A SIMPLE APPLICATION OF INTERVAL ARITHMETIC *

SIEGFRIED M. RUMP†

Abstract. In a recent paper by Campos and Mendoza an explicit formula was given for the limit probability distribution of the leading digit of a^n. Computations for $1 \le n \le N$, N large require a multiple precision arithmetic and are very slow. We use a method proposed by Campos and Mendoza to perform computations in ordinary floating point. With the help of interval arithmetic all results are rigorous.

1. The problem and its solution. For the following let fixed integers $a, b \ge 2$ be given. Denote the leading digit in the b-adic expansion of a^n by $\text{ldigit}(a^n, b)$. That means for

$$a^n = \sum_{\nu=m}^{-\infty} \beta_\nu \cdot b^\nu \quad \text{it is } \text{ldigit}(a^n, b) = \beta_m. \tag{1.1}$$

The usual conventions apply, i.e. $0 \le \beta_\nu < b$, $\beta_m \ne 0$ and $|\{\nu : \beta_\nu = b-1\}| < \infty$. This implies that the representation (1.1) is unique. Define

$$\varphi_N(k) := |\{1 \le n \le N : \quad \text{ldigit}(a^n, b) = k\}|.$$

Then obviously $\sum\limits_{k=1}^{b-1} \varphi_N(k) = N$, and in [3] it has been shown that

$$\lim_{N\to\infty} \varphi_N(k)/N = \log_b(1 + 1/k).$$

In order to run some numerical examples we need to compute, for given n, the leading digit of a^n in its b-adic expansion. Fortunately, this can be reduced to a simple test involving logarithms.

LEMMA 1.1. *[3] For given $a, b \ge 2$ and $n \in \mathbb{N}$ it is*

$$\textit{ldigit}(a^n, b) = k \quad \Leftrightarrow \quad \log_b k \le n\alpha - \lfloor n\alpha \rfloor < \log_b(k+1),$$

where $\alpha := \log_b a$.

PROOF. It is $b^m \le a^n < b^{m+1}$ iff $m = \lfloor \log_b a^n \rfloor = \lfloor n\alpha \rfloor$. Therefore

$$\text{ldigit}(a^n, b) = k \quad \text{iff} \quad k \cdot b^m \le a^n < (k+1)b^m \quad \text{iff} \quad \log_b k + \lfloor n\alpha \rfloor \le n\alpha < \log_b(k+1) + \lfloor n\alpha \rfloor. \qquad \blacksquare$$

Therefore, the problem of computing $\varphi_N(k)$ reduces to check

$$n\alpha - \lfloor n\alpha \rfloor \in [\log_b k, \log_b(k+1)] \quad \text{for} \quad 1 \le n \le N. \tag{1.2}$$

Rigorous results are produced if logarithms are computed with error bounds. Problems will occur if $n\alpha - \lfloor n\alpha \rfloor$ is exactly equal to $\log_b k$ because in that case (1.2) cannot be satisfied unless $n\alpha - \lfloor n\alpha \rfloor$ is computed without error. But $n\alpha - \lfloor n\alpha \rfloor = \log k$ implies $a^n = b^m \cdot k$. This is, for example, true if a^n has only digit in the b-adic expansion.

The following Matlab programm [2] uses the interval toolbox Intlab [4] to compute $\varphi_n(N)$. Special care is taken for the case $a^n = k < b$.

*PUBLISHED IN BRAZILIAN ELECTRONIC JOURNAL ON MATHEMATICS OF COMPUTATION (BEJMC), 2, 2000

† Inst. f. Informatik III, Technical University Hamburg-Harburg, Schwarzenbergstr. 95, 21071 Hamburg, Germany (rump@tu-harburg.de).

```
function f(a,b,N)

  p = zeros(b-1,1);
  logb = log(intval(b));
  alpha = log(intval(a))/logb;

  nmin = 1;
  while a^nmin<b
    nmin = nmin+1;
  end

  n = nmin:N;

  for k=1:b-1

    K = intval(k);

    x = n*alpha;
    x = x - floor(x.inf);
    logb_k = log(K)/logb;
    logb_k1 = log(K+1)/logb;

    I = ( logb_k <= x ) & ( x < logb_k1 );
    p(k) = sum(I);

  end

  I = a.^(1:nmin-1);
  p(I) = p(I) + 1;

  p, sum(p)
```

ALGORITHM 1.2. *Rigorous computation of* $\varphi_k(N)$

Finally, we sum the computed values $\varphi_N(k), 1 \leq k \leq b-1$ and check $\sum_{k=1}^{b-1} \varphi_n(k) = N$. Because the check for (1.2) is rigorous, this proves correctness of the result. Otherwise we use Matlab vector notation and think the program is self-explaining.

For $b = 10$, $N = 10^5$ and $a \in \{2, 3\}$ the program produces the following results.

k	$\varphi_N(k)$ for $a = 2$	$\varphi_N(k)$ for $a = 3$	$\log_b(1 + 1/k)$
1	30102	30101	0.30103
2	17611	17611	0.17609
3	12492	12492	0.12494
4	9692	9692	0.09691
5	7919	7917	0.07918
6	6695	6696	0.06695
7	5797	5799	0.05799
8	5116	5116	0.05115
9	4576	4576	0.04576

TABLE 1.3. $\varphi_N(k)$ *for* $a = 2$ *and* $a = 3$

Computing time was 5.1 seconds on a 300 MHz PentiumI Laptop. The numbers coincide with Figure 1 in [1] (except the typo $3^n \to 2^n$). The results in Table 1.3 are completely rigorous and demonstrate the ease of use of interval arithmetic and INTLAB.

REFERENCES

[1] M.A. Campos, A.A.R. de A. Maranhão, and R. Mendoza. Programming and Math, an Alliance that Works.

[2] MATLAB User's Guide, Version 5. The MathWorks Inc., 1997.

[3] R. Mendoza and M.A. Campos. The limit probability distribution of the first digit of a^n. (unpublished manuscript).

[4] S.M. Rump. INTLAB - INTerval LABoratory. In Tibor Csendes, editor, *Developements in Reliable Computing*, pages 77–104. Kluwer Academic Publishers, 1999.