

NECKLACE PROCESSES VIA PÓLYA URNS

TOSHIO NAKATA,* *Fukuoka University of Education*

Abstract

Mallows and Shepp (2008) developed the following *necklace processes*. Start with a necklace consisting of one white bead and one black bead, and insert, one at a time, under a deterministic rule, a white bead or a black bead between a randomly chosen adjacent pair. They studied the statistical properties of the number of white beads by investigating the nature of the moments and the expected number of *gaps* of given length between white beads. In this note we study the number of white beads via *Pólya urns* and give a classification of necklace processes for some general rules. Additionally, we discuss the number of *runs*, i.e. the number of consecutive same color beads, instead of the number of gaps.

Keywords: Necklace process; Pólya urn; law of large numbers; central limit theorem

2000 Mathematics Subject Classification: Primary 60C05
Secondary 60F05

1. Introduction

The necklace process was proposed by Mallows and Shepp [9] in the following way.

- (N1) Start with a necklace consisting of one white bead and one black bead.
- (N2) Add one new bead at a time between a randomly chosen adjacent pair of old beads with probability $1/n$, when there are n beads. The color of the newly inserted bead is decided by the following deterministic rule.
 - (R) If both colors of the chosen pair of adjacent beads are black then the color of the inserted bead is white. Otherwise, it is black.

Let W_n be the number of white beads when the total number of beads is n . Mallows and Shepp [9] gave the exact expectation of W_n and the exact variance of W_n . They also showed a central limit theorem with respect to W_n . Moreover, they derived the expected number of gaps of given length between white beads in the necklace. The results concerning the number of white beads are as follows.

Theorem 1. ([9].) *For the necklace process, the exact expectation of W_n and the exact variance of W_n are respectively given by*

$$E(W_n) = \frac{n}{3} \quad \text{and} \quad \text{var} W_n = \frac{2n}{45}. \quad (1)$$

Received 21 October 2008; revision received 11 December 2008.

* Postal address: Department of Mathematics, Fukuoka University of Education, Akama-Bunkyo-machi, Munakata, Fukuoka, 811-4192, Japan. Email address: nakata@fukuoka-edu.ac.jp

Dedicated to Professor Izumi Kubo on the occasion of his 70th birthday.

Moreover, the following strong law of large numbers and central limit theorem hold, respectively: as $n \rightarrow \infty$,

$$\frac{W_n}{n} \rightarrow \frac{1}{3} \text{ almost surely,} \tag{2}$$

$$\frac{W_n - n/3}{\sqrt{2n/45}} \rightarrow N(0, 1^2) \text{ in distribution,} \tag{3}$$

where $N(\mu, \sigma^2)$ denotes a normally distributed random variable with expectation μ and variance σ^2 .

Although (2) was not directly described in [9], we give another proof of Theorem 1 including (2) in a later section. Investigating the properties of the moments and cumulants of W_n , they obtained (3). They tried to determine the exact moments; however, it seemed to be difficult. By way of experiment, let us try to give exact forms of the distribution of W_n instead of the moments using a standard technique. Namely, the probability generating function of W_n is defined by

$$G_n(s) = E(s^{W_n}) = \sum_{l=1}^n P(W_n = l) s^l,$$

which satisfies

$$G_{n+1}(s) = \begin{cases} \frac{2s(1-s)}{n} G'_n(s) + sG_n(s) & \text{if } n \geq 3, \\ s & \text{if } n = 2, \end{cases} \tag{4}$$

because $W_2 = 1, W_3 = 1$, and, for $n \geq 3$,

$$\begin{aligned} P(W_{n+1} = W_n \mid W_n) &= \frac{2W_n}{n}, \\ P(W_{n+1} = W_n + 1 \mid W_n) &= 1 - \frac{2W_n}{n}. \end{aligned} \tag{5}$$

Note that (5) was given in [9, p. 272]. Since $G_n(s)$ is polynomial for each n , we can easily evaluate (4) using some symbolic computation software. For small n , we see that

$$\begin{aligned} G_4(s) &= \frac{s^2 + 2s}{3}, & G_5(s) &= \frac{2s^2 + s}{3}, & G_6(s) &= \frac{2s^3 + 11s^2 + 2s}{15}, \\ G_7(s) &= \frac{17s^3 + 26s^2 + 2s}{45}, & G_8(s) &= \frac{17s^4 + 180s^3 + 114s^2 + 4s}{315}, \end{aligned}$$

and so on. Now, Brennan and Prodinger [2] obtained closed-form solutions for a case as complicated as this. However, in this case it seems difficult to completely determine closed-form solutions for coefficients of s^k in $G_n(s)$. So we should stop trying to obtain closed-form solutions of the distributions. Hence, we will use another approach to investigate W_n .

In this note we deal with necklace processes via Pólya urns. Namely, some results of Pólya urns are applicable to the necklace process. In particular, (3) is easily obtained without moment calculations. Moreover, we consider some general rules which are not limited to rule (R). Namely, procedures (N1) and (N2) remain, and rule (R) is enhanced in the sense that the color of the newly inserted bead is generally determined by the colors of both beads which are incident with a randomly selected edge. Note that *edge* means a pair of adjacent beads in

the necklace. Hereafter, we use the terminology for simplicity. For each rule, we study the necklace process via Pólya urns.

Additionally, we discuss the number of runs, i.e. the number of consecutive same color beads in the necklace, instead of the number of gaps studied in [9, Theorem 3]. For $1 \leq j \leq n - 1$, let $R_j^W(n)$ and $R_j^B(n)$ be the number of white runs of length j and the number of black runs of length j in a necklace with n beads, respectively. We determine either the distributions of $R_j^W(n)$ and $R_j^B(n)$ or their expectations for simple cases.

The plan of this note is as follows. In Section 2 we classify some rules for constructing necklaces. In Section 3 we shortly review Pólya urns. In Section 4 we apply some results of Pólya urns to the necklace processes for each rule and give some discussion of runs. Finally, in Section 5 we give several concluding remarks.

2. Rules for constructing necklaces

When constructing necklaces, we assume the following rules. Procedures (N1) and (N2) remain, and rule (R) is enhanced in the sense that the color of the newly inserted bead is determined by the colors of both beads which are incident with a randomly selected edge. First, if the selected edge has two beads with different colors, by symmetry we have either (A) or (B):

$$(A): \begin{cases} \circ - \bullet \Rightarrow \circ - \bullet - \bullet \\ \bullet - \circ \Rightarrow \bullet - \bullet - \circ \end{cases} \quad (B): \begin{cases} \circ - \bullet \Rightarrow \circ - \circ - \bullet \\ \bullet - \circ \Rightarrow \bullet - \bullet - \circ \end{cases}$$

Throughout this note, ‘ \circ ’ and ‘ \bullet ’ denote a white bead and a black bead, respectively, and ‘ \Rightarrow ’ represents the transition by the insertion of a bead. After consideration of symmetry and exception of triviality, (A) and (B) are further classified in Table 1.

Since the processes are started with $\circ - \bullet$ or $\bullet - \circ$ by (N1), the edge $\circ - \circ$ does not appear for rules (A1), (A2), and (A3). Accordingly, transition (d) can be ignored for these three rules. Therefore, rule (A3) is trivial; in fact, the newly inserted bead is always black. Note that rules (A1) and (A2) are identified with rule (R), so that they are equivalent to the model of Mallows and Shepp. Hence, we only consider rules (A1), (B1), (B2), and (B3).

TABLE 1.

	(A1)	(A2)	(A3)
(a)	$\bullet - \bullet \Rightarrow \bullet - \circ - \bullet$	$\bullet - \bullet \Rightarrow \bullet - \circ - \bullet$	$\bullet - \bullet \Rightarrow \bullet - \bullet - \bullet$
(b)	$\circ - \bullet \Rightarrow \circ - \bullet - \bullet$	$\circ - \bullet \Rightarrow \circ - \bullet - \bullet$	$\circ - \bullet \Rightarrow \circ - \bullet - \bullet$
(c)	$\bullet - \circ \Rightarrow \bullet - \bullet - \circ$	$\bullet - \circ \Rightarrow \bullet - \bullet - \circ$	$\bullet - \circ \Rightarrow \bullet - \bullet - \circ$
(d)	$\circ - \circ \Rightarrow \circ - \bullet - \circ$	$\circ - \circ \Rightarrow \circ - \circ - \circ$	$\circ - \circ \Rightarrow \circ - \circ - \circ$
	(B1)	(B2)	(B3)
(a)	$\bullet - \bullet \Rightarrow \bullet - \bullet - \bullet$	$\bullet - \bullet \Rightarrow \bullet - \circ - \bullet$	$\bullet - \bullet \Rightarrow \bullet - \bullet - \bullet$
(b)	$\circ - \bullet \Rightarrow \circ - \circ - \bullet$	$\circ - \bullet \Rightarrow \circ - \circ - \bullet$	$\circ - \bullet \Rightarrow \circ - \circ - \bullet$
(c)	$\bullet - \circ \Rightarrow \bullet - \bullet - \circ$	$\bullet - \circ \Rightarrow \bullet - \bullet - \circ$	$\bullet - \circ \Rightarrow \bullet - \bullet - \circ$
(d)	$\circ - \circ \Rightarrow \circ - \circ - \circ$	$\circ - \circ \Rightarrow \circ - \bullet - \circ$	$\circ - \circ \Rightarrow \circ - \bullet - \circ$

3. A short review of Pólya urns

In this section we give a short review of Pólya urns to apply to necklace processes. The original model was proposed by Eggenberger and Pólya [4]. There is a large amount of literature concerning generalized Pólya urns, including recent results (see [6] and [8]). These results are especially important in the field of informatics and bioscience (see [8, Chapters 8 and 9]).

Consider an urn containing w_0 white balls and b_0 black balls. The urn evolves according to the following procedure. A ball is randomly chosen from the urn. The color of the ball is inspected as follows.

- If the color of the ball is black, a black balls and b white balls are placed into the urn.
- If the color of the ball is white, c black balls and d white balls are placed into the urn.

Note that the chosen ball is not removed. The above procedure is specified by the matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \tag{6}$$

which is called the *ball replacement matrix*. Some models are named according to special parameters of the matrix (see Table 2).

Throughout this section, let w_n be the number of white balls after n draws. Then some limit theorems concerning w_n/n are known for these models. We now refer to the following well-known basic result.

Lemma 1. ([4].) *Consider an urn containing one white ball and one black ball with ball replacement matrix*

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Then

$$P(w_n = k) = \frac{1}{n-1} \text{ for } 1 \leq k \leq n-1.$$

Moreover, as $n \rightarrow \infty$, $w_n/n \rightarrow X$ almost surely, where X is a random variable whose distribution is uniform on $[0, 1]$.

TABLE 2.

The Pólya Eggenberger urn [4], [8, Section 3.2]	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$
The Friedman urn [8, Section 3.3], [6, Section 2.2]	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
The coupon collector urn [6, Section 2.4]	$\begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix}$
The Ehrenfest urn [8, Section 3.5]	$\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$
The OK corral urn [6, Section 8]	$\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$
The Mabinogion urn [11, Section 15.3], [5]	$\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$

Lemma 1 is a simplification of [10, Theorem 2.1]. Under the condition of Lemma 1, we can obtain the closed-form expression of $P(w_n = k)$. However, in general, it is difficult to obtain the exact form of the distributions or the moments.

Moreover, we also refer to a strong law of large numbers and a central limit theorem for w_n .

Lemma 2. ([1].) *Consider an urn containing b_0 black balls and w_0 white balls with ball replacement matrix (6). Moreover, assume that $a + b = c + d =: s \geq 1$, $w_0 + b_0 \geq 1$, $w_0, b_0 \geq 0$, $a \geq 0$, $b, c > 0$, $a - c \leq s/2$, and $a \neq c$, and if $d < 0$ then d divides both b and b_0 . Then, as $n \rightarrow \infty$,*

$$\frac{w_n}{n} \rightarrow \frac{c}{b+c} \text{ almost surely,}$$

$$\frac{w_n - E(w_n)}{\sqrt{n}} \rightarrow N\left(0, \frac{bc}{(b+c)^2} \frac{(s-b-c)^2}{2b+2c-s}\right) \text{ in distribution.}$$

The above statement is a simplification of [3, Lemma 2] or [8, Theorem 3.5].

4. Necklace processes via Pólya urns

4.1. Rule (A1)

Remember that rule (A1) is identified with rule (R), so that it is equivalent to the model of Mallows and Shepp [9]. To determine the total number of white beads, it is sufficient to investigate the number of $\circ - \bullet$, $\bullet - \circ$, and $\bullet - \bullet$ edges in the necklace. Hence, we should investigate the change in the number of these edges at each step. Since the transitions of $\circ - \bullet$ and $\bullet - \circ$ are the same for rule (A1), we do not distinguish between $\circ - \bullet$ and $\bullet - \circ$. Hence, $\circ - \bullet$ is adopted for convenience. We argue that the necklace process of rule (A1) is regarded as a Pólya urn. In fact, a $\circ - \bullet$ edge in the necklace is identified with a white ball and a $\bullet - \bullet$ edge is identified with a black ball. For convenience, we call it the *Pólya urn induced by edges*. Throughout this subsection, let w_n be the number of edges of $\circ - \bullet$ in a necklace with n beads. Similarly, we define b_n as the number of edges of $\bullet - \bullet$ in a necklace with n beads.

Claim 1. *For rule (A1), we have $(w_2, b_2) = (2, 0)$ and the ball replacement matrix of the Pólya urn induced by edges is*

$$\begin{pmatrix} 0 & 1 \\ 2 & -1 \end{pmatrix}. \tag{7}$$

Proof. By the definition of w_n, b_n , and (N1), we have $(w_2, b_2) = (2, 0)$. Suppose that rule (A1) holds. If the chosen edge is $\circ - \bullet$ for (N2) then we have $\circ - \bullet - \bullet$. Namely, the number of edges of $\bullet - \bullet$ increases by 1 and the number of edges of $\circ - \bullet$ remains the same. If the chosen edge is $\bullet - \bullet$ for (N2) then we have $\bullet - \circ - \bullet$. Namely, the number of edges of $\circ - \bullet$ increases by 2 and the number of edges of $\bullet - \bullet$ decreases by 1. Hence, we have (7).

Moreover, by the definition of w_n , the number of white beads W_n satisfies

$$w_n = 2W_n. \tag{8}$$

Using Claim 1 and (8), we give another proof of Theorem 1.

Proof of Theorem 1. We first check (1). Considering the derivatives of both sides of (4), we have

$$G'_3(1) = 1, \quad G'_{n+1}(1) = \left(1 - \frac{2}{n}\right)G'_n(1) + 1 \quad \text{for } n \geq 3.$$

Therefore, $G'_n(1) = E(W_n) = n/3$ is the unique solution. Similarly, considering the second derivatives of both sides of (4), we have

$$G''_5(1) = \frac{4}{3}, \quad G''_{n+1}(1) = \left(1 - \frac{4}{n}\right)G''_n(1) + \frac{2}{3}n - \frac{4}{3} \quad \text{for } n \geq 5.$$

Therefore, $G''_n(1) = E(W_n(W_n - 1)) = n(5n - 13)/45$ is the unique solution. Hence, we also have $\text{var}W_n$.

Next we check (2) and (3). Set $a = 0, b = 1, c = 2, d = -1, w_2 = 2$, and $b_2 = 0$ in Lemma 2. Therefore, in this case, as $n \rightarrow \infty$,

$$\frac{w_n}{n} \rightarrow \frac{2}{3} \quad \text{almost surely,} \quad \frac{w_n - E(w_n)}{\sqrt{n}} \rightarrow N\left(0, \frac{8}{45}\right) \quad \text{in distribution.}$$

Hence, we have (2) and (3) because of (8).

Remark 1. The random variable W_n under rule (A) is equivalent to the number of leaves of a random binary search tree with n nodes (see [3, Theorem 2]). Actually, matrix (7) also appears in [3, p. 313], so these two models are indistinguishable.

Moreover, we investigate $R_j^W(n)$, the number of white runs of length j in a necklace with n beads, and $R_j^B(n)$, the number of black runs of length j in a necklace with n beads. By the definition of rule (A1), no two white beads are adjacent. Hence,

$$R_j^W(n) = \begin{cases} W_n & \text{if } j = 1, \\ 0 & \text{if } 2 \leq j \leq n - 1. \end{cases}$$

We directly apply the result of gaps studied in [9, Theorem 3] to this case.

Proposition 1. ([9, Theorem 3].) *For $n \geq 2$ and $1 \leq j \leq n - 1$, we have*

$$E(R_j^B(n)) = \begin{cases} \frac{j(j+3)2^{j+1}}{(j+4)!}n & \text{if } 1 \leq j \leq n - 5, \\ \frac{j2^{j+1}}{(j+2)!} & \text{if } j = n - 3, n - 4, \\ 0 & \text{if } j = n - 2, \\ \frac{2^{n-2}}{(n-1)!} & \text{if } j = n - 1. \end{cases}$$

Let j_{longest} be the length of the longest run in a necklace with n beads. Then, by [9, Section 7] and Proposition 1,

$$j_{\text{longest}} \sim \frac{\log n}{\log \log n}, \tag{9}$$

which is obtained by $E(R_{j_{\text{longest}}}^B(n)) \sim 1$.

4.2. Rule (B1)

Suppose that rule (B1) holds. We then have the observation that the color of the newly inserted bead is the same as its clockwise adjacent neighbor. When an edge is randomly chosen with respect to (N2), we should consider that the former bead of the edge in a clockwise direction is chosen instead of the edge. Hence, the necklace process of rule (B1) is regarded as a Pólya urn. For convenience, we call it the *Pólya urn induced by beads*.

Claim 2. For rule (B1), we have $(B_2, W_2) = (1, 1)$ and the ball replacement matrix of the Pólya urn induced by beads is

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

By the definition of rule (B1), Claim 2 is easily proved. By Claim 2, the necklace processes are the same as the Pólya–Eggenberger urn processes. Hence, we have the following statement, which corresponds to Lemma 1.

Proposition 2. For rule (B1), we have

$$P(W_n = k) = \frac{1}{n - 1} \quad \text{for } 1 \leq k \leq n - 1. \tag{10}$$

Moreover, as $n \rightarrow \infty$,

$$\frac{W_n}{n} \rightarrow X \quad \text{almost surely,}$$

where X is a random variable whose distribution is uniform on $[0, 1]$.

For the number of runs, we have the following corollary.

Corollary 1. For rule (B1), we have

$$P(R_j^W(n) = 1) = P(R_j^B(n) = 1) = \frac{1}{n - 1} \quad \text{for } 1 \leq j \leq n - 1.$$

Proof. By rule (B1), the necklace is constructed using one white run and one black run with probability 1 for each n . Using (10), we have

$$P(R_j^W(n) = 1) = P(W_n = j) = \frac{1}{n - 1} \quad \text{for } 1 \leq j \leq n - 1.$$

By symmetry, we have $P(R_j^W(n) = 1) = P(R_j^B(n) = 1)$.

4.3. Rule (B2)

Suppose that rule (B2) holds. We then have the observation that the color of the newly inserted bead is of opposite color to its clockwise adjacent neighbor. Hence, we deal with this case in the same manner as the rule (B1) case.

Claim 3. For rule (B2), we have $(B_2, W_2) = (1, 1)$ and the ball replacement matrix of the Pólya urn induced by beads is

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

By the definition of rule (B2), Claim 3 is easily proved. By Claim 3, the necklace processes are the same as the Friedman urn processes. Hence, we have the following statement, which corresponds to Lemma 2.

Proposition 3. For rule (B2), we have, as $n \rightarrow \infty$,

$$\frac{W_n}{n} \rightarrow \frac{1}{2} \quad \text{almost surely,} \quad \frac{W_n - n/2}{\sqrt{n}} \rightarrow N\left(0, \frac{1}{12}\right) \quad \text{in distribution.}$$

Now we discuss the number of runs. By definition,

$$\sum_{j=1}^{n-1} j\{R_j^W(n) + R_j^B(n)\} = n. \tag{11}$$

By the symmetry of rule (B2), $R_j^W(n)$ and $R_j^B(n)$ have the same distribution. For simplicity, we write $E(R_j(n))$ instead of $E(R_j^W(n))$ or $E(R_j^B(n))$. Considering the expectation of both sides of (11), we have

$$\sum_{j=1}^{n-1} j E(R_j(n)) = \frac{n}{2}. \tag{12}$$

We have the following result, which corresponds to Proposition 1.

Proposition 4. *For $n \geq 2$, we have*

$$E(R_j(n)) = \begin{cases} \frac{j^2 + 3j + 1}{(j + 3)!} n & \text{if } 1 \leq j \leq n - 3, \\ \frac{n - 2}{(n - 1)!} & \text{if } j = n - 2, \\ \frac{1}{(n - 1)!} & \text{if } j = n - 1. \end{cases} \tag{13}$$

By Proposition 4, the asymptotic behavior of the lengths of the longest runs is also (9) by the same argument as in [9, Section 7].

Proof of Proposition 4. We first prove that, for $n \geq 2$ and $1 \leq j \leq n$,

$$E(R_j(n+1)) = \begin{cases} \left(1 - \frac{3}{n}\right) E(R_1(n)) + \frac{1}{n} \sum_{i=1}^{n-1} E(R_i(n)) + \frac{1}{2} & \text{if } j = 1, \\ \frac{E(R_{j-1}(n))}{n} + \left(1 - \frac{j}{n}\right) E(R_j(n)) + \frac{2}{n} \sum_{i=j+1}^{n-1} E(R_i(n)) & \text{if } 2 \leq j \leq n - 2, \\ \frac{E(R_{n-2}(n)) + E(R_{n-1}(n))}{n} & \text{if } j = n - 1, \\ \frac{E(R_{n-1}(n))}{n} & \text{if } j = n. \end{cases} \tag{14}$$

By rule (B2), we have, for $1 \leq j \leq n$,

$$P(R_j(n) - 1 \leq R_j(n + 1) \leq R_j(n) + 2) = 1.$$

Consider $j = 1$. Then by the same technique as in [9, Section 10] we have, for $n \geq 4$,

$$\begin{aligned} P(R_1(n + 1) = R_1(n) - 1 \mid \mathcal{F}_n) &= \frac{R_1(n)}{n}, \\ P(R_1(n + 1) = R_1(n) + 1 \mid \mathcal{F}_n) &= \frac{2 \sum_{i=3}^{n-1} R_i(n) + \sum_{i=2}^{n-1} (i - 1) R_i(n)}{n}, \\ P(R_1(n + 1) = R_1(n) + 2 \mid \mathcal{F}_n) &= \frac{R_2(n)}{n}, \end{aligned}$$

where \mathcal{F}_n is the minimum σ -field generated by $\{R_j(n)\}_{j=1}^{n-1}$. Hence,

$$\begin{aligned} E(R_1(n+1)) &= E(R_1(n)) + \frac{1}{n} \left(-E(R_1(n)) + 2 \sum_{i=2}^{n-1} E(R_i(n)) + \sum_{i=2}^{n-1} (i-1) E(R_i(n)) \right) \\ &= \left(1 - \frac{3}{n} \right) E(R_1(n)) + \frac{1}{n} \sum_{i=1}^{n-1} (i+1) E(R_i(n)) \\ &= \left(1 - \frac{3}{n} \right) E(R_1(n)) + \frac{1}{n} \sum_{i=1}^{n-1} E(R_i(n)) + \frac{1}{2}. \end{aligned}$$

The last equality holds because of (12). Similarly, we have, for $2 \leq j \leq n-3$,

$$\begin{aligned} P(R_j(n+1) = R_j(n) - 1 \mid \mathcal{F}_n) &= \frac{jR_j(n)}{n}, \\ P(R_j(n+1) = R_j(n) + 1 \mid \mathcal{F}_n) &= \frac{1}{n} \left(R_{j-1}(n) + 2 \left(\sum_{i=j+1}^{n-1} R_i(n) - R_{2j}(n) \right) \right), \\ P(R_j(n+1) = R_j(n) + 2 \mid \mathcal{F}_n) &= \frac{R_{2j}(n)}{n}. \end{aligned}$$

Hence,

$$E(R_j(n+1)) = E(R_j(n)) + \frac{1}{n} \left(E(R_{j-1}(n)) - j E(R_j(n)) + 2 \sum_{i=j+1}^{n-1} E(R_i(n)) \right).$$

The other two cases, $j = n-1$ and $j = n$, can be treated similarly. Therefore, (14) holds. By (14) and $E(R_1(2)) = 1$, we have (13) with some calculation. Note that we obtain, for $n \geq 4$,

$$\sum_{i=1}^{n-1} E(R_i(n)) = \frac{n}{3},$$

while solving the difference equation (14).

4.4. Rule (B3)

The investigation under rule (B3) is more difficult than the previous cases. Hence, we only give some simple estimates. Observing rule (B3), we may think that the number of black beads is intuitively much larger than the number of white beads for sufficiently large n . Actually, this is true, and we have the following proposition.

Proposition 5. *For rule (B3), we have, as $n \rightarrow \infty$,*

$$\frac{W_n}{n} \rightarrow 0 \quad \text{in probability,}$$

which is equivalent to $B_n/n \rightarrow 1$ in probability as $n \rightarrow \infty$.

To prove this statement, we introduce some additional terminology and lemmas.

For $n \geq 2$, let a_n, b_n, c_n , and d_n be the number of $\bullet - \bullet, \circ - \bullet, \bullet - \circ$, and $\circ - \circ$ edges in a necklace with n beads, respectively. We abuse the notation slightly by using the same symbol b_n as used in Section 3; however, this should cause no confusion. Define $\alpha_n, \beta_n, \gamma_n$, and δ_n as

$$E(a_n) = \alpha_n, \quad E(b_n) = \beta_n, \quad E(c_n) = \gamma_n, \quad E(d_n) = \delta_n.$$

Lemma 3. *For rule (B3), we have $\beta_n = \gamma_n$ and*

$$E(W_n) = \beta_n + \delta_n, \quad E(B_n) = \alpha_n + \beta_n, \quad \text{for } n \geq 2. \tag{15}$$

Moreover,

$$\begin{pmatrix} \alpha_2 \\ \beta_2 \\ \delta_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} \alpha_{n+1} \\ \beta_{n+1} \\ \delta_{n+1} \end{pmatrix} = \begin{pmatrix} 1 + 1/n & 1/n & 0 \\ 0 & 1 & 1/n \\ 0 & 1/n & 1 - 1/n \end{pmatrix} \begin{pmatrix} \alpha_n \\ \beta_n \\ \delta_n \end{pmatrix} \quad \text{for } n \geq 2. \tag{16}$$

Proof. By the definition of rule (B3), we have

$$(a_2, b_2, c_2, d_2) = (0, 1, 1, 0), \tag{17}$$

$$a_{n+1} - a_n, b_{n+1} - b_n, c_{n+1} - c_n, |d_{n+1} - d_n| \in \{0, 1\} \quad \text{for } n \geq 2,$$

and, for $n \geq 2$,

$$\begin{aligned} P(a_{n+1} = a_n + 1 \mid \mathcal{F}_n) &= \frac{a_n + b_n}{n}, \\ P(b_{n+1} = b_n + 1 \mid \mathcal{F}_n) &= \frac{d_n}{n}, \\ P(c_{n+1} = c_n + 1 \mid \mathcal{F}_n) &= \frac{d_n}{n}, \\ P(d_{n+1} = d_n + 1 \mid \mathcal{F}_n) &= \frac{b_n}{n}, \\ P(d_{n+1} = d_n - 1 \mid \mathcal{F}_n) &= \frac{d_n}{n}, \end{aligned}$$

where \mathcal{F}_n is the minimum σ -field generated by a_n, b_n, c_n , and d_n . By considering expectations we obtain

$$\begin{aligned} P(a_{n+1} = a_n + 1) &= \frac{\alpha_n + \beta_n}{n}, & P(b_{n+1} = b_n + 1) &= \frac{\delta_n}{n}, \\ P(c_{n+1} = c_n + 1) &= \frac{\delta_n}{n}, & P(d_{n+1} = d_n + 1) &= \frac{\beta_n}{n}, & P(d_{n+1} = d_n - 1) &= \frac{\delta_n}{n}. \end{aligned} \tag{18}$$

By (17), the random variable b_n has the same distribution as the random variable c_n . Therefore, we obtain $\beta_n = \gamma_n$. Now, enumerating the white beads and the black beads, we have

$$W_n = \frac{b_n + c_n + 2d_n}{2} \quad \text{and} \quad B_n = \frac{2a_n + b_n + c_n}{2}.$$

Hence, (15) holds.

By (18), we have, for $n \geq 2$,

$$\begin{aligned} \alpha_{n+1} &= (\alpha_n + 1) \frac{\alpha_n + \beta_n}{n} + \alpha_n \left(1 - \frac{\alpha_n + \beta_n}{n}\right) = \left(1 + \frac{1}{n}\right) \alpha_n + \frac{1}{n} \beta_n, \\ \beta_{n+1} &= (\beta_n + 1) \frac{\delta_n}{n} + \beta_n \left(1 - \frac{\delta_n}{n}\right) = \beta_n + \frac{\delta_n}{n}, \\ \delta_{n+1} &= (\delta_n + 1) \frac{\beta_n}{n} + (\delta_n - 1) \frac{\delta_n}{n} + \delta_n \left(1 - \frac{\beta_n + \delta_n}{n}\right) = \frac{\beta_n}{n} + \left(1 - \frac{1}{n}\right) \delta_n. \end{aligned}$$

Hence, (16) holds.

Lemma 4. For $n \geq 2$, we have $\beta_n > \delta_n$.

Proof. We proceed by induction. If $n = 2$, we have $\beta_2 = 1 > 0 = \delta_2$. Assume that $\beta_k > \delta_k$. Then we have

$$\delta_{k+1} = \frac{\beta_k}{k} + \left(1 - \frac{1}{k}\right) \delta_k < \frac{\beta_k}{k} + \left(1 - \frac{1}{k}\right) \beta_k = \beta_k \leq \beta_{k+1}.$$

Note that the first equality holds because of (16).

Lemma 5. We have

$$\lim_{n \rightarrow \infty} \frac{\beta_n}{n} = 0. \tag{19}$$

Proof. Using (16), we have, for $n \geq 2$,

$$n(n + 1)\beta_{n+2} - 2n^2\beta_{n+1} + (n - 1)\beta_n = 0.$$

Namely,

$$\left(1 + \frac{1}{n}\right) \frac{\beta_{n+2}}{n} - 2 \frac{\beta_{n+1}}{n} + \frac{n - 1}{n^2} \frac{\beta_n}{n} = 0.$$

Since $0 < \beta_n \leq n$, there exists $\limsup_{n \rightarrow \infty} \beta_n/n =: A$, say. Then $0 \leq A \leq 1$ and $A - 2A = 0$. Therefore, $A = 0$, which implies (19).

Under these preparations, we prove Proposition 5.

Proof of Proposition 5. For $\varepsilon > 0$, we have

$$0 \leq P\left(\frac{W_n}{n} > \varepsilon\right) \leq \frac{E(W_n)}{\varepsilon n} = \frac{\beta_n + \delta_n}{\varepsilon n} \leq \frac{2\beta_n}{\varepsilon n} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

The equality holds by (15). The last inequality holds by Lemma 4. The convergence holds by Lemma 5.

5. Concluding remarks

In this note we studied necklace processes via Pólya urns. For simplicity, we just gave naive applications among the deeply investigated results of Pólya urns. When detailed properties of the necklace process are desired, we may proceed using some rich results of Pólya urns. For example, applying [7, Propositions 2.1 and 2.2], which are precise results of Lemma 2, we will have functional central limit theorems for necklace processes. On the other hand, with reasonable changes to rule (R), we may also obtain similar results by finding a corresponding

Pólya urn model. For example, consider the rule which states that if both adjacent beads of the chosen edge are black for (N2) then $\circ - \bullet$ is inserted between the edge and $\bullet - \bullet$ is inserted otherwise. Then the ball replacement matrix of the Pólya urn induced by edges is

$$\begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix},$$

by the same approach used in the proof of Claim 1, and we can also apply Lemma 2. Moreover, even if the number of inserted beads is an integer-valued random variable, or the number of colors is greater than 2, we may obtain some limit theorems which correspond to [8, Sections 6.4 and 6.5], respectively. Similarly, we can also consider some complicated models which correspond to some Pólya urn models. In conclusion, necklace processes can be considered to be a graph expression of Pólya urns with rings.

Acknowledgement

The author would like to thank the anonymous referee for helpful suggestions and for several improvements, including the investigations of the number of runs.

References

- [1] BAGCHI, A. AND PAL, A. K. (1985). Asymptotic normality in the generalized Pólya Eggenberger urn model, with an application to computer data structures. *SIAM J. Algebraic Discrete Methods* **6**, 394–405.
- [2] BRENNAN, C. A. C. AND PRODINGER, H. (2003). The pills problem revisited. *Quaest. Math.* **26**, 427–439.
- [3] DEVROYE, L. (1991). Limit laws for local counters in random binary search trees. *Random Structures Algorithms* **2**, 303–315.
- [4] EGGENBERGER, F. AND PÓLYA, G. (1923). Über die Statistik vorketter vorgänge. *Zeit. Angew. Math. Mech.* **3**, 279–289.
- [5] FLAJOLET, P. AND HUILLET, T. (2008). Analytic combinatorics of the Mabinogion urn. *Discrete Math. Theoret. Computer Sci.* **A1**, 549–572.
- [6] FLAJOLET, P., DUMAS, P. AND PUYHAUBERT, V. (2006). Some exactly solvable models of urn process theory. *Discrete Math. Theoret. Computer Sci.* **AG**, 59–118.
- [7] GOUET, R. (1993). Martingale functional central limit theorems for a generalized Pólya urn. *Ann. Prob.* **21**, 1624–1639.
- [8] MAHMOUD, H. (2008). *Pólya Urn Models*. Chapman and Hall/CRC, Boca Raton, FL.
- [9] MALLOWS, C. AND SHEPP, L. (2008). The necklace process. *J. Appl. Prob.* **45**, 271–278.
- [10] PEMANTLE, R. (2007). A survey of random processes with reinforcement. *Prob. Surveys* **4**, 1–79.
- [11] WILLIAMS, D. (1991). *Probability with Martingales*. Cambridge University Press.