If $\pi_t(\forall t \in T)$ are all degenerate distributions, that is to say, there exists a decision function sequence $(f_1, f_2, \ldots, f_N)$, satisfying

$$\pi_t(f_t(i_t)|i_t) = 1, \forall i_t \in S_t \qquad (8)$$

then π is called a deterministic Markov strategy, the universal set of which is denoted as Π_m^d.

(5) Stochastic Process and Principle function

For $t \in T$, Y_t and Δ_t denote state and action at the decision-point *t*, respectively. Obviously, they are random variables depending on strategy π. Now we have the random sequence

$$(Y_1, \Delta_1, Y_2, \Delta_2, \cdots, Y_N, \Delta_N)$$

denoted as $L(\pi)$. As it depends on the initial state of probability distribution of Y_1, transition probability and strategy π, we call it as stochastic process induced by strategy π. Accordingly, denote the probability of occurrence of an event in $L(\pi)$ as $P_\pi(\cdot)$, and corresponding expected utility as $E_\pi(\cdot)$. $\forall \pi \in \Pi$, define a random variable sequence in the stochastic process induced by π as

$$R(\pi) = (R_1(\pi), R_2(\pi), \ldots, R_N(\pi))$$

in which the random variable $R_t(\pi) = r(Y_t, \Delta_t)(t \in T)$ represents the immediate utility of game unit *t* at the decision-point *t*.

Now, our purpose is to define a principle function to discriminate between good and bad strategies, and then try to find an 'optimal' strategy in the sense of "equilibrium".

Definition 4

Finite game chain decision problem is a systematic decision process with finite stages in some degree. Therefore, we attempt to set up expected total utility as a criterion for this game chain. Suppose that there is an immediate utility at the decision point t as

$$r_t(i_t, a_t) = (r_t^{(G_{t1})}(i_t, a_t), r_t^{(G_{t2})}(i_t, a_t)), \forall i_t \in S_t, a_t \in A_t(i_t), t \in T$$

Then define the expected total utility obtained by selecting action $a_t \in A_t(i_t)$ at the decision-point t under the strategy π and state $i_t \in S_t$ as

$$V_{t,N}(i_t, a_t, \pi) = r_t(i_t, a_t) \oplus \sum_{n=t+1}^{N} E_\pi\{r(Y_n, \Delta_n)\}, \forall i_t \in S_t, a_t \in A_t(i_t) \quad (9)$$

where

$$E_\pi\{r(Y_n, \Delta_n)\} = \sum_{(i_n \in S_n, a_n \in A_n(i_n))} P_\pi\{Y_n = i_n, \Delta_n = a_n \,|\, h_n\} r_n(i_n, a_n)$$

$$n = t+1, \cdots, N \quad (10)$$

in which, h_n should meet the condition that it passes (i_t, a_t) at the decision-point t.

Notes:

1) Actually, $V_{t,N}(i_t, a_t, \pi)$ is a vector in the form of

$$V_{t,N}(i_t, a_t, \pi) = (V_{t,N}^{(G_{t1})}(i_t, a_t, \pi), V_{t,N}^{(G_{t2})}(i_t, a_t, \pi))$$

2) The expected total utility of decision-point N is

$$V_{N,N}(i_N, a_N) = r_N(i_N, a_N) = (r_N^{(G_{N1})}(i_N, a_N), r_N^{(G_{N2})}(i_N, a_N))$$

3) $\oplus$, $\sum$ are different from a summation notation in that only utilities of the identical player can be added.

4. SOLUTION FOR THE GAME-CHAIN DECISION PROBLEM

4.1 Existence of system equilibrium

Definition 5:

At the decision-point t, by employing expected total utility as payoff, we establishes a new game $G_t^*(i_t)(t \in T)$. Then a strategy

$$\pi^* = (\pi_1^*, \pi_2^*, \ldots, \pi_N^*) \in \Pi$$

is a system equilibrium strategy if satisfying

$$\pi_t^*(i_t) \in \arg Eq\{G_t^*(i_t)\}, \forall t \in T, i_t \in S_t \quad (11)$$

where $\arg Eq$ means to find Nash Equilibrium of the new game $G_t^*(i_t)$. $\pi_t^*(i_t)$ represents the decision or the selected action at the decision-point t under the state of $i_t \in S_t$.

Given a fixed initial state $i_1 \in S_1$, we can find the system equilibrium $E^* = (E_1^*, E_2^*, \ldots, E_N^*)$, in which $E_1^* = \pi_1^*(i_1)$, and $E_2^* = \pi_2^*(i_2)$ by combining i_1, $E_1^* = \pi_1^*(i_1)$ and state transition probability, and then $E_3^*, \cdots, E_N^*$.

From Assumption 1 and 2, we can see that the games at any decision-point are all finite games. As a result, they all have Nash Equilibrium, which means that $\forall t \in T$, $i_t \in S_t$

$$\arg Eq\{G_t^*(i_t)\} \neq \phi \quad (12)$$

Then we can approach a conclusion that finite game chain decision making process under complete information can be defined as a Markov decision process, and under the principle of expected total utility (9), the system equilibrium exists.

4.2 Solution for system equilibrium

4.2.1 Theoretical foundation

For generic strategies Π, as all possible pathways will be considered, it won't leave out any strategy while it is too difficult to operate when encountering with large sets of states and actions brought by numerous game units or a large strategy space.

(1) A large amount of calculation. When the game chain is in a large scale, "combinatorial explosion" will occur with an exponential growth of possible emergence of states, which will lead to inefficiency.

(2) A large calculation error. As the transition probability is dependent on experience on many occasions, so there will be errors inevitably. Then in the process to find the system strategy in a holistic view, with a larger time span, the accumulated error

will grow larger, and even result in that theoretical "optimal" solution is not practical at all.

Theorem 1: For any $\pi \in \Pi$ and $i \in S_1$, there exists a random Markov strategy $\pi' \in \Pi_m$, satisfying

$$P_{\pi'}\{Y_t = j, \Delta_t = a \mid Y_1 = i\} = P_{\pi}\{Y_t = j, \Delta_t = a \mid Y_1 = i\}, \forall t \in T \quad (13)$$

Brief proof: Fix the state $i \in S_1$, for any $j \in S_t, a \in A_t(j)$ and $t \in T$, define a random Markov policy as

$$\pi'_t(a \mid j) \equiv P_{\pi}\{\Delta_t = a \mid Y_t = j, Y_1 = i\}$$

Then we can prove that such strategy $\pi' \in \Pi_m$ can meet the condition above.

The meaning of this theorem lies in that searching for an optimal strategy from a generic strategy set can be equivalent to searching from a random Markov strategy set. The conclusion comes to existence only under the condition of fixing the initial state which indicates that the random Markov strategy π' is dependent on the initial state $i \in S_1$. Fortunately, it can be met in game chain decision problems described in this paper. Therefore, we confine the system equilibrium to the set of random Markov strategy, in accordance with which, design the corresponding pathway to find the equilibrium.

Theorem 2: Given $\pi = (\pi_1, \pi_2, \ldots, \pi_N) \in \Pi_m$, $L(\pi)$ is a non-homogeneous Markov chain with transition probability of the decision point *t* is

$$P_{\pi}\{Y_{t+1} = j, \Delta_{t+1} = b \mid Y_t = i, \Delta_t = a\} = p(j \mid i, a)\pi_{t+1}(b \mid j) \quad (14)$$

Brief proof: For $t \in T$ and $h_t \in H_t$, expand $P_{\pi}\{Y_{t+1} = j, \Delta_{t+1} = b \mid Y_t = i, \Delta_t = a\}$ by Total Probability Formula. Combining Markov property of π and property of single step transition probability, we can get the conclusion.

The significance of the theorem lies in that it shows the way to calculate the single step transition probability.

4.2.2 Pathway of the solution

Step1: For $t = N$ and $i_N \in S_N$, let

$$u_N(i_N, a) = r_N(i_N, a), \forall a \in A_N(i_N).$$

Based on the previous assumption, we can see that $A_N(i_N)$ is a finite set, $\forall i_N \in S_N$. Establish a new game $G_N^*(i_N)$ with the payoff $u_N(i_N, a)$ under the state $i_N \in S_N$. Then according to Game Theory, we can find decision policy $\pi_N^*(a \mid i_N), \forall a \in A_N(i_N)$ and the corresponding system equilibrium value $u_N^*(i_N)$, which are denoted as $\arg EqG_N^*(i_N)$ and $EqG_N^*(i_N)$, respectively. Find the equilibrium under all states $i_N \in S_N$ and then we can obtain π_N^*.

Step2: If $t = 1$, stop. Then $\pi^* = (\pi_1^*, \ldots, \pi_N^*)$ is the system equilibrium for this game chain under the given initial state. Or else, let $t-1 \Rightarrow t$, transfer to step3.

Step3: For every state $i_t \in S_t$, get $u_t(i_t, a), \forall a \in A_t(i_t)$ through

$$u_t(i_t, a) = r_t(i_t, a) \oplus \sum_{j \in S_{t+1}} p_t(j \mid i_t, a) u_{t+1}^*(j)$$

Then establish the new game $G_t^*(i_t)$ $\forall i_t \in S_t$ with the payoff $u_t(i_t, a)$. Find the decision policy $\pi_t^*(a \mid i_t), \forall a \in A_t(i_t)$ and corresponding equilibrium value $EqG_t^*(i_t) = u_t^*(i_t)$ by employing game theory and then obtain π_t^*.

Step4: Return to Step2.

5. APPLICATION TO DECISION-MAKING IN A TANK VIRTUAL GAME

5.1 Background

Interdependence of involved parties is usually overlooked in decision-making process. However, an excellent decision maker should always take his opponents' thinking into consideration just the same as what a wise man would do in a competitive or fair game. Moreover, on most occasions, decision making seems to be carried out by multi-stage operation, which indicates that a decision-maker is frequently faced with a multistep decision process or multiple games with his opponents which can be described as a game chain. It's unwise for him to make a decision just focusing on his own thinking or only partial stages. For example, in a campaign virtual game, before fighting the first battle one must have a general idea of how the second, third, fourth, and even the final battle will be fought, and consider what actions his opponent would probably take all the time, which means that decision should be made from holistic and game perspectives. In this part, we make up a scene of tank virtual game to demonstrate the pathway to solve game chain problem.

5.2 Game scenarios

5.2.1 Game stage scenario

Considering that there are many kinds of decisions to make during a tank game, we just take decision of force deployment into consideration in this part. Before the game starts, a player should make an overall deployment of forces on the basis of current situation of both sides. We simulate a game simply as follows (see Figure 5) in which two players are called Red and Blue. They may have different battles in different battlefields, which are called as an encounter battle, a position battle and a target battle successively in this part.

5.2.2 Force deployment assumptions

(1) At the beginning of the game, Red has 6 platoons at the station, which are 18 tanks in all. According to present intelligence investigation, there are 4 platoons at Blue's station, which are 12 tanks in all.

(2) Both players can only dispatch troops from their station to each battlefield by platoon as a unit. Either side is required to have no less than 1 platoon in any battlefield.

(3) In the encounter battle, Red must meet the condition that its probability of winning is no less than a pre-determined acceptable value θ. (Suppose $\theta = 0.6$)

(4) Remaining tanks of each battlefield are supposed to withdraw and never fight again at the following stages.

(5) Both players are rational indicating that both will pursue its maximization of utility on the given conditions.

5.3 A Game chain model

Simulated process of this tank game can be abstracted as a game chain (see Figure 6).

Before establishing a mathematical model of this game chain, we should first analyze three battles in details, respectively, by setting up sub-models which

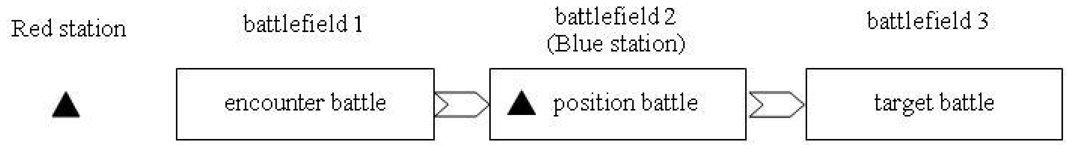

Figure 5 Game Stage scenario

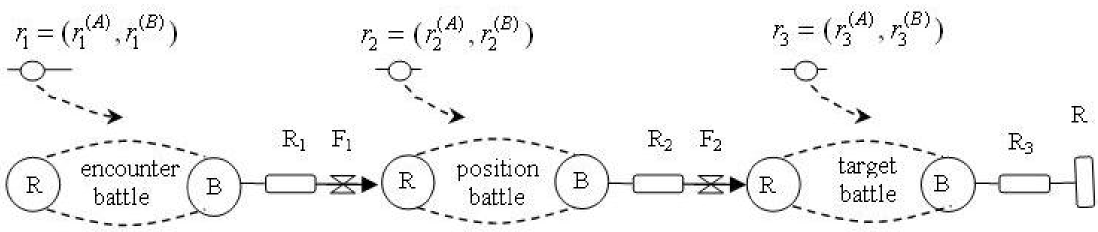

Figure 6 Tank game chain

have already been showed in my master thesis. Here, only some necessary formulas and results are given.

5.3.1 Sub-models of the virtue game

(1) Encounter battle

For Red

Initial number of tanks: $x_0 = m, m \in \{3,6,9,12\}$

Winning probability: $p_x \in [0,1]$

Number of remaining tanks after tth random fighting: $x_t, t \in N_+$

For Blue

Initial number of tanks: $y_0 = n, n \in \{3,6\}$

Winning probability: $p_y \in [0,1]$

Number of remaining tanks after tth random fighting: $y_t, t \in N_+$

Table 1 Winning probability and average remaining tanks

	Red p_x	Blue p_y	Average remaining tanks	
			Red	Blue
m=3 n=3	0.5000	0.5000	1.1333	1.1333
m=3 n=6	0.3603	0.6397	0.7724	2.7113
m=6 n=3	**0.6397**	0.3603	2.7113	0.7724
m=6 n=6	0.5000	0.5000	1.9295	1.9295
m=9 n=3	**0.7358**	0.2642	4.6583	0.5492
m=9 n=6	**0.6179**	0.3821	3.5120	1.3887
m=12 n=3	**0.8026**	0.1974	6.8444	0.4024
m=12 n=6	**0.7098**	0.2902	5.4134	1.0113

Utility function[10]:

Red: $$r_1^{(A)} = \min\left\{1, \frac{(p_x^2 / p_y) y_0 e^{y_t/y_0}}{x_0 e^{x_t/x_0}} e^{(x_0 - y_0)/30}\right\}$$

Blue: $$r_2^{(B)} = \min\left\{1, \frac{(p_y^2 / p_x) x_0 e^{x_t/x_0}}{y_0 e^{y_t/y_0}} e^{(y_0 - x_0)/30}\right\}$$

Table 2 Payoff matrix of the encounter battle

Red \ Blue	1	2
1	—	—
2	0.4230, 0.5449	0.5000, 0.5000
3	0.4004, 0.4855	0.5144, 0.4590
4	0.3907, 0.4056	0.5358, 0.3845

Note: The strategies of both sides are expressed by platoon as a unit in payoff matrix existing in this part.

(2) Position battle

Red is supposed to attack the Blue station.

For Red:

The number of platoons which can be used to attack: $\alpha_2, \alpha_2 \in \{1,2,3\}$

Attack strength: $\lambda = 1.0/\min$

For Blue:

The number of defense platoons which can be considered as defense lines: $\beta_2, \beta_2 \in \{1,2\}$

The average shooting time of each defense line for one target: $t_1 = 2min$ (the service strength of defense line is $u = 1/t_1$).

Shooting time t is subject to the negative exponential distribution.

The damage probability of target under the condition of being shot: $p = 0.7$

The time for a tank to pass through the target region is 2*min*.

If $\beta_2 = 2$, denote the shooting probability of ith defense lines as p_i. Then

$$\begin{cases} up_1 = \lambda p_0 \\ 2up_2 + \lambda p_0 = (\lambda + u) p_1 \\ 2up_3 = \lambda p_2 \end{cases}$$

Combining with $\sum_{i=0}^{2} p_i = 1$, we can find the solution as $p_0 = \frac{1}{5}, p_1 = \frac{2}{5}, p_2 = \frac{2}{5}$.

And for $\beta_2 = 1$, there is $p_0 = \frac{1}{3}, p_1 = \frac{2}{3}$.

Table 3 Penetration of Red in the position battle

strategy profile \ related parameters	Blue defense lines	penetration probability for Red p_c	expectation of penetration tanks for Red $N\cdot[1-(1-p_c)\cdot p]$
(3,1)	1	2/3	6.9000
(2,2)	2	2/5	3.4800
(2,1)	1	2/3	4.6000
(3,2)	2	2/5	5.2200
(1,1)	1	2/3	2.3000
(1,2)	2	2/5	1.7400

Note: N denotes the number of tanks of Red.

Utility functions:

Red:

$$r_2^{(A)} = \min\{\frac{N \cdot [1-(1-p_c) \cdot p] \cdot 3\beta_2}{N \cdot N}, 1\} = \min\{\frac{[1-(1-p_c) \cdot p] \cdot 3\beta_2}{N}, 1\}$$

Blue:

$$r_2^{(B)} = \min\{\frac{N(1-p_c)p \cdot N}{N \cdot 3\beta_2}, 1\} = \min\{\frac{N(1-p_c)p}{3\beta_2}, 1\}$$

Table 4 Payoff matrix of the position battle

Red \ Blue	1	2
1	0.2556, 0.7000	1.0000, 0.2100
2	0.3833, 0.4667	0.5800, 0.4200
3	0.7667, 0.2333	0.3867, 0.6300

(3) Target battle

Utility Functions

Red: $$r_3^{(A)} = \min\{\frac{Red\ force\ number}{Blue\ force\ number}, 1\}$$

Blue: $$r_3^{(B)} = \min\{\frac{Blue\ force\ number}{Red\ force\ number}, 1\}$$

Table 5 Payoff matrix of the target battle

Red \ Blue	1	2
1	1.0000, 1.0000	0.5000, 1.0000
2	1.0000, 0.5000	1.0000, 1.0000
3	1.0000, 0.3333	1.0000, 0.6667

5.3.2 The game chain model

On a basis of previous sub-models, a game chain model is established based on Markov Decision Process with decision point set as $T = \{1,2,3\}$.

(1) State and action

In this problem, state is defined as the platoons which can be assigned at the decision point t (t=1, 2, 3), while action represents the platoons dispatched to the corresponding battlefield which is related to state of the corresponding decision point. Then

$$S_1 = \{(6,4)\}$$

$$A_1(i) = \{(2,1),(2,2),(3,1),(3,2),(4,1),(4,2)\}$$
$$= \{a_1, a_2, \cdots, a_6\}, i = (6,4) \in S_1$$

$$S_2 = \{(4,3),(4,2),(3,3),(3,2),(2,3),(2,2)\}$$
$$= \{j_1, j_2, \cdots, j_6\}$$

$$A_2(j_1) = \{(1,1),(1,2),(2,1),(2,2),(3,1),(3,2)\}$$
$$A_2(j_2) = \{(1,1),(2,1),(3,1)\}$$
$$A_2(j_3) = A_2(j_2) = \{(1,1),(1,2),(2,1),(2,2)\}$$
$$A_2(j_4) = \{(1,1),(2,1)\}$$
$$A_2(j_5) = \{(1,1),(1,2)\}$$
$$A_2(j_6) = \{(1,1)\};$$

$$S_3 = \bigcup_{j \in S_2, b \in A_2(j)} (j-b)$$
$$= \{(3,2),(3,1),(2,2),(2,1),(1,2),(1,1)\}$$

$$A_3(k) = \{k\}, k \in S_3 \circ$$

(2) State transition probability and immediate utility

The state of decision point t+1 is decided by the state of decision point t and corresponding action which is selected. For example, if decision-maker selects action $a_1 = (2,1)$, then state of decision point 2 will be transferred to $(6,4)-(2,1)=(4,3)$. Hence, the state transition probability of decision point 1 is

$$p^{(1)}(j|i,a) = \begin{cases} 1, j = i-a \\ 0, \text{other} \end{cases} \quad \forall i \in S_1, a \in A_1(i)$$

The state transition probability of decision point 2 is

$$p^{(2)}(k|j,b) = \begin{cases} 1, k = j-b \\ 0, \text{other} \end{cases} \quad \forall j \in S_2, b \in A_2(j)$$

The immediate utilities are in the form of

$$r_1(i,a) = (r_1^{(A)}(i,a), r_1^{(B)}(i,a)), i \in S_1, a \in A_1(i)$$
$$r_2(j,b) = (r_2^{(A)}(j,b), r_2^{(B)}(j,b)), j \in S_2, b \in A_2(j)$$
$$r_3(k,k) = (r_1^{(A)}(k), r_1^{(B)}(k)), k \in S_3$$

Then we try to find system equilibrium of this game chain model.

5.4 Solution

Step1: For $t = 3$, expected total utility of each action is:

$$u_3(i,a) = r_3(i,a) = (r_3^{(A)}(i,a), r_3^{(B)}(i,a)), i \in S_3, a \in A_3(i)$$

Label the new game under any state $i \in S_1$ as $G_3^*(i)$. According to previous analysis, the equilibrium of new game $G_3^*(i)$ is

$$\arg Eq\{G_3^*(i)\} = i, \forall i \in S_1$$

Thus the decision function is $f_3^*(i) = i$ and the corresponding equilibrium value function is $u_3^*(i) = r_3(i,i)$ (see Table 5).

Step2: For $t = 2$,

$$u_2(i,a) = r_2(i,a) + u_3^*(i-a)$$
$$= (u_2^{(A)}(i,a), u_2^{(B)}(i,a)), i \in S_2, a \in A_2(i)$$

Label the new game in any state $i \in S_2$ as $G_2^*(i)$, and the corresponding results are listed as follows:

(1) For $i = (4,3)$ $a \in A_2(i) = \{(1,1),(1,2),(2,1),(2,2),(3,1),(3,2)\}$ and the corresponding pay-off matrix under the state $i = (4,3)$ is showed in Table 6.

Table 6 Pay-off Matrix under the state of $i = (4,3)$

Red \ Blue	1	2
1	1.2556, $\underline{1.3667}$	$\underline{2.0000}$, 0.5433
2	$\underline{1.3833}$, $\underline{1.4667}$	1.5800, 0.9200
3	1.2667, 1.2333	1.3867, $\underline{1.6300}$

Then $f_3^*(i) = (2,1)$, $u_3^*(i) = (1.3833, 1.4667)$.

(2) For $i = (4,2)$, $a \in A_2(i) = \{(1,1),(2,1),(3,1)\}$ and the corresponding pay-off matrix is showed in Table 7.

Table 7 Pay-off Matrix under the state of $i = (4,2)$

Red \ Blue	1
1	1.2556, $\underline{1.0333}$
2	1.3833, $\underline{0.9667}$
3	$\underline{1.7667}$, $\underline{1.2333}$

Then $f_3^*(i) = (3,1)$, $u_3^*(i) = (1.7667, 1.2333)$.

(3) For $i = (3,3)$, $a \in A_2(i) = \{(1,1),(1,2),(2,1),(2,2)\}$ and the corresponding pay-off matrix is showed in Table 8.

Table 8 Pay-off Matrix under the state of $i = (3,3)$

Red \ Blue	1	2
1	$\underline{1.2556}$, $\underline{1.7000}$	$\underline{2.0000}$, 0.7100
2	0.8833, $\underline{1.4667}$	1.5800, 1.4200

Then $f_3^*(i) = (1,1)$, $u_3^*(i) = (1.2556, 1.7000)$.

(4) For $i = (3,2)$, $a \in A_2(i) = \{(1,1),(2,1)\}$ and the corresponding pay-off matrix is showed in Table 9.

Table 9 Pay-off Matrix under the state of $i = (3,2)$

Red \ Blue	1
1	1.2556, $\underline{1.2000}$
$\underline{2}$	$\underline{1.3833}$, $\underline{1.4667}$

Then $f_3^*(i) = (2,1)$, $u_3^*(i) = (1.3833, 1.4667)$.

(5) For $i = (2,3)$, $a \in A_2(i) = \{(1,1),(1,2)\}$ and the corresponding pay-off matrix is showed in Table 10.

Table 10 Pay-off Matrix under the state of $i = (2,3)$

Red \ Blue	1
1	0.7556, $\underline{1.7000}$
2	$\underline{1.3833}$, $\underline{1.4667}$

Then $f_3^*(i) = (2,1)$, $u_3^*(i) = (1.3833, 1.4667)$.

(6) For $i = (2,2)$, $a \in A_2(i) = \{(1,1)\}$. Corresponding pay-off matrix is showed in Table 11.

Table 11 Pay-off Matrix under the state of $i = (2,2)$

Red \ Blue	1
1	1.2556, 1.7000

Then $f_3^*(i) = (1,1)$, $u_3^*(i) = (1.2556, 1.7000)$.

The decision function and corresponding value function are as follows:

$$f_2^*(i) = \begin{cases} (2,1), i = (4,3) \\ (3,1), i = (4,2) \\ (1,1), i = (3,3) \\ (2,1), i = (3,2) \\ (2,1), i = (2,3) \\ (1,1), i = (2,2) \end{cases}$$

$$u_2^*(i) = \begin{cases} (1.3833, 1.4667), i = (4,3) \\ (1.7667, 1.2333), i = (4,2) \\ (1.2556, 1.7000), i = (3,3) \\ (1.3833, 1.4667), i = (3,2) \\ (1.3833, 1.4667), i = (2,3) \\ (1.2556, 1.7000), i = (2,2) \end{cases}$$

Step3: For $t = 1$, the expected total utility is $u_1(i,a) = r_1(i,a) + u_2^*(i-a) = (u_1^{(A)}(i,a), u_1^{(B)}(i,a))$, such that $i \in S_1, a \in A_1(i)$. As $i = (6,4)$ and corresponding

$A_1(i)=\{(2,1),(2,2),(3,1),(3,2),(4,1),(4,2)\}$, we find the solution $Eq\{G_1^*(i)\}$ of new game $G_1^*(i)$ under the principle of expected total utility which is showed in Table 12.

Then $f_1^*((6,4))=(2,1),\ u_1^*((6,4))=(1.8132,2.01161)$.

Table 12 Pay-off Matrix of encounter battle under the state of $i=(6,4)$

Red \ Blue	1	2
2	<u>1.8132</u>, <u>2.01161</u>	<u>2.2667</u>, 1.7333
3	1.65596, <u>2.18551</u>	1.89768, 1.92568
4	1.77398, 1.87225	1.79316, <u>2.08447</u>

Above all, we find the system equilibrium of this game chain as $E^*=((2,1),(2,2),(2,1))$, which could provide proposal for decision-maker from holistic and game perspectives. The meaning of decision function also lies in that in actual decision-making process, it can provide the system equilibrium of the rest stages of the affair after combining with the feedback of environmental information.

6. CONCLUSION

Demonstrated by a virtue game, it seems feasible to find a solution for game chain decision problem through the pathway proposed in this paper. The result is more acceptable and reasonable from systematic and holistic perspectives. The most important purpose of this paper is that we want to draw more attention to adding systematic idea and game idea into complex decision making process in such fields as infrastructure management etc. in order to pursue a more reasonable scheme before taking action for decision makers.

REFERENCES

AN Tingyu, 2008. A Decision-Making Method of Game Chain Based On Markov Decision Processes, Master Thesis, Beijing Institute of Technology, Beijing.

Aumann, Robert J., 1987, *"game theory"*, *The New Palgrave: A Dictionary of Economics*, *2nd Edition*, 460–482 p.

Checkland, P. 1999. *System Thinking, System Practice: Including a 30-Year Retrospective*, John Wiley, Chichester.

Hou Guangming, 2006. *An Introduce to Organizational Systems Science*, Science Press, Beijing.

Hou Guangming, Game-chain and Its Application for Organizational Innovation for Defense R&D, *proceedings of the 288th XIANGSHAN Science Conferences*, Beijing, 84-101 p, 2007.

Liu Ke, 2004. *Applied Markov Decision Processes*, TsingHua University Press, Beijing.

Puterman M. L., *Markov Decision Processes*, John Wiley & Sons, New York, 1994.

Senge, P.M., 2006. *The Fifth Discipline: The Art and Practice of the Leaning Organazation*, Random House Business Books, London, 6-7 p.