Random deletion-right insertion and pattern avoiding permutations

Cyril Banderier, Jean-Luc Baril and Céline Moreira Dos Santos (Université Paris Nord & Université de Bourgogne)

Genomes can be modeled by unsigned permutations of $\{1, 2, \ldots, n\}$ where each gene is assigned to a unique number found once in the genome [6]. In this paper, we focus on the transposition mutation which consists in displacing an interval of the permutation. More precisely, our motivation is to find some combinatorial properties in terms of pattern avoiding permutations of the transposition mutation whenever we displace an interval of length one on its right, which is equivalent to delete one element and to insert it in a position located on its right. This operation will be called a *deletion-right insertion* (*dri* for short). Here is a *dri*-transformation on the permutation $\sigma = 1\ 2\ 3\ 4\ 5\ 6\ 7$:

$$1\ \ 2\ \ 3\ \ 4\ \ 5\ \ 6\ \ 7 \quad \rightsquigarrow \quad 1\ \ 2\ \ 4\ \ 5\ \ 6\ \ 3\ \ 7$$

This operation is a variant of the well-known genome duplication, which consists in copying a part of the original genome inserted into itself, followed by the loss of one copy of each of the duplicated genes. In particular, it is comparable to the whole duplication-random loss model [1, 2]. Although there are many connections between these models, it is surprising that the behavior of their combinatorial properties depends of different parameters: the *dri*-model reveals some links with *left-to-right maximum* statistics, while the whole duplication-random loss model reveals links with *descent* statistics.

In the literature, the *dri*-transformation is also found in the domain of sorting theory. Indeed, it corresponds (modulo a mirror symmetry) to the *insertion-sorting operator* [4] on permutations defined by

$$\sigma_1\sigma_2\ldots\sigma_n \longrightarrow \sigma_1\ldots\sigma_j\sigma_i\sigma_{j+1}\ldots\sigma_{i-1}\sigma_{i+1}\ldots\sigma_n.$$

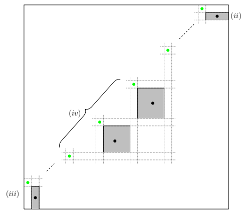
Magnùsson [5] proves that the set of permutations that can be sorted with one step of the insertion-sorting operator is the class of permutations avoiding the three patterns 321, 312 and 2143. We generalize his result by studying the problem of *dri*-model in terms of pattern avoiding permutations.We prove that the set of permutations obtained with this model after a given number of *dri*-operations from the identity is the class of permutations avoiding some patterns. We enumerate them by giving a bivariate exponential generating function, and give asymptotics and the corresponding limit law, via methods of analytic combinatorics [3]. It is a nice surprise that it involves an unusual algebraic exponent, and some unusual closed-form constants.

Theorem 1. *The class $\mathcal{C}(p)$ of permutations obtained from the identity after a given number p of dri-transformations is the class of permutations with at most p non-left-to-right-maxima.*

Corollary 2. *Let σ be a permutation and t_σ be the number of non-left-to-right-maxima. In the dri-model, t_σ transformations are necessary and sufficient to obtain σ from the identity.*

Corollary 3. *Let σ and π be two permutations and $t_{\sigma^{-1}.\pi}$ be the number of non-left-to-right-maxima in $\sigma^{-1}\cdot\pi$. In the dri-model, $t_{\sigma^{-1}.\pi}$ transformations are necessary to obtain π from σ. In particular, if $t_{\sigma^{-1}}$ is the number of non-left-to-right-maxima in σ^{-1}, then $t_{\sigma^{-1}}$ transformations are necessary and sufficient to sort by insertion the permutation σ into the identity.*

Theorem 4. *A permutation $\sigma \in S_n$ belongs to the basis B_p if and only if one has:*
(i) σ contains exactly $p+1$ non-left-to-right-maxima.
(ii) $n-1$ is a non-left-to-right-maxima,
(iii) σ_2 is a non-left-to-right-maxima,
(iv) For any three left-to-right maxima in σ, σ_i, σ_j and σ_k, $1 \le i < j < k \le n$, such that there is no left-to right maximum between σ_i and σ_j and between σ_j and σ_k, there exists a non-left-to-right-maxima σ_t in σ, $j < t < k$, satisfying $\sigma_t > \sigma_i$.



Lemma 5. *Let $\sigma \in S_n$ be a minimal permutation having $p \ge 1$ non-left-to-right-maxima and such that $\sigma_{\ell+1} = n$, $\ell \ge 0$. Let α be the subsequence $\sigma_1\sigma_2\ldots\sigma_\ell$ and π be the permutation in S_ℓ isomorphic to α. Then, π is a minimal permutation with $p-n+\ell+1$ non-left-to-right-maxima.*

Theorem 6. *The number $m_{n,p}$ of minimal permutations of length n having exactly p non-left-to-right-maxima, $p \ge 1$, is given by the following recurrence relation*

$$m_{n,p} = \sum_{\ell=0}^{p-1} (\ell+1)! \cdot \binom{n-2}{\ell} \cdot m_{n-\ell-2,p-\ell-1}$$

with $m_{n,p} = 0$ if $n > 2p$, and $m_{n,n-1} = (n-1)!$ for $n > 1$. Let $V(x,y) = \sum_{n\ge1,p\ge1} m_{n,n-p}\frac{x^n y^p}{n!}$ be the bivariate exponential generating function where the coefficient of $\frac{x^n y^p}{n!}$ is the number $m_{n,n-p}$ of minimal permutations of length n with $n-p$ non-left-to-right-maxima. Then, we have

$$V(x,y) = (\frac{1}{2} - \frac{1}{\sqrt{1+4y}}) \cdot (1-x)^{\frac{1}{2}(1+\sqrt{1+4y})} + (\frac{1}{2} + \frac{1}{\sqrt{1+4y}}) \cdot (1-x)^{\frac{1}{2}(1-\sqrt{1+4y})} - 1\,.$$

Theorem 7. *The mean is $\mu_n = \frac{[x^n]\partial_y V(x,1)}{[x^n]V(x,1)} \sim \frac{\ln n}{\sqrt{5}} + \frac{3}{5\sqrt{5}} - \frac{\Gamma(\frac{1+\sqrt{5}}{2})}{5\Gamma'(\frac{1+\sqrt{5}}{2})} + O(\frac{\ln n}{n})$. The m-th factorial moment is $\frac{[x^n]\partial_y^m V(x,1)}{[x^n]V(x,1)} \sim \frac{\ln(n)^m}{\sqrt{5}^m}$ and the corresponding limit law is Gaussian.*

References

[1] J.-L. Baril and R. Vernay, Whole mirror duplication-random loss model and pattern avoiding permutations, Information Processing Letters, 110(2010), 474-480.

[2] M. Bouvel, E. Pergola, Posets and permutations in the duplication-loss model: Minimal permutations with d descents, Theor. Comput. Sci., 411(26-28)(2010), 2487-2501.

[3] Ph. Flajolet and R. Sedgewick, Analytic Combinatorics, Cambridge U. Press, 2009.

[4] D.E. Knuth, The Art of Computer Programming, Addison-Wesley Publishing, Second Edition, 1998, Volume 3, Sorting and Searching.

[5] H. Magnùsson, Sorting Operators and Their Preimages, Thesis, Reykjavik University, May 2013.

[6] G.A. Watterson, W.J. Ewens, T.E. Hall, and A. Morgan, The Chromosome inversion problem, Journal of Theoretical Biology, 99(1982), 1-7.