On a Construction of Friedman

Jeffrey Shallit* and Ming-wei Wang
Department of Computer Science
University of Waterloo
Waterloo, Ontario, Canada N2L 3G1
shallit@graceland.uwaterloo.ca
m2wang@math.uwaterloo.ca

March 28, 2022

Abstract

H. Friedman obtained remarkable results about the longest finite sequence $x$ such that for all $i \neq j$ the word $x[i..2i]$ is not a subsequence of $x[j..2j]$. In this note we consider what happens when “subsequence” is replaced by “subword”.

1 Introduction

We say a word $y$ is a subsequence of a word $z$ if $y$ can be obtained by striking out 0 or more symbols from $z$. For example, “iron” is a subsequence of “introduction”. We say a word $y$ is a subword of a word $z$ if there exist words $w, x$ such that $z = wyx$. For example, “duct” is a subword of “introduction”.

We use the notation $x[k]$ to denote the $k$'th letter chosen from the string $x$. (The first letter of a string is $x[1]$.) We write $x[a..b]$ to denote the subword of $x$ of length $b - a + 1$ starting at position $a$ and ending at position $b$.

Recently H. Friedman has found a remarkable construction that generates extremely large numbers [1, 2]. Namely, consider words over a finite alphabet $\Sigma$ of cardinality $k$. If an infinite word $x$ has the property that for all $i, j$ with $0 < i < j$ the subword $x[i..2i]$ is not a subsequence of $x[j..2j]$, call it self-avoiding. We apply the same definition for a finite word $x$ of length $n$, imposing the additional restriction that $j \leq n/2$.

Friedman shows there are no infinite self-avoiding words over a finite alphabet. Furthermore, he shows that for each $k$ there exists a longest finite self-avoiding word $x$ over an alphabet of size $k$. Call $n(k)$ the length of such a word. Then clearly $n(1) = 3$ and

*Research supported in part by a grant from NSERC.

1Europeans sometimes use the term “factor” for what we have called “subword”, and they use the term “subword” for what we have called “subsequence”.
a simple argument shows that $n(2) = 11$. Friedman shows that $n(3)$ is greater than the incomprehensibly large number $A_{7198}(158386)$, where $A$ is the Ackermann function.

Jean-Paul Allouche asked what happens when “subsequence” is replaced by “subword”. A priori we do not expect results as strange as Friedman’s, since there are no infinite anti-chains for the partial order defined by “$x$ is a subsequence of $y$”, while there are infinite anti-chains for the partial order defined by “$x$ is a subword of $y$”.

2 Main Results

If an infinite word $x$ has the property that for all $i, j$ with $0 \leq i < j$ the subword $x[i..2i]$ is not a subword of $x[j..2j]$, we call it weakly self-avoiding. If $x$ is a finite word of length $n$, we apply the same definition with the additional restriction that $j \leq n/2$.

**Theorem 1** Let $\Sigma = \{0, 1, \ldots, k - 1\}$.

(a) If $k = 1$, the longest weakly self-avoiding word is of length 3, namely 000.

(b) If $k = 2$, there are no weakly self-avoiding words of length $> 13$. There are 8 longest weakly self-avoiding words, namely 001011111010, 001011111011, 0011110101010, 0011110101011 and the four words obtained by changing 0 to 1 and 1 to 0.

(c) If $k = 3$, there exists an infinite weakly self-avoiding word.

**Proof.**

(a) If a word $x$ over $\Sigma = \{0\}$ is of length $\geq 4$, then it must contain 0000 as a prefix. Then $x[1..2] = 00$ is a subword of $x[2..4] = 000$.

(b) To prove this result, we create a tree whose root is labeled with $\epsilon$, the empty word. If a node’s label $x$ is weakly self-avoiding, then it has two children labeled $x0$ and $x1$. This tree is finite if and only if there is a longest weakly self-avoiding word. In this case, the leaves of the tree represent non-weakly-self-avoiding words that are minimal in the sense that any proper prefix is weakly self-avoiding.

Now we use a classical breadth-first tree traversal technique, as follows: We maintain a queue, $Q$, and initialize it with the empty word $\epsilon$. If the queue is empty, we are done. Otherwise, we pop the first element $q$ from the queue and check to see if it is weakly self-avoiding. If not, the node is a leaf, and we print it out. If $q$ is weakly self-avoiding then we append $q0$ and $q1$ to the end of the queue.

If this algorithm terminates, we have proved that there is a longest weakly self-avoiding word. The proof may be concisely represented by listing the leaves in breadth-first order. We may shorten the tree by assuming, without loss of generality, that the root is labeled 0.

When we perform this procedure, we obtain a tree with 92 leaves, whose longest label is of length 14. The following list describes this tree:
where there are 0’s in positions 3

0000 00111100 0011010101 001011111011
0001 00111110 0011010110 00101111100
0101 00111111 0011010111 00101111110
001000 01000000 0111010000 00101111111
001001 01000001 0111010101 0011101000
001010 01000010 0111010111 0011101001
001100 01000011 0111010110 0011101010
010001 01100001 0111101110 0011101011
010010 01100010 0111101111 0011101010
010011 01100011 0110000000 0011101010
011010 01110010 0110000010 0111000000
011011 01110011 0110000011 0111000001
011101 0010111001 0111000001 0111000000
011110 0010111010 0111000010 0111000011
011111 0010111011 0111000011 0111011110
00101100 0010111011 0110111010 0111011110
00110100 0010111100 0010111011 0111011110
00110110 0010111101 0010111101 0111011111
00110111 0010111110 0010111101 0111101010
00111000 0010111111 0010111100 0111101011
00111001 0010111100 0010111110 0111101010
00111010 0010111000 0010111110 0111101011
00111011 0010111101 0010111111 0111101011

Figure 1: Leaves of the tree giving a proof of Theorem 1 (b)

(c) Consider the word

\[ x = 22010110101110111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111
$t' = 1^u0^v1^w$. For $t$ to be a subword of $t'$ we must have $u \leq u'$, $v = v'$, and $w \leq w'$. But since the blocks of 1’s in $x$ are distinct in size, this means that the middle block of 1’s in $t$ and $t'$ must occur in the same positions of $x$. Then $u \leq u'$ implies $i \geq j$, a contradiction. ■

3 Another construction

Friedman also has considered variations on his construction, such as the following: let $M_2(n)$ denote the length of the longest finite word $x$ over $\{0, 1\}$ such that $x[i..2i]$ is not a subsequence of $x[j..2j]$ for $n \leq i < j$. We can again consider this where “subsequence” is replaced by “subword”.

**Theorem 2** There exists an infinite word $x$ over $\{0, 1\}$ such that $x[i..2i]$ is not a subword of $x[j..2j]$ for all $i, j$ with $2 \leq i < j$.

**Proof.** Let

$$x = 001001301^201701^501^{15}01^{11}01^{31}01^{23} \ldots$$

$$= 001001^{g_1}01^{g_2}01^{g_3}0 \ldots$$

where $g_1 = 3$, $g_2 = 2$, and $g_n = 2g_{n-2} + 1$ for $n \geq 3$. Then a proof similar to that above shows that every subword of the form $x[i..2i]$ contains exactly two 0’s, and hence, since the $g_i$ are all distinct, we have $x[i..2i]$ is not a subword of $x[j..2j]$ for $j > i > 1$. ■

References

[1] H. Friedman. Long finite sequences. To appear, *J. Combinat. Theory A*. Also available at [http://www.math.ohio-state.edu/foundations/manuscripts.html](http://www.math.ohio-state.edu/foundations/manuscripts.html).

[2] H. Friedman. Enormous integers in real life. Manuscript, dated June 1 2000, available at [http://www.math.ohio-state.edu/foundations/manuscripts.html](http://www.math.ohio-state.edu/foundations/manuscripts.html).