Applicable Analysis and Discrete Mathematics, **1** (2007), 265–275.
Available electronically at **http://pefmath.etf.bg.ac.yu**

Presented at the conference: *Topics in Mathematical Analysis and Graph Theory*, Belgrade, September 1–4, 2006.

# OPTIMAL VENTCEL GRAPHS, MINIMAL COST SPANNING TREES AND ASYMPTOTIC PROBABILITIES

*Tzuu-Shuh Chiang, Yunshyong Chow*

For each $\epsilon > 0$, let $\{X_n^\epsilon\}$ be an irreducible, time-homogeneous Markov chain with a finite state space $S$ and transition function $p^\epsilon(i,j) = p_{i,j}\epsilon^{U(i,j)}(1 + o(1))$ where $0 \leq U(i,j) \leq \infty$ is a cost function. (We assume $p_{i,j} = 0$ iff $U(i,j) = \infty$.) It has been shown [**2**] that independent of the initial distribution, there are constants $h(i) \geq 0$ and $\beta_i > 0$ such that $\lim\limits_{\epsilon \downarrow 0} \mu^\epsilon(i)/\epsilon^{h(i)} = \beta_i$ for any $i \in S$, where $\mu^\epsilon$ is the invariant distribution of $\{X_n^\epsilon\}$. Let $\underline{S} = \{i \in S : h(i) = 0\}$, which is called the global minimum set. Various asymptotic probabilities related to $\underline{S}$ have been established in [3]. Among others, starting with the uniform or invariant distribution, the expected hitting time $E^\epsilon T$ of $\underline{S}$ is of order $\epsilon^{-\delta}$ and the constants $\delta$ and $h(i)$ above can be expressed in terms of a complicated hierarchy of "cycles" related to the cost function $U$. In this paper, we shall express these constants in terms of Ventcel graphs (minimum cost spanning trees) to simplify the concept and computation of these constants. We also establish some new properties of optimal Ventcel graphs.

## 1. INTRODUCTION

Let S be a finite set and $U : S \times S \to [0, \infty]$ be a cost function, where $U(i,j)$ is interpreted as the cost from the state $i$ to a different state $j$. Consider a family of irreducible, time-homogeneous Markov chains $\{X_n^\epsilon\}$ defined on S with transition probability

$$p^\epsilon(i,j) = p_{i,j} \cdot \epsilon^{U(i,j)}(1 + o(1)) \text{ for all } i \neq j, i, j \in S. \tag{1.1}$$

2000 Mathematics Subject Classification. 60K35, 68W20.
Key Words and Phrases. Ventcel graphs, cycles, asymptotic probability, minimal cost spanning tree.

Here we assume $\epsilon$ is a small parameter and $p_{i,j}=0$ iff $U(i,j)=\infty$. Note that $U(i,i)$ plays no role in (1.1) because $\sum\limits_j p^\epsilon(i,j)=1$. The purpose of running such Markov chains is to find the smallest set $\underline{S}\subseteq S$ such that for small $\epsilon$, $P(X_n^\epsilon\in\underline{S})\approx 1$ as $n\to\infty$ and the order estimates for the expected time of hitting $\underline{S}$. The set $\underline{S}$ is referred to as the global minimum set of the cost function U. In many physical models, $U(i,j)=(u(j)-u(i))^+$ if $j$ is a neighbor of $i$ and is $\infty$ otherwise, where $u$ is a potential function on $S$. It turns out that in this case, $\underline{S}=\{i\in S: u(i)=\min_S u\}$ as expected. However, it takes some efforts to determine $\underline{S}$ for a general cost function $U$.

Instead of running a family of Markov chains, one can have a single but time-inhomogeneous Markov chain. This is called simulated annealing process and readers are referred to [**7, 8**] for details.

Various properties related to $\{X_n^\epsilon\}$ have been obtained in [**3**]. Let $\mu^\epsilon$ be the invariant distribution of $\{X_n^\epsilon\}$. In this paper we shall be concerned with the following issues :

(1) For any state $i\in S$, the invariant distribution $\mu^\epsilon$ satisfies $\mu^\epsilon(i)\approx\epsilon^{h(i)}$ for $\epsilon$ small.
Hence, $\underline{S}=\{i\in S: h(i)=0\}$. Note that $\sum\limits_{j\in S}\mu^\epsilon(j)=1$.

(2) Starting from $\mu^\epsilon$, $E^\epsilon T\approx\epsilon^{-\delta_h}$ for $\epsilon$ small, where $T$ is the hitting time of $\underline{S}$.

(3) Furthermore, $E^\epsilon T_{i_0}\approx\epsilon^{-\delta_v}$ where $T_{i_0}$ is the hitting time of any fixed $i_0\in\underline{S}$.
The constants $h(i),\delta_h$ and $\delta_v$ are defined in [**2**] through a hierarchy of the so called "cycles". While conceptually it is easy to comprehend these constants, it is hard to actually compute them even through computers. The quantity $\mu^\epsilon(i)$ in (1) already appeared in [**6,10**] by solving linear equations. Related problems of (2) and (3) have been studied in [**4,5**].

Our aim of this paper is first to define these constants $h(i),\delta_h$ and $\delta_v$ in terms of optimal Ventcel graphs [**6,10**] and then simplify their computation by using minimum cost spanning trees. Indeed, optimal Ventcel graphs will be viewed as a kind of minimum cost spanning trees with pre-assigned roots.

One example is the potential case of the spin glass model. In this model, $S=\{-1,1\}^{\mathbf{D}_n}$, where $\mathbf{D}_n$ is the 2-dim lattice of size $n\times n$. For each state $i\in S$, its nearest-neighbor potential energy is defined as

$$u(i)=-\sum_{|x-y|=1}J_{x,y}\cdot i(x)\cdot i(y),$$

where the real number $J_{x,y}$ denotes the interaction strength between two neighboring sites $x,y$ in $\mathbf{D}_n$. Let $\mathrm{N}(i)=\{j\in S: i(x)=j(x)$ for all sites $x\in\mathbf{D}_n$ except one $\}$ be the neighborhood of state $i$. Then the transition probability in (1.1) is given by $p^\epsilon(i,j)=\frac{1}{|\mathrm{N}(i)|}\,\epsilon^{U(i,j)}$, where $U(i,j)=\big(u(j)-u(i)\big)^+$ if $j\in\mathrm{N}(i)$ and $\infty$ otherwise. The purpose of running the Markov chains with transition probability $p^\epsilon(i,j)$ is to find the states with the smallest potential energy in the spin-glass model.

Here is an example with non-potential cost $U$. In the well-known 2-person prisoner's dilemma game, each prisoner has to play a strategy from $\{C, N\}$. Here $C$ and $N$ stand for "confess" and "not-confess" respectively. The unique Nash equilibrium of the game requires each prisoner to play $C$. However, it is to both prisoners' favor if they both play $N$ [**9**]. It is always interesting to see how one can overcome such a dilemma. Recently, it was shown possible [**1**] in some evolutionary prisoner's dilemma games with local interaction, imitation and mutation. Instead of two players, there are N players sitting around a circle in such a model. At each time period, players first meet with each of their two neighbors once to play the prisoner's dilemma game, then imitate their neighbors or themselves whoever have the highest payoffs and finally, can make mistake independently with a small positive probability $\epsilon$ to choose the other strategy instead of the rational one. The dynamical process can be described by the above Markov chain $\{X_n^\epsilon\}$ with the state space $S = \{C, N\}^{\mathrm{N}}$ and the cost function $U(i, j)$ in (1.1) counts the number of player $x$ who makes mistake at the final stage by adopting the non-rational strategy $j(x)$. Since $|S| = 2^{\mathrm{N}}$ can be very huge, it will usually take a great effort to get $\underline{S}$ by penetrating the hierarchy of cycles. Besides $\underline{S}$, the order estimate like $E^\epsilon T \approx \epsilon^{-\delta_h}$ is important in applications. In this model, using Ventcel graphs turns out to be most efficient to get $\underline{S}$, $\delta_h$ and so on.

We now review the concepts of cycles and Ventcel W-graphs. One example is given at the end of this section to illustrate the process. For a subset $W \subseteq S$, a W-graph is a function g from $S \setminus W$ to S with no cycles, i.e., for any $i \in S \setminus W$, there exist $i_0 = i, i_1, \ldots, i_m$ in $S \setminus W$ such that $g(i_k) = i_{k+1}$ for $0 \leq k < m$, but $g(i_m) \in W$. For a Ventcel W-graph g, the cost of g is defined as follow :

$$V(g) = \sum_{i \in S \setminus W} U(i, g(i)).$$

A W-graph g is called W-optimal if

$$V(g) = v(W) \overset{def}{=} \min\{V(h) : h \text{ is a W-graph}\}. \tag{1.2}$$

Let G(k) be the set of all W-graphs with $\mid W \mid = k$. Define

$$v_k = \min\{V(g) \, : \, g \in G(k)\} \text{ for } k \geq 1. \tag{1.3}$$

A W-graph g is said $k$-optimal if $V(g) = v_k$ and $\mid W \mid = k$. We shall characterize optimal W-graphs, optimal k-graphs and $v_k$ in Sections 2, 3 and 4 respectively. We next define cycles. For $i \in S$, let

$$V(i) = \min\{U(i, j) \, : \, j \in S \text{ and } j \neq i\}$$

be the minimum cost for reaching out from i. For any two states $i, j \in S$, we say that $i \geq j$ if there exist $i_0 = i, i_1, \ldots, i_m = j$ such that $U(i_k, i_{k+1}) = V(i_k)$ for each k. This simply means there is a path from i to j such that each intermediate step has its minimum cost. A state $i$ is said minimal if $i \geq j$ implies $j \geq i$ for any other $j \in S$. Two different states $i, j$ are said equivalent ($i \sim j$) if

$$\text{(i)} \quad i \ \text{ is minimal,} \qquad \text{(ii)} \quad i \geq j \ \text{ and } \ j \geq i.$$

We always assume $i \sim i$ and thus "$\sim$" is an equivalence relation. The equivalent classes under "$\sim$" will be called cycles. A hierarchy of cycles can be established as follows. First, let $S^0 = S, U^0 = U$ and $V^0 = V$. Having defined $S^{n-1}, U^{n-1}$ and $V^{n-1}$, let $S^n = \{\text{cycles of } S^{n-1}\}$. Hence if $C^n \in S^n$ then $C^n = \{C_i^{n-1}\}_i$ where $C_i^{n-1} \in S^{n-1}$ for each i and $\{C_i^{n-1}\}_i$ forms a cycle under $U^{n-1}$. The depth of $C^n$ is defined as

$$d^{n-1}(C^n) = \max\{V^{n-1}(C_i^{n-1}) \, : \, C_i^{n-1} \in C^n\}.$$

For any two different states $C^n = \{C_i^{n-1}\}$ and $\bar{C}^n = \{\bar{C}_j^{n-1}\}$ in $S^n$, we now define

$$U^n(C^n, \bar{C}^n) = d^{n-1}(C^n) + \min_{i,j} \{U^{n-1}(C_i^{n-1}, \bar{C}_j^{n-1}) - V^{n-1}(C_i^{n-1})\} \tag{1.4}$$

and

$$V^n(C^n) = \min\{U^n(C^n, \bar{C}^n) \, : \, \bar{C}^n \in S^n \text{ and } \bar{C}^n \neq C^n\}.$$

This process will terminate first at some N, i.e., $\mid S^{N+1} \mid = 1$. For each state $i \in S$ we can find a unique sequence of cycles $i = C^0 \in C^1 \in \cdots \in C^{n-1} \in C^n \in \cdots \in C^N \in C^{N+1} = S^{N+1}$. Such a sequence will be referred to as the family tree of $i$. We shall abuse the notation a bit by saying that $C^k \in C^n$ if there are $C^j \in S^j$ for $k < j < n$ such that $C^k \in C^{k+1} \in \cdots \in C^n$ is part of some family tree. Finally, for a W-graph g, let

$$V(g; C^n) = \sum_{i \in C^n \setminus W} U(i, g(i)) \tag{1.5}$$

be the cost of g restricted to the cycle $C^n$. If $i$ has the family tree $i = C^0 \in C^1 \in \cdots \in C^n \in \cdots \in C^{N+1} = S^{N+1}$, then the global minimum set $\underline{S}$, the constants $h(i), \delta_h$ and $\delta_v$ are characterized in [2] as follows:

$$\begin{cases} \underline{S} = \{i \in S \, : \, h(i) = 0\} \text{ where } h(i) = \sum_{n=0}^{N} (d^n(C^{n+1}) - V^n(C^n)), \\ \delta_h = \max\{V^k(C^k) \, : \, \text{ all cycles } C^k \in S^k \text{ with } C^k \cap \underline{S} = \emptyset\}, \\ \delta_v = \max\{V^k(C^k) \, : \, \text{ all cycles } C^k \in S^k \text{ with } i_0 \notin C^k\}, \end{cases} \tag{1.6}$$

where $i_0 \in \underline{S}$ is fixed.

Note that $\delta_v$ above is in fact independent of the choice of state $i_0 \in \underline{S}$.

The main purpose of this paper is to represent the above constants in terms of Ventcel W-graphs. The following will be proved in Section 4.

**Main Theorem.** *For any $i \in S$, we have $h(i) = v(\{i\}) - v_1$, $\delta_v = v_1 - v_2$ and $\delta_h = v_{k_0-1} - v_{k_0}$, where $k_0 = \inf\{k \geq 2 : \exists$ an optimal k-graph $W$ with $W \not\subseteq \underline{S}\}$.*

EXAMPLE. Let $S = S^0 = \{1, 2, 3\}$ with the cost function $U = U^0$ on $S \times S$ given by $U(1, 2) = U(3, 1) = 4$, $U(1, 3) = U(2, 1) = 3$, $U(2, 3) = 1$ and $U(3, 2) = 0$. Note that the value of $U(i, i)$ is unimportant, but serves to make $\sum_{j=1}^{3} p^{\epsilon}(i, j) = 1$ in (1.1).

States 2 and 3 form a cycle and state 1 is itself a cycle in $S^1$. Thus $S^1 = \{\{1\}, \{2, 3\}\}$. A simple computation via (1.4) shows $U^1(\{1\}, \{2, 3\}) = U^1(\{2, 3\}, \{1\}) = 3$. Naturally, $\{1\}$ and $\{2, 3\}$ form a cycle in $S^2$ and the process terminates.

Based on (1.2), one can easily compute that

$v(\{1\}) = 3$ and the $\{1\}$-optimal graph is $g(3) = 2$ and $g(2) = 1$,

$v(\{2\}) = 3$ and the $\{2\}$-optimal graph is $g(1) = 3$ and $g(3) = 2$.

Similarly, $v(\{2\}) = 4$ and the $\{2\}$-optimal graph is $g(1) = 3$ and $g(3) = 2$. Thus, $v_1 = v(\{1\}) = v(\{2\}) = 3$ by (1.3). From the Main Theorem we have $h(1) = h(2) = 0$ and $h(3) = 1$. By (1.6), the global minimum set $\underline{S} = \{1, 2\}$. Obviously, $v_2 = 0$ and the 2-optimal graph is a $\{1, 2\}$-graph with $g(3) = 2$. Since $|S| = 3$ and $|\underline{S}| = 2$, $k_0 = 3$ in the Main Theorem and thus $\delta_v = v_1 - v_2 = 3$ and $\delta_h = v_2 - v_3 = 0 - 0 = 0$.

## 2. CONSTRUCTION OF OPTIMAL W-GRAPHS.

In this section, we shall identify the optimal Ventcel graphs for a fixed subset $W \in S$.

**Definition 2.1.** *Let $C^k, \bar{C}^k \in S^k$. For a W-graph $h$ we say $h \in (C^k \to \bar{C}^k)$ if there exist $i \in C^k$, $j \in \bar{C}^k$ such that $h(i) = j$. In the case that $\bar{C}^k$ satisfies $U^k(C^k, \bar{C}^k) = V^k(C^k)$, we simply write $h \in (C^k \to)$.*

For two cycles $C^k$ and $\bar{C}^k$ in $S^k$, we define the minimal cost on $C^k$ of W-graphs in $(C^k \to \bar{C}^k)$ as follows. For $k = 1$, $C^1 \cap W = \emptyset$ and any $\bar{C}^1 \neq C^1$, let

$$V_W(C^1 \to \bar{C}^1) = \sum_{i \in C^1} V(i) + U^1(C^1, \bar{C}^1) - d^0(C^1). \tag{2.1}$$

Note that $V_W(C^1 \to \bar{C}^1)$ is undefined if $C^1 \cap W \neq \emptyset$. For $C^1 \cap W \neq \emptyset$ we let

$$V_W(C^1 \to C^1) = \sum_{i \in C^1 \setminus W} V(i) \tag{2.2}$$

and $V_W(C^1 \to C^1)$ remains undefined if $C^1 \cap W = \emptyset$. We write $V_W(C^1 \to)$ for $V_W(C^1 \to \bar{C}^1)$ in (2.1) if $U^1(C^1, \bar{C}^1) = V^1(C^1)$. Suppose we have defined $V_W(C^{k-1} \to \bar{C}^{k-1})$ for any $C^{k-1}, \bar{C}^{k-1} \in S^{k-1}$ as in (2.1) and (2.2). Then for any $C^k = \{C_i^{k-1}\} \neq \bar{C}^k$ in $S^k$, let

$$V_W(C^k \to \bar{C}^k) = \sum_i V_W(C_i^{k-1} \to) + U^k(C^k, \bar{C}^k) - d^{k-1}(C^k) \text{ if } C^k \cap W = \emptyset. \tag{2.3}$$

Note that $V_W(C^k \to \bar{C}^k)$ is undefined if $C^k \cap W \neq \emptyset$. For $C^k \cap W \neq \emptyset$ we let

$$V_W(C^k \to C^k) = \sum_{C_i^{k-1} \cap W = \emptyset} V_W(C_i^{k-1} \to) + \sum_{C_i^{k-1} \cap W \neq \emptyset} V_W(C_i^{k-1} \to C_i^{k-1}) \tag{2.4}$$

and $V_W(C^k \to C^k)$ remains undefined if $C^k \cap W = \emptyset$. Similarly, we shall write $V_W(C^k \to)$ for $V_W(C^k \to \bar{C}^k)$ if $U^k(C^k, \bar{C}^k) = V^k(C^k)$ and $C^k \cap W = \emptyset$. Then by definition (2.3),

(2.5) $V_W(C^k \to \bar{C}^k) - V_W(C^k \to) = U^k(C^k, \bar{C}^k) - V^k(C^k) \geq 0$ for $\bar{C}^k \neq C^k \in S^k$.

**Theorem 2.2.** *Let $W \subseteq S$ and $C^k, \bar{C}^k \in S^k$ be different. Then for any $W$-graph $h$,*

$$V(h; C^k) \geq \begin{cases} V_W(C^k \to C^k) & \text{if } C^k \cap W \neq \emptyset, \\ V_W(C^k \to \bar{C}^k) & \text{if } C^k \cap W = \emptyset \text{ and } h \in (C^k \to \bar{C}^k). \end{cases}$$

**Proof.** We first consider $k = 1$. Let $C^1, \bar{C}^1 \in S^1$. If $C^1 \cap W \neq \emptyset$ then by (2.2),

$$V(h; C^1) = \sum_{i \in C^1 \setminus W} U(i, h(i)) \geq \sum_{i \in C^1 \setminus W} V(i) = V_W(C^1 \to C^1).$$

If $C^1 \cap W = \emptyset$ and $h \in (C^1 \to \bar{C}^1)$, then there are $i_0 \in C^1$ and $j \in \bar{C}^1$ such that $h(i_0) = j$. By definition (1.4) for $U^1(C^1, \bar{C}^1)$ and (2.1), we have

$$\begin{aligned} V(h; C^1) &= \sum_{i \in C^1} U\big(i, h(i)\big) \geq \sum_{i \in C^1 \setminus \{i_0\}} V(i) + U(i_0, j) \\ &\geq \sum_{i \in C^1 \setminus \{i_0\}} V(i) + U^1(C^1, \bar{C}^1) - d^0(C^1) + V(i_0) \\ &= \sum_{i \in C^1} V(i) + U^1(C^1, \bar{C}^1) - d^0(C^1) = V_W(C^1 \to \bar{C}^1) \end{aligned}$$

and the equalities hold iff $U\big(i, h(i)\big) = V(i)$ for all $i \neq i_0$ and $U\big(i_0, h(i_0)\big) = U(C^1, \bar{C}^1) - d^0(C^1) + V(i_0)$. Suppose we have proved the theorem up to $k$. Let $C^{k+1} = \{C_i^k\}$. If $C^{k+1} \cap W \neq \emptyset$ then by (1.5), (2.5) and the induction hypothesis,

$$\begin{aligned} V(h; C^{k+1}) &= \sum_{C_i^k \cap W \neq \emptyset} V(h; C_i^k) + \sum_{C_i^k \cap W = \emptyset} V(h; C_i^k) \\ &\geq \sum_{C_i^k \cap W \neq \emptyset} V_W(C_i^k \to C_i^k) + \sum_{C_i^k \cap W = \emptyset} V_W(C_i^k \to) = V_W(C^{k+1} \to C^{k+1}). \end{aligned}$$

If $C^{k+1} \cap W = \emptyset$ and $h \in (C^{k+1}, \bar{C}^{k+1})$, then there exist $C_{i_0}^k \in C^{k+1}$ and $\bar{C}^k \in \bar{C}^{k+1}$ such that $h \in (C_{i_0}^k \to \bar{C}^k)$. Using (2.5), (2.3) and the induction hypothesis again,

$$\begin{aligned} V(h; C^{k+1}) &= \sum_{i \neq i_0} V(h; C_i^k) + V(h; C_{i_0}^k) \geq \sum_{i \neq i_0} V_W(C_i^k \to) + V_W(C_{i_0}^k \to \bar{C}^k) \\ &\geq \sum_{i \neq i_0} V_W(C_i^k \to) + V_W(C_{i_0}^k \to) + U^k(C_{i_0}^k, \bar{C}^k) - V^k(C_{i_0}^k) \\ &\geq \sum_i V_W(C_i^k \to) + U^{k+1}(C^{k+1}, \bar{C}^{k+1}) - d^k(C^{k+1}) \\ &= V_W(C^{k+1} \to \bar{C}^{k+1}), \end{aligned}$$

where (1.4) is used in the last inequality. This completes the proof by induction.

Remark 2.3. Theorem 2.2 actually describes all the possible ways to construct optimal W-graphs. Indeed, if $C^k \cap W = \emptyset$ and $h$ is a W-graph, then it is obvious that $h \in (C^k \to \bar{C}^k)$ for some other $\bar{C}^k \in S^k$ and thus $h \in (C_i^{k-1} \to \bar{C}_j^{k-1})$ for some $C_i^{k-1} \in C^{k-1}$ and $\bar{C}_j^{k-1} \in \bar{C}^{k-1}$. Theorem 2.2 then dictates that $C_i^{k-1}$ and $\bar{C}_j^{k-1}$ must satisfy

$$U^k(C^k, \bar{C}^k) = U^{k-1}(C_i^{k-1}, \bar{C}_j^{k-1}) + d^{k-1}(C^k) - U^{k-1}(C_i^{k-1}) \tag{2.6}$$

in order for $h$ to be optimal, which simply means that the minimum in (1.4) is attained at the pair $(C_i^{k-1}, \bar{C}_j^{k-1})$. For the other $C_r^{k-1} \in C^k$, Theorem 2.2 forces $h$ to be in $(C_r^{k-1} \to)$. Different pairs $(C_i^{k-1}, \bar{C}_j^{k-1})$ satisfying (2.6) provide different W-graphs but they all have the same cost on $C^k$ and thus are optimal. Obviously this is the only option we have in constructing optimal W-graphs $h$ on $C^k$. If $C^k \cap W \neq \emptyset$ then Theorem 2.2 implies that $h \in (C^k \to C^k)$ and for each $C_i^{k-1} \in C^k$, $h \in (C_i^{k-1} \to)$ or $h \in (C_i^{k-1} \to C_i^{k-1})$ depending on $C_i^{k-1} \cap W = \emptyset$ or not. Since obviously $C^{N+1} \cap W \neq \emptyset$, an induction procedure can be initiated to construct all optimal W-graphs.

## 3. CONSTRUCTION OF $r$-OPTIMAL GRAPHS.

In this section we shall construct r-optimal graphs for any $1 \le r \le |S|$. We first make some notations. Recall that $N \ge 0$ is the first number that $|S^{N+1}| = |\{C^{N+1}\}| = 1$. For any $r \ge 1$ and $C^k \in S^k$ with $k \le N+1$, let

$$V_r(C^k \to C^k) = \inf\{V_W(C^k \to C^k) : |W \cap C^k| = r\}. \tag{3.1}$$

For $k \le N$ and $C^k \neq \bar{C}^k$, let

$$V_0(C^k \to \bar{C}^k) = V_W(C^k \to \bar{C}^k) \quad \text{for any } W \text{ with } \quad W \cap C^k = \emptyset. \tag{3.2}$$

Note that the right hand side of (3.2) is independent of $W$ as long as $W \cap C^k = \emptyset$. We use $V_0(C^k \to)$ for $V_0(C^k \to \bar{C}^k)$ if $U^k(C^k, \bar{C}^k) = V^k(C^k)$. Finally, for $C^k = \{C_i^{k-1}\}$ we define $V_0(C^k \to C^k)$ for $1 \le k \le N+1$ as follows :

$$V_0(C^k \to C^k) = \Sigma_i V_0(C_i^{k-1} \to). \tag{3.3}$$

In particular, (3.3) for $k = 1$ can be written as $V_0(C^1 \to C^1) = \sum_{i \in C^1} V(i)$.

For any $C^k \in S^k$, let $\mathrm{P}(C^k) = \{\text{all cycles } \; C^i \in C^k\}$. A sequence $C^0 \in C^1 \in \cdots \in C^i \in \cdots \in C^k$ is called a principal sequence of $C^k$ if it has the property that $V^i(C^i) = d^i(C^{i+1})$ for each $i \le k-1$. We use the notation $\mathrm{PS}(C^k)$ to denote such a sequence. Principal sequences of $C^k$ may not be unique. Finally, let

$$m_1(C^k) = \max\{V^i(C^i) : C^i \in \mathrm{P}(C^k) \backslash \mathrm{PS}(C^k)\}.$$

Here $\mathrm{P}(C^k) \setminus \mathrm{PS}(C^k)$ denote the collection of cycles in $C^k$ except one and any principal sequence of $C^k$. It is easy to see that $m_1(C^k)$ is independent of the choice of such a principal sequence. If $m_1(C^k)$ is attained at some $C^{i_1} \in \mathrm{P}(C^k) \setminus \mathrm{PS}(C^k)$, i.e., $m_1(C^k) = V^{i_1}(C^{i_1})$, let

$$m_2(C^k) = \max\{V^i(C^i) : C^i \in \mathrm{P}(C^k) \setminus \mathrm{PS}(C^k) \cup \mathrm{PS}(C^{i_1})\}.$$

Similarly, we can define $m_r(C^k)$ until $\mathrm{P}(C^k)\setminus\mathrm{PS}(C^k)\cup\mathrm{PS}(C^{i_1})\cup\cdots\cup\mathrm{PS}(C^{i_{r-1}}) = \emptyset$. We now prove the main result of this section.

**Theorem 3.1.** *For any different* $C^k, \bar{C}^k \in S^k$ *and* $r \geq 1$, *we have*

$$V_0(C^k \to \bar{C}^k) - V_r(C^k \to C^k) = U^k(C^k, \bar{C}^k) + \sum_{i=1}^{r-1} m_i(C^k) \text{ for } 1 \leq k \leq N \tag{3.4}$$

*and*

$$V_0(C^k \to C^k) - V_r(C^k \to C^k) = d^{k-1}(C^k) + \sum_{i=1}^{r-1} m_i(C^k) \text{ for } 1 \leq k \leq N+1. \tag{3.5}$$

**Proof.** We first prove (3.4) by induction on k. Let $k = 1$ and $C^1 \neq \bar{C}^1 \in S^1$. By (3.2) and (2.1), $V_0(C^1 \to \bar{C}^1) = \sum_{i \in C^1} V(i) + U^1(C^1, \bar{C}^1) - d^0(C^1)$. By using (3.1), (2.2) and the definitions of $m_i(C^1)$, $V_r(C^1 \to C^1) = \sum_{i \in C^1} V(i) - \sum_{i=1}^{r-1} m_i(C^1) - d^0(C^1)$. A simple arithmetic verifies (3.6) for $k = 1$. Suppose (3.6) holds true up to $k-1 \leq N-1$. For $k \leq N$ and any different $C^k = \{C_i^{k-1}\}$, $\bar{C}_k \in S^k$, (3.2) and (2.5) imply that for any $W \cap C^k = \emptyset$,

$$\begin{aligned} V_0(C^k \to \bar{C}^k) &= \sum_i V_W(C_i^{k-1} \to) + U^k(C^k, \bar{C}^k) - d^{k-1}(C^k) \\ &= \sum_i V_0(C_i^{k-1} \to) + U^k(C^k, \bar{C}^k) - d^{k-1}(C^k). \end{aligned}$$

For some $W$ fulfilling (3.1) with $|W \cap C^k| = r$, (2.4) and the induction hypothesis imply that

$$\begin{aligned} &V_r(C^k \to C^k) = \sum_{C_i^k \cap W = \emptyset} V_W(C_i^{k-1} \to) + \sum_{C_i^k \cap W \neq \emptyset} V_W(C_i^{k-1} \to C_i^{k-1}) \\ &= \sum_{C_i^k \cap W = \emptyset} V_0(C_i^{k-1} \to) + \sum_{|C_i^k \cap W| = r_i} V_{r_i}(C_i^{k-1} \to C_i^{k-1}) \quad \text{where } \sum_i r_i = r \\ &= \sum_{C_i^k \cap W = \emptyset} V_0(C_i^{k-1} \to) + \sum_{|C_i^k \cap W| = r_i} \Big( V_0(C_i^{k-1} \to) - V^{k-1}(C_i^{k-1}) - \sum_{j=1}^{r_i - 1} m_j(C_i^{k-1}) \Big) \\ &= \sum_i V_0(C_i^{k-1} \to) - \sum_{|C_i^{k-1} \cap W| = r_i} \Big( V^{k-1}(C_i^{k-1}) + \sum_{j=1}^{r_i - 1} m_j(C_i^{k-1}) \Big). \end{aligned}$$

Taking the difference of the two equations above, (3.4) follows as

$$U^k(C^k,\bar{C}^k) - d^{k-1}(C^k) + \sum_{|C_i^{k-1}\cap W|=r_i} \Big(V^{k-1}(C_i^{k-1}) + \sum_{j=1}^{r_i-1} m_j(C_i^{k-1})\Big)$$
$$= U^k(C^k,\bar{C}^k) + \sum_{i=1}^{r-1} m_i(C^k)$$

by the definitions of of $m_i(C^k)$. The proof of (3.5) is similar and thus omitted.

## 4. PROOF OF THE MAIN THEOREM

The proof is done in three parts. We first consider $\delta_v$.

**Part (i).** $\delta_v = v_1 - v_2$.

The proof is almost obvious. Let $i_0$ be a state in $\underline{S}$ and $i_0 = C^0(i_0) \in C^1(i_0) \in \cdots \in C^N(i_0) \in C^{N+1}$ be the family tree of $i_0$. It is a principal sequence of $C^{N+1}$ because $i_0 \in \underline{S}$. By (1.3), (3.1) and (3.5), $v_1 - v_2 = V_1(C^{N+1} \to C^{N+1}) - V_2(C^{N+1} \to C^{N+1}) = m_1(C^{N+1}) = \max\{V^k(C^k) : C^k \in S^k \text{ but } C^k \neq C^k(i_0)\} = \delta_v$ in view of its definition in (1.6).

**Part (ii).** $\delta_h = v_{k_0-1} - v_{k_0}$, where

$$k_0 = \inf\{k \geq 2 : \exists \text{ an optimal } k\text{-graph W with } W \not\subseteq \underline{S}\}.$$

By (1.3), (3.1) and (3.5) again,

$$\begin{aligned} &v_{k_0-1} - v_{k_0} = m_{k_0-1}(C^{N+1}) \\ &\quad = \max\{V^k(C^k) : C^k \in \mathrm{P}(C^{N+1}) \setminus \mathrm{PS}(C^{i_1}) \setminus \mathrm{PS}(C^{i_2}) \setminus \cdots \setminus \mathrm{PS}(C^{i_{k_0-1}})\}. \end{aligned}$$

Since every $\mathrm{PS}(C^{i_r})$ is a part of the family tree of a state in $\underline{S}$, we obviously have

$$\{C^k : C^k \in \mathrm{P}(C^{N+1}) \setminus \mathrm{PS}(C^{i_1}) \setminus \mathrm{PS}(C^{i_2}) \setminus \cdots \setminus \mathrm{PS}(C^{i_{k_0-1}})\} \supseteq \{C^k : C^k \cap \underline{S} = \emptyset\}.$$

Thus $v_{k_0-1} - v_{k_0} \geq \delta_h$ by its definition in (1.6). On the other hand, if $v_{k_0-1} - v_{k_0} = m_{k_0-1}(C^{N+1}) > \max\{V^k(C^k) : C^k \cap \underline{S} = \emptyset\}$ then $m_{k_0-1} = V^k(C^k)$ for some $C^k$ with $C^k \cap \underline{S} \neq \emptyset$. This implies that for any $W$-optimal graph with $|W| = k_0$, we must have $W \subseteq \underline{S}$ which contradicts the definition of $k_0$. The proof of Part (ii) is completed.

**Part (iii).** For any $i \in S$, $h(i) = v(\{i\}) - v_1$.

For $i \in S$, let $i = C^0 \in C^1(i) \in \cdots \in C^k(i) \in \cdots \in C^N(i) \in C^{N+1}$ be the family tree of $i$. Suppose temporarily that the following holds.

**Lemma 4.1.** *Let $W = \{i\}$. Then for $1 \le k \le N+1$,*

$$(3.6) \qquad V_W\big(C^k(i) \to C^k(i)\big) - V_1\big(C^k(i) \to C^k(i)\big) = \sum_{r=1}^{k} \Big(d^{r-1}\big(C^r(i)\big) - V^{r-1}\big(C^{r-1}(i)\big)\Big).$$

*The conclusion follows then from* (1.6) *and the lemma with $k = N+1$ as shown below*:

$$\begin{aligned} h(i) &= \sum_{r=1}^{N+1} d^{r-1}\big(C^r(i)\big) - V^{r-1}\big(C^{r-1}(i)\big) \\ &= V_W\big(C^{N+1}(i) \to C^{N+1}(i)\big) - V_1\big(C^{N+1}(i) \to C^{N+1}(i)\big) = v(\{i\}) - v_1. \end{aligned}$$

It remains to verify Lemma 4.1, which is done by induction on $k$.

**Proof of Lemma 4.1.** For $k = 1$, we have from (2.2), (3.5) and (3.3) that

$$\begin{aligned} V_W\big(C^1(i) \to C^1(i)\big) = \sum_{j \in C^1(i) \setminus \{i\}} V(j) \text{ and } V_1\big(C^1(i) \to C^1(i)\big) \\ = \sum_{j \in C^1(i)} V(j) - d^0\big(C^1(i)\big). \end{aligned}$$

By taking the difference, (3.6) for $k = 1$ is verified. Suppose the lemma is proved up to $k-1$. Let $C^k(i) = \{C_j^{k-1}\}$. By (2.4), the induction hypothesis and (3.5),

$$\begin{aligned} (3.7) \quad V_W\big(C^k(i) \to C^k(i)\big) &= \sum_{C_j^{k-1} \neq C^{k-1}(i)} V_0(C_j^{k-1} \to) + V_W\big(C^{k-1}(i) \to C^{k-1}(i)\big) \\ &= \sum_j V_0(C_j^{k-1} \to) - V_0\big(C^{k-1}(i) \to \big) + V_1\big(C^{k-1}(i) \to C^{k-1}(i)\big) \\ &\qquad + \sum_{r=1}^{k-1} d^{r-1}\big(C^r(i)\big) - V^{r-1}\big(C^{r-1}(i)\big) \\ &= \sum_j V_0(C_j^{k-1} \to) - V^{k-1}\big(C^{k-1}(i)\big) + \sum_{r=1}^{k-1} d^{r-1}\big(C^r(i)\big) - V^{r-1}\big(C^{r-1}(i)\big). \end{aligned}$$

But $V_1\big(C^k(i) \to C^k(i)\big) = \sum_j V_0(C_j^{k-1} \to) - d^{k-1}\big(C^k(i)\big)$ by (3.5) and (3.3). By subtracting it from (3.7), the proof of (3.6) is completed by induction and thus so does Part (iii).

## REFERENCES


1. H. C. Chen, Y. Chow: *Cooperation in prisoner's dilemma games with local interaction and imitation.* (2005) Preprint.
2. T. S. Chiang, Y. Chow: *A limit theorem for a class of inhomogeneous Markov processes.* Ann. Probab., **17** (1989), 1483–1502.
3. T. S. Chiang, Y. Chow; *Asymptotic behavior of eigenvalues and random updating schemes.* Appl. Math. Optim., **28** (1993), 259–275.
4. T. S. Chiang, Y. Chow: *The asymptotic behavior of simulated annealing processes with absorption.* SIAM J. Control Optim., **32** (1994), 1247–1265.
5. T. S. Chiang, Y. Chow: *On the exit problem from cycles for simulated annealing processes-a backward equation approach.* Ann. Appl. Probability, **8** (1998), 896–916.
6. M. I. Freidlin, A. D. Wentzell: *Random Perturbations of Dynamical Systems.* Springer-Verlag, New York (1984).
7. S. Geman, D. Geman: *Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images.* IEEE Trans. Pattern Anal. Mach. Intell., **6** (1984), 721–741.
8. S. Kirkpatrick, C. Gebatt, M. Vecchi: *Optimizations by simulated annealing.* Science, **220** (1983), 671–680.
9. J. H. Wang: *The Theory of Games.* Oxford University Press (1988).
10. A. D. Ventcel: *On the asymptotics of eigenvalues of matrices with elements of order* $\exp(-V_{ij}/2\epsilon^2)$. Soviet Math. Dokl., **13** (1972), 65–68.


Institute of Mathematics,
Academia Sinica,
Taipei,
Taiwan 115
E–mails: matsch@math.sinica.edu.tw
chow@math.sinica.edu.tw

(Received September 25, 2006)