The number of steps in the Robinson-Schensted algorithm

Dan Romik [1]

Abstract. Suppose a permutation $\sigma \in \mathcal{S}_n$ is chosen at random (n large) and the Robinson-Schensted algorithm is applied to compute its associated Young diagram. Then for almost all permutations, the number of *bumping* operations performed is about $(128/27\pi^2)n^{3/2}$, and the number of comparison operations performed is about $(64/27\pi^2)n^{3/2}\log_2 n$.

1. Introduction

The Robinson-Schensted algorithm (also known in a more general version as the Robinson-Schensted-Knuth, or RSK, algorithm) is an important algorithm in combinatorics. It takes a permutation $\sigma \in \mathcal{S}_n$, and computes a partition λ of n,

$$n = \lambda_1 + \lambda_2 + ... + \lambda_k, \qquad \lambda_1 \geq \lambda_2 \geq ... \geq \lambda_k$$

together with two associated Young tableaux P, Q of shape λ (fillings of the Young diagram of λ with the numbers $1, 2, ..., n$ which are increasing along rows and columns). The correspondence between permutations in $\mathcal{S}_n$ and pairs of Young tableaux of equal shape and area n is a bijection, known as the Robinson-Schensted correspondence. The partition λ encodes the *increasing/decreasing subsequence profile* of the permutation σ, in the following sense: for any $1 \leq j \leq \lambda'_1$, $\lambda_1 + \lambda_2 + ... + \lambda_j$ is equal to the maximal cardinality of a union of j disjoint increasing subsequences of σ, and for any $1 \leq j \leq \lambda_1$, $\lambda'_1 + ... + \lambda'_j$ is equal to the maximal cardinality of a union of j disjoint decreasing subsequences of σ (where λ' is the partition conjugate to λ, i.e., $\lambda'_j = \#\{i : \lambda_i \geq j\}$). In particular, λ_1 is the length of the longest increasing subsequence of σ, an important statistic of permutations. For a full description of the algorithm and its remarkable properties, see [4], section 5.1.4, and [8], chapter 7. Figure 1 shows a sample input-output pair.

$$\lambda = 4, 4, 1$$

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 5 & 2 & 7 & 4 & 6 & 9 & 8 & 1 \end{pmatrix} \xrightarrow{\text{R-S}} \quad P = \begin{array}{|c|c|c|c|} \hline 1 & 4 & 6 & 8 \\ \hline 2 & 5 & 7 & 9 \\ \hline 3 \\ \cline{1-1} \end{array} \qquad Q = \begin{array}{|c|c|c|c|} \hline 1 & 2 & 4 & 7 \\ \hline 3 & 5 & 6 & 8 \\ \hline 9 \\ \cline{1-1} \end{array}$$

Figure 1: A permutation and its associated partition and Young tableaux

A natural question is, How many steps are performed during the application of the algorithm? Certainly, there are always exactly n steps of laying down a new square in the Young diagram λ being constructed. The other steps performed are of two types: comparisons, and executions of the so-called *bumping* operation, which replaces a number with another number and pushes the former down to the next row. For every bumping operation performed, a binary search of a sorted list must be performed to find the place in the next row where a bumping will take place

[1] Department of Mathematics, The Weizmann Institute of Science, Rehovot 76100, Israel. email: `romik@wisdom.weizmann.ac.il`

next (or a new square laid down). It is the steps of these two types which are interesting, as their number is not fixed.

Note that the number of bumping steps can range between 0, in the best case, and $n(n-1)/2$, in the worst case. We prove

Theorem. As $n \to \infty$, for almost all (i.e. a $(1-o(1))$-fraction) $\sigma \in \mathcal{S}_n$, the number of bumping operations performed is

$$(1) \qquad (1+o(1))\,\frac{128}{27\pi^2}\,n^{3/2} = (1+o(1))\;0.480337...\;n^{3/2},$$

and the number of comparison operations performed is

$$(2) \qquad (1+o(1))\,\frac{64}{27\pi^2}\,n^{3/2}\log_2 n = (1+o(1))\;0.240169...\;n^{3/2}\log_2 n$$

The *average* numbers of bumping operations and comparison operations over all permutations in $\mathcal{S}_n$, are also given by (1) and (2), respectively.

2. The proof

The number N of bumping operations is a function of the partition λ: The λ_1 squares in the first row were not bumped at all; the λ_2 squares in the second row were bumped down one row, etc. Therefore N has the expression

$$N = 0 \cdot \lambda_1 + 1 \cdot \lambda_2 + ... + (k-1) \cdot \lambda_k = \sum_{i=1}^{k} (i-1)\lambda_i$$

Now, if the permutation $\sigma \in \mathcal{S}_n$ is chosen at random, then the distribution of the resulting partition λ is the so-called *Plancherel measure* $\mathbb{P}_n$, namely

$$\mathbb{P}_n(\lambda) = \frac{d_\lambda^2}{n!},$$

where d_λ is the number of standard Young tableaux of shape λ - this follows from the bijection mentioned above. Our objective is therefore to analyze the statistic $N = N(\lambda)$ of a Plancherel-distributed random partition of n. Fortunately, Plancherel random partitions are an extensively studied object, and our goal can be achieved quite easily using known results.

First, note the following geometric interpretation of $N(\lambda)$: If we choose a coordinate system such that row i in the Young diagram of λ has $y-$coordinate $i-1$, then $N(\lambda)/n$ is simply the $y-$coordinate of the *center of mass* of λ.

Let $\mathcal{P}_n$ be the set of partitions of n. For a partition $\lambda \in \mathcal{P}_n$, let $f_\lambda(x)$ be the graph of the Young diagram of λ, i.e.

$$f_\lambda(x) = \lambda_{\lceil x \rceil}, \qquad (x > 0)$$

Let $\tilde{f}_\lambda(x)$ be the *rescaled* Young diagram, given by

$$\tilde{f}_\lambda(x) = \frac{1}{\sqrt{n}} f_\lambda(\sqrt{n}x)$$

Let $\tilde{g}_\lambda(u)$ be the rescaled Young diagram of λ in the rotated coordinate system $u = (x-y)/\sqrt{2}$, $v = (x+y)/\sqrt{2}$. That is, $\tilde{g}_\lambda(u)$ satisfies

$$\tilde{g}_\lambda\left(\frac{x - \tilde{f}_\lambda(x)}{\sqrt{2}}\right) = \frac{x + \tilde{f}_\lambda(x)}{\sqrt{2}}, \qquad (0 < x < \tilde{f}_\lambda^{-1}(0))$$

$$\tilde{g}_\lambda(u) = -u, \qquad (u \leq -\tilde{f}_\lambda(0)/\sqrt{2})$$

$$\tilde{g}_\lambda(u) = u, \qquad (u \geq \tilde{f}_\lambda^{-1}(0)/\sqrt{2})$$

Recall the *limit shape theorem for Plancherel random partitions*:

Theorem. (Logan-Shepp [7], Kerov-Vershik [5, 6]) Define the function $\Omega : \mathbb{R} \to \mathbb{R}$ by

$$\Omega(u) = \begin{cases} \frac{2}{\pi}(u \arcsin(u/\sqrt{2}) + \sqrt{2-u^2}) & |u| \leq \sqrt{2} \\ |u| & |u| > \sqrt{2} \end{cases}$$

Then for every $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}_n\left(\lambda \in \mathcal{P}_n : \sup_{u\in\mathbb{R}} |\tilde{g}_\lambda(u) - \Omega(u)| > \epsilon\right) = 0$$

The proof of our Theorem is now essentially contained in the following picture:

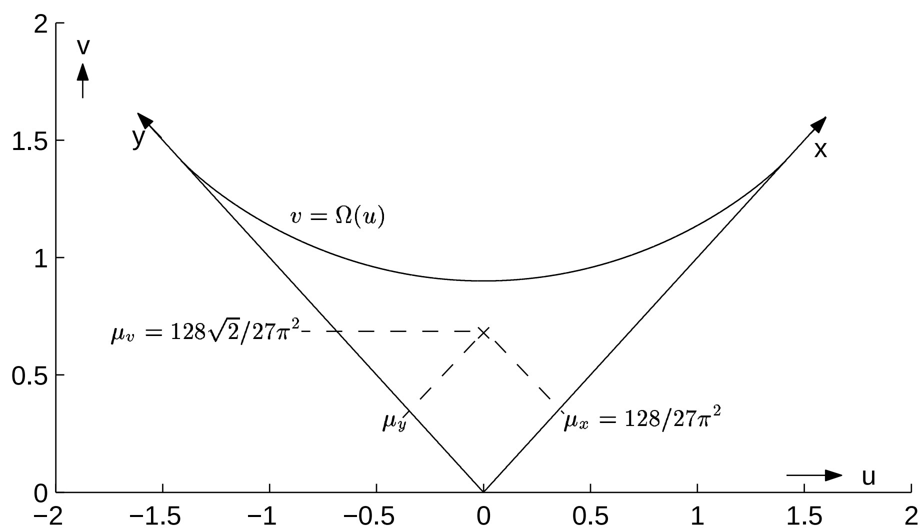

Figure 2: The limit shape $v = \Omega(u)$ and the center of mass

The limit shape theorem asserts that for large n, $\mathbb{P}_n$-almost all partitions of n have the graph of their rescaled rotated Young diagrams lying within a stripe of thickness ϵ around the curve $\Omega(u)$. This is almost sufficient to imply that the center of mass of those (rescaled) Young diagrams is close to the ideal center of mass (μ_u, μ_v) of the area lying between the curve $\Omega(u)$ and the curve $|u|$; one need only mention the additional simple fact that the support of $\tilde{g}_\lambda(u) - |u|$ is known to be contained in the interval $[-e/\sqrt{2}, e/\sqrt{2}]$ for $\mathbb{P}_n$-almost all partitions of n as $n \to \infty$ (see e.g. [4], p. 68, exercise 29. In fact, the number e in the above assertion can be replaced by 2, but this is much more difficult to prove. Also, the $\mathbb{P}_n$ measure of the set of partitions where the support is not contained in the interval $[-e/\sqrt{2}, e/\sqrt{2}]$, decreases to 0 exponentially fast in n. This justifies the assertion that the *average* number of bumping operations over all permutations is the same as the number of operations for a *typical* permutation.)

It remains to calculate the coordinates of the center of mass (μ_u, μ_v). Clearly $\mu_u = 0$ by the symmetry $\Omega(u) = \Omega(-u)$. μ_v is evaluated as

$$\mu_v = \int_{-\sqrt{2}}^{\sqrt{2}} (\Omega(u) - |u|) \left(\frac{\Omega(u) + |u|}{2} \right) du = \frac{1}{2} \int_{-\sqrt{2}}^{\sqrt{2}} (\Omega(u)^2 - u^2) du = \frac{128\sqrt{2}}{27\pi^2},$$

as can be easily verified by hand or using a symbolic algebra system. This gives the values asserted in Figure 2

$$\mu_x = \mu_y = \frac{\mu_v}{\sqrt{2}} = \frac{128}{27\pi^2}$$

This finishes the proof of the claim about the number of bumping operations, since $N(\lambda)$ is n times the y-coordinate of the center of mass of the diagram λ, or $n^{3/2}$ times the y-coordinate of the center of mass of the rescaled diagram.

Next, we turn to the number T of comparison steps. This number is not a function of the partition λ, but depends on the order by which the elements are inserted into the tableau. However, we can bound T from both sides by functions of λ, and these bounds are asymptotically tight.

Our claim is, roughly, that the vast majority of binary search operations performed during the application of the Robinson-Schensted algorithm for a typical permutation require about $\log_2(\sqrt{n}) = \frac{1}{2}\log_2 n$ comparison operations. Indeed, the number of comparisons differs by at most one from the logarithm to base 2 of the length of the row in the tableau where the search is being performed. This gives immediately the upper bound

$$T \leq (N(\lambda) + \lambda_1) \cdot (\log_2 \lambda_1 + 1)$$

(the additional λ_1 accounts for binary searches in the first row, *before* the first bumping in each insertion.) Since, as was mentioned above, we have $\lambda_1 \leq e\sqrt{n}$ $\mathbb{P}_n$-almost-surely, this proves that (2) is an upper bound for the number of comparison steps in $\mathbb{P}_n$-almost all diagrams.

The lower bound is obtained in a similar fashion: Fix i, $1 \leq i \leq \lambda_1$. Observe that, over all permutations leading to the diagram λ, the number of comparisons is minimized if, during the construction of the tableau, the inserted elements are pushed down as fast as possible, so the first λ'_1 elements form a column of height λ'_1, the next λ'_2 elements form a column of height λ'_2, etc. After $\lambda'_1 + ... + \lambda'_i$ insertions, this minimal tableau will have columns of heights $\lambda'_1, ..., \lambda'_i$, and any binary search performed thereafter will require at least $\log_2 i$ comparisons. This gives us the lower bound

$$T \geq \left(N(\lambda) - \sum_{j=1}^{i} \frac{\lambda'_j(\lambda'_j - 1)}{2} \right) \cdot \log_2 i \geq \left(N(\lambda) - i\frac{\lambda_1'^2}{2} \right) \cdot \log_2 i$$

This is enough to prove our claim, since for a typical diagram, i.e. one for which $N(\lambda)$ is close to its expected value (1) and λ_1' is no more than $e\sqrt{n}$, we may take $i = \lfloor \epsilon\sqrt{n} \rfloor$ for some $\epsilon > 0$ to obtain finally

$$T \geq N(\lambda)\log_2(\sqrt{n}) - \frac{\epsilon e^2}{2} n^{3/2} \log_2(\sqrt{n}) - N(\lambda)\log_2(\epsilon)$$

Since $\epsilon > 0$ is arbitrary, this shows that (2) is a lower bound for T.

3. Concluding remarks

In this paper, we initiated the probabilistic analysis of the Robinson-Schensted algorithm, by obtaining the asymptotics of the average and typical number of steps performed during the application of this algorithm. It would be interesting to pursue this study further by trying to determine asymptotics of the next order, namely the variance of the number of steps and its limiting distribution. For the number of bumping operations at least, this does not seem a completely hopeless venture, as it consists of analyzing the statistic $N(\lambda)$ of a Plancherel-distributed random partition. The current state of the art concerning asymptotics of Plancherel measure, such as the paper [3], is not sufficient to solve this problem (although it does suggest that the limiting distribution is almost certainly Gaussian), but it is to be hoped that future developments might make this problem more accessible.

We remark further that we are not aware of many previous works dealing with computational aspects in the theory of Young tableaux. The interested reader may consult the papers [1], [9], [2], of which only [1] deals with questions of a probabilistic nature (we thank the referee for pointing out these references). In the theory of combinatorial structures such as Young tableaux, plane partitions etc., there appear many fascinating and beautiful algorithms, and the analysis of such algorithms seems like a promising unexplored direction.

References

[1] A. Barvinok, S. Fomin, Sparse interpolation of symmetric polynomials. *Adv. Appl. Math.* 18 (1997), no. 3, 271–285.

[2] D. Bernstein, The computational complexity of rules for the character table of S_n. Preprint, arxiv:CO/0309225.

[3] V. Ivanov, G. Olshanski, Kerov's central limit theorem for the Plancherel measure on Young diagrams. In: Symmetric Functions 2001: Surveys of Developments and Perspectives. Proc. NATO Advanced Study Institute, ed. S. Fomin, Kluwer, 2002.

[4] D. E. Knuth, The Art of Computer Programming, vol. 3: Sorting and Searching, 2nd ed. Addison Wesley, 1998.

[5] S. V. Kerov, A. M. Vershik, Asymptotics of the Plancherel measure of the symmetric group and the limiting shape of Young tableaux. *Soviet Math. Dokl.* 18 (1977), 527–531.

[6] S. V. Kerov, A. M. Vershik, The asymptotics of maximal and typical dimensions of irreducible representations of the symmetric group. *Funct. Anal. Appl.* 19 (1985), 21–31.

[7] B. F. Logan, L. A. Shepp, A variational problem for random Young tableaux. *Adv. Math.* 26 (1977), no. 2, 206–222.

[8] R. P. Stanley, Enumerative Combinatorics, vol. 2. Cambridge University Press, 1999.

[9] J. R. Stembridge, Computational aspects of root systems, Coxeter groups, and Weyl characters. In: Interaction of Combinatorics and Representation Theory, 1–38, *MSJ Memoirs* 11, Math. Soc. Japan, Tokyo, 2001.