A Note on the First Occurrence of Strings

Ying-Chao Hung

Department of Statistics
National Chengchi University, Taipei 11605, Taiwan

hungy@nccu.edu.tw

Robert W. Chen

Department of Mathematics
University of Miami, Coral Gables, FL 33124, USA

chen@math.miami.edu

Alan Zame

Department of Mathematics
University of Miami, Coral Gables, FL 33124, USA

a.zame@math.miami.edu

May-Ru Chen

Department of Applied Mathematics
National Sun Yat-sen University, Kaohsiung 80424, Taiwan

chenmr@math.nsysu.edu.tw

Submitted: Jun 20, 2009; Accepted: Jan 3, 2010; Published: Jan 14, 2010
Mathematics Subject Classification: 65C50

Abstract

We consider the context of a three-person game in which each player selects strings over $\{0, 1\}$ and observe a series of fair coin tosses. The winner of the game is the player whose selected string appears first. Recently, Chen et al. [4] showed that if the string length is greater and equal to three, two players can collude to attain an advantage by choosing the pair of strings $11 \dots 10$ and $00 \dots 01$. We call these two strings "complement strings", since each bit of one string is the complement bit of the corresponding bit of the other string. In this note, we further study the property of complement strings for three-person games. We prove that if the string length is greater than five and two players choose any pair of complement strings (except for the pair $11 \dots 10$ and $00 \dots 01$), then the third player can always attain an advantage by choosing a particular string.

1 Introduction and Preliminaries

Consider a game in which players select strings over $\{0,1\}$ and observe a series of fair coin tosses, i.e., a string $\sigma = s_1s_2\ldots$ where each s_i is chosen independently and randomly from $\{0,1\}$. The winner of the game is the player whose selected string appears first. This problem has been formulated as a game or studied as a classical probabilistic problem by Chen [1], Chen and Lin [2], Chen and Zame [3], Chen et al. [4], Guibas and Odlyzko [6], Li [7], Gerber and Li [5], and Mori [8]. In [3], Chen and Zame proved that for two-person games, public knowledge of the opponent's string leads to an advantage. In [4], Chen et al. established the results for three-person games. In particular, they showed that if the string length is greater and equal to three, two players can collude to attain an advantage by choosing the pair of strings $11\ldots10$ and $00\ldots01$. We call these two strings "complement strings", since each bit of one string is the complement bit of the corresponding bit of the other string.

In this note, we further study the property of complement strings for three-person games. We prove that if the string length is greater than five and two players choose any pair of complement strings (except for the pair $11\ldots10$ and $00\ldots01$), then the third player can always attain an advantage by choosing a particular string. Before we proceed, we first introduce the following notations and some useful results obtained in [4].

Let $\{0,1\}^n$ be the set of all finite strings of length n over $\{0,1\}$. A string $\sigma \in \{0,1\}^n$ can be written as $\sigma = s_1s_2\ldots s_n$, with each bit $s_i \in \{0,1\}$. Given two strings σ, τ, their concatenation is denoted by $\sigma\tau$. The length of string σ is denoted by $|\sigma|$; for example, $|\sigma| = n$ if $\sigma \in \{0,1\}^n$. The empty string ϵ is the unique string of length zero. Given a string σ, its prefixes $\pi(\sigma)$ are all strings π such that $\sigma = \pi\tau$ for some string τ; its suffixes $\lambda(\sigma)$ are all strings λ such that $\sigma = \tau\lambda$ for some string τ.

Let $\{X_i\}$ be a sequence of random variables having values in $\{0,1\}$. Define the probability space Ω which is such that the X_i are i.i.d. with $P(X_i = s_j) = p_j$ for all i and j. The space Ω can be identified with the space of semi-infinite strings over $\{0,1\}$ by $\sigma = s_1s_2\ldots$ with $s_i = X_i(\omega)$. The definition of the prefix operation $\pi(\omega)$ is extended to apply to semi-infinite $\omega \in \Omega$ under this identification. For each string $\sigma \in \{0,1\}^n$, let T_σ be the waiting time for the first occurrence of σ in a randomly chosen $\omega \in \Omega$, i.e.,

$$T_\sigma(\omega) = \min\{|\tau| : \tau \in \pi(\omega) \text{ and } \sigma \in \lambda(\tau)\},$$

or $T_\sigma(\omega) = \infty$ if σ never appears in ω.

For strings $\sigma = s_1s_2\ldots s_n$, define $P(\sigma) = \prod_{i=1}^n P(X_i = s_i)$, i.e., the probability that a randomly chosen $\omega \in \Omega$ begins with σ. For strings $\sigma, \tau \in \{0,1\}^n$, define the operation

$$\sigma \circ \tau = \sum_{\substack{\rho \in \lambda(\sigma) \bigcap \pi(\tau) \\ \rho \neq \epsilon}} P(\rho)^{-1}.$$

For example, if $\sigma = 1111$, $\tau = 1101$, and $P(X_i = 0) = P(X_i = 1) = 1/2$, then $\lambda(\sigma) \bigcap \pi(\tau) = \{1, 11\}$ and $\sigma \circ \tau = 2 + 2^2 = 6$. The *complement string* of $\sigma = s_1s_2\ldots s_n$ is defined as $\bar{\sigma} = \bar{s}_1\bar{s}_2\ldots\bar{s}_n$, where $\bar{s}_i = 1 - s_i$ is the complement bit of s_i. For example, $\sigma_1 = 00\ldots01$ is clearly the complement string of $\sigma_2 = 11\ldots10$.

We cite Lemma 5 in [4] as Lemma 1 in this note, since it is essential for proving our main theorem. For comparison purposes, we also cite Theorem 3 in [4] as Theorem 1 in this note.

Lemma 1 *Let $\sigma_1, \sigma_2, \ldots, \sigma_k$ be k distinct strings in $\{0,1\}^n$. We have the following system of $k+1$ linear equations, where $p_i = P(T_{\sigma_i} = N_k)$ for $i = 1, \ldots, k$,*

$$\begin{pmatrix} 0 & 1 & \cdots & 1 \\ 1 & & & \\ \vdots & & (\sigma_i \circ \sigma_i - \sigma_j \circ \sigma_i)_{i+1,j+1} & \\ 1 & & & \end{pmatrix} \begin{pmatrix} E(N_k) \\ p_1 \\ \vdots \\ p_k \end{pmatrix} = \begin{pmatrix} 1 \\ \sigma_1 \circ \sigma_1 \\ \vdots \\ \sigma_k \circ \sigma_k \end{pmatrix}$$

Note that for the remaining of this note, we assume that $P(X_i = 0) = P(X_i = 1) = 1/2$ and σ_2 is always treated as the complement string of σ_1. This means that $\sigma_1 \circ \sigma_1 = \sigma_2 \circ \sigma_2$, and $\sigma_1 \circ \sigma_2 = \sigma_2 \circ \sigma_1$. To simplify the notations, we denote $\sigma_1 \circ \sigma_1$ and $\sigma_2 \circ \sigma_2$ by $2^n + \alpha$, $\sigma_1 \circ \sigma_2$ and $\sigma_2 \circ \sigma_1$ by β, $\sigma_3 \circ \sigma_1$ by γ, $\sigma_3 \circ \sigma_2$ by δ, $\sigma_1 \circ \sigma_3$ by a, $\sigma_2 \circ \sigma_3$ by b, and $\sigma_3 \circ \sigma_3$ by $2^n + c$, respectively. Thus, we have the following facts.

Fact 1 *By the preceding definitions, we have $0 \leqslant \alpha < 2^n$ and $0 \leqslant \beta < 2^n$.*

Proof. The result is straightforward from Lemma 1, so the proof is omitted.

Fact 2 *By the preceding definitions, we have $\gamma \neq \delta$. Further, if $\gamma > \delta$,*

$$(2^n + \alpha - \beta)p_1 - (2^n + \alpha - \beta)p_2 + (\gamma - \delta)p_3 = 0$$

and

$$(\beta - a)p_1 + (2^n + \alpha - b)p_2 - (2^n + c - \delta)p_3 = 0;$$

while if $\gamma < \delta$,

$$(2^n + \alpha - \beta)p_1 - (2^n + \alpha - \beta)p_2 - (\delta - \gamma)p_3 = 0$$

and

$$(2^n + \alpha - a)p_1 + (\beta - b)p_2 - (2^n + c - \gamma)p_3 = 0.$$

Proof. Due to the property of symmetry, here we assume that $s_1 = 0$. The result can be directly obtained from Lemma 1.

For notational convenience, a repeating string such as $\sigma\sigma\ldots\sigma$ is written as $[\sigma]^*$.

Theorem 1 *For $n \geqslant 3$, let σ_1, σ_2, and σ_3 be three distinct strings in $\{0,1\}^n$, where $\sigma_1 = [0]^*1$, $\sigma_2 = [1]^*0$, and σ_3 is arbitrary. Let $p_i = P(T_{\sigma_i} = N_3)$ be the probability that σ_i appears first among the three. Then $p_3 < \max(p_1, p_2)$.*

2 Main Results

Lemma 2 *Let $\sigma_1 = s_1 s_2 \ldots s_n$ and $\sigma_2 = \bar{s}_1 \bar{s}_2 \ldots \bar{s}_n$ satisfy $\sigma_1, \sigma_2 \in \{0,1\}^n \setminus \{[0]^*1, [1]^*0\}$. If $s_1 = s_2$ and $n > 5$, then there exists a string $\sigma_3 \in \{0,1\}^n \setminus \{\sigma_1, \sigma_2\}$ such that $p_3 > \max(p_1, p_2)$.*

Proof. We consider the following four cases.
Case 1: $s_1 = s_2 = \ldots = s_{n-1} = s_n = 0$.

In this case, let $\sigma_3 = 1[0]^*$. By Fact 2, we then have $(2^{n+1}-2)p_1-(2^{n+1}-2)p_2+(2^n-2)p_3 = 0$ and $(2^{n+1}-4)p_2 - 2^n p_3 = 0$. Therefore, $0 < p_1 < p_2 < p_3$ since $n \geqslant 6$.
Case 2: $s_1 = s_2 = s_{n-1} = s_n = 0$ and $\sigma_1 \neq [0]^*$.

In this case, let $\sigma_3 = [01]^*00$ if n is even; otherwise let $\sigma_3 = [10]^*100$. Thus, we have $a = 0$ or 2, $b = 0$ or 2, $a + b = 2$, $c = 0$ or 2, $\gamma = 6$, and $\delta = 0$. Since $\gamma > \delta$, by Fact 2, we then have $(2^n + \alpha - \beta)p_1 - (2^n + \alpha - \beta)p_2 + 6p_3 = 0$ and $(\beta - a)p_1 + (2^n + \alpha - b)p_2 - (2^n + c)p_3 = 0$. The last equation can be written as $(\beta - a)p_1 + (2^n + \alpha - 2 + a)p_2 - (2^n + c)p_3 = 0$, and thus $(2^n + c)(p_2 - p_3) + \beta p_1 + \alpha(p_2 - p_1) + (\alpha - 2 - c)p_2 = 0$. Therefore, $0 < p_1 < p_2 < p_3$ since $\alpha \geqslant 6$, $0 \leqslant \beta < 2^n$ (by Fact 1), $a \leqslant 2$, $b \leqslant 2$, and $c \leqslant 2$.
Case 3: $s_1 = s_2 = 0$ and $s_{n-1} = s_n = 1$.

In this case, let $\sigma_3 = [01]^*00$ if n is even; otherwise let $\sigma_3 = [10]^*100$. Thus, we have $a = 0$ or 2, $b = 0$ or 2, $c = 0$ or 2, $a + b = 2$, $\gamma = 6$, and $\delta = 0$. By Fact 2, we then have $(2^n + \alpha - \beta)p_1 - (2^n + \alpha - \beta)p_2 + 6p_3 = 0$ and $(\beta - a)p_1 + (2^n + \alpha - b)p_2 - (2^n + c)p_3 = 0$. Hence $p_1 > p_2 - 0.2p_3$ since $n \geqslant 6$, $\alpha \geqslant 0$, and $\beta < 2^{n-1}$. Further, since $\beta \geqslant 6$ and $a = 0$ or 2 (i.e., $\beta > a$), we then have $(2^n + \alpha + \beta - a - b)p_2 - (2^n + c + 0.2(\beta - a))p_3 < 0$. Since $a + b = 2$ and $c \leqslant 2$, we conclude that $(2^n + \alpha + \beta - a - b) > (2^n + c + 0.2(\beta - a))$, and thus $0 < p_1 < p_2 < p_3$.
Case 4: $s_1 = s_2 = s_n = 0$, $s_{n-1} = 1$ or $s_1 = s_2 = s_{n-1} = 0$, $s_n = 1$, and $\sigma_1 \neq [0]^*1$.

In this case, let $\sigma_3 = [0]^*1$. Thus, we have $a = 0$ or 2, $b = 0$ or 2, $c = 0$, $a + b = 2$, $\gamma \geqslant 8$, and $\delta = 2$. By Fact 2, we then have $(2^n + \alpha - \beta)p_1 - (2^n + \alpha - \beta)p_2 + (\gamma - 2)p_3 = 0$ and $(\beta - a)p_1 + (2^n + \alpha - b)p_2 - (2^n - 2)p_3 = 0$. The last equation can be written as $(2^n - 2)(p_2 - p_3) + (\alpha + 2 - b)p_2 + (\beta - a)p_1 = 0$, and thus $(2^n - 2)(p_2 - p_3) + \alpha p_2 + \beta p_1 + a(p_2 - p_1) = 0$ since $2 - b = a$. Therefore, $0 < p_1 < p_2 < p_3$ since $\gamma \geqslant 8$, $\alpha \geqslant 0$, and $0 \leqslant \beta < 2^n$ (by Fact 1).

The proof of Lemma 2 is complete by summarizing the results from Case 1 - Case 4.

Lemma 3 *Let $\sigma_1 = s_1 s_2 \ldots s_n$ and $\sigma_2 = \bar{s}_1 \bar{s}_2 \ldots \bar{s}_n$ satisfy $\sigma_1, \sigma_2 \in \{0,1\}^n \setminus \{[0]^*1, [1]^*0\}$. If $s_1 \neq s_2 = s_3$ and $n > 5$, then there exists a string $\sigma_3 \in \{0,1\}^n \setminus \{\sigma_1, \sigma_2\}$ such that $p_3 > \max(p_1, p_2)$.*

Proof. We consider the following three cases.
Case 1: $s_1 = s_{n-1} = s_n = 0$, $s_2 = s_3 = 1$ or $s_1 = 0$, $s_2 = s_3 = s_{n-2} = s_{n-1} = s_n = 1$.

In this case, let $\sigma_3 = [01]^*1$ if n is odd; otherwise let $\sigma_3 = [01]^*0011$. Thus, $a = 0$ or 2, $b = 0$ or 2, $a + b = 2$, $c = 0$, $\gamma = 8$, and $\delta = 2$ or 34. If $\delta = 2$, then by Fact 2, we have $(2^n + \alpha - \beta)p_1 - (2^n + \alpha - \beta)p_2 + 6p_3 = 0$ and $(\beta - a)p_1 + (2^n + \alpha - b)p_2 - (2^n - 2)p_3 = 0$. Note that the first equation directly implies $p_1 < p_2$, while the second equation implies $(2^n - 2)(p_2 - p_3) + \alpha p_2 + \beta p_1 + a(p_2 - p_1) = 0$ since $b = 2 - a$. Therefore, $0 < p_1 < p_2 < p_3$ since $\alpha \geqslant 0$, $\beta \geqslant 0$, and $a \geqslant 0$. If $\delta = 34$, then by Fact 2, we have $(2^n + \alpha - \beta)p_1 - (2^n + \alpha - \beta)p_2 - 26p_3 = 0$

and similarly $(2^n-8)(p_1-p_3)+(6+\alpha)p_1+b(p_1-p_2)+\beta p_2=0$. Therefore, $0<p_2<p_1<p_3$ since $\alpha \geqslant 0$, $\beta \geqslant 0$, and $b \geqslant 0$.

Case 2: $s_1=s_{n-1}=0$, $s_n=1$ or $s_1=s_n=0$, $s_{n-1}=1$.

In this case, let $\sigma_3=[0]^*1$. Thus, $a=0$ or 2, $b=0$ or 2, $a+b=2$, $c=0$, $\gamma=4$, and $\delta=2$. By Fact 2, we then have $(2^n+\alpha-\beta)p_1-(2^n+\alpha-\beta)p_2+2p_3=0$ and $(2^n-2)(p_2-p_3)+(\alpha+2-b)p_2+(\beta-a)p_1=0$. Therefore, $0<p_1<p_2<p_3$ since $\alpha \geqslant 2$, $\beta \geqslant 2$, $a \leqslant 2$, and $b \leqslant 2$.

Case 3: $s_1=s_{n-2}=0$ and $s_2=s_3=s_{n-1}=s_n=1$.

In this case, let $\sigma_3=[0]^*1$. Thus, $a=0$, $b=6$, $c=0$, $\alpha \geqslant 8$, $\beta \geqslant 0$, $\gamma=4$, and $\delta=2$. By Fact 2, we then have $(2^n+\alpha-\beta)p_1-(2^n+\alpha-\beta)p_2+2p_3=0$ and $\beta p_1+(2^n+\alpha-6)p_2-(2^n-2)p_3=0$. Therefore, $0<p_1<p_2<p_3$ since $\alpha \geqslant 8$ and $\beta \geqslant 0$.

The proof of Lemma 3 is complete by summarizing the results from Case 1 - Case 3.

Lemma 4 *Let $\sigma_1=s_1s_2\ldots s_n$ and $\sigma_2=\bar{s}_1\bar{s}_2\ldots\bar{s}_n$ satisfy $\sigma_1,\sigma_2 \in \{0,1\}^n \setminus \{[0]^*1,[1]^*0\}$. If $s_1=s_3=0$, $s_2=s_4=1$, and $n>6$, then there exists a string $\sigma_3 \in \{0,1\}^n \setminus \{\sigma_1,\sigma_2\}$ such that $p_3>\max(p_1,p_2)$.*

Proof. We consider the following four cases.

Case 1: $s_1=s_3=s_{n-2}=s_{n-1}=s_n=0$, $s_2=s_4=1$.

In this case, let $\sigma_3=10[01]^*$ if n is even; otherwise let $\sigma_3=10011[01]^*$. Thus, $a=0$, $b=2$, $c=2$, $\alpha \geqslant 2$, $\beta \geqslant 0$, $\gamma \geqslant 4$, $\delta \geqslant 8$, and $\gamma \neq \delta$. If $\gamma>\delta$, then by Fact 2, we have $(2^n+\alpha-\beta)p_1-(2^n+\alpha-\beta)p_2+(\gamma-\delta)p_3=0$ and $\beta p_1+(2^n+\alpha-2)p_2-(2^n+2-\delta)p_3=0$. Therefore, $0<p_1<p_2<p_3$ since $\alpha \geqslant 2$, $\beta \geqslant 0$, and $\delta \geqslant 8$. If $\gamma<\delta$, then by Fact 2, we have $(2^n+\alpha-\beta)p_1-(2^n+\alpha-\beta)p_2-(\delta-\gamma)p_3=0$ and $(2^n+\alpha)p_1+(\beta-2)p_2-(2^n+2-\gamma)p_3=0$. Therefore, $0<p_2<p_1<p_3$ since $\alpha \geqslant 2$, $\beta \geqslant 0$, and $\gamma \geqslant 4$.

Case 2: $s_1=s_3=s_{n-1}=s_n=0$ and $s_2=s_4=s_{n-2}=1$.

In this case, let $\sigma_3=[0]^*101$. Thus, $a=6$, $b=0$, $c=0$, $\alpha \geqslant 2$, $\beta \geqslant 0$, $\gamma=20$, and $\delta=10$. By Fact 2, we then have $(2^n+\alpha-\beta)p_1-(2^n+\alpha-\beta)p_2+10p_3=0$ and $(\beta-6)p_1+(2^n+\alpha)p_2-(2^n-10)p_3=0$. Note that the first equation directly implies $p_1<p_2$. The second equation implies $p_3=\left(\frac{\beta-6}{2^n-10}\right)p_1+\left(\frac{2^n+\alpha}{2^n-10}\right)p_2$, thus $p_3 \geqslant \left(\frac{-6}{2^n-10}\right)p_1+\left(\frac{2^n+\alpha}{2^n-10}\right)p_2$ since $\beta \geqslant 0$. Since $p_1<p_2$ and $\alpha \geqslant 2$, we then have $p_3>\left(\frac{2^n-4}{2^n-10}\right)p_2$, and thus $0<p_1<p_2<p_3$.

Case 3: $s_1=s_3=s_{n-1}=0$, $s_2=s_4=s_n=1$ or $s_1=s_3=s_n=0$, $s_2=s_4=s_{n-1}=1$.

In this case, let $\sigma_3=[0]^*1$. Thus, $a=0$ or 2, $b=0$ or 2, $a+b=2$, $c=0$, $\alpha \geqslant 2$, $\beta \geqslant 2$, $\gamma=4$, and $\delta=2$. By Fact 2, we then have $(2^n+\alpha-\beta)p_1-(2^n+\alpha-\beta)p_2+2p_3=0$ and $(\beta-a)p_1+(2^n+\alpha-b)p_2-(2^n-2)p_3=0$. Therefore, $0<p_1<p_2<p_3$ since $\alpha \geqslant 2$, $\beta \geqslant 2$, $a \leqslant 2$, and $b \leqslant 2$.

Case 4: $s_1=s_3=0$ and $s_2=s_4=s_{n-1}=s_n=1$.

In this case, let $\sigma_3=0[01]^*$ if n is odd; otherwise let $\sigma_3=00[10]^*$. Thus, $a=0$, $b=6$, $c=0$ or 2, $\alpha \geqslant 0$, $\beta \geqslant 2$, $\gamma \geqslant 10$, $\delta \geqslant 10$, and $\gamma \neq \delta$. If $\gamma>\delta$, then by Fact 2, we have $(2^n+\alpha-\beta)p_1-(2^n+\alpha-\beta)p_2+(\gamma-\delta)p_3=0$ and $\beta p_1+(2^n+\alpha-6)p_2-(2^n+c-\delta)p_3=0$. Therefore, $0<p_1<p_2<p_3$ since $\alpha \geqslant 0$, $\beta \geqslant 2$, $c \leqslant 2$, and $\delta \geqslant 10$. If $\gamma<\delta$, then by Fact 2, we have $(2^n+\alpha-\beta)p_1-(2^n+\alpha-\beta)p_2-(\delta-\gamma)p_3=0$ and $(2^n+\alpha)p_1+(\beta-6)p_2-(2^n+c-\gamma)p_3=0$. Therefore, $0<p_2<p_1<p_3$ since $\alpha \geqslant 0$, $\beta \geqslant 2$, $c \leqslant 2$, and $\gamma \geqslant 10$.

The proof of Lemma 4 is complete by summarizing the results from Case 1 - Case 4.

Lemma 5 *Let* $\sigma_1 = s_1 s_2 \ldots s_n$ *and* $\sigma_2 = \bar{s}_1 \bar{s}_2 \ldots \bar{s}_n$ *satisfy* $\sigma_1, \sigma_2 \in \{0,1\}^n \setminus \{[0]^*1, [1]^*0\}$. *If* $s_1 = s_3 = s_4 \neq s_2$ *and* $n > 6$, *then there exists a string* $\sigma_3 \in \{0,1\}^n \setminus \{\sigma_1, \sigma_2\}$ *such that* $p_3 > \max(p_1, p_2)$.

Proof. We consider the following four cases.

Case 1: $s_1 = s_3 = s_4 = s_{n-2} = s_{n-1} = s_n = 0$, $s_2 = 1$.

In this case, let $\sigma_3 = 0111010$ when $n = 7$. Thus, by Fact 2, it is easy to see that $0 < p_1 < p_2 < p_3$. When $n \geqslant 8$, let $\sigma_3 = 01[10]^*$ if n is even; otherwise let $\sigma_3 = 01101[10]^*$. Thus, $a = 2$, $b = 0$, $c = 2$, $\alpha \geqslant 2$, $\beta \geqslant 0$, $\gamma = 10$, and $\delta = 4$. By Fact 2, we then have $(2^n+\alpha-\beta)p_1-(2^n+\alpha-\beta)p_2+6p_3 = 0$ and $(\beta-2)p_1+(2^n+\alpha)p_2-(2^n-2)p_3 = 0$. Note that the first equation directly implies $p_1 < p_2$. The second equation implies $p_3 = \left(\frac{\beta-2}{2^n-2}\right)p_1+\left(\frac{2^n+\alpha}{2^n-2}\right)p_2$, thus $p_3 \geqslant \left(\frac{-2}{2^n-2}\right)p_1 + \left(\frac{2^n+\alpha}{2^n-2}\right)p_2$ since $\beta \geqslant 0$. Since $p_1 < p_2$ and $\alpha \geqslant 2$, we then have $p_3 > \left(\frac{2^n}{2^n-2}\right)p_2$, and thus $0 < p_1 < p_2 < p_3$.

Case 2: $s_1 = s_3 = s_4 = 0$, $s_2 = s_{n-2} = s_{n-1} = s_n = 1$.

In this case, let $\sigma_3 = 1[10]^*$ if n is odd; otherwise let $\sigma_3 = 1100[10]^*$. Thus, $a = 6$, $b = 0$, $c = 0$, $\alpha \geqslant 0$, $\beta \geqslant 2$, $\gamma = 10$, and $\delta = 4$. By Fact 2, we then have $(2^n + \alpha - \beta)p_1 - (2^n + \alpha - \beta)p_2 + 6p_3 = 0$ and $(\beta - 6)p_1 + (2^n + \alpha)p_2 - (2^n - 4)p_3 = 0$. Note that the first equation directly implies $p_1 < p_2$, while the second equation implies $p_3 = \left(\frac{\beta-6}{2^n-4}\right)p_1 + \left(\frac{2^n+\alpha}{2^n-4}\right)p_2$, thus $p_3 \geqslant \left(\frac{-4}{2^n-4}\right)p_1 + \left(\frac{2^n+\alpha}{2^n-4}\right)p_2$ since $\beta \geqslant 2$. Since $p_1 < p_2$ and $\alpha \geqslant 0$, we then have $p_3 > \left(\frac{2^n-4}{2^n-4}\right)p_2 = p_2$, and thus $0 < p_1 < p_2 < p_3$.

Case 3: $s_1 = s_3 = s_4 = s_{n-2} = 0$, $s_2 = s_{n-1} = s_n = 1$.

In this case, let $\sigma_3 = 110[1]^*010$ if n is odd; otherwise let $\sigma_3 = 11[10]^*$. The proof for this case is the same as that for **Case 2**, so it is omitted.

Case 4: $s_1 = s_3 = s_4 = s_{n-1} = 0$, $s_2 = s_n = 1$ or $s_1 = s_3 = s_4 = s_n = 0$, $s_2 = s_{n-1} = 1$.

In this case, let $\sigma_3 = [0]^*1$. Thus, $a = 0$ or 2, $b = 0$ or 2, $c = 0$, $\alpha \geqslant 2$, $\beta \geqslant 2$, $\gamma = 4$, and $\delta = 2$. By Fact 2, we then have $(2^n + \alpha - \beta)p_1 - (2^n + \alpha - \beta)p_2 + 2p_3 = 0$ and $(\beta - a)p_1 + (2^n + \alpha - b)p_2 - (2^n - 2)p_3 = 0$. Therefore, $0 < p_1 < p_2 < p_3$ since $\alpha \geqslant 2$, $\beta \geqslant 2$, $a \leqslant 2$, and $b \leqslant 2$.

Case 5: $s_1 = s_3 = s_4 = s_{n-1} = s_n = 0$, $s_2 = s_{n-2} = 1$.

It suffices to consider the following two sub-cases.

Sub-Case 5-1: $\alpha+\beta \geqslant 4$. In this case, let $\sigma_3 = 0[01]^*0$ if n is even; otherwise let $\sigma_3 = 00[01]^*0$. Thus, $a = 6$, $b = 0$, $c = 2$, $\alpha \geqslant 2$, $\beta \geqslant 0$, $\gamma = 10$, and $\delta = 4$. By Fact 2, we then have $(2^n + \alpha - \beta)p_1 - (2^n + \alpha - \beta)p_2 + 6p_3 = 0$ and $(\beta - 6)p_1 + (2^n + \alpha)p_2 - (2^n - 2)p_3 = 0$. Note that the first equation directly implies $p_1 < p_2$, while the second equation implies $(2^n - 2)(p_2 - p_3) + \alpha p_2 + 2(p_2 - p_1) + \beta p_1 - 4p_1 = 0$. Therefore, $0 < p_1 < p_2 < p_3$ since $\alpha + \beta \geqslant 4$.

Sub-Case 5-2: $\alpha = 2$ and $\beta = 0$. The fact that $\alpha = 2$ implies that $s_{n-3} = 1$, since $s_1 = s_3 = s_4 = s_{n-1} = s_n = 0$ and $s_2 = s_{n-2} = 1$. It also implies that $s_1 s_2 \ldots s_i \neq s_{n-i+1} s_{n-i+2} \ldots s_n$ for all $i = 2, 3, \ldots, n-1$. The fact that $\beta = 0$ implies that $s_1 s_2 \ldots s_i \neq \bar{s}_{n-i+1} \bar{s}_{n-i+2} \ldots \bar{s}_n$ and $\bar{s}_1 \bar{s}_2 \ldots \bar{s}_i \neq s_{n-i+1} s_{n-i+2} \ldots s_n$ for all $i = 1, 2, \ldots, n$. Since $s_{n-3} = s_{n-2} = 1$ and $s_3 = s_4 = s_{n-1} = 0$, we then have that $n \geqslant 8$. To select σ_3 for each possible σ_1, we consider a substring $s_1s_2\cdots s_{n-1}$ of σ_1. Let $\sigma_3 = 0s_1s_2\cdots s_{n-1}$, we then have that $a = 6$, $b = 0$, $c = 2$, $\alpha = 2$, $\beta = 0$, $\gamma = 2^{n-1}+2$, and $4 \leqslant \delta < 4+2^{n-5}$. By Fact 2, we then have $(2^n+2)p_1 - (2^n+2)p_2 + (2^{n-1}+2-\delta)p_3 = 0$ and $-6p_1 + (2^n+2)p_2 - (2^n+2-\delta)p_3 = 0$. Note that the first equation implies $p_1 < p_2$ (since $2^{n-1}+2-\delta > 0$) and $p_1 = p_2 - \left(\frac{2^{n-1}+2-\delta}{2^n+2}\right)p_3$. Adding these results to the second equation, we then have $-6p_2 + 6\left(\frac{2^{n-1}+2-\delta}{2^n+2}\right)p_3 + (2^n+2)p_2 - (2^n+2-\delta)p_3 = 0$, thus $(2^n-4)p_2 = \left[-3\left(\frac{2^{n-1}+2-\delta}{2^{n-1}+1}\right) + 2^n + 2 - \delta\right]p_3 = \left[2^n - 1 - \delta + 3\left(\frac{\delta-1}{2^{n-1}+1}\right)\right]p_3$. Since $n \geqslant 8$ and $4 \leqslant \delta < 4+2^{n-5}$, we have that $2^n - 1 - \delta + 3\left(\frac{\delta-1}{2^{n-1}+1}\right) < 2^n - 4$. Thus, we conclude that $0 < p_1 < p_2 < p_3$. The proof of Lemma 5 is complete by summarizing the results from Case 1 - Case 5.

Note that when $n = 6$, there are eight strings that are not included in the cases of Lemma 2 - Lemma 5. Now we show that how to choose σ_3 so that $p_3 > \max(p_1, p_2)$ for each of these eight strings. When $\sigma_1 = 010000$ and $\sigma_2 = 101111$, choose $\sigma_3 = 011010$; when $\sigma_1 = 010001$ and $\sigma_2 = 101110$, choose $\sigma_3 = 111100$; when $\sigma_1 = 010010$ and $\sigma_2 = 101101$, choose $\sigma_3 = 100000$; when $\sigma_1 = 010011$ and $\sigma_2 = 101100$, choose $\sigma_3 = 111010$; when $\sigma_1 = 010100$ and $\sigma_2 = 101011$, choose $\sigma_3 = 001010$; when $\sigma_1 = 010101$ and $\sigma_2 = 101010$, choose $\sigma_3 = 000000$; when $\sigma_1 = 010110$ and $\sigma_2 = 101001$, choose $\sigma_3 = 111000$; when $\sigma_1 = 010111$ and $\sigma_2 = 101000$, choose $\sigma_3 = 110010$. Combining this result with those from Lemma 2 - Lemma 5, we have the following main theorem:

Theorem 2 *For any string σ_1 and its complement string σ_2 in $\{0,1\}^n \setminus \{[0]^*1, [1]^*0\}$, there always exists a string σ_3 in $\{0,1\}^n \setminus \{\sigma_1, \sigma_2\}$ such that $p_3 > \max(p_1, p_2)$ when $n > 5$.*

Remark 1. Note that Theorem 2 does not hold when $n = 4$ or 5. For example, if $\sigma_1 = 0011$ and $\sigma_2 = 1100$, then $p_3 < \max(p_1, p_2)$ for any string σ_3 in $\{0,1\}^4 \setminus \{\sigma_1, \sigma_2\}$. In addition, if $\sigma_1 = 0100$ and $\sigma_2 = 1011$, then $p_3 \leqslant \max(p_1, p_2)$ for any string σ_3 in $\{0,1\}^4 \setminus \{\sigma_1, \sigma_2\}$. In summary, numerical results show that for any string σ_1 and its complement string σ_2 in $\{0,1\}^4 \setminus \{0001, 1110, 0011, 1100, 0100, 1011, 0111, 1000\}$, there always exists a string σ_3 in $\{0,1\}^4 \setminus \{\sigma_1, \sigma_2\}$ such that $p_3 > \max(p_1, p_2)$. Analogously, numerical results show that for any string σ_1 and its complement string σ_2 in $\{0,1\}^5 \setminus \{00001, 11110, 01000, 10111\}$, there always exists a string σ_3 in $\{0,1\}^5 \setminus \{\sigma_1, \sigma_2\}$ such that $p_3 > \max(p_1, p_2)$.

We next present some other interesting results regarding to the complement strings. These results are summarized in the following Theorem 3 and Theorem 4.

Theorem 3 *Let σ_1, σ_2, and σ_3 be three distinct strings in $\{0,1\}^n$, where $\sigma_1 = [0]^*1$, $\sigma_2 = [1]^*0$, and σ_3 is arbitrary. When $n \geqslant 3$, we have that either $P(T_{\sigma_1} < T_{\sigma_3}) > P(T_{\sigma_3} < T_{\sigma_1})$ or $P(T_{\sigma_2} < T_{\sigma_3}) > P(T_{\sigma_3} < T_{\sigma_2})$, i.e., either σ_1 or σ_2 has the better chance of occurring before σ_3.*

Proof. The proof is similar to that of Theorem 1 and therefore omitted.

Theorem 4 *Let $\sigma_1 = s_1 s_2 \dots s_n$ and $\sigma_2 = \bar{s}_1 \bar{s}_2 \dots \bar{s}_n$ satisfy $\sigma_1, \sigma_2 \in \{0,1\}^n \setminus \{[0]^*1, [1]^*0\}$. When $n > 5$, there always exists a string $\sigma_3 \in \{0,1\}^n \setminus \{\sigma_1, \sigma_2\}$ such that $P(T_{\sigma_3} < T_{\sigma_1}) \geqslant P(T_{\sigma_1} < T_{\sigma_3})$ and $P(T_{\sigma_3} < T_{\sigma_2}) \geqslant P(T_{\sigma_2} < T_{\sigma_3})$, i.e., σ_3 has the same or better chance of occurring before σ_1 and σ_2.*

Proof. The proof can be shown in a similar fashion to that of Theorem 2, so it is omitted.

Remark 2. It should be noted that the inequalities in Theorem 4 can not be replaced by the strict inequalities. In addition, the string σ_3 chosen in Theorem 4 may not work in Theorem 2. To illustrate, let us consider the pair of complement strings $\sigma_1 = 101001000$ and $\sigma_2 = 010110111$. Let $\sigma_3 = 011110101$, some algebra shows that $P(T_{\sigma_3} < T_{\sigma_1}) = \frac{510}{1018} > \frac{1}{2}$ and $P(T_{\sigma_3} < T_{\sigma_2}) = \frac{1}{2}$, which clearly satisfy the result of Theorem 4. However, in this case we have that $p_3 < 0.331 < \max(p_1, p_2)$, which contradicts the result of Theorem 2.

References

[1] Chen, R., *A circular property of the occurrence of sequence patterns in the fair coin-tossing process,* Adv. Appl. Prob., Vol. 21, 1989. pp. 938–940.

[2] Chen, R. and Lin, H. E., *On the fair coin-tossing processes,* J. Multivariate Anal., Vol. 15, 1984. pp. 222–227.

[3] Chen, R. and Zame, A., *On the fair coin-tossing games*, J. Multivariate Anal., Vol. 9, 1979. pp. 150–157.

[4] Chen, R., Zame, A. and Rosenberg, B., *On the first occurrence of strings*, The Electronic J. of Combin., Vol. 16, 2009. # R29.

[5] Gerber, H. V. and Li, S. Y. R., *The occurrence of sequence patterns in repeated experiments and hitting times in a Markov chain,* Stoch. Processes and their Appli., Vol. 11, 1981. pp. 101–108.

[6] Guibas, L. and Odlyzko, A. M., *String overlaps, pattern matching, and nontransitive games*, J. Combin. Theory Ser.A, Vol. 30, 1981. pp. 183–208.

[7] Li, S. Y. R., *A martingale approach to the study of occurrence of sequence patterns in repeated experiments*, Ann. Probab. Vol. 8, 1980. pp. 1171–1176.

[8] Mori, T. F., *On the waiting times till each of some given patterns occurs as a run*, Probab. Th. Rel. Fields, Vol. 87, 1991. pp. 313–323.