AN ASYMPTOTIC STUDY OF A RECURSION OCCURRING IN THE ANALYSIS OF AN ALGORITHM ON BROADCAST COMMUNICATION

PETER J. GRABNER AND HELMUT PRODINGER

Abstract

In [7] it was proved that a certain quantity T_n occurring in the analysis of an algorithm on broadcast communication satisfies $4 \ll \frac{T_n}{\log n} \ll 5$. We give an explicit expression for T_n involving Bernoulli numbers and a precise asymptotic formula showing in particular that $\frac{T_n}{\log n} \to \frac{\pi^2}{3\log 2} = 4.74627644\ldots$.

In [7] a maximum finding algorithm was presented that we rephrase in our own words for the reader's convenience.

Assume that each of n persons thinks about a number, and the maximum of these should be determined. The party selects a "loser," as in [5], by (recursive) coin flippings. This person's number will be the next current maximum (initially the maximum was set to $-\infty$), and will be officially announced. Then, all people who have a number smaller or equal to the current maximum drop out of the game, and the procedure is repeated with the reduced party. The interest is of course in the expected time T_n that this procedure takes in order to determine the maximum.

For the analysis of the algorithm the authors of [7] come up with the recursion

$$\left(1 - \frac{1}{2^{n-1}}\right) T_n = 1 + \frac{1}{2^n}\left[1 + \sum_{k=1}^{n-1}\binom{n}{k} T_k + \sum_{k=0}^{n-1} 2^{n-1-k} T_k\right] ,$$

($n \ge 2$; $T_0 = T_1 = 1$). Such a recursion is a typical "probabilistic divide–and–conquer recurrence" and directly reflects the recursive structure of the algorithm. As general references, we mention [4] and [6].

By elementary methods it was shown in [7] that $T_n = \Theta(\log n)$. In this note it will be shown that we can achieve much better results, since we will prove that

Theorem 1. *(i) An explicit formula for T_n is given by*

$$T_n = 1 + \sum_{k=0}^{n-1} \frac{2B_k}{1 - 2^{-k-1}}\left[\binom{n}{k+1} - 1\right] ,$$

where B_n is the n–th Bernoulli number.

(ii) For $n \to \infty$ we have

$$T_n = \frac{\pi^2}{3\log 2}\log n + K + \delta(\log_2 n) + O\left(\frac{1}{n}\right).$$

Date: November 6, 1997.

Key words and phrases. Broadcast communication, Recursion, Bernoulli numbers, Rice's method, Mellin–Perron formula.

The constant K is given by

$$K=\frac{\gamma\pi^2}{3\log 2}+1-\frac{\pi^2}{3\log 2}-\frac{2\zeta'(2)}{\log 2}+\frac{\pi^2}{6}-2\sum_{n=1}^{\infty}\frac{\lceil\log_2 n\rceil}{n^2}=.34118435603469732225\dots$$

and $\delta(x)$ is a periodic function with period 1, mean 0, $|\delta(x)|\le 0.852\cdot 10^{-5}$, and the Fourier expansion

$$\delta(x)=\frac{2}{\log 2}\sum_{k\neq 0}\frac{\Gamma(2-\chi_k)\,\zeta(2-\chi_k)}{\chi_k}\,e^{2k\pi ix}\ .$$

Observe that $\frac{\pi^2}{3\log 2}=4.74627644\dots$, which "explains" the elementary bounds $4\ll\frac{T_n}{\log n}\ll 5$ from [7]*.* □

By elementary rearrangements we write the recursion as

$$\left(1-\frac{1}{2^{n-1}}\right)T_n+\frac{1}{2^n}T_n+\frac{1}{2^{n+1}}T_n=1+\frac{1}{2^n}\left[\sum_{k=0}^{n}\binom{n}{k}T_k+\sum_{k=0}^{n}2^{n-1-k}T_k\right]\ ;$$

taking the difference between this equation and the one where n is substituted by $n-1$ we find

$$\left(1-\frac{1}{2^n}\right)(T_n-T_{n-1})=\frac{1}{2^n}\sum_{k=0}^{n-1}\binom{n-1}{k}(T_{k+1}-T_k)\ .$$

It is natural to introduce the auxiliary quantities $U_k=T_{k+1}-T_k$ and replace n by $n+1$ to find

$$\left(1-\frac{1}{2^{n+1}}\right)U_n=\frac{1}{2^{n+1}}\sum_{k=0}^{n}\binom{n}{k}U_k\ .$$

It holds for $n\ge 2$ and $U_0=0$, $U_1=4$. Now such recursions are fairly well understood, see e.g. [6]: One introduces the exponential generating function

$$U(z)=\sum_{n\ge 0}U_n\frac{z^n}{n!}$$

and translates the recursion into the functional equation

$$U(z)=\tfrac12(1+e^{z/2})U(\tfrac z2)+2z\ .$$

This equation can be solved using an old trick of Knuth's, [4, pp. 502, 686 (Ex. 6.3.33)], that was used to analyze a variant of digital search trees called Patricia trees: We define

$$\widehat U(z)=\frac{U(z)}{e^z-1}\tag{1}$$

and get the simpler equation

$$\widehat U(z)=\tfrac12\widehat U(\tfrac z2)+\frac{2z}{e^z-1}\ .$$

Inserting

$$\widehat U(z)=\sum_{n\ge 0}\widehat U_n\frac{z^n}{n!}\ ,$$

into (1) yields

$$\sum_{n\ge 0}\frac{z^n}{n!}\widehat{U}_n\left(1-\frac{1}{2^{n+1}}\right)=2\sum_{n\ge 0}\frac{z^n}{n!}B_n,$$

where B_n denotes the n-th Bernoulli number. Comparing coefficients we deduce

$$\widehat{U}_n\left(1-\frac{1}{2^{n+1}}\right)=2B_n \quad \text{or} \quad \widehat{U}_n=\frac{2B_n}{1-\frac{1}{2^{n+1}}}\ .$$

Thus we have

$$U_n=\sum_{k=0}^{n-1}\binom{n}{k}\frac{2B_k}{1-2^{-k-1}}\ .$$

Summing up this formula, we find the explicit formula for T_n, mentioned in the theorem. It is interesting to compare this expression with "the hardest asymptotic nut" Knuth ever had to crack, [4, p. 499 and Ex. 6.3.34]. In [5], related quantities appear in the context of the election of a loser.

Now we turn to asymptotics. Following the method described in [3] (see also [8, 5]) we can write U_n as a complex contour integral;

$$U_n=\frac{1}{2\pi i}\int_{-\frac12-i\infty}^{-\frac12+i\infty}\frac{\Gamma(n+1)\,\Gamma(1-s)}{\Gamma(n+1-s)}\frac{2\zeta(1-s)}{1-2^{-s-1}}ds\ . \tag{2}$$

Since

$$T_n=1+U_0+U_1+\sum_{m=2}^{n-1}U_m\ ,$$

we find

$$T_n=5-\frac{1}{2\pi i}\int_{-\frac12-i\infty}^{-\frac12+i\infty}\frac{4\zeta(1-s)}{(1-s)(s+1)(1-2^{-s-1})}ds \tag{3}$$

$$+\frac{1}{2\pi i}\int_{-\frac12-i\infty}^{-\frac12+i\infty}\frac{\Gamma(n+1)\,\Gamma(1-s)}{\Gamma(n-s)}\frac{2\zeta(1-s)}{(s+1)(1-2^{-s-1})}ds\ . \tag{4}$$

Observe that the first two terms were taken out in order to make the first integral convergent. We can compute (3) as follows,

$$\frac{1}{2\pi i}\int_{-\frac12-i\infty}^{-\frac12+i\infty}\frac{4\zeta(1-s)}{(1-s)(s+1)(1-2^{-s-1})}ds=\frac{1}{2\pi i}\int_{\frac32-i\infty}^{\frac32+i\infty}\frac{4\zeta(s)}{s(2-s)(1-2^{s-2})}ds \tag{5}$$

$$=\sum_{k\ge 0}\sum_{n\ge 1}4^{1-k}\frac{1}{2\pi i}\int_{\frac32-i\infty}^{\frac32+i\infty}\frac{1}{s(2-s)}\left(\frac{2^k}{n}\right)^s ds\ . \tag{6}$$

Here, we first used the substitution $s:=1-s$ and then the the geometric series and the expansion of $\zeta(s)$, both valid for $\Re z=\frac32$. A simple application of residue calculus, as it is often used in the context of the Mellin–Perron summation formula (see [9, 2]) evaluates the integrals inside the summation:

$$\frac{1}{2\pi i}\int_{\frac{3}{2}-i\infty}^{\frac{3}{2}+i\infty}\frac{1}{s(2-s)}t^s ds = \begin{cases}\frac{1}{2} & \text{for } t\geq 1\\ \frac{t^2}{2} & \text{for } t<1\end{cases}.$$

Now the series according to the first alternative is

$$\sum_{k\geq 0}4^{1-k}\sum_{n=1}^{2^k}\frac{1}{2}=2\sum_{k\geq 0}4^{-k}2^k=2\sum_{k\geq 0}2^{-k}=4$$

and the other one is

$$2\sum_{k\geq 0}\sum_{n>2^k}\frac{1}{n^2}=2\sum_{k=0}^{\infty}(k+1)\sum_{n=2^k+1}^{2^{k+1}}\frac{1}{n^2}=3.0022908154228746627\ldots\ . \tag{7}$$

Notice that in the range of n we have $k+1=\lceil\log_2 n\rceil$, which gives the series in the statement of the theorem. For numerical purposes the expression in (7) is preferred, since Maple computes the partial sums $\sum_{a\leq n\leq b}\frac{1}{n^2}$ efficiently using the Euler–MacLaurin summation formula. Observe that the (slowly converging) integral has been transformed into a (better converging) series. We mention here that another possibility to compute the integral (5) would be to move the line of integration to the right. Here residues at the points $2+\frac{2k\pi i}{\log 2}$ would have to be collected, which would form a slowly convergent sum. This sum corresponds to (7) by replacing $\lceil\log_2 n\rceil$ with $\log_2 n+1+\{\log_2 n\}$ and inserting the Fourier expansion for $\{x\}$ (the fractional part of x). Certainly, the values at $n=2^k$ have to be adjusted accordingly. Since we could not find a closed form expression for this sum and (7) could be computed easily numerically, we preferred this form of the sum.

The strategy to evaluate the second integral is to shift the line of integration to the left, and to collect the residues; they constitute the terms in the asymptotic expansion (see [3]). A similar computation as the one used before to derive the series expansion for the first integral shows that

$$\frac{1}{2\pi i}\int_{-\frac{3}{2}-i\infty}^{-\frac{3}{2}+i\infty}\frac{\Gamma(n+1)\,\Gamma(1-s)}{\Gamma(n-s)}\frac{2\zeta(1-s)}{(s+1)(1-2^{-s-1})}ds=0.$$

At $s=-1$, we find the residue

$$\frac{\pi^2 H_n}{3\log 2}-\frac{\pi^2}{3\log 2}-\frac{2\zeta'(2)}{\log 2}+\frac{\pi^2}{6},$$

with a harmonic number $H_n=\sum_{1\leq k\leq n}\frac{1}{k}=\log n+\gamma+O(\frac{1}{n})$. At $s=-1+\chi_k$ with $\chi_k=\frac{2k\pi i}{\log 2}$ (for $k\in\mathbb{Z}$, $k\neq 0$) we find the residue

$$\frac{\Gamma(n+1)\,\Gamma(2-\chi_k)}{\Gamma(n+1-\chi_k)}\frac{2\zeta(2-\chi_k)}{\chi_k\log 2};$$

collecting them gives the *explicit formula*

$$T_n=\frac{\pi^2}{3\log 2}H_n+\left(K-\frac{\gamma\pi^2}{3\log 2}\right)+\sum_{k\in\mathbb{Z}\setminus\{0\}}\frac{\Gamma(n+1)\,\Gamma(2-\chi_k)}{\Gamma(n+1-\chi_k)}\frac{2\zeta(2-\chi_k)}{\chi_k\log 2}.$$

The periodic function mentioned in the theorem can be computed from the asymptotic expansion of the quotients of Γ-functions (compare [1])

$$\frac{\Gamma(n+1)}{\Gamma(n+1-\chi_k)} = n^{\chi_k}\Big(1 + O\Big(\frac{k^2}{n}\Big)\Big) .$$

Combining known estimates for the Γ- and ζ-function (cf. [1]) yields

$$\frac{2\Gamma(2-\chi_k)\,\zeta(2-\chi_k)}{\chi_k \log 2} \le 13|k|^{\frac{1}{2}} e^{-\frac{\pi^2}{\log 2}|k|} \quad \text{for } |k| \ge 1$$

which (by uniform convergence) justifies interchanging summation and asymptotics. Furthermore, this yields the estimate for $\delta(x)$ stated in the theorem. By taking more terms in the above formula and in the asymptotic expansion of the harmonic numbers, a remainder of the form $O(\frac{1}{n^m})$ for any $m > 0$ could be achieved. Consequently, one could give as many terms as desired in the asymptotic formula of the theorem.

References

[1] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions.* Dover, 1973. A reprint of the tenth National Bureau of Standards edition, 1964.

[2] P. Flajolet, P. Grabner, P. Kirschenhofer, H. Prodinger, and R.F. Tichy. Mellin transforms and asymptotics: Digital sums. *Theor. Comput. Sci.*, 123:291–314, 1994.

[3] P. Flajolet and R. Sedgewick. Mellin transforms and asymptotics: Finite differences and Rice's integrals. *Theor. Comput. Sci.*, 144:101–124, 1995.

[4] D. E. Knuth. *The Art of Computer Programming*, volume 3: Sorting and Searching. Addison-Wesley, 1973.

[5] H. Prodinger. How to select a loser. *Discrete Math.*, 120:149–159, 1993.

[6] R. Sedgewick and P. Flajolet. *An Introduction to the Analysis of Algorithms.* Addison-Wesley, 1996.

[7] S.-H. Shiau and C.-B. Yang. A fast maximum finding algorithm on broadcast communication. *Inf. Process. Lett.*, 60:81–89, 1996.

[8] W. Szpankowski. The evaluation of an alternating sum with applications to the analysis of algorithms. *Inf. Process. Lett.*, 28:13–19, 1988.

[9] G. Tenenbaum. *Introduction to analytic and probabilistic number theory.* Cambridge University Press, 1995.

Institut für Mathematik. Technical University of Graz. Steyrergasse 30, A-8010 Graz, Austria.

Institut für Algebra und Diskrete Mathematik. Technical University of Vienna. Wiedner Hauptstrasse 8–10, A-1040 Vienna, Austria.

E-mail address: grabner@weyl.math.tu-graz.ac.at, Helmut.Prodinger@tuwien.ac.at