Lower Bounds on Universal Traversal Sequences Based on Chains of Length Five

Jonathan F. Buss

Department of Computer Science

University of Waterloo

Waterloo, Ontario, Canada N2L 3G1

jfbuss@math.uwaterloo.ca

Martin Tompa*

Department of Computer Science and Engineering, FR-35

University of Washington

Seattle, Washington, U.S.A. 98195

tompa@cs.washington.edu

March 24, 1995

Abstract

Universal traversal sequences for cycles require length $\Omega(n^{1.43})$, improving the previous bound of $\Omega(n^{1.33})$. For $d \geq 3$, universal traversal sequences for d-regular graphs require length $\Omega(d^{0.57}n^{2.43})$. For constant d, the best previous bound was $\Omega(n^{2.33})$.

*This material is based upon work supported in part by the National Science Foundation under Grant CCR-9002891 and by IBM under Research Contract 16980043.

1. Reflecting and Universal Traversal Sequences

Universal traversal sequences were introduced by Cook (see Aleliunas *et al.* [2]) as a particularly simple method for traversing graphs. Universal traversal sequences are defined as follows.

For $2 \leq d < n$, let $\mathcal{G}(d,n)$ be the set of all connected, d-regular, n-vertex, edge-labeled, undirected graphs $G = (V, E)$. For this definition, edges are labeled as follows. For every edge $\{u,v\} \in E$ there are two labels $l_{u,v}$ and $l_{v,u}$ with the property that, for every $u \in V$, $\{l_{u,v} \mid \{u,v\} \in E\} = \{0, 1, \ldots, d-1\}$. For such labeled graphs, a string over $\{0, 1, \ldots, d-1\}$ can be thought of as a sequence of edge traversal commands. In particular, any $U = U_1 U_2 \ldots U_{|U|} \in \{0, 1, \ldots, d-1\}^*$ and $v_0 \in V$ determine a unique sequence $(v_0, v_1, \ldots, v_{|U|}) \in V^{|U|+1}$ such that $l_{v_{i-1},v_i} = U_i$, for all $i \in \{1, 2, \ldots, |U|\}$. Such a sequence U is said to *traverse* G starting at v_0 if and only if every vertex in G appears at least once in the sequence $v_0, v_1, \ldots, v_{|U|}$. U is a *universal traversal sequence* for $\mathcal{G}(d,n)$ if and only if U traverses each $G \in \mathcal{G}(d,n)$ starting at any vertex in G. $U(d,n)$ denotes the length of a shortest universal traversal sequence for $\mathcal{G}(d,n)$. To avoid trivialities, define $U(d,n) = U(d,n+1)$ in case $\mathcal{G}(d,n)$ is empty, which occurs exactly when both d and n are odd (see [4, Proposition 1]).

Good bounds on $U(d,n)$ translate into good bounds on the time required by certain simple undirected graph traversal algorithms running in very limited space. In particular, determining good lower bounds on $U(d,n)$ is a prerequisite to proving time-space tradeoffs for traversing undirected graphs. (See Borodin, Ruzzo, and Tompa [4] for more detailed discussion.) Proving such lower bounds is the emphasis of this paper.

The current best upper and lower bounds on $U(d,n)$ are summarized in Table 1. (See Borodin, Ruzzo, and Tompa [4] for more background.) Prior to the current work, the best lower bound for $d = 2$ was

$$U(2,n) = \Omega(n^{1.33}),$$

due to Coppersmith (see Tompa [7]). This bound is improved in Corollary 7 to

$$U(2,n) = \Omega(n^{\log_5 10}) = \Omega(n^{1.43}).$$

Table 1: Best Known Bounds on Length of Universal Traversal Sequences

Bound	*Relevant Range*	*Source*
$U(d,n) = O(n^3)$	$d = 2$	Aleliunas [1]
$U(d,n) = O(dn^3 \log n)$	$3 \leq d \leq n/2 - 1$	Kahn *et al.* [6]
$U(d,n) = O(n^3 \log n)$	$n/2 - 1 < d$	Chandra *et al.* [5]
$U(d,n) = \Omega(n^{1.43})$	$d = 2$	This paper
$U(d,n) = \Omega(d^{0.57} n^{2.43})$	$3 \leq d < n^{\log_{10} 2}$	This paper
$U(d,n) = \Omega(d^2 n^2)$	$n^{\log_{10} 2} \leq d \leq n/3 - 2$	Borodin *et al.* [4]
$U(d,n) = \Omega(n^2)$	$n/3 - 2 < d$	Alon *et al.* [3]

For $3 \leq d \leq n/3 - 2$, the best previous lower bound was

$$U(d,n) = \Omega(d^{0.67} n^{2.33} + d^2 n^2),$$

the first term due to Coppersmith (see Tompa [7]) and the second term due to Borodin, Ruzzo, and Tompa [4]. This bound is improved in Corollary 8 to

$$U(d,n) = \Omega(d^{2-\log_5 10} n^{1+\log_5 10} + d^2 n^2) = \Omega(d^{0.57} n^{2.43} + d^2 n^2).$$

The value of d at which the second term begins to dominate is $d = n^{\log_{10} 2} \approx n^{0.3}$.

As in Tompa [7], the method underlying the new lower bounds depends on analyzing traversals of "labeled chains". A *labeled chain of length* n is an undirected graph $G = (V, E)$, where $V = \{0, 1, \ldots, n\}$ and $E = \{\{i, i+1\} \mid 0 \leq i < n\}$, with edge labels as follows. Every edge $\{i, i+1\}$ has two labels $l_{i,i+1}$ and $l_{i+1,i}$, each a nonempty subset of $\{0, 1\}$, restricted as follows:

1. $l_{0,1} = l_{n,n-1} = \{0, 1\}$, and

2. one of $l_{i,i-1}$ and $l_{i,i+1}$ is $\{0\}$ and the other is $\{1\}$, for all $0 < i < n$.

Let $\mathcal{L}(n)$ be the set of labeled chains of length n. Thus, $|\mathcal{L}(n)| = 2^{n-1}$. A labeled chain G will often be identified with the string $\beta = \beta_1\beta_2\cdots\beta_{n-1} \in \{0,1\}^{n-1}$, where $l_{i,i+1} = \{\beta_i\}$ for $0 < i < n$. The string β is called the *label* of G.

Given a labeled chain G of length n, a string $U \in \{0,1\}^*$ can be considered as a sequence of edge traversal commands starting at vertex 0. In particular, if $U = U_1U_2\cdots U_{|U|}$, where $U_i \in \{0,1\}$ for $1 \le i \le |U|$, U determines a unique sequence $(v_0 = 0, v_1, v_2, \ldots, v_{|U|}) \in \{0,1,\ldots,n\}^{|U|+1}$ such that $U_i \in l_{v_{i-1},v_i}$, for all $1 \le i \le |U|$. Such a sequence U is said to *reflect t times* on G if and only if there exist $0 < j_1 < j_2 < \cdots < j_t \le |U|$ such that $v_{j_{2k-1}} = n$ for all $1 \le k \le \lceil t/2 \rceil$ and $v_{j_{2k}} = 0$ for all $1 \le k \le \lfloor t/2 \rfloor$. U is a *t-reflecting sequence* for $\mathcal{L}(n)$ if and only if U reflects t times on each $G \in \mathcal{L}(n)$. $R(t,n)$ denotes the length of a shortest t-reflecting sequence for $\mathcal{L}(n)$.

The following two theorems show that lower bounds on the lengths of reflecting sequences for short chains translate into lower bounds on the lengths of universal traversal sequences. In particular, Theorem 1 provides a lower bound on the length of universal traversal sequences on the cycle, and Theorem 2 provides a lower bound on the length of universal traversal sequences for all higher degree graphs.

Theorem 1 (Tompa [7, Corollary 7]): Let c and r be positive integers such that, for all positive integers t, $R(t,c) \ge rt$. Then $U(2,n) = \Omega(n^{\log_c r})$.

Theorem 2 (Tompa [7, Corollary 8]): Let c and r be positive integers such that, for all positive integers t, $R(t,c) \ge rt$. Then for any integer $3 \le d \le n/17 + 1$,

$$U(d,n) = \Omega(d^{2-\log_c r} n^{1+\log_c r}).$$

By analyzing t-reflecting sequences on $\mathcal{L}(4)$, Tompa [7] demonstrated that $R(t,4) \ge 6t$, yielding the exponent $\log_c r = \log_4 6 \approx 1.29$. He left open the analysis of t-reflecting sequences on $\mathcal{L}(5)$. The main result of this paper is that $R(t,5) \ge 10t$, yielding the improved exponent $\log_c r = \log_5 10 \approx 1.43$.

2. Reflecting Sequences on $\mathcal{L}(5)$

Consider a $(2t-1)$-reflecting sequence U for $\mathcal{L}(5)$. Consider the chain with label 0000, and the t indices in U at which the traversal leaves vertex 1 for the last time before arriving at vertex 5. Notice that the traversal only leaves vertex 1 at even-numbered indices of U. Beginning at each of these t even indices, U must contain a substring matching $00(01+10)^*00$, and these t substrings must be nonoverlapping.

Definition: A substring of length 2 beginning at an even index in U will be called simply a *pair*. The pair matching the first 00 in $00(01+10)^*00$ is said to *start* a (left-to-right) traversal of 0000, and the second 00 pair is said to *finish* the traversal.

Thus, U must contain at least $2t$ occurrences of the pair 00 beginning at even indices. Similarly, there must also be $2t$ occurrences of each of the other pairs 01, 10, and 11, beginning at even indices. These pairs, which correspond to starting and finishing left-to-right traversals of the chains with labels 0000, 0101, 1010, and 1111, are called *base pairs*.

This simple argument shows that $R(2t-1,5) \geq 16t$ ($8t$ base pairs, each with 2 symbols), or $R(t,5) \geq 8t$, but since $\log_5 8$ is slightly less than $\log_4 6$, there is no improvement over the result of Tompa [7]. This section will show that there must be at least $2t$ other pairs beginning at even indices, by considering the additional chains with labels 0011, 0110, 1001, and 1100. The idea underlying the argument is to show that the $8t$ base pairs alone cannot simultaneously serve all the necessary functions on all of these chains.

In what follows, any reference to a traversal of a certain chain will always refer to one of the first t left-to-right traversals of that chain. Thus, for example, there are *exactly* $2t$ base 00 pairs, even if U should happen to traverse the chain with label 0000 more than t times from left to right.

Definition: A *mark* on a 00 pair is either (1) a *left mark*, which is on a 00 pair that finishes a traversal of either 0000 or 1100, or (2) a *right mark*, which is on a 00 pair that either starts a traversal of 0000, or is the last 00 pair during a traversal of 0011. Marks on 01, 10, and 11 pairs are defined analogously. Marks may occur on any pair, not just base pairs.

Definition: A *debt* on a 00 pair is either (1) a *left debt*, which is on a 00

Table 2: Substrings of U traversing chains from left to right

To traverse left to right on the chain with label:	The substring of U must match:
0000	$00(01+10)^*00$
0011	$00(00+10)^*11$
0101	$01(00+11)^*01$
0110	$01(01+11)^*10$
1001	$10(00+10)^*01$
1010	$10(00+11)^*10$
1100	$11(01+11)^*00$
1111	$11(01+10)^*11$

pair that does not have a left mark, or (2) a *right debt*, which is on a 00 pair that does not have a right mark. Debts on 01, 10, and 11 pairs are defined analogously. Debts may occur on any pair, not just base pairs.

Lemma 3: The number of nonbase pairs is exactly half the number of debts.

Proof: As shown earlier, a substring traversing the chain with label 0000 from left to right must match $00(01+10)^*00$; similarly, a substring traversing 0011 must match $00(00+10)^*11$. (See Table 2.) Hence, no 00 pair can both start a traversal of 0000 and be the last 00 pair during a traversal of 0011, since finishing either traversal "spoils" the other. A substring traversing 1100 must match $11(01+11)^*00$. Hence, no 00 pair can finish traversals of both 0000 and 1100.

In summary, a 00 pair has at most one left mark and one right mark. Hence, the number of marks plus the number of debts on each 00 pair is exactly two.

Suppose there are n nonbase 00 pairs and a total of d debts on 00 pairs. Since there are $2t$ base 00 pairs and a total of $4t$ marks on 00 pairs, $4t + d = 2(n + 2t)$. That is, $n = d/2$.

Analogous proofs hold for 01, 10, and 11 pairs. □

Lemma 4: If a 00 pair has a left mark, then the next 00 or 11 pair to its left must have a right mark. If a 00 pair has a right mark, then the next 00 or 11 pair to its right must have a left mark. Analogous results hold for 01, 10, and 11 pairs.

Proof: This follows directly from the definition of a mark, and Table 2. □

Definition: A *0000 interval* is a substring of U that begins with a 00 pair that finishes some traversal of 0000, and ends with the 00 pair that starts the next traversal of 0000. 0101, 1010, and 1111 traversals are defined similarly.

Lemma 5: Each 0000 interval and each 1111 interval contains at least two debts on 00 and/or 11 pairs. Each 0101 interval and each 1010 interval contains at least two debts on 01 and/or 10 pairs.

Proof: Each 0000 interval must contain a substring that matches $11(01 + 10)^*11$ beginning at an *odd* index: this substring accounts for the right-to-left traversal of the chain with label 0000. Expressed as pairs each beginning at an even index, this regular expression is equivalent to $x1(10)^*11((01)^*00(10)^*11)^*(01)^*1y$, where each of x and y may be either 0 or 1. (See Table 3.) Consider the leftmost 11 pair in the match, not including the pair that matches $x1$. If it has a left debt then, by Lemma 4, the next 00 or 11 pair to its left must have a right debt, and these two debts both occur within the 0000 interval. If, on the other hand, the leftmost 11 has a left mark, it must finish a traversal of 1111, because $x1(10)^*11$ cannot be a suffix of $00(00 + 10)^*11$.

The rightmost 11 pair in the match is similar. If it has a right debt, then there is a left debt on the next 00 or 11 pair to its right, and both debts occur

Table 3: Substrings of U traversing chains from right to left

To traverse right to left on the chain with label:	The substring of U must match:
0000	$x1(10)^*11((01)^*00(10)^*11)^*(01)^*1y$
0101	$x0(11)^*10((00)^*01(11)^*10)^*(00)^*1y$
1010	$x1(00)^*01((11)^*10(00)^*01)^*(11)^*0y$
1111	$x0(01)^*00((10)^*11(01)^*00)^*(10)^*0y$

within the 0000 interval. If it has a right mark, it must start a traversal of 1111, because $11(01)^*1y$ cannot be a prefix of $11(01)^*00$.

Suppose the 0000 interval contains fewer than two debts on 00 and/or 11 pairs. From the discussion above, there must be a 1111 interval within the 0000 interval. An analogous argument shows that this 1111 interval must either contain at least two debts on 00 and/or 11 pairs, or contain a 0000 interval. Since no 0000 interval can contain another 0000 interval, this is a contradiction. Therefore, the original 0000 interval contains at least two debts on 00 and/or 11 pairs.

Analogous proofs hold for 0101, 1010, and 1111 intervals. □

Theorem 6: $R(t,5) \geq 10t$ for all positive integers t.

Proof: Consider a sequence that is $(2t-1)$-reflecting on $\mathcal{L}(5)$. This sequence contains $t-1$ 0000 intervals so, by Lemma 5, there are at least $2t-2$ debts on 00 and/or 11 pairs. By an analogous argument, there are at least $2t-2$ debts on 01 and/or 10 pairs. These $4t-4$ debts account for $2t-2$ nonbase pairs, by Lemma 3. Since there are $8t$ base pairs, there are at least $10t-2$ pairs, for a total length of $20t-4$. Thus, $R(2t-1,5) \geq 20t-4$. By the following case analysis, $R(t,5) \geq 10t+1$.

Case 1: $t = 2k-1$ for some integer k. Then

$$R(t,5) = R(2k-1,5) \geq 20k-4 = 10t+6.$$

CASE 2: $t = 2k$ for some integer k. Then

$$R(t,5) \geq R(t-1,5) + 5 = R(2k-1,5) + 5 \geq 20k + 1 = 10t + 1.$$

$\square$

Combining Theorem 6 with Theorems 1 and 2 yields the following lower bounds.

Corollary 7: $U(2,n) = \Omega(n^{\log_5 10}) = \Omega(n^{1.43})$.

Corollary 8: For any integer $3 \leq d \leq n/17 + 1$,

$$U(d,n) = \Omega(d^{2-\log_5 10} n^{1+\log_5 10}) = \Omega(d^{0.57} n^{2.43}).$$

3. Directions for Improvement

It is an open question whether the bound in Theorem 6 is tight. The best upper bound we have is given in Theorem 9.

Theorem 9: $R(t,5) \leq 12t + O(1)$ for all positive integers t.

Proof: The string $(000010110111001001011110)^{t+1}$ is a $(2t)$-reflecting sequence for $\mathcal{L}(5)$, as can be verified by applying it to the 16 labeled chains of length 5. $\square$

This upper bound cannot be improved with the type of simply repeating string used in the proof of Theorem 9. That is, if $|\alpha|$ is even and α^{t+1} is $(2t)$-reflecting for $\mathcal{L}(5)$, then $|\alpha| \geq 24$. In fact, Theorem 10 provides a stronger result.

Theorem 10: Let $c < 4/3$. If $|\alpha|$ is even and $\alpha^{\lfloor ct \rfloor}$ is a $(2t-1)$-reflecting sequence for $\mathcal{L}(5)$ for all positive integers t, then $|\alpha| \geq 24$.

Proof: Suppose $\alpha^{\lfloor ct \rfloor}$ is $(2t-1)$-reflecting on $\mathcal{L}(5)$, and α contains only two 00 pairs. There are t left-to-right traversals of each of the chains with labels 0000, 0011, and 1100, so there are $4t$ marks on 00 pairs altogether. Therefore, some occurrence of α has at least $\lceil 4t/\lfloor ct \rfloor \rceil \geq \lceil 4/c \rceil \geq 4$ marks on 00. In order to have 4 marks on only two 00 pairs, α must be a rotation (by an odd number of positions) of $(01+10+11)^*11(01)^*00(01+10)^*00(10)^*11$. (See Table 2.) However, in that case α^* has no occurrence of $00(01+10)^*00$ beginning at an *odd* index; hence it does not reflect even two times on the chain with label 1111. Because of this contradiction, α must contain at least three 00 pairs.

Similarly, α must contain at least three occurrences of the pairs 01, 10, and 11. (In the case of 01 pairs, for example, the contradiction is that there is no occurrence of $10(00+11)^*10$ beginning at an odd index.) Therefore, α contains at least 12 pairs, from which the theorem follows. □

Acknowledgement

We thank Donald Chinn for helpful discussions.

References

[1] R. Aleliunas. A simple graph traversing problem. Master's thesis, University of Toronto, Apr. 1978. Technical Report No. 120.

[2] R. Aleliunas, R. M. Karp, R. J. Lipton, L. Lovász, and C. Rackoff. Random walks, universal traversal sequences, and the complexity of maze problems. In *20th Annual Symposium on Foundations of Computer Science*, pages 218–223, San Juan, Puerto Rico, Oct. 1979. IEEE.

[3] N. Alon, Y. Azar, and Y. Ravid. Universal sequences for complete graphs. *Discrete Applied Mathematics*, 27:25–28, 1990.

[4] A. Borodin, W. L. Ruzzo, and M. Tompa. Lower bounds on the length of universal traversal sequences. *Journal of Computer and System Sciences*, 45(2):180–203, Oct. 1992.

[5] A. K. Chandra, P. Raghavan, W. L. Ruzzo, R. Smolensky, and P. Tiwari. The electrical resistance of a graph captures its commute and cover times. In *Proceedings of the Twenty First Annual ACM Symposium on Theory of Computing*, pages 574–586, Seattle, WA, May 1989.

[6] J. D. Kahn, N. Linial, N. Nisan, and M. E. Saks. On the cover time of random walks on graphs. *Journal of Theoretical Probability*, 2(1):121–128, Jan. 1989.

[7] M. Tompa. Lower bounds on universal traversal sequences for cycles and other low degree graphs. *SIAM Journal on Computing*, 21(6):1153–1160, Dec. 1992.