Chamchuri Journal of Mathematics
Volume 11(2019), 1–9
http://www.math.sc.chula.ac.th/cjm

Classes of equally likely outcomes of a riffle shuffle on a deck of alternating cards

S. Kotchasarn, P. Khampheng*,
T. Thanodomdech and G. Tassanasophon

Received 13 June 2019
Revised 14 August 2019
Accepted 14 August 2019

Abstract: The mathematical model of riffle shuffle has been a subject of some studies. Whereas most of these regard the cards as all different, in 2006 Conger and Viswanath treated some of them as identical and investigated the implications. When the initial deck is arranged in alternating reds and blacks, they showed that two outcomes are equally likely if a number of particular transformations can turn one of them into the other. This transformation, which may be viewed as a reversible string rewriting system, partitions the set of outcomes into equivalence classes. They conjectured that the number of such classes is precisely $(n+3)2^{n-2}$, where n is the number of cards of each color. In this paper, the assertion is proven true by the method of invariant and derivation of canonical forms.

Keywords: Alternating cards, Equivalence class, Riffle shuffle, Transformation

2000 Mathematics Subject Classification: 05A05

1 Introduction

Riffle shuffle, one of the most common methods of card shuffling and its mathematical model, usually called Gilbert-Shannon-Reeds or GSR model, has been

* *Corresponding author*

a subject of some studies (see, e.g. [1],[3]). In 2006, Conger and Viswanath [2] investigated a new aspect of this shuffle whereby some of the cards are regarded as the same. One particular case is when there are only two types of cards, red and black, and the initial deck is arranged in an alternating manner (i.e. red, black, red, black, and so on). When the initial deck consists of n reds and n blacks, each outcome can be viewed as a binary string of n zeros and n ones and we denote the set of all these strings by S_n. It was shown that any two members of S_n are the equally likely outcomes of riffle shuffle if they are equivalent in the following sense.

First, we define a *transformation* R which can be applied at any local contiguous stretch (i.e. substring) τ of $\sigma \in S_n$ that contains equal numbers of two digits. The transformation proceeds by reversing the substring, then inverting every zero and one of this substring into one and zero respectively. We say $\sigma \xrightarrow{R} \sigma'$ if σ becomes σ' when some substring of σ is applied by a transformation R. For example, when R is applied at the middle six digits of 00111001, namely the substring 011100, it becomes 01000111 after the first step and the resulting string is 01100011; hence, we may write $00111001 \xrightarrow{R} 01100011$. The equivalence relation defined as the transitive, reflexive closure of $\xrightarrow{R}$ ($\xrightarrow{R}$ is obviously symmetric) will be denoted by $\leftrightarrow$.

As a result, S_n can be partitioned into equivalence classes of equally likely outcomes by $\leftrightarrow$. Conger and Viswanath [2] observed that the number of these classes appears to be given by a simple formula $(n+3)2^{n-2}$ but offered no proof. Our goal here is to prove this hypothesis, restated as follows:

Theorem 1.1. *The number of equivalence classes defined by* $\leftrightarrow$ *is* $(n+3)2^{n-2}$.

Note that this relation can also be realized under the notion of a reversible string rewriting system.

2 Proof of Theorem 1.1

The proof is divided into 3 parts. First we give an invariant called $F(\sigma)$ for each $\sigma \in S_n$ that is preserved through each transformation R. The second part provides an algorithm of how to apply R so that each σ arrives at one of the *canonical forms*. These forms are shown to have pairwise distinct $F(\sigma)$ and thus cannot be turned into one another by R; moreover, the result also implies that two strings with the same $F(\sigma)$ are equivalent. The final step enumerates the numbers of these forms by elementary methods.

2.1 The Invariant $F(\sigma)$

We give an invariant preserved by transformation R that perfectly characterizes each equivalence class i.e. one that is different for each class. For any $\sigma \in S_n$, $f_\sigma(i)$ denotes the number of 0's preceding $\sigma(i)$, the i^{th} digit of σ, subtracted by the number of 1's preceding $\sigma(i)$, and define $F(\sigma)$ to be a multiset:

$$F(\sigma) = \{f_\sigma(i) | \sigma(i) = 0\}.$$

For example, in the case $n = 4$, $F(00111001) = \{0, 1, -1, 0\}$. We will show that if $\sigma \xrightarrow{R} \sigma'$ then $F(\sigma) = F(\sigma')$. Suppose that R is applied from i^{th} to j^{th} digit of σ. Since this segment of σ contains the same number of 0's and 1's, for any $k \notin [i, j]$, we have $f_\sigma(k) = f_{\sigma'}(k)$. It is therefore sufficient to consider only the substring τ representing the segment $[i, j]$ that is turned into τ' by R.

We will pair each 0 at position i_0 to some 1 at position i_1 such that $f_\tau(i_0) = f_\tau(i_1) - 1$. Consider the following algorithm applied to a binary string containing the same number of 0's and 1's: From left to right, whenever we encounter two adjacent unequal digits, delete both of them until we reach the end of the string. Then we proceed again and again until there is no digit left. When the starting string is 0011110010, the steps are as follows:

$$0\underline{01}1110010 \Longrightarrow 011\underline{10}010 \Longrightarrow 011\underline{01}0 \Longrightarrow \underline{01}10 \Longrightarrow \underline{10} \Longrightarrow -.$$

This can always be done towards the end since every step leaves equal numbers of 0's and 1's, and there are always 2 adjacent unequal digits if the length is nonzero. Notice that it follows directly from definition that any two digits deleted together satisfy our condition, namely if 0 at position i_0 in the starting string is deleted with 1 at position i_1, we have $f_\tau(i_0) = f_\tau(i_1) - 1$. This pairing is a bijection between digits 0's and 1's. Hence, if $G(\tau) := \{f_\tau(i) | \tau(i) = 1\}$, we have

$$G(\tau) = \{\lambda + 1 | \lambda \in F(\tau)\}.$$

Denote the length of string τ by N_0. If the numbers of 0's and 1's before the digit 1 at position i_1 of τ are A and B, respectively, then the numbers of 0's and 1's after digit i_1 are $\frac{N_0}{2} - A$ and $\frac{N_0}{2} - B - 1$. Therefore, in τ_0, a string resulted from reversing τ, the numbers of 0's and 1's before digit $N_0 + 1 - i_1$ are $\frac{N_0}{2} - A$ and $\frac{N_0}{2} - B - 1$, that is,

$$f_{\tau_0}(N_0 + 1 - i_1) = (\tfrac{N_0}{2} - A) - (\tfrac{N_0}{2} - B - 1) = -(A - B) + 1 = -f_\tau(i_1) + 1.$$

Define $G(\tau_0) := \{f_{\tau_0}(i) |\, \tau_0(i) = 1\}$. We then have

$$G(\tau_0) = \{-\lambda + 1 |\, \lambda \in G(\tau)\} = \{-\lambda | \lambda \in F(\tau)\}.$$

Then, after all zeroes and ones of τ_0 are switched to ones and zeroes, it becomes τ' which has the property $f_{\tau'}(k) = -f_{\tau_0}(k)$. It follows that $F(\tau') = \{-\lambda | \lambda \in G(\tau_0)\} = \{\lambda | \lambda \in F(\tau)\} = F(\tau)$. We conclude $F(\sigma)$ is invariant with respect to transformation R.

2.2 The Algorithm

From now on, let a, b represent substrings 01 and 10, respectively. Each $\sigma \in S_n$ can be abbreviated by turning two unequal contiguous digits into a and b sequentially until there are no two contiguous 01 and 10 left in this representative string. We call this new string of 4 letters a, b, 0, 1 a *reduced form* of σ. For example, both $0ab1$ and $0a1a$ are reduced forms of 001101.

$$001101 \Longrightarrow 0a101 \Longrightarrow 0ab1 \text{ or } \; 0a1a$$

Note that $0b$ and $a0$ abbreviate the same string, and similarly for $1a$ and $b1$.

When R is applied to a reduced form, we can easily see that R really performs as normal; we only need to treat a and b as new letters to be included in the reversal step. This is because a becomes b after the first step but then switches back to a again after reversal, and analogously for b.

If we remove all a's and b's of a reduced form π, the resulting string will be binary. We call the number of adjacent pairs of unequal digits (01 and 10) in this string, *the number of transitions of* π. It will be shown next that every member of S_n has an equivalent reduced form with at most 2 transitions. Let π be a reduced form of σ and suppose that π has more than 2 transitions.

When all a's and b's in π are removed, the resulting string must have a substring which provides 3 transitions i.e. it looks like one of the following:

$$01 \ldots 10 \ldots 01 \quad \text{or} \quad 10 \ldots 01 \ldots 10.$$

We only consider the first case as it proceeds analogously in the second. This means π has a substring in the following form:

$$0\underline{\quad\;}1 \ldots 10 \ldots 0\underline{\quad\;}1,$$

where both intervals can only be filled by a's or b's. The substring between these intervals, π_0, can also contain a's and b's, but they are irrelevant to our consideration here.

Let A_1 and A_0 be the numbers of ones and zeros in π_0. If $A_1 \leq A_0$, there will be a substring of π_0 beginning at the leftmost 1 that contains the equal numbers of ones and zeros. We can now apply R to the stretch comprising this substring and the interval filled by a, b before it.

$$0\boxed{\_\_\_1\ldots 10\ldots 0}\ldots 0\_\_\_1 \xrightarrow{R} 0\boxed{1\ldots 10\ldots 0\_\_\_}\ldots 0\_\_\_1.$$

The result is that all a's and b's in the interval is flung into π_0, creating a new pair of adjacent unequal digits 01, which is abbreviated into a. This gives a reduced form equivalent to π but with length decreased by 1.

Similarly, in the case $A_1 \geq A_0$ there will be a substring of π_0 ending at the rightmost zero that contains ones and zeros equally. R can then be applied to the stretch comprising this substring and the interval filled by a's and b's after it, creating adjacent pair 01, which is abbreviated into a. This also gives a reduced form equivalent to π with length decreased by 1.

The above procedure can always be performed as long as the reduced form's number of transitions stays above 2. Since the length of a reduced form is finite, the process must terminate, at which point the number of transitions must have dropped to at most 2, as desired.

Let us consider possible transformations to a reduced form π^* with at most 2 transitions. First, we note that R permits a substring consisting entirely of a's and b's to be permuted in any way since any adjacent a and b can be swapped. Three possible numbers of transitions are considered separately.

Case 1 : π^* has zero transition.

Thus, it consists entirely of a's and b's. We permute π^* so that all a's come before the first b.

Case 2 : π^* has one transition.

In this case, when all of a's and b's of π^* are removed, it becomes k' 0's followed by k' 1's, or vice versa, where $1 \leq k' \leq n-1$. We consider only the first case, i.e. when π^* is of the following form:

$$\_\overbrace{0\_0\_0\ldots\_0}^{k' \text{ 0's}}\_\overbrace{1\_1\_1\ldots\_1}^{k' \text{ 1's}}\_,$$

where the intervals are to be filled by a's and b's.

By applying a single R in the location shown below, all a's and b's in each of the first k' intervals can be sent toward the last $k'+1$ intervals adjacent to 1's by a single R:

$$\ldots\boxed{\_0\_0\_0\ldots\_0\_1\_1\_1\ldots\_1}\_\ldots$$

We perform such R at most k' times until all the first k' intervals become empty. Then, except for those in the last interval, every b is associated with a digit 1 immediately to the right of its interval. We can make each of these b's come into contact with its associated 1, in succession, in order to apply rewriting rule $b1 \Rightarrow 1a$, and thereby making all b's outside the last interval disappeared. If π^* has not become a reduced form with zero transition (every 01 we encounter will be turned to a), it will be of the form

$$00\ldots0\_1\_1\_1\ldots\_1\_\ ,$$

where the first k' intervals are filled by a's and the last is filled by a's and b's; moreover at least one a must be between the last 0 and the first 1 to accommodate the transition. The other case may be treated analogously, so that the resulting string is of the form

$$11\ldots1\_0\_0\_0\ldots\_0\_\ ,$$

where the last k' intervals are filled by a's and the first interval is filled by a's and b's; moreover at least one b must be in the first interval.

<u>Case 3</u> : π^* has two transitions.

If π^* is not of the following form, we may simply apply R to the whole string so that it becomes

$$\_\overbrace{0\_0\_0\ldots\_0}^{k'\ \text{0's}}\_\overbrace{1\_1\_1\ldots\_1}^{k'+l'\ \text{1's}}\_\overbrace{0\_0\_0\ldots\_0}^{l'\ \text{0's}}\_.$$

By applying a single R in the location shown below, all a or b in each of the first k' intervals and the last l' intervals can be sent toward $k'+l'+1$ intervals next to 1's.

$$\ldots\boxed{\_0\_0\_0\ldots\_0\_1\_1\_1\ldots\_1}\_\ldots \text{ or } \ldots\boxed{1\_1\_1\_\ldots1\_0\_0\_0\_\ldots0\_}\ldots$$

We perform such R at most $k+l$ times until all those intervals become empty i.e. the result is of the form

$$0\ldots0\_1\_1\_1\ldots\_1\_0\ldots0.$$

Similar to Case 2, the string can be further transformed into a reduced form that b only exists in the last interval. Furthermore, in the case that no b exists in the last interval, one can replace $a0$ by $0b$ until a new pair of adjacent digits 10 appears, which can then be abbreviated as b, reducing the string's length. We only consider the final result when this can no longer be done. Hence, if the reduced form still has 2 transitions, at least one a must be in the first interval to accommodate the transition, and at least one b must be in the last interval.

What we have done amounts to showing that every $\sigma \in S_n$ has an equivalent reduced form in one of the following three types. It will be straightforward to calculate $F(\sigma)$ for each form.

Type I : Reduced forms with zero transition

$$\overbrace{aa\ldots a}^{k\geq 0}\overbrace{bb\ldots b}^{n-k}$$

Let T_1 be the set of reduced forms in Type I. If $\sigma \in T_1$, $F(\sigma)$ is

$$\{\underbrace{0,\ldots,0}_{k},\underbrace{-1,\ldots,-1}_{n-k}\}.$$

Type II: Reduced forms with one transition

$$\overbrace{00\ldots 0}^{k\geq 1}a\ \overbrace{\_\_\_}^{\text{only } a}\ 1\ \overbrace{\_\_\_}^{\text{only } a}\ 1\ldots\ \overbrace{\_\_\_}^{\text{only } a}\ 1\ \overbrace{\_\_\_}^{a \text{ or } b}$$

or

$$\overbrace{11\ldots 1}^{k\geq 1}b\ \overbrace{\_\_\_}^{a \text{ or } b}\ 0\ \overbrace{\_\_\_}^{\text{only } a}\ 0\ldots\ \overbrace{\_\_\_}^{\text{only } a}\ 0\ \overbrace{\_\_\_}^{\text{only } a}$$

Let T_2 be the set of reduced forms in Type II. If $\sigma \in T_2$, for $1 \leq i \leq k+1$, denote the number of a's in the i^{th} interval by u_i and the number of b's by u_{k+2}. By simple calculation, we find that in the first case, $F(\sigma)$ is

$$\{\underbrace{k,\ldots,k}_{u_1},\underbrace{k-1,\ldots,k-1}_{u_2+1},\ldots,\underbrace{1,\ldots,1}_{u_k+1},\underbrace{0,\ldots,0}_{u_{k+1}+1},\underbrace{-1,\ldots,-1}_{u_{k+2}}\}.$$

and in the second case, $F(\sigma)$ is

$$\{\underbrace{-k-1,\ldots,-k-1}_{u_{k+2}},\underbrace{-k,\ldots,-k}_{u_1+1},\underbrace{-k+1,\ldots,-k+1}_{u_2+1},\ldots,\underbrace{-1,\ldots,-1}_{u_k+1},\underbrace{0,\ldots,0}_{u_{k+1}}\}.$$

Type III : Reduced forms with two transitions

$$\overbrace{00\ldots0}^{k\geq1}a\overbrace{\underline{\quad}}^{\text{only } a}1\overbrace{\underline{\quad}}^{\text{only } a}1\ldots\overbrace{\underline{\quad}}^{\text{only } a}1b\overbrace{\underline{\quad}}^{a \text{ or } b}\overbrace{00\ldots0}^{l\geq1}$$

Let T_3 be the set of reduced forms in Type III. If $\sigma \in T_3$, for $1 \leq i \leq k+l+1$, denote the number of a's in the i^{th} interval by v_i and denote the number of b's by v_{k+l+2}. By simple calculation, $F(\sigma)$ can be shown to be

$$\{\underbrace{k,\ldots,k}_{v_1},\underbrace{k-1,\ldots,k-1}_{v_2+1},\underbrace{k-2,\ldots,k-2}_{v_3+1},\ldots,\underbrace{-l,\ldots,-l}_{v_{k+l+1}+1},\underbrace{-l-1,\ldots,-l-1}_{v_{k+l+2}}\}.$$

These are the *canonical forms* that each $\sigma \in S_n$ can be turned into. Hence, any two strings with the same $F(\sigma)$ can be turned toward one another. Furthermore, notice that all the possible $F(\sigma)$ are pairwise distinct, implying that two distinct forms are not equivalent. The number of equivalence classes is indeed the number of the forms in these types.

2.3 Enumeration

The number of forms in each type is counted by elementary techniques. For Type I, since the number of a's can range from 0 to n, $|T_1| = n+1$. In Type II, for each k, there are $n-k-1$ a's and b's left to be filled. We have $u_i \geq 0$ for each $1 \leq i \leq k+2$ and $\sum_{i=1}^{k+2} u_i = n-k-1$. There are $\binom{n}{k+1}$ solutions under this condition; hence,

$$|T_2| = 2\sum_{k=1}^{n-1}\binom{n}{k+1} = 2(2^n - n - 1).$$

Lastly, for Type III, and for each k, l, we have $v_i \geq 0$ for each $1 \leq i \leq k+2$ and $\sum_{i=1}^{k+l+2} v_i = n-(k+l+2)$. The number of solutions is therefore $\binom{n-1}{k+l+1}$. Thus,

$$|T_3| = \sum_{k,l\geq1}\binom{n-1}{k+l+1} = \sum_{c=2}^{n-2}(c-1)\binom{n-1}{c+1} = 2^{n-2}(n-1) - 2^n + n + 1.$$

Since $|T_1| + |T_2| + |T_3| = (n+3)2^{n-2}$, Theorem 1.1 is hereby proven.

3 Conclusion

We show that the number of equivalent classes determining equally likely outcomes after riffle shuffles on an alternating is $(n+3)2^{n-2}$, where n is the number of cards of each type. In doing so, we have essentially solved the word problem for this specific reversible string rewriting system by: (i) finding an invariant preserved by the rules of transformation, and (ii) providing canonical forms with pairwise distinct invariants.

References

[1] D. Bayer and P. Diaconis. Trailing the dovetail shuffle to its lair. *The Annals of Applied Probability 2*, 2:294–313, 1992.

[2] M. Conger and D. Viswanath. Riffle shuffles of decks with repeated cards. *The Annals of Probability*, 34:804–819, 2006.

[3] P. Diaconis. *Group Representations in Probability and Statistics*. Inst of Mathematical Statistic, 1998.

Suphakrit Kotchasarn
Kamnoetvidya Science Academy, Rayong
Email: `suphakrit_k@kvis.ac.th`

Phunnawat Khampheng
Kamnoetvidya Science Academy, Rayong
Email: `phunnawat_k@kvis.ac.th`

Thanaporn Thanodomdech
Kamnoetvidya Science Academy, Rayong
Email: `thanaporn.t@kvis.ac.th`

Guntaphon Tassanasophon
Kamnoetvidya Science Academy, Rayong
Email: `guntaphon.t@kvis.ac.th`