ON THE EVALUATION OF POWERS*

ANDREW CHI-CHIH YAO†

Abstract. It is shown that for any set of positive integers \( \{n_1, n_2, \ldots, n_p\} \), there exists a procedure which computes \( \{x^{n_1}, x^{n_2}, \ldots, x^{n_p}\} \) for any input \( x \) in less than \( \lg N + c \sum_{i=1}^{p} [\lg n_i/\lg (n_i + 2)] \) multiplications for some constant \( c \), where \( N = \max_i \{n_i\} \). This gives a partial solution to an open problem in Knuth [3, §4.6.3, Ex. 32] and generalizes Brauer’s theorem on addition chains.

Key words. addition chains, Brauer’s theorem

1. Introduction. An addition chain (of length \( r \)) is a sequence of \( r + 1 \) integers \( a_0, a_1, a_2, \ldots, a_r \) such that (i) \( a_0 = 1 \) and (ii) for each \( i \), \( a_i = a_j + a_k \) for some \( j \leq k < i \). It is clear that, for any \( r \) and any set of integers \( \{n_1, n_2, \ldots, n_p\} \), there exists an addition chain of length \( r \) which contains the values \( n_1, n_2, \ldots, n_p \) if and only if there exists a procedure which, for any input \( x \), computes \( \{x^{n_1}, x^{n_2}, \ldots, x^{n_p}\} \) in \( r \) operations using only multiplications. A theorem by Brauer [1], [3, pp. 398–418] states that, for any \( n \), there exists an addition chain of length \( \lg n + O(\lg n/\lg \lg n) \) which contains the value \( n \); this implies the existence of a corresponding procedure to compute \( x^n \) in \( \lg n + O(\lg n/\lg \lg n) \) multiplications. Furthermore, it was shown by Erdős [2], [3, pp. 398–418] that the above result is asymptotically with probability 1 nearly the best possible. In an open problem posed in Knuth [3, §4.6.3, Ex. 32], it is asked if there are fast procedures to compute \( \{x^{n_1}, x^{n_2}, \ldots, x^{n_p}\} \) for \( p \geq 2 \). This problem cannot be solved by a direct extension of the technique used by Brauer in the proof of his theorem.

In this paper we show that for any positive integers \( n_1, n_2, \ldots, n_p \), there exists a procedure using only multiplications which, for any input \( x \), computes \( \{x^{n_1}, x^{n_2}, \ldots, x^{n_p}\} \) in \( \lg N + \text{constant} \times \sum_{i=1}^{p} [\lg n_i/\lg (n_i + 2)] \) multiplications where \( N = \max_i \{n_i\} \). This gives a solution to Knuth’s problem and leads to a corresponding theorem on addition chains which generalizes Brauer’s theorem mentioned earlier.

2. Definition. Let \( e_i, 1 \leq i \leq p \), and \( f_j, 1 \leq j \leq q \), be positive integers. We shall say that \( \{x^{e_i}, \ldots, x^{e_q}\} \) is computable from \( \{x^{f_1}, \ldots, x^{f_q}\} \) in \( r \) multiplications \( (r \geq 0) \) if there exists a set of \( r \) positive integers, \( \{f_{q+1}, \ldots, f_{q+r}\} \), such that

(i) for all \( i = q + 1, \ldots, q + r \),

\[ x^{f_i} = x^{f_j} \cdot x^{f_k} \text{ for some } j \leq k < i. \]

(ii) \[ \{x^{e_1}, \ldots, x^{e_q}\} \subseteq \{x^{f_1}, \ldots, x^{f_{q+r}}\}. \]

Received by the editors August 29, 1974.
† Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801. This research was supported by the National Science Foundation under Grant GJ-41538.

\(^1\) \( \lg \) is logarithm to the base 2.
Since the exponents are added when two powers of \(x\) are multiplied, the above definition is a natural generalization of the definition of addition chains (cf. §1). The exponents appearing in \(\{x^{f_1}, \cdots, x^{f_s}\}\) correspond to a set of numbers initially available in the chain, as opposed to a single number, 1, in the earlier definition.

3. The computation of \(\{x^n, \cdots, x^{n^p}\}\). The following lemma is well known [3, pp. 398-418].

**Lemma 1.** For any integer \(i > 0\), \(\{y^i\}\) is computable from \(\{y\}\) in at most \(2\lfloor \log i \rfloor\) multiplications.

**Proof.** Let the binary representation of \(i\) be

\[
i = \sum_{j=0}^{v} b_j \cdot 2^j,
\]

where \(v = \lfloor \log i \rfloor\). Then,

\[
y^i = \prod_{b_j=1} y^{2^j}.
\]

Thus, we can first compute \(y^2, y^4, y^8, \cdots, y^{2^v}\) sequentially in \(v\) multiplications and then compute \(y^i\) by (2) in no more than \(v\) multiplications. The total number of multiplications is no greater than \(2v\). \(\square\)

**Theorem 2.** For any integers \(m, n\) where \(0 < m \leq n\), \(\{x^m\}\) is computable from \(\{x, x^2, x^4, x^8, \cdots, x^{2^{\lfloor \log n \rfloor}}\}\) in less than \(c \log n / \log (n + 2)\) multiplications for some constant \(c\).

**Proof.** Assume \(n \geq 4\). Define the following quantities:

\[
k = \lfloor \log \log n / 2 \rfloor,
\]

\[
D = 2^k,
\]

\[
t = \lfloor \log_D n \rfloor,
\]

Let the \(D\)-ary representation of \(m\) be

\[
m = \sum_{j=0}^{t} a_j D^j,
\]

where

\[
0 \leq a_j \leq D - 1 \quad \text{for} \quad j = 0, 1, \cdots, t.
\]

We partition the set of integers \(\{0, 1, \cdots, t\}\) into \(D\) disjoint subsets \(S(0), S(1), \cdots, S(D - 1)\) by letting

\[
S(i) = \{a | a_i = i\} \quad \text{for} \quad i = 0, 1, \cdots, D - 1.
\]

It follows from (6) that

\[
m = \sum_{i=1}^{D-1} i \cdot \left[ \sum_{\ell \in S(i)} D^\ell \right] = \sum_{i=1}^{D-1} i \cdot m_i,
\]

where

\[
m_i = \sum_{\ell \in S(i)} D^\ell.
\]
From (7) and (8), we obtain the following two equations:

\[ x^{m_i} = \prod_{i \in S(i)} x^{D'} \quad \text{for } i = 1, 2, \ldots, D - 1, \]

\[ x^m = \prod_{i=1}^{D-1} (x^m)^i. \]

Since all the \(x^{D'}\) in (9) are available in the set \(\{x, x^2, x^4, x^8, \ldots, x^{2^{\lg n}}\}\), we can construct a procedure to compute \(x^m\) as follows.

**Step 1.** For \(i = 1, 2, \ldots, D -1\) do the following:

(a) Compute \(x^{m_i}\) from (9) in fewer than \(|S(i)|\) multiplications.

(b) Compute \((x^{m_i})^i\) in at most \(2|\lg i|\) multiplications (by Lemma 1).

**Step 2.** Compute \(x^m\) from (10) in \(D - 2\) multiplications.

Let \(M\) be the total number of multiplications in the above procedure. Then,

\[ M < \sum_{i=1}^{D-1} (|S(i)| + 2|\lg i|) + D - 2 \]

Noting that the \(S(i)\)'s form a partition of the set \(\{0, 1, \ldots, t\}\), we obtain from (11) that

\[ M < t - 1 + 2(D - 1)\lg(D - 1) + D - 2, \]

which together with equations (3), (4) and (5), implies that

\[ M < 2(\lg n/l; \lg n) + 1 + 4(\lg n)^{1/2} \lg \lg n + 2(\lg n)^{1/2}. \]

It follows from (13) that there exists a constant \(c\) such that

\[ M < c \lg n/l \lg \lg n \lg (n + 2). \]

Thus the theorem is true if \(n \geq 4\). Obviously we can choose \(c\) so that the theorem is also true for \(n = 1, 2, 3\).

**Theorem 3.** For any set of positive integers \(\{n_1, n_2, \ldots, n_p\}\), \(\{x^{n_1}, x^{n_2}, \ldots, x^{n_p}\}\) is computable from input \(\{x\}\) in less than \(\lg N + c \sum_{i=1}^{p} [\lg n_i/|\lg \lg (n_i + 2)|]\) multiplications for some constant \(c\), where \(N = \max_i \{n_i\}\).

**Corollary.** \(\{x^n, x^{n_2}, \ldots, x^{n_p}\}\) is computable from \(\{x\}\) in less than \(\lg N + c p \lg n /|\lg \lg (N + 2)|\) multiplications.

**Proof of Theorem 3 and Corollary.** First we compute \(\{x, x^2, x^4, x^8, \ldots, x^{2^{\lg N}}\}\) from input \(x\) in \(\lg N\) multiplications. For each \(i\), according to Theorem 2, \(x^{n_i}\) is computable from \(\{x, x^2, x^4, \ldots, x^{2^{\lg N}}\}\) in \(c \lg N/|\lg \lg (N + 2)|\) multiplications for some constant \(c\). The theorem and corollary then follow immediately.

In terms of addition chains, Theorem 3 and its corollary give the following generalization of Brauer's theorem [1], [3, pp. 398–418].

**Theorem 4.** For any positive integers \(n_1, n_2, \ldots, n_p\), there exists an addition chain of length less than \(\lg N + c \sum_{i=1}^{p} \lg n_i/|\lg \lg (n_i + 2)|\) containing \(n_1, n_2, \ldots, n_p\) for some constant \(c\), where \(N = \max_i \{n_i\}\).

**Corollary.** For positive integers \(n_1, n_2, \ldots, n_p\), there exists an addition chain of length less than \(\lg N + c p \lg N/|\lg \lg (N + 2)|\) containing \(n_1, n_2, \ldots, n_p\)
4. Conclusion. We have shown that \( \{x^{n_1}, x^{n_2}, \cdots, x^{n_p}\} \) can be computed in \( \lg N + cp \lg N/\lg \lg (N + 2) \) multiplications for input \( x \) where \( N = \max_i \{n_i\} \) and \( c \) is a constant. On the other hand, it is well known that to evaluate \( \{x^{n_1}, x^{n_2}, \cdots, x^{n_p}\} \) by arithmetic operations, at least \( \lg N \) operations are necessary. Thus our procedures for evaluating \( \{x^{n_1}, x^{n_2}, \cdots, x^{n_p}\} \) are nearly the best possible when \( p \ll \lg \lg (N + 2) \). It remains an interesting open problem to determine the complexity of computing \( \{x^{n_1}, x^{n_2}, \cdots, x^{n_p}\} \) for general \( p \).

Note added in proof. (A) By choosing the value of \( k \) in (3) more carefully, say \( k = \lfloor \lg \lg n - 3 \lg \lg \lg n \rfloor \), our algorithm in Theorem 3 takes at most \( \lg N + plg N/\lg \lg N \) (smaller terms) multiplications as \( N \to \infty \). For fixed \( p \), these leading terms are almost the best possible since, as observed by Larry Stockmeyer (private communication), the lower bound of Erdős [2] can be generalized straightforwardly. (B) Nicholas Pippenger proved the following (private communication): \( \{x^{n_1}, x^{n_2}, \cdots, x^{n_p}\} \) can be computed from \( x \) in \( \min \{p + 2l \lfloor \lg N/l \rfloor \mid l \) is a positive integer\} multiplications, and for some \( c_1 > 0 \) and every \( N, p, c_1p \lg N/(lg P + \lg \lg N) \) multiplications are needed for some set of \( \{n_1, n_2, \cdots, n_p\} \) with \( \max \{n_i\} \leq N \). For large \( p \) \( (p \geq \lg N) \), this determines the worst-case complexity to be \( p \lg N/\lg p \) up to a constant factor. (C) A related theorem on power evaluation may be found in Schönhage [4].

REFERENCES