THE TOILET PAPER PROBLEM

DONALD E. KNUTH
Computer Science Department, Stanford University, Stanford, CA 94305

1. Introduction. The toilet paper dispensers in a certain building are designed to hold two rolls of tissues, and a person can use either roll.

There are two kinds of people who use the rest rooms in the building: *big-choosers* and *little-choosers*. A big-chooser always takes a piece of toilet paper from the roll that is currently larger; a little-chooser always does the opposite. However, when the two rolls are the same size, or when only one roll is nonempty, everybody chooses the nearest nonempty roll. When both rolls are empty, everybody has a problem.

Let us assume that people enter the toilet stalls independently at random, with probability p that they are big-choosers and probability $q = 1 - p$ that they are little-choosers. If the janitor supplies a particular stall with two fresh rolls of toilet paper, both of length n, let $M_n(p)$ be the average number of portions left on one roll when the other roll first empties. (We assume that everyone uses the same amount of paper, and that the lengths are expressed in terms of this unit.) For example, it is easy to establish that

$$M_1(p) = 1, \quad M_2(p) = 2 - p, \quad M_3(p) = 3 - 2p - p^2 + p^3; \quad M_n(0) = n; \quad M_n(1) = 1.$$

The purpose of this paper is to study the asymptotic value of $M_n(p)$ for fixed p as $n \to \infty$. We will see that the generating function $\Sigma_n M_n(p) z^n$ has a surprisingly simple form, from which the asymptotic behavior can readily be deduced. Along the way we will encounter several other interesting facts.

2. Recurrence Relations. Let us begin by generalizing the problem slightly, using the notation $M_{mn}(p)$ to stand for the mean number of portions left when one roll empties, if we start with m on one roll and n on the other. Thus

$$\begin{aligned} M_n(p) &= M_{nn}(p); \\ M_{m0}(p) &= m; \\ M_{nn}(p) &= M_{n(n-1)}(p), \qquad \text{if } n > 0; \\ M_{mn}(p) &= pM_{(m-1)n}(p) + qM_{m(n-1)}(p), \qquad \text{if } m > n > 0. \end{aligned}$$

The value of $M_n(p)$ can be computed for all n from these recurrence relations, since no pairs (m', n') with $m' < n'$ will arise.

It is convenient to visualize the recurrence by drawing certain arcs between adjacent lattice points in the plane, where the arc from (n, n) to $(m - 1, n)$ has weight p and from (m, n) to $(m, n - 1)$ has weight q, for all $0 < n < m$; the arc from (m, n) to $(n, n - 1)$ has weight 1 for all $n > 0$; and there are no other arcs. Then $M_{mn}(p)$ is the sum, over all $k \geqslant 1$, of k times the sum of the weights of all paths from (m, n) to $(k, 0)$, where the weight of a path is the product of the individual arc weights.

A path that starts at the diagonal point (n, n) must go first to $(n, n - 1)$; then it either returns to the diagonal at $(n - 1, n - 1)$ or goes to $(n, n - 2)$, etc. Let c_k be the number of paths from (n, n) to $(n - k, n - k)$ whose intermediate points do not touch the diagonal, and let d_{nk} be the number of paths from $(n, n - 1)$ to $(k, 1)$ whose points do not ever touch the diagonal. A path that starts at (n, n) either returns to the diagonal for the first time at some point $(n - k, n - k)$, or never returns to the diagonal at all; it follows that

The author is Fletcher Jones Professor of Computer Science at Stanford University. "My main life's work is what I like to call the *analysis of algorithms*, but I also enjoy doing research in supporting disciplines such as combinatorial and discrete mathematics, programming languages, and digital typography."

$$\begin{aligned} M_n(p) &= c_1 p M_{n-1}(p) + c_2 p^2 q M_{n-2}(p) + \cdots + c_{n-1} p^{n-1} q^{n-2} M_1(p) + L_n(p) \\ &= \sum_{0<k<n} c_k p^k q^{k-1} M_{n-k}(p) + L_n(p); \end{aligned}$$

$$L_n(p) = \sum_{2 \leqslant k \leqslant n} k d_{nk} p^{n-k} q^{n-1}, \quad \text{for } n \geqslant 2; \qquad L_1(p) = 1.$$

(Each path from (n, n) to $(n-k, n-k)$ has weight $p^k q^{k-1}$ if no intermediate diagonal points are involved, since the step to $(n, n-1)$ has weight 1 and then there are k steps of weight p and $k-1$ of weight q, in some order. Similarly, each diagonal-avoiding path from $(n, n-1)$ to $(k, 1)$ has weight $p^{n-k} q^{n-2}$.)

The coefficients c_k are the well-known Catalan numbers, and the coefficients d_{nk} are the well-known numbers that arise in the classical ballot problem; see, for example, [**2**, III.1], [**3**, exercise 2.2.1–4]. We can discover the required values by observing that d_{nk} is the number of decreasing paths from $(n, n-1)$ to $(k, 1)$ minus the number of decreasing paths from $(n, n-1)$ to $(1, k)$, where a "decreasing path" is any path that decreases either the left component or the right component by unity at each step. This follows because there is a 1-1 correspondence between all decreasing paths from $(n, n-1)$ to $(k, 1)$ that do touch the diagonal and all decreasing paths from $(n, n-1)$ to $(1, k)$; the idea [**1**] is to reflect the path about the diagonal, starting after the place where it first touches a diagonal point. Since the number of decreasing paths from (a, b) to (c, d) is $\binom{a+b-c-d}{a-c} = \binom{a+b-c-d}{b-d}$ for all $a \geqslant c$ and $b \geqslant d$, we have

$$d_{nk} = \binom{2n-k-2}{n-2} - \binom{2n-k-2}{n-1} = \binom{2n-k-2}{n-2} \frac{k-1}{n-1}.$$

Furthermore $c_{n-1} = d_{n2}$, hence

$$c_n = \binom{2n-2}{n-1} \frac{1}{n}.$$

3. Special power series. The generating function for Catalan numbers

$$C(z) = c_1 z + c_2 z^2 + \cdots = \sum_{n \geqslant 1} \binom{2n-2}{n-1} \frac{1}{n} z^n = \frac{1 - \sqrt{1-4z}}{2}$$

can be derived in many ways. For our purposes it seems best to make use of the general identity

$$(*) \qquad \sum_{k \geqslant 0} \binom{2k+w}{k} z^k = \frac{1}{\sqrt{1-4z}} \left(\frac{1 - \sqrt{1-4z}}{2z} \right)^w.$$

This well-known identity holds for all complex numbers w; it can be proved easily by contour integration: The coefficient of z^k in the Maclaurin expansion of the right-hand side is

$$\frac{1}{2\pi i} \oint \frac{1}{\sqrt{1-4z}} \left(\frac{1-\sqrt{1-4z}}{2z} \right)^w \frac{dz}{z^{k+1}} = \frac{1}{2\pi i} \oint \frac{dt}{(1-t)^{w+k+1} t^{k+1}}$$

if we make the substitutions $t = \frac{1}{2}(1 - \sqrt{1-4z})$, $z = t - t^2$, $dz = (1-2t)\,dt$. The latter integral is the residue of the integrand, i.e., the coefficient of t^k in $(1-t)^{-w-k-1}$, namely $\binom{-w-k-1}{k}(-1)^k = \binom{2k+w}{k}$. (A more elementary proof can be found in [**3**, exercise 1.2.6–26].)

The derivative of $C(z)/z$ with respect to z is $C(z)^2/(z^2\sqrt{1-4z})$; hence we can replace w by $w+1$ in $(*)$ and integrate, obtaining the companion formula

$$\sum_{k \geqslant 0} \frac{w}{k+w} \binom{2k+w-1}{k} z^k = \left(\frac{1-\sqrt{1-4z}}{2z} \right)^w.$$

Again, this result is valid for all complex w, if we evaluate the coefficient by continuity when $k + w = 0$. The case $w = 1$ of this formula reduces to the generating function for Catalan numbers stated earlier.

These power series converge for $|z| < 1/4$, because the righthand side is singular only when z is infinite or $\sqrt{1-4z}$ is singular. It is interesting to consider what happens when $z = pq$, $p \geqslant 0$, $q \geqslant 0$, and $p + q = 1$: We have $1 - 4z = (p-q)^2$, hence

$$\sqrt{1-4z} = |p-q| = \max(p,q) - \min(p,q),$$

and we obtain the interesting formula

$$C(pq) = \sum_{n\geqslant 1}\binom{2n-2}{n-1}\frac{1}{n}p^n q^n = \min(p,q).$$

We have $pq < 1/4$ unless $p = q = 1/2$; the formula holds also in the latter case, by Abel's limit theorem.

4. Generating functions. Let us now set

$$M(z) = \sum_{n\geqslant 1} M_n(p)z^n; \qquad L(z) = \sum_{n\geqslant 1} L_n(p)z^n.$$

The recurrence relation for $M_n(p)$ in section 2 is equivalent to

$$M(z) - L(z) = q^{-1}C(pqz)M(z),$$

and we also have

$$\begin{aligned} L(z) &= z + \sum_{2\leqslant k\leqslant n}\frac{q^{n-1}}{n-1}k(k-1)p^{n-k}\binom{2n-k-2}{n-2}z^n \\ &= \sum_{j,k\geqslant 0}\frac{q^{j+k-1}}{j+k-1}k(k-1)p^j\binom{2j+k-2}{j}z^{j+k} \\ &= \sum_{k\geqslant 0} q^{k-1}kz^k\sum_{j\geqslant 0}\frac{k-1}{j+k-1}\binom{2j+k-2}{j}(pqz)^j. \end{aligned}$$

By the identity in section 3, the latter sum is

$$\begin{aligned} &= \sum_{k\geqslant 0} q^{k-1}kz^k\left(\frac{1-\sqrt{1-4pqz}}{2pqz}\right)^{k-1} \\ &= z\sum_{k\geqslant 0} kp^{1-k}C(pqz)^{k-1} = \frac{p^2 z}{(p-C(pqz))^2}. \end{aligned}$$

We can now eliminate $L(z)$ and solve for $M(z)$, obtaining a "closed form" for the desired generating function:

$$M(z) = z\left(\frac{p}{p-C(pqz)}\right)^2\left(\frac{q}{q-C(pqz)}\right).$$

Such a simple form for $M(z)$ is unexpected; but in fact, we can do even more! We have

$$(p - C(pqz))(q - C(pqz)) = pq - C(pqz) + C(pqz)^2 = pq(1-z),$$

because $C(z) - C(z)^2 = z$. Hence the denominator of $M(z)$ can be vastly simplified:

$$M(z) = \frac{z}{(1-z)^2}\left(\frac{q - C(pqz)}{q}\right).$$

This is the product $(z + 2z^2 + 3z^3 + \cdots)\cdot(1 - c_1pz - c_2p^2qz^2 - c_3p^3q^2z^3 - \cdots)$, so the coefficient of z^n can be written

$$M_n(p) = n - (n-1)c_1p - (n-2)c_2p^2q - \cdots - 1\cdot c_{n-1}p^{n-1}p^{n-2}.$$

When a formula turns out to be so simple, it must have a simple explanation. But the author hasn't been able to think of any direct proof. For some reason, $M_n(p)$ is not only the expected size of the remaining roll when one roll empties, it is also the expected value of the "first return to the diagonal," in the following sense: Suppose the two toilet paper rolls start in the full state (n, n), and they are used by big-choosers and little-choosers until the empty state $(0,0)$ is reached; and suppose that the rolls first become equal in size again at state $(n-k, n-k)$. Then the average value of k is $M_n(p)$. (This follows from our formula for $M_n(p)$, because $c_k p^k q^{k-1}$ is the probability of first return to $(n-k, n-k)$ for each $k < n$, and $1 - c_1 p - \cdots - c_{n-1} p^{n-1} q^{n-2}$ is the probability that the diagonal is not encountered until state $(0,0)$ is reached.)

Is there an easy way to prove that the same expected value occurs in both problems? The distributions are different, but the mean values are the same.

5. The limiting behavior. Now that $M(z)$ has been put into a fairly simple form, we are ready to deduce the asymptotic value of $M_n(p)$ for fixed p as $n \to \infty$.

Let's assume first that $p \neq q$. Then $4pq < 1$, and the function $C(pqz) = \frac{1}{2}(1 - \sqrt{1 - 4pqz})$ is analytic for $|z| < 1/(4pqz)$; so it is analytic in a neighborhood of $z = 1$. In fact, a simple computation proves that its Taylor series at the point $z = 1$ involves the Catalan numbers once again:

$$C(pqz) = \min(p,q) + (\max(p,q) - \min(p,q))C\left(\frac{pq(z-1)}{(p-q)^2}\right).$$

(This formula generalizes our previous observation that $C(pq) = \min(p,q)$.)

If $q < p$, our formula for $M(z)$ reduces to

$$\begin{aligned} M(z) &= \frac{z}{(1-z)^2}\frac{q-p}{q}C\left(\frac{pq(z-1)}{(p-q)^2}\right) \\ &= \frac{z}{1-z}\frac{p}{p-q} - z\left(c_2\frac{p^2q}{(p-q)^3} + c_3\frac{p^3q^2}{(p-q)^5}(z-1) + \cdots\right) \\ &= \frac{z}{1-z}\frac{p}{p-q} + f(z), \end{aligned}$$

where $f(z)$ is analytic in the region $|z| < 1/(4pq)$. This determines the value of $M_n(p)$ quite accurately:

THEOREM 1. *Let r be any value greater than $4pq$. Then*

$$M_n(p) = \begin{cases} p/(p-q) + O(r^n), & \text{if } q < p; \\ ((q-p)/q)n + p/(q-p) + O(r^n), & \text{if } q > p. \end{cases}$$

(The constants implied by O in these formulas depend on p and r, but not on n.)

Proof. If $q < p$, the value of $M_n(p)$ is the coefficient of z^n in $M(z)$, which is $p/(p-q)$ plus the coefficient of z^n in $f(z)$. But $f(z)$ converges absolutely when $z = 1/r$, hence its nth coefficient is $O(r^n)$.

If $q > p$, the stated result follows from the formula for $q < p$, using the identity

$$qM_n(p) + pn = pM_n(q) + qn$$

which is an immediate consequence of the formula for $M_n(p)$ in section 4. QED.

For example, if $p = 2/3$ and $q = 1/3$, so that big-choosers outnumber little-choosers by 2 to 1, the average size of the remaining roll will be very close to 2, when n is large; but when $p = 1/3$ and $q = 2/3$ the average will be approximately $\frac{1}{2}n + 1$.

This agrees with our intuition: If little-choosers predominate, the size of the larger roll will tend to be proportional to n, when the smaller roll is used up. But if big-choosers are in the majority, the larger roll will tend to be reduced to a bounded size, independent of the initial size n.

6. The transition point. But what about the boundary case, $p = q$? Does it lead to lengths of order n, or order 1, or something in between?

This is actually the simplest case to analyze, because $p = q = 1/2$ is equivalent to saying that everybody is a random-chooser; the problem reduces to a fairly simple "random walk." In fact, we are essentially dealing here with "Banach's match box problem" as discussed by Feller [**2**, IX.3(f)]. According to our general formula, the generating function in this case is simply

$$M(z) = \frac{z}{(1-z)^{3/2}},$$

so there is a solution in closed form:

$$M_n\left(\frac{1}{2}\right) = \binom{-3/2}{n-1}(-1)^{n-1} = \frac{2n}{4^n}\binom{2n}{n}.$$

By Stirling's approximation we have the following result:

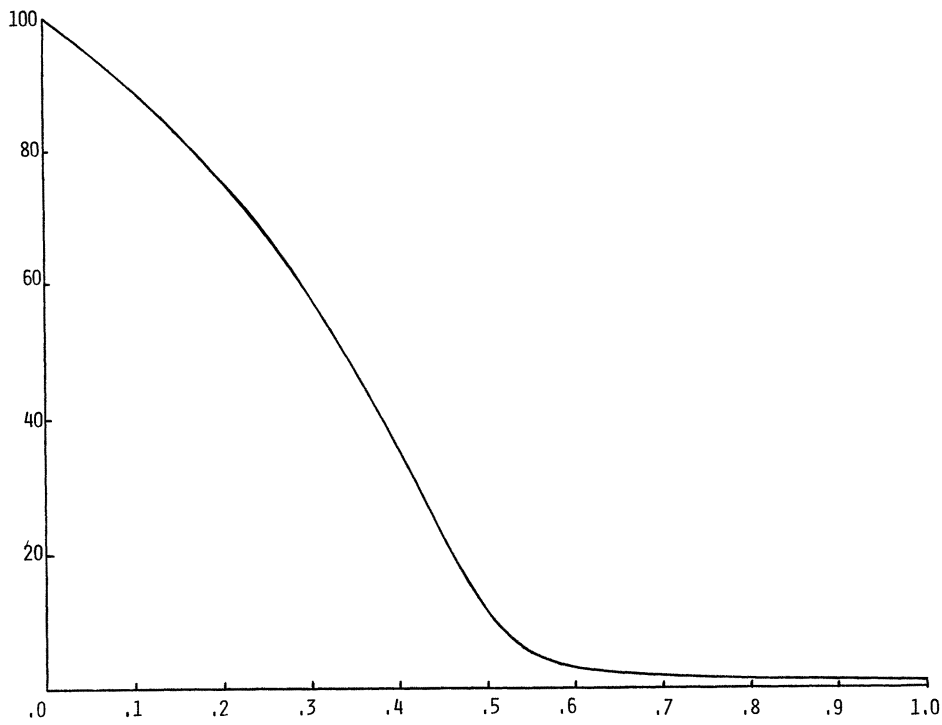
THEOREM 2.

$$M_n(p) = 2\sqrt{\frac{n}{\pi}} - \frac{1}{4}\sqrt{\frac{1}{\pi n}} + O(n^{-3/2}),$$

when $p = q$.

The function $M_n(p)$ is a polynomial in p of degree $2n - 3$, for $n \geqslant 2$, and it decreases monotonically from n down to 1 as p increases from 0 to 1. The remarkable thing about this decrease is that it changes in character rather suddenly when p passes $1/2$.

We can't use the formulas of Theorem 1 when p is too close to $1/2$, even if n is extremely



large. For example, if $n = 10^{10}$ and $p = \frac{1}{2} \pm 10^{-20}$, both approximations in Theorem 1 give the ridiculous estimate $M_n(p) \approx \frac{1}{4} \times 10^{20}$. Indeed, we know that $M_n(1/2)$ is of order $\sqrt{n}$, so the approximations can be valid only when $|p - \frac{1}{2}|$ is of order $1/\sqrt{n}$ at least.

The slope of $M_n(p)$ at $p = 1/2$ can be calculated by differentiating $M(z)$ with respect to p and extracting the coefficient of z^n. The derivative is

$$-\frac{z}{(1-z)^2}\frac{d}{dp}\left(\frac{C(p(1-p)z)}{1-p}\right) = -\frac{z}{(1-z)^2} \times\left(\frac{(1-2p)zC'(p(1-p)z)}{1-p} + \frac{C(p(1-p)z)}{(1-p)^2}\right)$$

and at $p = 1/2$ this equals $-2z(1-z)^{-2} + 2z(1-z)^{-3/2}$. Hence

$$M_n'(1/2) = -2n + 2M_n(1/2);$$

this is consistent with $M_n(p)$ dropping from n to a small value as p goes from 0 to 1/2. The graph of $M_{100}(p)$ is shown on page 469.

Acknowledgements. I wish to thank the architect of the computer science building at Stanford University for implicitly suggesting this problem, and Richard Beigel for the insight that led to the closed form of $L(z)$. Andrei Broder offered helpful comments and computed the graph of $M_{100}(p)$. My research was supported in part by National Science Foundation grant MCS-83-00984.

References

1. Désiré André, Solution directe du problème résolu par M. Bertrand, Comptes Rendus, Acad. Sci. Paris, 105 (1887) 436–437.
2. William Feller, An Introduction to Probability Theory and its Applications, vol. 1, 2nd ed., Wiley, New York, 1957.
3. Donald E. Knuth, The Art of Computer Programming, vol. 1: Fundamental Algorithms, Addison-Wesley, Reading, MA, 1968, xxii + 634 pp.

THE LOGIC OF PROVABILITY

GEORGE BOOLOS
Department of Linguistics and Philosophy, Massachusetts Institute of Technology, Cambridge, MA 02139

The subject of this article is the way in which an ancient branch of logic, first investigated by Aristotle and known as *modal logic*, has recently been found to shed light on a branch of logic of much later date, the mathematical study of mathematics itself, a study begun by David Hilbert and brought to fruition by Kurt Gödel.

The fundamental concepts studied in modal logic are those of *necessity* and *possibility*: a statement is called "necessary" if it *must* be true, and "possible" if it *might* be true. Thus, since there might be a war in the year 2000, the statement "there will be a war in 2000" is possible, but it is not necessary, as there might not be a war then. On the other hand, the statement "there will

George Boolos: After an undergraduate degree in mathematics at Princeton, where my supervisor was Raymond Smullyan, and a graduate degree in philosophy at Oxford, I became the first person ever to receive a Ph.D. in philosophy from MIT, writing a thesis on hierarchy theory under Hilary Putnam. I taught for three years at Columbia and in 1969 returned to MIT, where I am now a professor of philosophy. In addition to a book on the topic of this article, *The Unprovability of Consistency*, and a textbook, *Computability and Logic* (co-authored with Richard Jeffrey), I have written a number of articles in logic and philosophy.