JOURNAL OF ALGORITHMS 10, 403–412 (1989)

Addition Chains Using Continued Fractions

F. BERGERON*, J. BERSTEL†, S. BRLEK*, AND C. DUBOC

Département de mathématiques et d'informatique, Université du Québec à Montréal, C.P. 8888, Succ. "A", Montréal, Canada H3C 3P8

Received February 29, 1988; accepted August 18, 1988

This paper introduces a new algorithm for the evaluation of monomials in two variables $x^a y^b$ based upon the continued fraction expansion of a/b. A method for fast explicit generation of addition chains of small length for a positive integer n is deduced from this Algorithm. As an illustration of the properties of the method, a Scholz–Brauer-like inequality $p(N) \leq nb + k + p(n+1)$, is shown to be true whenever N is an integer of the form $2^k(1 + 2^b + \cdots + 2^{nb})$. Computer experimentation has shown that the length of the chains constructed are of optimal length for all integers up to 1000, with 29 exceptions for which the length is equal to the optimal length plus one.

1. INTRODUCTION

Let n be a positive integer. The problem of how to evaluate most efficiently a monomial x^n (considered as early as 1894) has been considered many times. For an history of this problem and further results, see Knuth [K]. The main contribution of this paper is an efficient algorithm for the generation of short addition chains for ordered pairs (a, b) as a tool for the efficient evaluation of monomials $x^a y^b$ in two variables. As a corollary, one obtains an easy method for the generation of addition chains for integers (see Olivos [O]). Surprisingly enough, this algorithm produces shorter chains than the binary (or n-ary) algorithm, without involving too many computation steps.

Both concepts of addition chains are special cases of the concept of word chains (see [BB]) in a semi-group, introduced for the study of the efficient evaluation of monomials in non-commutative variables.

*With the support of the NSERC of Canada Grant A9041 and of the FCAR of Québec Grant EQ1608.

†With the support of the PRC de Mathématiques et Informatique, France.

0196-6774/89 $3.00

Let n be a positive integer. Following Knuth an *addition chain* c for n is a sequence $c = (n_0, n_1, n_2, \ldots, n_r)$ of nonnegative integers such that:

(i) $n_0 = 1$ and $n_r = n$,

(ii) for each i, $1 \leq i \leq r$, $n_i = n_j + n_k$, for some $k \leq j < i$.

Similarly, define an addition chain for an ordered pair of positive integers (a, b) to be sequence $c = ((a_{-1}, b_{-1}), (a_0, b_0), (a_1, b_1), \ldots, (a_r, b_r))$ of ordered pairs such that:

(i) $(a_{-1}, b_{-1}) = (0,1), (a_0, b_0) = (1,0)$ and $(a_r, b_r) = (a, b)$,

(ii) for each i, $1 \leq i \leq r$, there exist j and k, $-1 \leq j, k < i$, such that $(a_i, b_i) = (a_j + a_k, b_j + b_k)$.

The integer r appearing in both these definitions, is called the *length* $|c|$, of addition chain c. The *chain length* $l(n)$ (resp. $l(a, b)$) of an integer n (resp. an ordered pair (a, b)) is the minimal length of all possible chains for n (resp. (a, b)). It is well known that $l(n) \geq \lfloor \log_2 (n) \rfloor$, where $\lfloor x \rfloor$ is the lower integer part of x, and that equality holds when n is a power of 2. The addition chain for a obtained from an addition chain $c = ((a_{-1}, b_{-1}), (a_0, b_0), (a_1, b_1), \ldots, (a_r, b_r))$ for (a, b), by deletion of zeros and repetitions (if any) in the first projection $(1, a_0, a_1, a_2, \ldots, a_r)$ of c will be simply called the *projection* of c in the remainder of this paper.

There is no simple way to compute an explicit chain of minimal length, or even the chain length of an integer, but there are simple algorithms that produce addition chains of near minimal length. Recall that the most commonly known algorithm uses the binary decomposition bin(n) of n; if for instance bin$(n) = 101011$ ($n = 43$) this algorithm produces the addition chain $(1, 2, 4, 5, 10, 20, 21, 42, 43)$ of length 8. Which is not a minimal chain for 43 since $(1, 2, 4, 8, 9, 17, 34, 43)$ is a (minimal) chain of length 7. The chains produced by this usual algorithm have length $(\lambda(n) + \nu(n) - 1)$, where $\lambda(n) = \lfloor \log_2 (n) \rfloor$, and $\nu(n)$ is the number of 1 in the binary decomposition bin(n) of n. It follows immediately that $l(n) \leq 2\lambda(n)$. More generally, Brauer has shown (see [B]) that:

$$\operatorname*{Lim}_{n \to \infty} \frac{l(n)}{\lambda(n)} = 1.$$

We consider now the computation of addition chains for pairs (a, b). A simple way of producing an addition chain for (a, b) is by concatenating a chain for a with a chain for b. Thus for (43, 5), one obtains the following addition chain of length 12:

$$((\mathbf{0,1}), (\mathbf{1,0}), (2,0), (4,0), (8,0), (16,0), (32,0), (40,0), (42,0), (\mathbf{43,0}), (0,2), (0,4), (0,5), (\mathbf{43,5}))$$

from the two addition chains $(\mathbf{1}, 2, 4, 8, 16, 32, 40, 42, \mathbf{43})$ and $(\mathbf{1}, 2, 4, \mathbf{5})$. Since 43 and 5 have been computed *independently*, this method may not be the best one. It is clear that the smallest length of an addition chain obtained this way is $l(a) + l(b) + 1$.

Another method uses the *euclidean division* of a by b: $a = nb + s$, with $0 \leq s < b$. First, one constructs an addition chain for n, thus producing an addition chain for $(n, 1)$ in $l(n) + 1$ steps. Any chain for b can then be used to extend this chain to a chain for (nb, b), in $l(b)$ more steps. One completes this chain for (nb, b) to a chain for (a, b) using a chain for $(s, 0)$ easily obtained from a chain for s. In this last portion of the construction, one has already constructed $(2, 0)$, hence only $l(s)$ steps are needed in order to complete the global chain. Thus the minimal length of the addition chain constructed by this method is $l(n) + l(b) + l(s) + 1$. Let us work out one example in the case: $43 = 8 * 5 + 3$. Using the chain $(1, 2, 4, 8)$ for 8, and the chain $(1, 2, 4, 5)$ for 5, one obtains the chain

$$((\mathbf{0}, \mathbf{1}), (\mathbf{1}, \mathbf{0}), (2, 0), (4, 0), (8, 0), (8, 1), (16, 2), (32, 4), (\mathbf{40}, \mathbf{5}))$$

for $(5 * 8, 5)$. Then with the chain $(1, 2, 3)$ for 3, one finally produces the chain:

$$((\mathbf{0}, \mathbf{1}), (\mathbf{1}, \mathbf{0}), (2, 0), (4, 0), (8, 0), (8, 1), (16, 2), (32, 4), (40, 5), (3, 0), (\mathbf{43}, \mathbf{5}))$$

of global length 9 for $(43, 5)$.

But better yet, one can iterate this approach, naturally introducing *continued fractions*. From now on, the continued fraction expansion of ab^{-1}:

$$\frac{a}{b} = u_1 + \cfrac{1}{u_2 + \cfrac{1}{\ddots + \cfrac{1}{u_{r-1} + \cfrac{1}{u_r}}}}$$

will be denoted by $[u_1, u_2, u_3, \ldots, u_r]$. For our running example, this expansion is

$$\frac{43}{5} = [8, 1, 1, 2].$$

As will be shown in the rest of this paper, one can produce a nice addition chain for (43, 5) from this continued fraction expansion, namely:

$$((\mathbf{0},\mathbf{1}),(\mathbf{1},\mathbf{0}),(2,0),(4,0),(8,0),(8,1),(9,1),(17,2),(34,4),(\mathbf{43},\mathbf{5}))$$

which is of length 8. The first projection of this addition chain gives a minimal length addition chain for 43: (1, 2, 4, 8, 9, 17, 34, 43). The length of this last chain for 43 is clearly the length of the chain for (43, 5) minus 1, because of the single repetition of the 8.

2. The Algorithms

In this section, we will show how to obtain an explicit addition chain for (a, b), with almost shortest length. From this method we will deduce an effective algorithm for the generation of addition chains for integers. The properties of these algorithms, and related results will be studied in the next section.

The first algorithm computes an addition chain $((a_{-1}, b_{-1}), (a_0, b_0), (a_1, b_1), \ldots, (a_t, b_t))$ for (a, b), where $a \geq b$ and the previous conventions that $(a_{-1}, b_{-1}) = (0, 1)$, and $(a_0, b_0) = (1, 0)$ still hold. It is a parametrized algorithm depending on the choice of another algorithm, A, for the generation of addition chains for "small" integers.

Algorithm $P(A)$.

- *Let d be the greatest common divisor of a and b,*
- *let* $[u_1, u_2, u_3, \ldots, u_r]$ *be the continued fraction expansion of* ab^{-1},
- *set* $q := 1$ *and* $j := 0$,
- *for i going from* 1 *to r do:*
  - *produce (by Algorithm A) an addition chain* $(1, v_1, v_2, \ldots, v_{\kappa(i)})$ *for* u_i,
  - *for s going from* 1 *to* $\kappa(i)$: *set* $(a_{j+s}b_{j+s}) := (v_s a_j, v_s b_j)$,
  - *set* $(a_{j+\kappa(i)+1}, b_{j+\kappa(i)+1}) := (a_{j+\kappa(i)} + a_q, b_{j+\kappa(i)} + b_q)$,
  - *set* $q := j$ *and* $j := j + \kappa(i) + 1$;
- *produce (by A) an addition chain* $(1, \delta_1, \delta_2, \ldots, \delta_\mu)$ *for d,*
- *for s going from* 1 *to* μ: *set* $(a_{j+s}, b_{j+s}) := (\delta_s a_j, \delta_s b_j)$.

Whatever Algorithm A might be, the chain for 2^k should always be $(1, 2, 4, \ldots, 2^k)$, of length k. Thus in the case of $a = 43$ and $b = 5$, the chain produced by Algorithm $P(A)$ is exactly the last chain considered in Section 1.

Let $l_A(n)$ denote the length of the chains constructed by Algorithm A. Then the length of the chain for (a, b) produced by Algorithm $P(A)$ is

$$L_{P(A)}(a, b) = l_A(d) + \sum_i (1 + l_A(u_i));$$

hence for this example $L_{P(A)}(43, 5) = 8$.

It has been observed earlier that for a given integer a, one can obtain addition chains for a by projection of addition chains for (a, b). The choice of b will certainly influence the length of the resulting chain. Thus one is lead to consider the following algorithm for the generation of addition chains for a.

ALGORITHM B.

- *If* $a = 2^k$ *then* $(1, 2, 4, \ldots, 2^k)$,

  else
  - *choose* b *between* 2 *and* $a - 1$ *minimizing* $L_{P(B)}(a, b)$,
  - *compute the chain for* (a, b) *by Algorithm* $P(B)$,
  - *the result is the first projection of this chain.*

Let $q(a)$ denote the length of the chain for the integer a generated by Algorithm B, then:

$$q(a) = \min\{L_{P(B)}(a, b) | 1 < b < a\} - 1,$$

where

$$L_{P(B)}(a, b) = q(d) + \sum_i (1 + q(u_i)),$$

with $d = \gcd(a, b)$ and $ab^{-1} = [u_1, u_2, \ldots, u_r]$.

In the definition of $q(a)$, the -1 is clearly needed because of the single repetition that always occurs in the first projection of a chain for (a, b). Computer calculations (using an implicit table for l given by Knuth in [K]) have shown that $q(a) = l(a)$ for all a's between 1 and 1000, with the exception of the 13 following integers: 367, 371, 381, 571, 623, 659, 667, 691, 734, 739, 742, 749, and 762, in which cases $q(a) = 1 + l(a)$.

One problem with Algorithm B in its present form, is its running time. By choosing a small family of b's susceptible of minimizing the length of a chain for (a, b), one can greatly improve the efficiency of this algorithm at the possible risk of losing a little on the shortness of the chains produced.

ALGORITHM B'.

- *If* $a = 2^k$ *then* $(1, 2, 4, \ldots, 2^k)$,

  else
  - *choose* b *of the form* $\lfloor a2^{-k} \rfloor$, $2 \le k \le \lambda(a)$, *minimizing* $L_{P(B)}(a, b)$,
  - *compute the chain for* (a, b) *by Algorithm* $P(B')$,
  - *the result is the first projection of this chain.*

The reason for this choice of b's (of the form $\lfloor a2^{-k} \rfloor$) is that one would like to obtain continued fractions $[u_1, u_2, u_3, \ldots, u_r]$, with $v(u_i)$ small. Note that Knuth [K] has characterized minimal chains for all integers u with $v(u) \leq 4$. We shall denote by $p(a)$ the length of the chain constructed by Algorithm B'. Thus one has

$$p(a) = \min\{ L_{P(B')}(a, \lfloor a2^{-k} \rfloor) | 1 < k < \lambda(a) \} - 1,$$

and

$$L_{P(B')}(a, b) = p(d) + \sum_i (1 + p(u_i)),$$

where $d = \gcd(a, b)$ and $ab^{-1} = [u_1, u_2, \ldots, u_r]$.

Computer calculations have shown that $p(a) = l(a)$ for all a's between 1 and 1000, with the exception of 29 integers (evidently including the previous 13) for which cases $p(a) = 1 + l(a)$. For instance, the only values for b that produce an optimal chain for 631 are 13 and 97. The corresponding continued fraction expansion are $\frac{631}{13} = [48, 1, 1, 6]$ and $\frac{631}{97} = [6, 1, 1, 48]$, but neither 13 nor 97 are of the specified form.

3. Main Results

Probably the best known open problem about addition chains, is Scholz–Brauer's conjecture stating that for all n:

$$l(2^n - 1) \leq n - 1 + l(n).$$

The object of this section is to prove a generalization of the analogous inequality for p.

Lemma. *For all integers n and k, $p(2^n k) \leq n + p(k)$.*

Proof. Let $b = k$ and $a = 2^n k$, then $b = a/2^n$, $\gcd(a, b) = k$ and $a/b = [2^n]$. Therefore by definition of p, $p(a) \leq p(k) + p(2^n)$. The result follows from the obvious equality $p(2^n) = n$. □

Theorem. *For all integers a and b:*

$$P\left(\left\lfloor \frac{2^a - 1}{2^b - 1} \right\rfloor\right) \leq a - b + p\left(\left\lfloor \frac{a}{b} \right\rfloor\right).$$

Proof. The proof is by induction. Suppose that for all $N < \left\lfloor (2^a - 1)(2^b - 1)^{-1} \right\rfloor$ of the form in question, the inequality is true.

Observe that

$$\frac{2^a-1}{2^b-1} = 2^r\left(\frac{2^{nb}-1}{2^b-1}\right) + \left(\frac{2^r-1}{2^b-1}\right)$$

when $a = nb + r$ and $0 \le r < b$. Thus the largest integer contained in $(2^a - 1)(2^b - 1)^{-1}$ is clearly

$$2^r\left(\frac{2^{nb}-1}{2^b-1}\right),$$

and one immediately concludes by Lemma that

$$p\left(\left\lfloor\frac{2^a-1}{2^b-1}\right\rfloor\right) \le r + p\left(\frac{2^{nb}-1}{2^b-1}\right).$$

If $r \neq 0$, then by the induction hypothesis,

$$p\left(\left\lfloor\frac{2^a-1}{2^b-1}\right\rfloor\right) \le r + nb - b + p(n),$$

which is clearly the identity mentioned above, since $a = nb + r$ and $n = \lfloor a/b \rfloor$.

Now when a is of the form nb, the argument is a little more complicated. First of all, by definition of p, there exist a c of the form $\lfloor n/2^k \rfloor$, for which

$$p(n) = p(d) + \sum_i (1 + p(u_i)),$$

where $d = \gcd(n, c)$ and $n/c = [u_1, u_2, \ldots, u_r]$. But the continued fraction expansion of

$$\left(\frac{2^{nb}-1}{2^b-1}\right)\left(\frac{2^{cb}-1}{2^b-1}\right)^{-1}$$

is

$$\left[2^{r_1 b}\left(\frac{2^{u_1 cb}-1}{2^{cb}-1}\right), 2^{r_2 b}\left(\frac{2^{u_2 r_1 b}-1}{2^{r_1 b}-1}\right), \ldots, 2^{r_s b}\left(\frac{2^{u_s r_{s-1} b}-1}{2^{r_{s-1} b}-1}\right)\right],$$

where the u_k and r_k are obtained by Euclid's algorithm:

$$\begin{aligned}
n &= u_1 c + r_1, && \text{with } 0 < r_1 < c,\\
c &= u_2 r_1 + r_2, && \text{with } 0 < r_2 < r_1,\\
r_1 &= u_3 r_2 + r_3, && \text{with } 0 < r_3 < r_2,\\
&\vdots\\
r_{s-2} &= u_s r_{s-1} + r_s, && \text{with } r_s = 0.
\end{aligned}$$

Recall that r_{s-1} is the greatest common divisor d of n and c. The verification of

$$\gcd\left(\frac{2^{nb}-1}{2^b-1}, \frac{2^{cb}-1}{2^b-1}\right) = \frac{2^{\gcd(n,c)b}-1}{2^b-1}$$

and

$$\frac{2^{cb}-1}{2^b-1} = \left\lfloor \frac{A}{2^x} \right\rfloor, \qquad \text{where } A = \frac{2^{nb}-1}{2^b-1} \text{ and } x = (n-c)b$$

is easy. The conclusion comes once more from an application of the definition of p, since

$$p\left(\frac{2^{nb}-1}{2^b-1}\right) \le p\left(\frac{2^{db}-1}{2^b-1}\right) + \sum_{k=1}^{s}\left(1 + p\left(2^{r_k b}\left(\frac{2^{u_k r_{k-1} b}-1}{2^{r_{k-1}b}-1}\right)\right)\right) - 1,$$

where $r_0 = c$. By the induction hypothesis and the above lemma, the right-hand term of this inequality is less than or equal to

$$db - b + p(d) + \sum_{k=1}^{s}\left(1 + r_k b + u_k r_{k-1} b - r_{k-1} b + p(u_k)\right) - 1,$$

which can also be rearranged in the form

$$nb - b + p(n) + db + \left(\sum_{k=2}^{s}\left((u_k r_{k-1} + r_k) - r_{k-2}\right)b\right) - r_{s-1} b.$$

But all the terms of the summation in parentheses are zero and $r_{s-1} = d$. Thus we have obtained the inequality of the theorem. □

In fact, a careful reading of the proof of the theorem shows that

$$p\left(\left\lfloor \frac{2^a-1}{2^b-1} \right\rfloor\right) = a - b + q\left(\left\lfloor \frac{a}{b} \right\rfloor\right),$$

thus equality in the theorem holds whenever $p(\lfloor a/b \rfloor) = q(\lfloor a/b \rfloor)$. But $p(n) = q(n)$ for all n between 1 and 1000 with the exception of the 16 integers: 135, 270, 319, 437, 540, 559, 629, 631, 638, 697, 699, 731, 747, 809, 869, and 874 for which Algorithm B' produces a chain longer than the chain produced by Algorithm B.

TABLE 1

n	Running time (s)
9	0.12
99	1.13
999	6.02
9999	14.45
99999	29.70
999999	73.43
9999999	221.61

4. Conclusion

As an indication of the running time for a straightforward implementation of algorithms $P(B')$ and B', Table 1 gives the time needed for the construction of an explicit addition chain for some integers n. This implementation was written in MAPLE on a SUN 3/50 workstation.

Many interesting observations can be made on the addition chains produced by either algorithms B or B'. For instance, it is easy to observe (see [T1]) that good candidates for integers N such that $p(N) = p(2N)$, are those of the form $n2^k + j$ with n, j odd, k large enough (≥ 6 or 7) and $L(n, j) = L(n, 2j) + 1$, where $L(n, j)$ denotes the minimal length of an addition chain for n containing j. Eight such pairs (n, j) have been proposed by Thurber (in [T1]) with one pair (69, 7) which does not satisfy $L(n, j) = L(n, 2j) + 1$. The following list has been easily generated by computer: **(23, 7)**, **(35, 11)**, **(37, 7)**, **(43, 13)**, (47, 15), (57, 11), (59, 19), (63, 19), **(67, 21)**, **(69, 21)**, (69, 13), (71, 23), (79, 11), (79, 7), (79, 23), (79, 15), **(83, 25)**, (87, 23), (91, 23), (91, 29), (91, 25), (93, 7), (95, 13), (101, 11), (101, 19), (103, 31), (105, 33), (107, 35), (109, 21), (111, 21), (115, 35), (115, 37), (121, 7), (121, 13), (123, 37), (127, 37), (127, 39), (127, 33), (131, 41), (133, 25), (139, 43), (139, 19), (139, 15), (139, 45), (141, 45), (141, 25), (143, 43), and (151, 21), where the pairs in bold are those proposed by Thurber.

It has been mentioned that an addition chain for n characterizes a specific method for the evaluation of x^n. But the memory requirement for this evaluation varies with the choice of this addition chain. For instance, while evaluating x^{13} by means of the addition chain (1, 2, 4, 8, 12, 13), one needs to remember at some point, the three values x, x^4, and x^{12}. But the chain (1, 2, 3, 6, 12, 13) gives a means of doing the same evaluation with a maximum of two values remembered at the same time. The algorithms of Section 2 and the theorem of Section 3 can easily be adapted to include a measure of this register requirement need. One defines the *width* $\omega(c)$ of a chain c to be the number of registers needed in order to compute x^n with

the chain c. And in Algorithm B, choose b that minimizes first $L_{P(B)}(a, b)$, and then $\omega(a)$. Once more, computer calculations, with these modified algorithms, have shown that the number of registers needed in order to evaluate x^n with the chains produced by Algorithm B is less than or equal to two for all integers between 1 and 512, with the exception of 107, 173, 179, 203, 211, 214, 241, 271, 307, 317, 346, 347, 355, 373, 395, 403, and 406. The first integer for which at least four registers are needed is 1187.

Another interesting variation on the method presented in this paper, is to modify the algorithm generating continued fractions, so that the approximation used for a given rational number q is the closest integer, instead of the floor of q. This produces continued fractions with both positive and negative u_i's. The chains then obtained are addition–substraction chains. This last variation has been suggested to us by J. Vuillemin.

REFERENCES

[B] A. BRAUER, On addition chains, *Bull. Amer. Math. Soc.* **45** (1939), 736–739.

[BB] J. BERSTEL AND S. BRLEK, On the length of word chains, *Inform. Process Lett.* **26**, September 87, 23–28.

[DL] D. DOBKIN AND R. J. LIPTON, Addition chain methods for the evaluation of specific polynomials, *SIAM J. Comput.* **9** (1980), 121–125.

[GSS] A. A. GIOIA, M. V. SUBBARAO AND M. SUGUNAMA, The Scholz–Brauer problem in addition chains, *Duke Math. J.* **29** (1962), 481–487.

[K] D. E. KNUTH, "The Art of Computer Programming," Vol. 2, Addison–Wesley, Reading, MA, 1981.

[O] J. A. O. OLIVOS ARAVENA, Ph.D. thesis, Paris, 1979.

[T1] E. G. THURBER, Addition chains and solutions of $l(2n) = l(n)$ and $l(2^n - 1) = n + l(n) - 1$, *Discrete Math.* **16** (1976), 279–289.

[T2] E. G. THURBER, The Scholz–Brauer problem on addition chains, *Pacific J. Math.* **39**, No. 1 (1973), 229–242.