7. HARARY, F. *Graph theory* (Teoriya grafov), "Mir", Moscow, 1973.

8. BERGE, C. *The theory of graphs and its applications* (Teoriya grafov i ee primenenie), Izd-vo in. lit., Moscow, 1962.

9. SACHKOV, V. N. *Combinatorial methods of discrete mathematics* (Kombinatornye metody diskretnoi matematiki), "Nauka", Moscow, 1977.

10. HU, T. C. *Integer programming and network flows* (Tselochislennoe programmirovanie i potoki v setyakh), "Mir", Moscow, 1974.

11. PRIM, R. K. The shortest connecting networks and some generalizations. In: *Cybernetic collection* (Kibernetich. sb.), No. 2, 95–107, Izd-vo in. lit., Moscow, 1961.

12. COOKE, K. L. and HALSEY, E. The shortest path through a network with variable time of travelling along the arcs. In: *Cybernetic collection* (Kibernetich. sb.), No. 5, 189–194, "Mir", Moscow, 1968.

13. ADEL'SON-VEL'SKII, G. M., DINITS, E. A. and KARZANOV, A. V. *Flow algorithms* (Potokovye algoritmy), "Nauka", Moscow, 1975.

14. HOPCROFT, J. E. and KARP, R. M. A $n^{5/2}$ algorithm for maximum matchings in bipartite graphs. *SIAM J. Comput.*, 2, 4, 225–231, 1973.

U.S.S.R. Comput. Maths. Math. Phys. Vol. 19, pp. 207–212 0041-5553/79/0601-0207$07.50/0
© Pergamon Press Ltd. 1980. Printed in Great Britain.

SHORT COMMUNICATIONS

SOME ESTIMATES CONNECTED WITH EUCLID'S ALGORITHM*

S. A. ABRAMOV

Moscow

(*Received* 3 *May* 1978)

THE NUMBER of divisions in Euclid's algorithm can be overbounded by the number of digits in writing down the lesser of two given numbers in the position system of calculation with some base q. The dependence of the properties of these estimates on the value of q is investigated. The release of the memory in the process of using Euclid's algorithm is also studied.

1. Let the application of Euclid's algorithm to the natural numbers $a_0, a_1, a_0 > a_1$, require m divisions, including the last division giving a zero remainder. We denote the resulting partial quotients by $q_1, \dots, q_m$, and the non-zero remainders by $a_2, \dots, a_m$. We put $a_{m+1} = 0$. By a well-known theorem of Lamé [1],

$$a_1 \geqslant u_{m+1}, \tag{1}$$

where u_{m+1} is the $(m + 1)$-th Fibonacci number. From the inequality (1) it is easy to deduce

$$m+1<\log_{(1+\sqrt{5})/2}(1+a_1)+\log_{(1+\sqrt{5})/2}\sqrt{5}=\log_{(1+\sqrt{5})/2}(1+a_1)+1.672275\dots \tag{2}$$

*Zh. vȳchisl. Mat. mat. Fiz., 19, 3, 756–760, 1979.

For the practical application of the estimate (2) we replace the logarithm by the number of digits of the natural number a_1 in some position system. If an integer q greater than unity is the base of the system, then we obtain

$$m \leqslant ([\log_{(1+\sqrt{5})/2} q]+1)([\log_q a_1]+1) = ([\log_{(1+\sqrt{5})/2} q]+1)\, n_q(a_1), \tag{3}$$

where $n_q(a_1)$ is the number of digits of a_1 in the q-ary system.

For the most widespread systems – binary and denary – we obtain the following:

$$\log_{(1+\sqrt{5})/2} 2 = 1.440\ldots, \qquad m \leqslant 2n_2(a_1),$$
$$\log_{(1+\sqrt{5})/2} 10 = 4.784\ldots, \qquad m \leqslant 5n_{10}(a_1).$$

To simplify the writing we use the notation $c_q = [\log_{(1+\sqrt{5})/2} q]+1$. The estimate (3) is written in the form $m \leqslant c_q n_q(a_1)$.

For any q the estimate (3) is cruder than (2). Choosing a criterion of quality of the estimate, we can consider which of the estimates (3) is preferable for two different values of q. The fact that, having chosen two values of q, we can obtain estimates greatly differing from each other, is already demonstrated by comparing the estimates given for $q = 2$ and $q = 10$.

For fairly large a_1 the denary estimate is better than the binary estimate. But the binary estimate admits of localization in the following sense: for $i = 1, 2, \ldots, m-2$ we have $a_i \geqslant 2a_{i+2}$ (and even $a_i > 2a_{i+2}$; proof: $a_i = q_{i+1}a_{i+1} + a_{i+2} \geqslant a_{i+2} + a_{i+2} > 2a_{i+2}$). But the statement $a_i \geqslant 10a_{i+5}$ is in general untrue (example $a_0 = 49, a_1 = 30, i = 1$). The binary estimate can be improved: indeed the inequality $m < 2n_2(a_1)$ is satisfied (this will be proved below; but the example $a_0 = 13$ shows that in the denary estimate it is impossible to replace the sign $\leqslant$ by $<$.

It is easy to see that for fairly large a_1 we have $c_q n_q(a_1) < c_p n_p(a_1)$, if $(1 - \{\log_{(1+\sqrt{5})/2} q\})/\log_{(1+\sqrt{5})/2} q$ is less than the similar quantity for p (in this case for fairly large a_1 the inequality $c_q(\log_q a_1 + 1) < c_p \log_p a_1)$ will be satisfied). Using this we arrange the values of $c_2, \ldots, c_{10}$ in order of decrease of accuracy of the estimate $m \leqslant c_q n_q(a_1)$ for fairly large a_1: $c_4=3$, $c_{10}=5$, $c_9=5$, $c_6=4$, $c_8=5$, $c_5=4$, $c_7=5$, $c_3=3$, $c_2=2$.

It is clear that for any q we can find a Q such that for any natural $p > Q$ the estimates $m \leqslant c_p n_p(a_1)$ is for large a_1 more accurate than $m \leqslant c_q n_q(a_1)$.

It is not so simple in the general case to establish the possibility of localizing and improving the estimates. In this section it will be shown that the class of estimates of the form (3) admitting of localization, is identical with the class of estimates of the same form admitting of improvement. We will formulate the necessary and sufficient conditions which q must satisfy for the corresponding estimates to belong to this class.

By induction using the equation $a_i = q_{i+1}a_{i+1} + a_{i+2}$ for $i = 1, 2, \ldots, m$ it can be proved that

$$a_1 = Q_{i-1}(q_2, \ldots, q_i)a_i + Q_{i-2}(q_2, \ldots, q_{i-1})a_{i+1}, \tag{4}$$

where $Q_0, Q_1, \ldots$ are the Eulerian polynomials:

$$Q_0=1, \qquad Q_1(x_1)=x_1,$$
$$Q_n(x_1,\dots,x_n)=x_nQ_{n-1}(x_1,\dots,x_{n-1})+Q_{n-2}(x_1,\dots,x_{n-2}).$$

Since for any collection of natural $q_2,\dots,q_i$ we have $Q_{i-1}(q_2,\dots,q_i)\geqslant Q_{i-1}(1,\dots,1)=u_i$, we obtain as a corollary that for $i=1,2,\dots,m$ the inequality $a_1\geqslant u_ia_i+u_{i-1}a_{i+1}$ holds and all the more

$$a_1\geqslant u_ia_i. \tag{5}$$

Remark. If $a_0=u_{m+2}$, $a_1=u_{m+1}$, then (4) is transformed into the well-known identity for Fibonacci numbers:

$$u_{m+1}=u_iu_{m-i+2}+u_{i-1}u_{m-i+1}.$$

Relations (4), (5) permit us to describe all the q for which the estimate (3) permits of localization.

Theorem 1

Let $i\geqslant 2$. For every real $\epsilon>0$ natural numbers a_0 and a_1 can be chosen such that Euclid's algorithm applied to them requires not less than $i-1$ divisions and thereby $a_1/a_i<u_i+\epsilon$ will be satisfied.

Proof. We choose N so large that $u_{i-1}/N<\epsilon$ is satisfied, and consider the sequence of numbers defined recursively:

$$v_1=1, \qquad v_2=N, \qquad v_n=v_{n-1}+v_{n-2}.$$

We now put $a_0=v_{i+2}$, $a_1=v_{i+1}$. We have $q_1=\dots=q_{i-1}=1$ and by (4),

$$a_1=u_ia_i+u_{i-1} \text{ and } \frac{a_1}{a_i}=u_i+\frac{u_{i-1}}{N}<u_i+\varepsilon.$$

Comparison of the statement of Theorem 1 with inequality (5) permits two corollaries to be obtained.

Corollary 1. For $i\geqslant 2$ and for any a_0, a_1 such that Euclid's algorithm does not end after $i-1$ divisions the inequality $a_1/a_i\geqslant u_i$ holds. At the same time it is possible to choose a_0 and a_1 such that Euclid's algorithm does not end after $i-1$ divisions and $a_1/a_i<u_i+1$.

Corollary 2. The estimate for the number of divisions $m\leqslant c_qn_q(a_1)$ admits of localization, that is, $a_1/a_{c_q+1}\geqslant q$ if and only if

$$u_{c_q+1}\geqslant q. \tag{6}$$

Returning to the estimate $m\leqslant 5n_{10}(q_1)$, we see that it does not permit localization: $u_6=8<10$. The estimate $m\leqslant 2n_2(a_1)$ permits localization since $u_3=2$. We give a list of all the values of q not exceeding 10 for which estimate (3) permits localization: 2, 3, 5, 7, 8.

Theorem 2

The estimate $m \leqslant c_q n_q$ permits localization if and only if for some natural t

$$\left(\frac{1+\sqrt{5}}{2}\right)^{t-1} < q \leqslant u_{t+1}.$$

In this case $c_q = t - 1$, that is, $m \leqslant (t-1)n_q(a_1)$.

The proof follows from corollary 2 and the inequality

$$\left(\frac{1+\sqrt{5}}{2}\right)^{t-1} < u_{t+1} < \left(\frac{1+\sqrt{5}}{2}\right)^{t},$$

valid for all natural t.

Theorem 3

Let the estimate $m \leqslant c_q n_q(a_1)$ not permit localization. Then it is unimprovable in the following sense: we can choose a_0, a_1 such that the application to them of Euclid's algorithm requires exactly $c_q n_q(a_1)$ divisions.

Proof. By Theorem 2 we can find a natural t such that

$$u_t < q < \left(\frac{1+\sqrt{5}}{2}\right)^{t-1}.$$

We choose $a_0 = u_{t+1}$, $a_1 = u_t$. Then $n_q(a_1) = 1$. $c_q = [\log_{(1+\sqrt{5})/2} q] + 1 = t-1$; the number of divisions is $t-1$.

Theorem 4

Let the estimate $m \leqslant c_q n_q(a_1)$ permit localization. Then it is improvable: the inequality $m < c_q n_q(a_1)$ is satisfied.

Proof. Let $q^{s-1} \leqslant a_1 < q^s$ for some natural s. We proceed by induction with respect to s:

1) $s = 1$; using (6) and (1), we obtain

$$u_{c_q+1} \geqslant q > a_1 \geqslant u_{m+1},$$

from which $c_q > m$, which is what is required;

2) $s > 1$; if $c_q > m$, then there is nothing to prove; otherwise, since the estimate permits localization, we have $a_{c_q+1} < q^{s-1}$.

By the inductive hypothesis

$$m - c_q < c_q n_q(a_{c_q+1}),$$

from which we obtain

$$m<c_q(n_q(a_{c_q+1})+1)\leqslant c_q n_q(a_1).$$

The proof is complete.

2. From the inequality $a_i > 2a_{i+2}$ it follows that the passage from the pair of numbers a_i, a_{i+1} to the pair a_{i+1}, a_{i+2} reduces the total number of digits in the binary formula for the numbers by at least 1, that is, at each step of Euclid's algorithm clearly 1 bit of the memory carrying out this algorithm is freed. The memory freed can be occupied by other calculations. Some algorithms are known which are executed step by step in parallel with Euclid's algorithm, for example the construction for two integers a_0 and a_1 of integers s and t such that $a_0s + a_1t =$ GCD (a_0, a_1). To obtain s and t we construct in succession $s_0, t_0; s_1, t_1; \ldots$ such that $a_0s_i + a_1t_i = a_i$. It is obvious that $s_0 = t_1 = 1, s_1 = t_0 = 0$ and that for $i = 2, 3, \ldots, m$ we have

$$s_i=(-1)^iQ_{i-2}(q_2,\ldots,q_{i-1}),\qquad t_i=(-1)^{i+1}Q_{i-1}(q_1,\ldots,q_{i-1}).$$

A particular case of this algorithm is the algorithm for the inversion of a natural a in the field of residues with respect to the prime modulus p for $a<p$: $ps + at = 1$, implies $at \equiv 1 \pmod p$, here the value of s is of no interest.

We estimate the memory capacity necessary for the application of the last algorithm. Let the binary system be chosen for recording the numbers. It will be shown that $2n_2(p) + 4$ places are sufficient to solve the problem. Retaining the system of notation introduced at the beginning of the paper, we consider that $a_0 = p$, $a_1 = a$. Theorem 5 plays a basic role in deriving the estimate.

Theorem 5

For $i = 0, 1, \ldots, m$ the inequalities $n_2(|t_i|)+n_2(a_i)\leqslant n_2(a_0)+1$, $n_2(|t_i|)+n_2(q_i)+n_2(a_{i+1})\leqslant n_2(a_0)+2$ are satisfied.

Proof. Using the fact that $a_0=Q_m(q_1,\ldots,q_m)$, $|t_i|=Q_{i-1}(q_1,\ldots,q_{i-1})$, $a_{i+1}=Q_{m-i-1}(q_{i+2},\ldots,q_m)$, we have

$$n_2(|t_i|)+n_2(a_i)=([\log_2|t_i|]+1)+([\log_2 a_i]+1)$$
$$\leqslant\log_2|t_i|a_i+2=\log_2 Q_{i-1}(q_1,\ldots,q_{i-1})Q_{m-i}(q_{i+1},\ldots,q_m)+2<\log_2 a_0+2.$$

Since $n_2(|t_i|) + n_2(a_i)$ is an integer, $n_2(|t_i|)+n_2(a_i)\leqslant[\log_2 a_0]+2=n_2(a_0)+1$ is satisfied. Also

$$n_2(|t_i|)+n_2(q_i)+n_2(a_{i+1})\leqslant\log_2|t_i|q_ia_{i+1}+3$$
$$=\log_2 Q_{i-1}(q_1,\ldots,q_{i-1})q_iQ_{m-i-1}(q_{i+2},\ldots,q_m)\leqslant\log_2 a_0+3,$$

and as before we obtain

$$n_2(|t_i|)+n_2(q_i)+n_2(a_{i-1})\leqslant[\log_2 a_0]+3=n_2(a_0)+2.$$

The theorem is proved.

Dividing "at the corner" a_i by a_{i+1} it is possible to obtain the partial quotient q_{i+1} and the remainder a_{i+2} at the place where a_i was situated: the places occupied by a_i are obviously sufficient for this. For the calculation of $|t_{i+2}|=q_{i+1}|t_{i+1}|+|t_i|$ we use the location occupied by $|t_i|$ and q_{i+2}: the digits of the number q_{i+1} beginning with the lowest in the calculation of the product "by the column" become unused one after the other. If the storage of the algorithm being satisfied is this order sequence of binary digits, then it is desirable that each of the numbers considered occupy several consecutive places. For even i the numbers considered may be arranged as follows:

$$\underbrace{|\overleftarrow{t_i}|\,\alpha\overrightarrow{a_i}}_{\substack{n_2(a_0)+2\\ \text{places}}}\ \underbrace{\overleftarrow{a}_{i+1}\,\alpha\,|\overrightarrow{t_{i+1}}|}_{\substack{n_2(a_0)+2\\ \text{places}}} \tag{7}$$

The superscript arrow $\rightarrow$ means that the digits of the number go in succession (from left to right) beginning with the highest order digit, the arrow $\leftarrow$ denotes the opposite order; the letter α everywhere denotes groups of figures of no interest. For odd i the quantities with subscripts i and $i + 1$ change places.

The transformation of the content of the first $n_2\,(a_0) + 2$ binary digits in (7) for the transition from quantities with subscript i to quantities with subscript $i + 2$ proceeds as follows:

$$|\overleftarrow{t_i}|\alpha\overrightarrow{q_{i+1}}\alpha\overrightarrow{a_{i+2}}$$

$$|\overleftarrow{t_i}|\alpha\overleftarrow{q_{i+1}}\overrightarrow{a_{i+2}}$$

$$|\overleftarrow{t_{i+2}}|\alpha\overrightarrow{a_{i+2}}$$

Theorem 5 guarantees that there are sufficient places for this transformation.

If the last value of $|\,t_m\,|$ is on the right (that is, m is odd), then t_m is a negative number and to find the element inverse to a it is also necessary to calculate $|\,t_m\,|$ from p. However, this fact by no means compels us to remember the value of p to the end of the calculations: it is easy to see that $p = t_{m+1}$.

The author thanks V. D. Podderyugin for advice.

Translated by J. Berry.

REFERENCES

1. KNUTH, D. E. *The art of computer programming* (Iskusstvo programmirovaniya dlya EVM), Vol. 2, "Mir", Moscow, 1977.