Where the Monotone Pattern (Mostly) Rules

Miklós Bóna

June 17, 2007

1 Introduction

1.1 The Context

1. Let $S_n(q)$ denote the number of permutations of length n that avoid the pattern q, and consider the numbers $S_n(q)$ for each pattern q of length k. A very interesting and counter-intuitive phenomenon is that in this multiset of $k!$ numbers, the number corresponding to the *monotone case*, that is, $S_n(12 \cdots k)$, will, in general, not be the largest or the smallest number. There are several results on this fact but the phenomenon is still not perfectly well understood.

2. In 2001, Elizalde and Noy proposed another definition of pattern containment. We say that $p = p_1p_2\cdots p_n$ *tightly* contains $q = q_1q_2\cdots q_k$ if there exists an index $0 \leq i \leq n-k$ so that $q_j < q_r$ if and only if $p_{i+j} < p_{i+r}$. In other words, for p to contain q, we require that p has a *consecutive* string of entries that relate to each other the same way the entries of q do.

   Example 1 *Permutation 246351 contains 132 (take the second, third, and fifth entries, for instance), but it does not* tightly contain *132 since there are no three entries in consecutive positions in 246351 that would form a 132-pattern.*

   If p does not tightly contain q, then we say that p *tightly avoids* q. Let $T_n(q)$ denote the number of n-permutations that tightly avoid q. Elizalde and Noy conjectured that no pattern of length k is tightly avoided by more n-permutations than the monotone pattern. In other words, if q is a pattern of length k, then

$$T_n(q) \leq T_n(12\cdots k). \quad (1)$$

   This conjecture is still open in the general case.

   We point out that changing the definition of pattern avoidance changed the status of the monotone pattern among all patterns of the same length. With this definition, it is believed that the monotone pattern is the easiest pattern to avoid.

3. Let us take the idea of Elizalde and Noy one step further, by restricting the notion of pattern containment further as follows. Let $p = p_1 p_2 \cdots p_n$ be a permutation, let $k < n$, and let $q = q_1 q_2 \cdots q_k$ be another permutation. We say that p *very tightly* contains q if there is an index $0 \leq i \leq n - k$ and an integer $0 \leq a \leq n - k$ so that $q_j < q_r$ if and only if $p_{i+j} < p_{i+r}$, and,

$$\{p_{i+1}, p_{i+2}, \cdots, p_{i+k}\} = \{a + 1, a + 2, \cdots, a + k\}.$$

That is, p very tightly contains q if p tightly contains q and the entries of p that form a copy of q are not just in consecutive positions, but they are also consecutive as integers (in the sense that their set is an interval).

Example 2 *Permutation 15324 tightly contains 132 (consider the first three entries), but does not very tightly contain 132. On the other hand, 15324 very tightly contains 213, as can be seen by considering the last three entries. If p does not very tightly contain q, then we will say that p* very tightly avoids *q.*

Note that in the special case when q is the monotone pattern, this notion was studied before pattern avoidance became widely known. The literature of permutations very tightly avoiding *monotone* patterns goes back at least to the 1940s, and to papers by Kaplansky, Riordan, and Wolfowitz.

1.2 Our contribution

Let $V_n(q)$ denote the number of n-permutations that very tightly avoid q. While we cannot prove that $V_n(q) \leq V_n(12\cdots k)$ for all patterns q of length k, we will be able to prove that this inequality holds for *most* patterns q of length k. As a byproduct, we will prove that for all k, there exists a set W_k of patterns of length k so that $\lim_{k\to\infty} \frac{|W_k|}{k!} = 1$, and $V_n(q)$ is identical for all patterns $q \in W_k(q)$. In other words, almost all patterns of length k are equally difficult to very tightly avoid. There are no comparable statements known for the other two discussed notions of pattern avoidance.

Our argument is a probabilistic one. Once the framework is set up, the computation is elementary. However, this is the first time we know of that expectations are successfully used to compare the number of permutations avoiding a given pattern (admittedly, with a very restrictive definition of pattern avoidance). We wonder whether more sophisticated methods of enumeration could extend the reach of this technique to less restrictive definitions of pattern avoidance.

2 A Probabilistic Argument

2.1 The outline of the argument

For the rest of this section, let $k \geq 3$ be a fixed positive integer. Let $\alpha = 12 \cdots k$, the monotone pattern of length k. Recall that $V_n(\alpha)$ is the number of n-permutations very tightly avoiding α. Our goal is to prove that

$$V_n(q) \leq V_n(\alpha)$$

for any pattern q of length k.

Let q be any pattern of length k. For a fixed positive integer n, let $X_{n,q}$ be the random variable counting the occurrences of q in a randomly selected n-permutation. As the following straightforward proposition shows, the expectation of $X_{n,q}$ does not depend on q; it only depends on n, and the length k of q.

Proposition 1 *For any fixed n, and $q \in S_k$, we have*

$$E(X_{n,q}) = \frac{(n-k+1)^2}{\binom{n}{k} k!}.$$

Proof: Let X_i be the indicator random variable of the event that the string $p_{i+1}\cdots p_{i+k}$ is a q-pattern in the very tight sense. Then $E(X_i) = P(p_{i+1}\cdots p_{i+k} \simeq q) = \frac{n-k+1}{\binom{n}{k}} \cdot \frac{1}{k!}$, since there are $n-k+1$ favorable choices for the set of the entries $p_{i+1}, \cdots, p_{i+k}$, and there is $1/k!$ chance that their pattern is q. Now note that $E(X_{n,q}) = \sum_{i=0}^{n-k} X_i$, and the statement is proved by the linearity of expectation. $\diamond$

Let $p_{n,i,q}$ be the probability that a randomly selected n-permutation contains *exactly* i copies of q, and let $P_{n,i,q}$ be the probability that a randomly selected n-permutation contains *at least* i copies of q.

Set $m = n - k + 1$, and observe that no n-permutation can very tightly contain more than m copies of any given pattern q of length k. By the definition of expectation

$$
\begin{aligned}
E(X_n, q) &= \sum_{i=1}^{m} i p_{n,i,q} \\
&= \sum_{j=0}^{m-1} \sum_{i=0}^{j} p_{n,m-i,q} \\
&= p_{n,m,q} + (p_{n,m,q} + p_{n,m-1,q}) + \cdots + (p_{n,m,q} + \cdots + p_{n,1,q}) \\
&= \sum_{i=1}^{m} P(n, i, q).
\end{aligned}
$$

By Proposition **??**, we know that $E(X_{n,q}) = E(X_{n,\alpha})$, and then previous displayed equation implies that

$$
\sum_{i=1}^{m} P(n, i, q) = \sum_{i=1}^{m} P(n, i, \alpha). \tag{2}
$$

So if we could show that for $i \geq 2$, the inequality

$$
P(n, i, q) \leq P(n, i, \alpha) \tag{3}
$$

holds, then (**??**) would imply that $P(n, 1, q) \geq P(n, 1, \alpha)$, which is just what we set out to prove.

The simple counting argument that we present will not prove (**??**) for every pattern q. However, it will prove (**??**) for most patterns q. We describe these patterns in the next subsection.

Assume that the permutation $p = p_1 p_2 \cdots p_n$ very tightly contains two *non-disjoint* copies of the pattern $q = q_1 q_2 \cdots q_k$. Let these two copies be $q^{(1)}$ and $q^{(2)}$, so that $q^{(1)} = p_{i+1} p_{i+2} \cdots p_{i+k}$ and $q^{(2)} = p_{i+j+1} p_{i+j+2} \cdots p_{i+j+k}$ for some $j \in [1, k-1]$.

Then $|q^{(1)} \cap q^{(2)}| = k - j + 1 = s$. Furthermore, since the set of entries of $q^{(1)}$ is an interval, and the set of entries of $q^{(2)}$ is an interval, it follows that the set of entries of $q^{(1)} \cap q^{(2)}$ is also an interval. So the rightmost s entries of q, and the leftmost s entries of q must form identical patterns, and the respective sets of these entries must both be intervals.

For obvious symmetry reasons, we can assume that $q_1 < q_k$. We claim that then the *rightmost* s entries of q must also be the *largest* s entries of q. This can be seen by considering $q^{(1)}$. Indeed, the set of these entries of $q^{(1)}$ is the intersection of two intervals of the same length, and therefore, must be an ending segment of the interval that starts on the left of the other. An analogous argument, applied for $q^{(2)}$, shows that the leftmost s entries of q must also be the *smallest* s entries of q.

The following Proposition collects the observations made in this subsection.

Proposition 2 *Let p be a permutation that very tightly contains copies $q^{(1)}$ and $q^{(2)}$ of the pattern $q = q_1 q_2 \cdots q_k$. Let us assume that $q_1 < q_k$. Then $q^{(1)}$ and $q^{(2)}$ are disjoint unless all of the following hold.*

There exists a positive integer $s \leq k-1$ so that

1. *the rightmost s entries of q are also the largest s entries of q, and the leftmost s entries of q are also the smallest s entries of q, and*
2. *the pattern of the leftmost s entries of q is identical to the pattern of the rightmost s entries of q.*

If q satisfies both of these criteria, then two very tightly contained copies of q in p may indeed intersect. For example, $q = 2143$ satisfies both of the above criteria with $s = 2$, and indeed, 214365 very tightly contains two intersecting copies of q, namely 2143 and 4365.

Definition 1 *Let q be a pattern that satisfies both conditions of Proposition* **??**. *Then we say that q is condensible.*

It is not difficult to prove that almost all patterns of length k are non-condensible.

2.2 The Computational Part of the Proof

The following Lemma is the heart of our main result.

Lemma 1 *Let q be a non-condensible pattern, and let $i > 1$. Then*

$$P(n, i, q) \leq P(n, i, \alpha).$$

Proof: If $n < ik$, then the statement is clearly true. Indeed, $P(n, i, q) = 0$ since any two copies of q in an n-permutation p would have to be disjoint, and n is too small for that. In the rest of the proof, we assume that $n \geq ik$. For $k \leq 2$, the statement is trivial, so we assume $k \geq 3$ as well.

First, we prove a lower bound on $P(n, i, \alpha)$. The number of n-permutations very tightly containing i copies of α is at least as large as the number of n-permutations very tightly containing the pattern $12\cdots(i+k-1)$. The latter is at least as large as the number of n-permutations that very tightly contain a $12\cdots(i+k-1)$-pattern in their first $i+k-1$ positions, that is, $(n-k-i+2)\cdot(n-k-i+1)! = (n-k-i+2)!$. Therefore,

$$\frac{(n-k-i+2)!}{n!} \leq P(n, i, \alpha). \tag{4}$$

We are now going to find an upper bound for $P(n, i, q)$. Let S be an i-element subset of $[n]$ so that the elements of S can be the starting positions of i (necessarily disjoint) very tight copies of q in an n-permutation. If $S = \{s_1, s_2, \cdots, s_i\}$, then this is equivalent to saying that

$$1 \leq s_1 < s_2-k+1 \leq s_3-2k+2 \leq \cdots \leq s_i-(i-1)(k-1) \leq n-i(k-1).$$

Therefore, there are $\binom{n-i(k-1)}{i}$ possibilities for S. Now let A_S be the event that in a random permutation $p = p_1 \cdots p_n$, the subsequence $p_j p_{j+1} \cdots p_{j+k-1}$ is a very tight q-subsequence for all $j \in S$. Let $A_{i,q}$ be the event that p contains at least i very tight copies of q. Then $P(A_{i,q}) = P(n, i, q)$. Furthermore,

$$A_{i,q} = \cup_S A_S,$$

where the union is taken over all $\binom{n-i(k-1)}{i}$ possible subsets for S. Therefore,

$$P(n, i, q) = P(A_{i,q}) \leq \sum_S P(A_S). \tag{5}$$

Let us now compute $P(A_S)$. This probability does not depend on the choice of S. Indeed, just as there are $\binom{n-i(k-1)}{i}$ possibilities for S, there are $\binom{n-i(k-1)}{i}$ possibilities for the *entries* in the positions belonging to

S. Once those entries are known, the rest of the q-patterns starting in those entries are determined, and there are $(n-ik)!$ possibilities for the rest of the permutation. This shows that $P(A_S) = \binom{n-i(k-1)}{i}(n-ik)!\frac{1}{n!}$ for all S. Therefore, (**??**) implies

$$P(n,i,q) \leq \binom{n-i(k-1)}{i}^2 (n-ik)!\frac{1}{n!}. \tag{6}$$

Comparing (**??**) and (**??**), we see that our lemma will be proved if we show that for $i > 1$, the inequality

$$\binom{n-i(k-1)}{i}^2 (n-ik)! \leq (n-i-k+2)!,$$

or, equivalently,

$$(n-i(k-1))_i \leq i!^2(n-k-i+2)(n-k-i+1)\cdots(n-i(k-1)+1)) \tag{7}$$

holds. Here $(z)_j = z(z-1)\cdots(z-j+1)$. The left-hand side has i factors, while the right-hand side, not counting $i!^2$, has $(k-1)(i-1) > i$ factors, each of which are larger than the factors of the left-hand side. Therefore, (**??**) holds, and the Lemma is proved. $\diamond$

The proof of our main result is now immediate.

Theorem 1 *Let q be any non-condensible pattern of length k. Then*

$$V_n(q) \leq V_n(\alpha).$$

Proof: Lemma **??** and formula (**??**) together imply that $P(n,1,q) \geq P(n,1,\alpha)$, which means that there are at least as many n-permutations that very tightly contain q as n-permutations that very tightly contain α. $\diamond$

2.3 A Result on Non-condensible Patterns

We have seen in Proposition **??** that if q is non-condensible and $q_1 < q_k$, then any two copies of q contained in a given permutation p are disjoint. Therefore, the number of n-permutations that very tightly avoid q can be computed by the Principle of Inclusion-Exclusion. Indeed, in this case, the following holds.

Proposition 3 *Let q be a non-condensible pattern. Then*

$$V_n(q) = n! - \sum_{i=1}^{\lfloor n/k \rfloor} \binom{n-i(k-1)}{i}^2 (n-ik)!.$$

In particular, $V_n(q)$ does not depend on the choice of q.

Proof: In the proof of Lemma **??**, more precisely, in our argument showing that (**??**) holds, we showed that there are $\binom{n-i(k-1)}{i}$ ways to choose an i-element set of positions that can be the starting positions of i disjoint very tight copies of q, and there are $\binom{n-i(k-1)}{i}$ ways to choose the sets of entries forming these same copies. Once these choices are made, the rest of the permutation can be chosen in $(n-ik)!$ ways. The statement now follows by the Principle of Inclusion-Exclusion. $\diamond$

3 Further Directions

Question 1 *Is it possible to apply our method to compare the numbers $T_n(q)$ and $T_n(\alpha)$, or the numbers $S_n(q)$ and $S_n(\alpha)$, for at least some patterns q?*

Question 2 *Are there other patterns q for which $V_n(q)$ can be explicitly determined?*

Let us call patterns q and q' *very tightly equivalent* if $V_n(q) = V_n(q')$ for all n. We have seen that almost all patterns of length k are very tightly equivalent. This raises the following questions.

Question 3 *How many equivalence classes are there for very tight patterns of length k?*