Discrete Mathematics 23 (1978) 115–119.


# ADDITION CHAINS WITH MULTIPLICATIVE COST*

R.L. GRAHAM
*Bell Laboratories, Murray Hill, NJ 07974, U.S.A.*

A.C.-C. YAO
*Massachusetts Institute of Technology, Cambridge, MA 02139, U.S.A.*

F.-F. YAO
*Brown University, Providence, Rhode Island, U.S.A.*

Received 30 December 1975
Revised 20 February 1978

If each step in an addition chain is assigned a cost equal to the product of the numbers at that step, "binary" addition chains are shown to minimize total cost.

## Introduction

For a positive integer $n$, by a *chain to* $n$ we mean a sequence $C = ((a_1, b_1), (a_2, b_2), \ldots, (a_r, b_r))$ where $a_k$ and $b_k$ are positive integers satisfying:

(i) $a_r + b_r = n$,

(ii) for all $k$, either $a_k = 1$ or $a_k = a_i + b_i$ for some $i < k$, with the same also holding for $b_k$.

The *cost* of $C$, denoted by $\$(C)$, is defined by

$$\$(C) = \sum_{k=1}^{r} a_k b_k.$$

The minimum cost required among all chains to $n$ is denoted by $f(n)$. (In the case of ordinary addition chains $\$(C)$ is just equal to $r$; e.g., see [1].) A few small values of $f(n)$ are given in Table 1.

Table 1

| $n=$ 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| $f(n)=$ 0 | 1 | 3 | 5 | 9 | 12 | 18 | 21 | 29 | 34 |

The function $f$ arises in connection with determining the optimal multiplication

* The work on this paper was done by all three authors while visiting Stanford University, Stanford, CA 94306. Partially supported by National Science Foundation grant DCR 72-03752 A02, by the Office of Naval Research contract NR 044-402, and by IBM Corporation. Reproduction in whole or in part is permitted for any purpose of the United States Government.

chain for computing the $n$th power of a number by ordinary multiplication. If a number $x$ has $d$ digits, then computing $x^{a_k}$ from $x^{a_i}$ and $x^{b_i}$ requires $(a_i b_i) \cdot d^2$ digitwise multiplications in general. Let $g$ be defined by

$$\left.\begin{aligned} &g(1)=0, \\ &g(2n)=g(n)+n^2 \\ &g(2n+1)=g(n)+n^2+2n \end{aligned}\right\} \quad n\geqslant 1$$

It was conjectured by McCarthy [2] that $f(n)=g(n)$ for all $n$. In this note we prove his conjecture.

## Two properties of $g$

We first establish several facts concerning the function $g$ which will be used later.

**Fact 1.** For $m, t\geqslant 0$ with $m$ odd we have

$$g(2^t m)-g(2^t m-1)=t+m-1. \tag{1}$$

**Proof.** For $t=0$, (1) follows at once from the definition of $g$. Assume $t>0$. Then

$$\begin{aligned} g(2^t m)&=g(2^{t-1}m)+(2^{t-1}m)^2, \\ g(2^t m-1)&=g(2^{t-1}m-1)+(2^{t-1}m-1)^2+2(2^{t-1}m-1) \\ &=g(2^{t-1}m-1)+(2^{t-1}m)^2-1. \end{aligned}$$

Thus

$$g(2^t m)-g(2^t m-1)=g(2^{t-1}m)-g(2^{t-1}m-1)+1$$

and consequently, (1) holds by induction on $t$.

**Fact 2.**

$$g(n)-g(x)\geqslant (n-x)^2+2x-n, \quad \text{for} \quad x+2\leqslant n\leqslant 2x+1. \tag{2}$$

**Proof.** Note that for $n=2x$ and $n=2x+1$, this is just the definition of $g$. The validity of (2) for $x=1,2,3$ is immediate. We assume by induction on $x$ that (2) holds for all values less than some $x>3$. The proof of (2) can be most easily accomplished by splitting it into 4 cases, depending on the parity of $n$ and $x$.

*Case 1.* $n=2N$, $x=2X$.

By hypothesis

$$2X+2\leqslant 2N\leqslant 4X+1$$

i.e.,

$$X+1 \leqslant N \leqslant 2X.$$

For $N = X+1$,

$$\begin{aligned} g(2N) - g(2X) &= g(X+1) + (X+1)^2 - g(X) - X^2 \\ &= g(X+1) - g(X) + 2X + 1 \\ &\geqslant 2X + 2 = (2X+2-2X)^2 + 4X - 2(X+1). \end{aligned}$$

by Fact 1 and (2) is proved in this case. For $N \geqslant X+2$, the induction hypothesis applies and

$$\begin{aligned} g(2N) - g(2X) &= g(N) - g(X) + N^2 - X^2 \\ &\geqslant (N-X)^2 + 2X - N + N^2 - X^2 \end{aligned}$$

and so (2) will hold in this case provided

$$(N-X)^2 + N^2 - X^2 + 2X - N \geqslant (2N-2X)^2 + 4X - 2N.$$

However, this equality can be rewritten as

$$(2N-2X-1)(2X-N) \geqslant 0$$

which certainly holds for $X+2 \leqslant N \leqslant 2X$.

The other three cases are similar and will be omitted.

## The main result

**Theorem.** *For all $n$,*

$$f(n) = g(n).$$

**Proof.** It is clear that $f(n) \leqslant g(n)$ for all $n$ since the definition of $g(n)$ determines a unique chain to $n$ with cost $g(n)$. Hence, it will suffice to show that $f(n) \geqslant g(n)$. In fact, it will be enough to establish the following analogue of (2) for $f$:

$$f(n) - f(x) \geqslant (n-x)^2 + 2x - n, \quad \text{for} \quad x+2 \leqslant n \leqslant 2x+1. \tag{2'}$$

For this implies

$$f(2x) - f(x) \geqslant x^2, \qquad f(2x+1) - f(x) \geqslant x^2 + 2x,$$

and so, by induction,

$$\begin{aligned} &f(2x) \geqslant f(x) + x^2 \geqslant g(x) + x^2 = g(2x), \\ &f(2x+1) \geqslant f(x) + x^2 + 2x \geqslant g(x) + x^2 + 2x = g(2x+1). \end{aligned}$$

From Table 1, (2′) certainly holds for $x = 1, 2, 3$. Assume that for some $X > 3$, (2′) holds for all $x < X$ and all $n$ with $x+2 \leqslant n \leqslant 2x+1$. In particular, this implies $f(m) = g(m)$ for $1 \leqslant m \leqslant 2X-1$. Suppose $N$ satisfies $X+2 \leqslant N \leqslant 2X+1$. If

$N \leqslant 2X-1$ then in fact,

$$f(N)-f(X) \geqslant (N-X)^2+2X-N$$

holds by applying (2′) with $x=X-1$. Hence, we are left with the two cases $N=2X$ and $N=2X+1$.

(i) $N=2X$. Suppose the last step in some arbitrary chain $C$ to $N$ is $(a, b)$ with $a+b=N$ and $X \leqslant b < 2X$.

Thus,

$$\$(C) \geqslant f(b)+ab = f(b)+b(2X-b) \geqslant f(X)+X^2$$

since the last inequality is immediate for $b=X$, and follows by induction from (2) for $b \geqslant X+1$. Since $C$ was arbitrary then

$$f(2X) \geqslant f(X)+X^2$$

which is the desired inequality.

(ii) $N=2X+1$. Again, assume the last step in some chain $C$ to $N$ is $(a, b)$ with $a+b=N$ and $X+1 \leqslant b < 2X+1$.

(a) If $b > X+1$ then

$$\begin{aligned}\$(C) &\geqslant f(b)+b(2X+1-b)\\ &\geqslant f(X)+X^2+2X\end{aligned}$$

since

$$f(b)-f(X) \geqslant (b-X)^2+2X-b$$

holds for $X+2 \leqslant b \leqslant 2X-1$ by induction and for $b=2X$ by the preceding case (i).

(b) If $b=X+1$ then $a=X$. Consider the step $(a', b')$ of $C$ for which $a'+b'=b$. We have

$$\begin{aligned}\$(C) &\geqslant f(X)+a'b'+ab\\ &= f(X)+b'(X+1-b')+X^2+X\\ &\geqslant f(X)+X^2+2X\end{aligned}$$

since for $1 \leqslant b' \leqslant X$,

$$b'(X+1-b') \geqslant X.$$

Hence

$$f(2X+1) \geqslant f(X)+X^2+2X.$$

This completes the induction step and the Theorem is proved.

## Concluding remarks

We should note that the optimal chains to $n$ are not unique. This is due to the fact that

$$f(2n+1)=f(n)+n^2+2n$$

can be realized in going from $n$ to $2n+1$ by either

$$(n, n), (2n, 1) \text{ with additional cost } n \cdot n+2n \cdot 1=n^2+2n$$

or

$$(n, 1), (n+1, n) \text{ with additional cost } n \cdot 1+(n+1) \cdot n=n^2+2n.$$

One might consider generalizations of the problem in which the cost of a chain $C=((a_1, b_1), \ldots, (a_r, b_r))$ is given by

$$\$_\lambda(C)=\sum_{k=1}^{r} \lambda(a_k, b_k),$$

where $\lambda$ maps $Z \times Z \to R$. It would be interesting to know for which $\lambda$ the "binary representation" chain to $n$ is always optimal. This is the case for example for $\lambda(x, y)=(x+1)(y+1)$ (see [2]), but it is not the case for $\lambda(x, y)=x+y$.

## References

[1] D.E. Knuth, The Art of Computer Programming, Volume II, Seminumerical Algorithms (Addison-Wesley, Reading, MA, 1969).

[2] D.P. McCarthy, An optimal algorithm to evaluate $x^n$ over integers and polynomials modulo $M$, Mathematics of Computation (to appear).