ADDITION CHAINS WITH MULTIPLICATIVE COST

by

R. L. Graham
A. C-C. Yao
F-F. Yao

STAN-CS-76-540
JANUARY 1976

COMPUTER SCIENCE DEPARTMENT
School of Humanities and Sciences
STANFORD UNIVERSITY

Addition Chains with Multiplicative Cost

by

R. L. Graham */
Bell Laboratories
Murray Hill, New Jersey

A. C-C. Yao */
Massachusetts Institute of Technology
Cambridge, Massachusetts

and

F-F. Yao */
Brown University
Providence, Rhode Island

Abstract

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

*/ The work on this-paper was done by all three authors while visiting Stanford University, Stanford, California 94306. Partially supported by National Science Foundation grant DCR 72-03752 A02, by the Office of Naval Research contract NR044-402, and by IBM Corporation.

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

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

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.

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

Table 1

The function f arises in connection with determining the optimal multiplication 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

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

It was conjectured by D. P. 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 $m, t \geq 0$ with m odd we have

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

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

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

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

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

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.

(2) $$g(n) - g(x) \geq (n-x)^2 + 2x-n \ , \quad \text{for } x+2 \leq n \leq 2x+1 \ .$$

Proof. Note that for $n = 2x$ and $2x+1$, 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

$$2X+2 \leq 2N < 4X+1$$

i.e.,

$$X+1 \leq N < 2X \ .$$

For $N = X+1$,

$$g(2N) - g(2X) = g(X+1) + (X+1)^2 - g(X) - X^2$$
$$= g(X+1) - g(x) + 2X+1$$
$$\geq 2x+2 = (2x+2 - 2X)^2 + 4X-2(X+1).$$

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

$$g(2N) - g(2X) = g(N) - g(X) + N^2 - X^2$$
$$\geq (N-X)^2 + 2X-N + N^2 - X^2$$

and so (2) will hold in this case provided

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

However, this equality can be rewritten as

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

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

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 $f(n) \leq g(n)$ 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 $f(n) \geq g(n)$. In fact, it will be enough to establish the following analogue of (2) for f :

(2') $\quad f(n) - f(x) \geq (n-x)^2 + 2x-n , \quad$ for $x+2 \leq n \leq 2x+1$.

For this implies

$$f(2x) - f(x) \geq x^2 , \quad f(2x+1) - f(x) \geq x^2+2x ,$$

and so, by induction,

$$f(2x) \geq f(x) + x^2 \geq g(x) + x^2 = g(2x) ,$$

$$f(2x+1) \geq f(x) + x^2 + 2x \geq g(x) + x^2 + 2x = g(2x+1) .$$

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 $x+2 \leq n \leq 2x+1$. In particular, this implies $f(m) = g(m)$ for $1 \leq m \leq 2X-1$. Suppose N satisfies $X+2 \leq N \leq 2X+1$. If $N \leq 2X-1$ then in fact,

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

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

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

Thus,

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

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

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

which is the desired inequality.

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

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

$$\$(C) \geq f(b) + b(2X+1-b)$$
$$> f(X) + X^2 + 2x$$

since

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

holds for $X+2 \leq b \leq 2X-1$ 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

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

since for $1 \leq b' \leq X-1$,

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

Hence

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

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

References

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

[2] McCarthy, D. P., "An Optimal Algorithm to Evaluate x^n over Integers and Polynomials Module M ," to appear in Mathematics of Computation.