LEFT-TO-RIGHT MAXIMA IN
WORDS AND MULTISET PERMUTATIONS

BY

Amy N. Myers

Bryn Mawr College in Bryn Mawr
Philadelphia, PA 19010, USA USA
e-mail: amyers@sju.edu

AND

Herbert S. Wilf

Department of Mathematics, University of Pennsylvania
Philadelphia, PA 19104, USA
e-mail: wilf@math.upenn.edu

ABSTRACT

We extend classical theorems of Rényi by finding the distributions of the numbers of both weak and strong left-to-right maxima (a.k.a. **outstanding elements**) in words over a given alphabet and in permutations of a given multiset.

1. Introduction

Given a sequence $w = w_1 w_2 \ldots w_n$ of members of a totally ordered set, we say j is a **strongly outstanding element** of w if whenever $i < j$ we have $w_i < w_j$. In this case we call w_j a **strongly outstanding value**. We say j is a **weakly outstanding element** if $w_i \leq w_j$ whenever $i < j$, and call w_j a **weakly outstanding value**. In this paper we will explore the contexts in which w is a permutation, a multiset permutation, or a word over some finite alphabet. A famous theorem of Rényi [12] (see also [1]) states that the number of permutations of $[k]$ with r strongly outstanding elements is equal

to the number of such permutations with r cycles, the latter being given by $\begin{bmatrix} k \\ r \end{bmatrix}$, the unsigned Stirling number of the first kind. In this paper we investigate additional properties of the outstanding elements and values of permutations, and extend them to multiset permutations and words on $[k] = \{1, 2, \ldots, k\}$. An interesting sidelight to our results is that we obtain a proof of Gauss's celebrated $_2F_1$ evaluation by comparing two forms of one of our generating functions, in Section 5 below.

2. Summary of results

2.1. MULTISET PERMUTATIONS. For permutations[1] we have the following results.

THEOREM 1: Let $M = \{1^{a_1}, 2^{a_2}, \ldots, k^{a_k}\}$ be a multiset with $N = a_1 + a_2 + \cdots + a_k$, and let $f(M, r)$ denote the number of permutations of M that contain exactly r strongly outstanding elements. Then

(1)
$$F_M(x) \overset{\text{def}}{=} \sum_r f(M, r) x^r$$
$$= \frac{(N-1)! a_k x}{a_1! a_2! \cdots a_k!} \prod_{i=1}^{k-1} \left(1 + \frac{a_i x}{N - (a_1 + a_2 + \cdots + a_i)}\right).$$

This result is also found in [4].

COROLLARY 1: The generating function for the probability $p(M, r)$ that a randomly selected permutation of M has exactly r strongly outstanding elements is

(2) $$P_M(x) \overset{\text{def}}{=} \sum_r p(M, r) x^r = \frac{a_k x}{N} \prod_{i=1}^{k-1} \left(1 + \frac{a_i x}{N - (a_1 + a_2 + \cdots + a_i)}\right),$$

and the average number of strongly outstanding elements among permutations of M is

(3) $$P'_M(1) = \sum_{i=1}^{k} \frac{a_i}{a_i + a_{i+1} + \cdots + a_k}.$$

---

[1] Note that for permutations, all outstanding elements and values are strongly outstanding.

THEOREM 2: *The generating function for the number $g(M,t)$ of permutations of M that contain exactly t weakly outstanding elements is given by*

(4)
$$G_M(x) \stackrel{def}{=} \sum_t g(M,t) = \phi_{N,a_1}(x)\phi_{N-a_1,a_2}(x) \cdots \phi_{N-a_1-\cdots-a_{k-2},a_{k-1}}(x)x^{a_k}$$

where

$$\phi_{N,a}(x) = \sum_{m=0}^{a} \binom{N-m-1}{a-m} x^m.$$

COROLLARY 2: *The average number of weakly outstanding elements among permutations of M is*

(5)
$$\sum_{i=1}^{k} \frac{a_i}{a_{i+1} + a_{i+2} + \cdots + a_k + 1}.$$

COROLLARY 3: *Let $A = \max_i\{a_i\}$. The amount by which the average number of weakly outstanding elements exceeds the average number of strongly outstanding elements is $\leq \frac{\pi^2}{6} A(A-1)$.*

2.2. WORDS. The next theorem involves Stirling numbers of the second kind, denoted by $\{{n \atop m}\}$. This is defined as the number of ways to partition a set of n elements into m nonempty subsets.

THEOREM 3: *The number $f(n,k,r)$ of n-letter words over an alphabet of k letters which have exactly r strongly outstanding elements is given by*

(6)
$$f(n,k,r) = \sum_m \binom{k}{m} \begin{bmatrix} m \\ r \end{bmatrix} \begin{Bmatrix} n \\ m \end{Bmatrix},$$

and the average number of strongly outstanding elements among such words is

$$H_k = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{k} + o(1) \quad (n \to \infty).$$

THEOREM 4: *The generating function for the number $g(n,k,t)$ of n-letter words over an alphabet of k letters, which have exactly t weakly outstanding elements, is given by*

(7) $$G_k(n,x) \stackrel{def}{=} \sum_t g(n,k,t)x^t = \sum_{t=0}^{k-1} (-1)^{k-1-t}(x+t)^n \binom{k-1}{t}\binom{x+t-1}{k-1},$$

and the average number of weakly outstanding elements among these words is

$$\frac{n}{k} + H_{k-1} + O\left(\left(\frac{k-1}{k}\right)^n\right).$$

Next we introduce the notion of a **template** for words on $[k]$. A permutation of 5 letters **matches the template** '$YN*YY$', for example, if 1, 4, and 5 are outstanding elements, 2 is not an outstanding element, and 3 is unconstrained. For example the permutation 21457 matches this template. In this case we think of the letters of the template Y, N, * as representing yes, no, and unconstrained, respectively. We generalize the Y and N constraints to S, W, $\overline{S}$, and $\overline{O}$; where S indicates a strongly outstanding element, W a weakly outstanding element, $\overline{S}$ indicates the absence of a strongly outstanding, and $\overline{O}$ the absence of an outstanding element. We provide an algorithm for producing the generating function for the number of words that match a given template. When the word is a permutation, we have

THEOREM 5: *Let τ be a given template of length at most n, and let τ_j denote the letter that appears in position j, counting from the left, of the template τ. Since every element of a permutation is either strongly outstanding or not outstanding, the letters of τ are chosen from $\{Y, N, *\}$. The probability that a permutation of at least n letters matches the template τ is*

$$(8) \qquad \prod_{j:\tau_j=\text{`N'}}\left(1-\frac{1}{j}\right) \prod_{j:\tau_j=\text{`Y'}}\frac{1}{j}.$$

The corresponding result for words is

THEOREM 6: *Let τ be a word on $\{S, W, *, \overline{S}, \overline{O}\}$. Suppose*

$$F(\tau, x) = \sum_{k\geq 1} f(k, \tau)x^k$$

is the ordinary generating function for $f(k, \tau)$, the number of words over the alphabet $[k]$ that match the template τ. Consider the adjunction of one new symbol, $A \in \{S, W, *, \overline{S}, \overline{O}\}$, at the right end of τ. The generating function, $F(\tau A, x)$ can be obtained from $F(\tau, x)$ by applying an operator Ω_A, i.e.,

$$F(\tau A, x) = \Omega_A F(\tau, x),$$

where

$$\Omega_S F(\tau, x) = xF(\tau, x)/(1 - x), \tag{9}$$

$$\Omega_W F(\tau, x) = F(\tau, x)/(1 - x), \tag{10}$$

$$\Omega_* F(\tau, x) = x\frac{d}{dx}F(\tau, x), \tag{11}$$

$$\Omega_{\overline{S}} F(\tau, x) = \left(x\frac{d}{dx} - \frac{x}{1-x}\right)F(\tau, x), \text{ and} \tag{12}$$

$$\Omega_{\overline{O}} F(\tau, x) = \left(x\frac{d}{dx} - \frac{1}{1-x}\right)F(\tau, x). \tag{13}$$

3. Strongly outstanding elements of multiset permutations

In this section we establish Theorem 1 and its corollaries.

Given a multiset $M = \{1^{a_1}, 2^{a_2}, \ldots, k^{a_k}\}$, let $N = \sum_j a_j$. We construct the permutations of M that have exactly r strongly outstanding elements as follows. We have N slots into which we will put the N elements of M to make these permutations.

Take the a_1 1's that are available and place them in some a_1-subset of the N slots that are available. There are two cases now. If the set of slots that we chose for the 1's did not include the first (leftmost) slot, then we can fill in the remaining slots with any permutation of the multiset $M/1^{a_1}$ that has exactly r strongly outstanding elements. On the other hand, if we did place a 1 into the first slot, then after placing all of the 1's, the remaining slots can be filled in with any permutation of the multiset $M/1^{a_1}$ that has exactly $r - 1$ strongly outstanding elements.

Let $f(M, r)$ denote the number of permutations of the multiset M that have exactly r strongly outstanding elements. The argument in the preceding paragraph shows that

$$f(M, r) = \left(\binom{N}{a_1} - \binom{N-1}{a_1-1}\right)f(M/1^{a_1}, r) + \binom{N-1}{a_1-1}f(M/1^{a_1}, r-1).$$

When we define

$$F_M(x) = \sum_r f(M, r)x^r,$$

we have the recurrence

$$F_M(x) = \left(\left(\binom{N}{a_1} - \binom{N-1}{a_1-1} \right) + \binom{N-1}{a_1-1} x \right) F_{M/1^{a_1}}(x)$$

$$= \left(\binom{N-1}{a_1} + \binom{N-1}{a_1-1} x \right) F_{M/1^{a_1}}(x)$$

This shows that the generating polynomial resolves into linear factors over the integers. Indeed we get the explicit form

$$F_M(x) = \left(\binom{N-1}{a_1} + \binom{N-1}{a_1-1} x \right) \left(\binom{N-a_1-1}{a_2} + \binom{N-a_1-1}{a_2-1} x \right) \cdots$$

$$= \prod_{i=1}^{k} \left(\binom{N-a_1-a_2-\cdots-a_{i-1}-1}{a_i} \right.$$

$$\left. + \binom{N-a_1-a_2-\cdots-a_{i-1}-1}{a_i-1} x \right)$$

$$= \frac{(N-1)! a_k x}{a_1! a_2! \ldots a_k!} \prod_{i=1}^{k-1} \left(1 + \frac{a_i x}{N-(a_1+a_2+\cdots+a_i)} \right)$$

This gives us Theorem 1. Its corollaries follow by obvious calculations. We note that this result was previously obtained by Clément, Flajolet, and Vallée [4] in the context of binary search trees.

Since the generating polynomial has real zeros only, the probabilities $\{p(M,r)\}$ are unimodal and log concave. By Darroch's Theorem [5], the value of r for which $p(M,r)$ is maximum differs from $P'_M(1)$ of (3) above by at most 1.

4. Weakly outstanding elements of multiset permutations

Next we prove Theorem 2.

Consider $g(M,t)$, the number of permutations of $M = \{1^{a_1}, 2^{a_2}, \ldots, k^{a_k}\}$ that have exactly t weakly outstanding elements, and $G_M(x) = \sum_t g(M,t) x^t$. To find $g(M,t)$, suppose the permutation begins with a block of exactly $m \geq 0$ 1's. Since the value that follows the last 1 in the block is not available for a 1, there remain $N-m-1$ slots into which the remaining 1's can be put, in $\binom{N-m-1}{a_1-m}$ ways. Once all of the 1's have been placed, if the remaining permutation of the multiset $M/1^{a_1}$ has exactly $t-m$ weakly outstanding elements then the whole

thing will have t weakly outstanding elements. Hence we have

$$g(M, t) = \sum_{m \geq 0} \binom{N - m - 1}{a_1 - m} g(M/1^{a_1}, t - m),$$

if $M/1^{a_1}$ is nonempty, whereas if only $M = 1^{a_1}$, then $g(M, t) = \delta_{t, a_1}$. If we multiply by x^t and sum on t we find that

$$G_M(x) = \Big\{ \sum_{m \geq 0} \binom{N - m - 1}{a_1 - m} x^m \Big\} G_{M/1^{a_1}}(x),$$

except that if only $M = 1^{a_1}$, then $G_M(x) = x^{a_1}$. So if we write

$$\phi_{N,a}(x) = \sum_{m=0}^{a} \binom{N - m - 1}{a - m} x^m,$$

then we have

(14) $\qquad G_M(x) = \phi_{N,a_1}(x) \phi_{N-a_1,a_2}(x) \cdots \phi_{N-a_1-\cdots-a_{k-2},a_{k-1}}(x) x^{a_k}.$

This gives (4).

The two sums $\phi_{N,a}(1) = \binom{n}{a}$ and $\phi'_{N,a}(1) = \binom{n}{a}/(n-a+1)$ are elementary, and imply that $\phi'_{N,a}(1)/\phi_{N,a}(1) = 1/(n - a + 1)$. Then logarithmic differentiation of (14) and evaluation at $x = 1$ shows that the average number of weakly outstanding elements in permutations of M is

(15) $\qquad\qquad\qquad \displaystyle\sum_{i=1}^{k} \frac{a_i}{a_{i+1} + a_{i+2} + \cdots + a_k + 1}.$

Let $A = \max_i\{a_i\}$. If we compare (15) and (3) we find that the amount by which the average number of weakly outstanding elements exceeds the average number of strongly outstanding elements is

$$\sum_{i=1}^{k} \left(\frac{a_i}{a_{i+1} + a_{i+2} + \cdots + a_k + 1} - \frac{a_i}{a_i + a_{i+1} + \cdots + a_k} \right)$$

$$= \sum_{i=1}^{k} \frac{a_i(a_i - 1)}{(a_{i+1} + a_{i+2} + \cdots + a_k + 1)(a_i + a_{i+1} + \cdots + a_k)}$$

$$\leq A(A - 1) \sum_{i=1}^{k} \frac{1}{(k - i + 1)^2} \leq \frac{\pi^2}{6} A(A - 1).$$

This estimate is best possible when $A = 1$, i.e., when every element occurs just once.

5. Strongly outstanding elements of words

Next we investigate $f(n, k, r)$, the number of n-letter words over an alphabet of k letters that have exactly r strongly outstanding elements. In so doing we will prove Theorem 3.

Note first that the number of strongly outstanding elements of such a word depends only on the permutation of the distinct letters appearing in the word that is achieved by the **first** appearances of each of those letters, because a value j can be strongly outstanding in a word w only if it is the first (i.e., leftmost) occurrence of j in w.

Hence, associated with each n-letter word w over an alphabet of k letters which has exactly r strongly outstanding elements there is a triple $(S, \mathcal{P}, \sigma)$ consisting of

1. a subset $S \subseteq [k]$, which is the set of all of the distinct letters that actually appear in w, and
2. a partition $\mathcal{P}$ of the set $[n]$ into $m = |S|$ classes, namely, the i-th class of $\mathcal{P}$ consists of the set of positions in the word w that contain the i-th letter of S, and
3. a permutation $\sigma \in S_m$, $m = |S|$, with r strongly outstanding elements. This is the sequence of **first appearances** in w of each of the k letters that occur in w.

Conversely, if we are given such a triple $(S, \mathcal{P}, \sigma)$, we uniquely construct an n-letter word w over $[k]$ with exactly r strongly outstanding elements as follows.

First arrange the classes of the partition $\mathcal{P}$ in ascending order of their smallest elements. Then permute the set S according to the permutation σ, yielding a list $\tilde{S}$. In all of the positions of w that are described by the first class of the partition $\mathcal{P}$ (i.e., the class in which the letter '1' lives) we put the first letter of $\tilde{S}$, etc., to obtain the required word w.

Thus the number of words that we are counting is equal to the number of these triples, viz.

$$(16) \qquad f(n, k, r) = \sum_m \binom{k}{m} \begin{bmatrix} m \\ r \end{bmatrix} \begin{Bmatrix} n \\ m \end{Bmatrix}.$$

It is noteworthy that three flavors of "Pascal-triangle-like" numbers occur in this formula.

Let $\rho(n, k)$ denote the average number of strongly outstanding elements among the n-letter words that can be formed from an alphabet of k letters. To find $\rho(n, k)$ for large n, we have first that

$$\sum_r f(n, k, r) = \sum_m \binom{k}{m} \left\{ {n \atop m} \right\} \sum_r \left[{m \atop r} \right] = \sum_m \binom{k}{m} \left\{ {n \atop m} \right\} m!$$

$$\sim \sum_m \binom{k}{m} \left\{ {n \atop m} \right\} m! \sim \sum_m \binom{k}{m} n^m \sim n^k,$$

for large n, where we have used the facts that $\left\{ {n \atop m} \right\} \sim n^m/m!$ and $\sum_r \left[{m \atop r} \right] = m!$. Similarly,

$$\sum_r r f(n, k, r) = \sum_m \binom{k}{m} \left\{ {n \atop m} \right\} \sum_r r \left[{m \atop r} \right] = \sum_m \binom{k}{m} \left\{ {n \atop m} \right\} m! H_m$$

$$\sim \sum_m \binom{k}{m} \left\{ {n \atop m} \right\} m! H_m \sim \sum_m \binom{k}{m} n^m H_m \sim n^k H_k,$$

where we have used the additional fact that $\sum_r r \left[{m \atop r} \right] = m! H_m$. If we divide these last two equations we find that

$$\lim_{n \to \infty} \rho(n, k) = H_k = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{k}.$$

The proof of Theorem 3 is complete.

From (16) we can use the standard generating functions for the two kinds of Stirling numbers to show that

$$\sum_{r,n} f(n, k, r) y^r t^{n-r} = \prod_{j=1}^k \left(1 + \frac{y}{1 - jt} \right).$$

But from (16) we also have

$$\sum_{r,n} f(n, k, r) y^r t^{n-r} = \sum_{\ell, r} \binom{k}{\ell} \left[{\ell \atop r} \right] \left(\frac{y}{t} \right)^r \sum_n \left\{ {n \atop \ell} \right\} t^n$$

$$= \sum_{\ell, r} \binom{k}{\ell} \left[{\ell \atop r} \right] \left(\frac{y}{t} \right)^r \frac{t^\ell}{(1 - t)(1 - 2t) \ldots (1 - \ell t)}$$

$$= \sum_\ell \binom{k}{\ell} \prod_{j=0}^{\ell - 1} \left(\frac{y}{t} + j \right) \frac{t^\ell}{(1 - t)(1 - 2t) \ldots (1 - \ell t)}$$

$$= \sum_\ell \binom{k}{\ell} \prod_{j=1}^\ell \frac{y + (j - 1)t}{1 - jt}.$$

Comparison of these two evaluations shows that we have found and proved the following identity:

$$(17) \qquad \sum_{\ell} \binom{k}{\ell} \prod_{j=1}^{\ell} \frac{y + (j-1)t}{1 - jt} = \prod_{j=1}^{k} \left(1 + \frac{y}{1 - jt} \right).$$

But Gauss had done it earlier, since it is the evaluation of his well known

$$ {}_2F_1 \left[\begin{array}{cc} -k & y/t \\ 1 - 1/t & \end{array} \middle| 1 \right]. $$

6. A calculus of templates

6.1. TEMPLATES AND PERMUTATIONS. In this section we prove Theorems 5 and 6. We begin by establishing (8), which appeared first in [13].

This result is a generalization of a theorem of R. V. Kadison [9], who discovered the case where the template is 'NN$\cdots$NY', and proved it by the sieve method. To make this paper self-contained, we include a proof of (8).

Our proof is by induction on n. Suppose it has been proved that, for all templates τ of at most $n-1$ letters, the number of permutations of $n-1$ letters that match τ is correctly given by (8), and let τ be some template of $\leq n$ letters. If in fact the length of τ is $< n$ then the formula (8) gives the same result as it did when applied to $(n-1)$-permutations, which is the correct probability.

In the case where the length of τ is n and the rightmost letter of τ is 'Y', every matching permutation σ must have $\sigma(n) = n$. Hence the number of matching permutations is $(n-1)!p_{n-1}(\tau')$, where τ' consists of the first $n-1$ letters of τ, which is equal to $n!p_n(\tau)$, proving the result in this case.

In the last case, where the length of τ is n and the rightmost letter of τ is 'N', the probability of a permutation match must be $p_{n-1}(\tau')(1 - 1/n)$, since this case and the preceding one are exhaustive of the possibilities and the preceding one had a probability of $p_{n-1}(\tau')/n$. But this agrees with formula (8) for this case, completing the proof of the theorem.

6.2. TEMPLATES ON WORDS. Next we include results for words that are analogous to those we found for permutations. A preview of the kind of results that we will get is the following. Suppose $F(\tau, x) = \sum_{k \geq 1} f(k, \tau) x^k$, where $f(k, \tau)$ denotes the number of words over the alphabet $[k]$ whose length is equal to the length of the template, and which match the template. Then consider the

adjunction of one new symbol, let's call it A, at the right end of τ. Then we will show that the new generating function, $F(\tau A, x)$ can be obtained from $F(\tau, x)$ by applying a certain operator Ω_A. That is,

$$F(\tau A, x) = \Omega_A F(\tau, x).$$

The operator Ω_A will depend only on the letter A that is being adjoined to the template τ.

Hence, to find the generating function for a complete template τ, we begin with $F(\emptyset, x) = 1/(1-x)$, and we read the template τ from left to right. Corresponding to each letter in τ we apply the appropriate operator Ω. When we have finished scanning the entire template the result will be the desired generating function for τ.

The letters that we will allow in a template $\tau = \tau_1 \tau_2 \ldots \tau_l$ are $\{S, W, *, \overline{S}, \overline{O}\}$. Their meanings are that if $w = w_1 \ldots w_l$ is a word of length l over the alphabet $[k]$ then for w to match the template it must be that w_i is

1. a strongly outstanding value whenever $\tau_i = S$, or
2. a weakly outstanding value whenever $\tau_i = W$, or
3. unrestricted whenever $\tau_i = *$, or
4. not a strongly outstanding value whenever $\tau_i = \overline{S}$, or
5. neither a weakly nor a strongly outstanding value whenever $\tau_i = \overline{O}$.

Let's consider what happens to the count of matching words when we adjoin one of these letters to a template whose counting function is known. Let $f(k, \tau)$ denote the number of words of length $l = \text{length}(\tau)$, on the alphabet $[k]$ that match the template τ.

1. If w is one of the words counted by $f(k, \tau S)$, and if we delete its last letter, we obtain one of the words that is counted by $f(i, \tau)$ for some $1 \leq i < k$, and consequently

$$f(k, \tau S) = \sum_{i=1}^{k-1} f(i, \tau).$$

If $F(\tau, x) = \sum_{k \geq 1} f(k, \tau) x^k$, then we have

$$F(\tau S, x) = \frac{x}{1-x} F(\tau, x).$$

Thus we have found the operator Ω_S, and it is defined by $\Omega_S F(x) = x \frac{F(x)}{1-x}$. This is equation (9).

2. Similarly, if w is one of the words counted by $f(k, \tau W)$, and if we delete its last letter, we obtain one of the words that is counted by $f(i, \tau)$ for some $1 \le i \le k$, and consequently

$$f(k, \tau W) = \sum_{i=1}^{k} f(i, \tau).$$

If $F(\tau, x) = \sum_{k \ge 1} f(k, \tau) x^k$, then we have

$$F(\tau W, x) = \frac{F(\tau, x)}{1 - x}.$$

Thus we have found the operator Ω_W, and it is defined by $\Omega_W F(x) = F(x)/(1-x)$. This is (10).

Since the argument in each case is easy and similar to the above we will simply list the remaining three operators, equations (11), (12), and (13), as follows:

$$\Omega_* F(x) = x \frac{d}{dx} F(x)$$

$$\Omega_{\overline{S}} F(x) = \left(x \frac{d}{dx} - \frac{x}{1-x} \right) F(x)$$

$$\Omega_{\overline{O}} F(x) = \left(x \frac{d}{dx} - \frac{1}{1-x} \right) F(x)$$

The successive applications of these operators can be started with the generating function for the empty template,

$$F(\emptyset, x) = \frac{1}{1-x}.$$

Thus, to find the generating function for some given template τ, we begin with the function $1/(1-x)$, and then we read one letter at a time from τ, from left to right, and apply the appropriate one of the five operators that are defined above.

For example, how many words of length 3 over the alphabet $[k]$ match the template $\tau = S * \overline{S}$? This is the coefficient of x^k in

$$F(S * \overline{S}, x) = \Omega_{\overline{S}} \Omega_* \Omega_S \frac{1}{1-x} = \left(x \frac{d}{dx} - \frac{x}{1-x} \right) \left(x \frac{d}{dx} \right) \left(\frac{x}{1-x} \right) \frac{1}{1-x}$$

$$= \frac{x + 3x^2}{(1-x)^4}$$

The required number of 3-letter words over a k letter alphabet that match the template τ is the coefficient of x^k in the above. Since these generating functions

will always be of the form $P(x)/(1-x)^r$, with P a polynomial, we note for ready reference that

$$[x^k]\left\{\frac{\sum_j a_j x^j}{(1-x)^r}\right\} = \sum_j a_j \binom{r+k-j-1}{r-1}.$$

In the example above we have $r = 4$, $a_1 = 1$, $a_2 = 3$, so

$$f(k, S * \overline{S}) = \binom{k+2}{3} + 3\binom{k+1}{3}.$$

7. Weakly outstanding elements of words

Finally we prove Theorem 4.

Let $g(n, k, t)$ denote the number of n-letter words over the alphabet $[k]$ which have exactly t weakly outstanding elements. Consider just those words w that contain exactly m 1's, $m < n$. If w begins with a block of exactly l 1's, $0 \le l \le m$ then by deleting all m of the 1's in w we find that the remaining word is one with $n - m$ letters over an alphabet of $k - 1$ letters and it has exactly $t - l$ weakly outstanding elements. Thus we have the recurrence

$$g(n, k, t) = \sum_{m=0}^{n-1} \binom{n-l-1}{m-l} g(n-m, k-1, t-l) + \delta_{t,n}.$$

When we set $G_k(n, x) = \sum_t g(n, k, t) x^t$, we find that

(18)
$$G_k(n, x) = \sum_{m=0}^{n-1} \sum_{l=0}^{k} \binom{n-l-1}{m-l} x^l G_{k-1}(n-m, x) + x^n \quad (n, k \ge 1; G_0(n, x) = 0).$$

To discuss this recurrence, let $\Delta = \Delta_x$ be the usual forward difference operator on x, i.e., $\Delta_x f(x) = f(x+1) - f(x)$. Then from the recurrence above we discover that

$$G_1(n, x) = x^n; \quad G_2(n, x) = \Delta_x x^n (x - 1); \quad G_3(n, x) = \Delta_x^2 x^n \binom{x-1}{2}.$$

This leads to the conjecture that

$$G_k(n, x) = \Delta_x^{k-1} \left\{ x^n \binom{x-1}{k-1} \right\}$$

$$= \sum_{t=0}^{k-1} (-1)^{k-1-t} (x+t)^n \binom{k-1}{t} \binom{x+t-1}{k-1}.$$

To prove this it would suffice to show that the function $G_k(n,x)$ above satisfies the recurrence (18). If we substitute the conjectured form of G_k into the right side of (18) we find that the sums over l and m can easily be done, and the identity to be proved now reads as

$$\sum_{t=0}^{k-2}(-1)^{k-t}\binom{k-2}{t}\binom{x+t-1}{k-2}\frac{(x+t)}{t+1}((x+t+1)^n - x^n) + x^n$$

$$= \sum_{t=0}^{k-1}(-1)^{k-1-t}(x+t)^n\binom{k-1}{t}\binom{x+t-1}{k-1}.$$

If we replace the dummy index of summation t by $t-1$ on the left side, the identity to be proved becomes

$$\sum_{t=1}^{k-1}(-1)^{k-t-1}\binom{k-2}{t-1}\binom{x+t-2}{k-2}\frac{(x+t-1)}{t}((x+t)^n - x^n) + x^n$$

$$= \sum_{t=0}^{k-1}(-1)^{k-1-t}(x+t)^n\binom{k-1}{t}\binom{x+t-1}{k-1}.$$

It is now trivial to check that for $1 \leq t \leq k-1$, the coefficient of $(x+t)^n$ on the left side is equal to that coefficient on the right. If we cancel those terms and divide out a factor x^n from what remains, the identity to be proved becomes

$$-\sum_{t=1}^{k-1}(-1)^{k-t-1}\binom{k-2}{t-1}\binom{x+t-2}{k-2}\frac{(x+t-1)}{t} + 1 = (-1)^{k-1}\binom{x-1}{k-1}.$$

The sum that appears above is a special case of Gauss's original $_2F_1[a,b;c|1]$ evaluation, and the proof of (7) is complete. A straightforward calculation now shows that the average number of weakly outstanding elements among n-letter words over the alphabet $[k]$ is

$$= \frac{n}{k} + H_{k-1} + O\left(\left(\frac{k-1}{k}\right)^n\right).$$

8. Related results

In the literature, the outstanding elements and values of sequences go by various names: **éléments salients** (Rényi), **left-to-right maxima**, **records**, and others; and appear in several different contexts and results, for example:

It is well-known that the probability of obtaining at least r strongly outstanding elements in a sequence $X_1, X_2, \ldots, X_n$ of n independent, identically distributed, continuous random variables approaches 1 as $n \to \infty$. (Glick's survey [7], for example, contains this result and those that follow in this paragraph and the next.) Let $Y_i = 1$ if i is a strongly outstanding element of such a sequence, and 0 otherwise. The expected value is $E[Y_i] = 1/i$ and the variance is $V[Y_i] = 1/i - 1/i^2$. The number of strongly outstanding elements in a sequence of continuous iid random variables is therefore $R_i = \sum_i Y_i$ with expectation $E[R_i] = \sum_{i=1}^n 1/i$ and variance $V[R_i] = \sum_{i=1}^n 1/i - \sum_{i=1}^n 1/i^2$. Note $\sum_{i=1}^n 1/i - \ln(n) \to$ Euler's constant $= 0.5772\ldots$, and $\sum_{i=1}^n 1/i^2 \to \pi^2/6 = 1.6449\ldots$ for $n \to \infty$.

Let N_r denote the r-th outstanding element in a sequence of n iid continuous random variables. Then $N_1 = 1$, $E[N_r] = \infty$ for $r \geq 2$, and $E[N_{r+1} - N_r] = \infty$ as well. The probability $P[N_2 = i_2, N_3 = i_3, \ldots, N_r = i_r] = 1/(i_2 - 1)(i_3 - 1)\ldots(i_r - 1)i_r$ for $1 < i_2 < i_3 < \cdots < i_r$. The probability that an iid sequence of n continuous random variables has exactly r strongly outstanding elements is $\begin{bmatrix} k \\ r \end{bmatrix}/(r-1)!$ for large n.

Chern and Hwang [3] consider the number $f_{n,k}$ of k consecutive records (strongly outstanding elements) in a sequence of n iid continuous random variables. They improve upon known results for the limiting distribution of $f_{n,2}$. In particular, they show $f_{n,k}$ is asymptotically Poisson for $k = 1, 2$, and this is not the case for $k \geq 3$. They give the probability generating function for $f_{n,2}$, and observe that the distribution of $f_{n,2}$ is identical to that for the number of fixed points j in a random permutation of $[n]$ for $1 \leq j < n$. They give a recurrence for the probability generating function for $f_{n,k}$, and compute the mean and variance for this number.

Are there similar results for discrete distributions? Prodinger [11] considers left-to-right maxima in both the strict (strongly outstanding) and loose (weakly outstanding) senses for geometric random variables. He finds the generating function for the probability that a sequence of n independent geometric random variables (each with probability pq^{v-1} of taking the positive integer value v, where $q = p - 1$) has k strict left-to-right maxima. This probability is the coefficient of $z^n y^k$ in

$$F(z, y) = \prod_{k \geq 1} \left(1 + \frac{yzpq^{k-1}}{1 - z(1 - q^k)} \right).$$

For loose left-to-right maxima the analogous generating function is

$$\prod_{k \geq 0} \frac{1 - z(1 - q^k)}{1 - z + zq^k(1 - py)}.$$

In this paper, Prodinger also finds the asymptotic expansions for both the expected numbers of strict and loose left-to-right maxima, and the variances for these numbers. He does all of the above for uniform random variables as well.

Knopfmacher and Prodinger [10] consider the **value** and **position** of the r-th left-to-right maximum for n geometric random variables. The **position** is the r-th strongly outstanding element, and the **value** is that taken by the random variable in that position (again the value v is taken with probability pq^{v-1}, where $q = 1 - p$). For the r-th strong left-to-right maximum, the asymptotic formulas for value and position are $\frac{r}{p}$ and $\frac{1}{(r-1)!}\left(\frac{p}{q}\log_{\frac{1}{q}} n\right)^{r-1}$, respectively. For the **r**th weak left-to-right maximum, the value and position are asymptotically $\frac{rq}{p}$ and $\frac{1}{(r-1)!}\left(p\log_{\frac{1}{q}} n\right)^{r-1}$. These results are obtained by first computing the relevant generating functions.

A number of additional properties of outstanding elements of permutations are in Wilf [13].

Key [8] describes the asymptotic behavior of the number of records (strongly outstanding elements) and weak records (weakly outstanding elements) that occur in an iid sequence of integer valued random variables.

When the sequence $w_1, w_2, \ldots, w_n$ under consideration is a random permutation on n letters, then the expected number of strongly outstanding elments is the **n**th harmonic number $H_n = 1 + \frac{1}{2} + \frac{1}{2} + \cdots + \frac{1}{n}$. The variance is $H_n = H_n^{(2)}$, where $H_n^{(2)} = 1 + \frac{1}{4} + \cdots + \frac{1}{n^2}$ denotes the n-th harmonic number of the second order.

Banderier, Mehlhorn, and Brier [2] show that the average number of left-to-right maxima (strongly outstanding elements) in a **partial permutation** is $\log(pn) + \gamma + 2\frac{1-p}{p} + \left(\frac{1}{2} + \frac{2(1-p)}{p^2}\right)\frac{1}{n} + O\left(\frac{1}{n^2}\right)$ where $\gamma = 0.5772\ldots$ is Euler's constant. To obtain a **partial permutation**, we begin with the sequence $1, 2, \ldots, k$ and select each element with probability p; then take one of the $(pn)!$ permutations of $[pn]$ uniformly at random and let it act on the selected elements, while the nonselected elements stay in place.

Foata and Han [6] investigate the (right-to-left) **lower records** of **signed permutations**. A **signed permutation** is a word $w = w_1 w_2 \ldots w_n$ for which

the letters w_i are positive or negative integers, and $|w_1||w_2|\dots|w_n|$ is a permutation of $\{1, 2, \dots, n\}$. A **lower record** of such a word is a letter w_i such that $w_i < w_j$ for all j with $i+1 \leq j \leq n$. The authors consider the **signed subword** obtained by reading the lower records of w from left to right; and derive generating functions for signed permutations in terms of signed subwords, numbers of positive and negative letters in signed subwords, and other statistics.

References

[1] L. Comtet, *Advanced Combinatorics*, D. Reidel Publishing Co., Dordrecht, 1974.

[2] C. Banderier, K. Mehlhorn and R. eier, *Smoothed analysis of three combinatorial algorithms*, in *Mathematical Foundations of Computer Science 2003: 28th International Symposium*, Proceedings, Lecture Notes in Computer Science, vol. 2747, Springer, Berlin, 2003, pp. 198–207.

[3] H.-H. Chern and H.-K. Hwang, *Limit distribution of the number of consecutive records*, Random Structures Algorithms **26** (2005), 404–417.

[4] J. Clément, P. Flajolet and B. Vallée, *The analysis of hybrid trie structures*, in *Proceedings of the Ninth Annual ACM-SIAM Symposium on Discrete Algorithms* (San Francisco, CA, 1998), ACM, New York, 1998, pp. 531–539.

[5] J. N. Darroch, On the distribution of the number of successes in independent trials, Annals of Mathematical Statistics **35** (1964), 1317–1321.

[6] D. Foata and G.-N. Han, *Signed words and permutations II: The Euler-Mahonian polynomials*, Electronical Journal of Combinatorics **11** (2004/06), Research Paper 22, 18 pp. (electronic).

[7] N. Glick, *Breaking records and breaking boards*, American Mathematical Monthly **85** (1978), 2–26.

[8] E. S. Key, *On the number of recordes in an iid discrete sequence*, Journal of Theoretical Probability **18** (2005), 99–108.

[9] R. V. Kadison, *Strategies in the secretary problem*, Expositiones Mathematicae **12** (1994), 125–144.

[10] A. Knopfmacher and H. Prodinger, *Combinatorics of geometrically distributed random variables: Value and position of the rth left-to-right maximum*, Discrete Mathematics **226** (2001), 255–267.

[11] H. Prodinger, *Combinatorics of geometrically distributed random variables: Left-to-right maxima*, Discrete Mathematics **153** (1996), 253–270.

[12] A. Rényi, *Théorie des éléments saillants d'une suite d'observations*, in *Colloquium Aarhus*, 1962, pp. 104–117.

[13] H. S. Wilf, *On the outstanding elements of permutations*, `<http://www.cis.upenn.edu/ wilf>`, 1995.