The Restricted Word Shadow Problem

R. Ahlswede[1] and V. Lebedev[2*]

[1] Department of Mathematics, University of Bielefeld

[2] IPPI (Institute for Information Transmission Problems
of the Russian Academy of Sciences (Kharkevich Institute))

September 21, 2012

Abstract

Recently we introduced and studied the shadow minimization problem under word-subword relation. In this paper we consider this problem for the restricted case and give optimal solution.

1 Introduction

In [1], [2] the minimal shadow problem for the word-subword relation was introduced. The shadow problem for words has not been studied before, whereas its analogs for sets ([3], [4], [5], [6]), sequences ([7]), and vector spaces for finite fields ([8]) are well-known.

For an alphabet $\mathcal{X} = \{0, 1, \cdots, q-1\}$ we consider the set $\mathcal{X}^k$ of words $x^k = x_1 x_2 \cdots x_k$ of length k. A word x^n is an n-subword of y^k if there exist a^i and b^{k-n-i} such that $y^k = a^i x^n b^{k-n-i}$, where $i \in \{0, 1, \cdots, k-n\}$.

Definition 1 [2]. *The shadow of y^k is the set of all its n-subword:*

$$shad_{k,n}(y^k) = \{x^n : x^n \textit{is an n-subword of } y^k\} \tag{1}$$

and for any subset $A \subset \mathcal{X}^k$ we define its shadow

*Supported in part by the Russian Foundation for Basic Research, project no 12-01-00905 and by the DFG, project AH46/7-1 "General Theory of Information Transfer".

$$shad_{k,n}(A) = \bigcup_{a^k \in A} shad_{k,n}(a^k). \tag{2}$$

In [2] we studied the problem of finding optimal or at least asymptotically optimal lower bounds on the cardinality of N-sets $A \subset \mathcal{X}^k$, that is the function

$$\triangle_{k,n}(q, N) = \min\{|\text{shad}_{k,n}(A)| : A \subset \mathcal{X}^k, |A| = N\}. \tag{3}$$

Theorem 1 [2]. *For integers $N = q^{l+v} + q^{l+v-1}(l-v)(q-1)$ and $k = l + m + v > 2l \geq 2v$, where $v = k - n$, we have*

$$\frac{1}{q^v} N \leq \triangle_{k,n}(q, N) \leq \frac{1}{q^v}\left(1 + \frac{v}{l - v + 1}\right) N. \tag{4}$$

In this paper we are interested in the restricted word shadow problem.

Let us denote by $\mathcal{X}_w^k$ the set of words $a^k \in \mathcal{X}^k$ of weight (the number of nonzero symbols) $wt(a^k) = w$.

We consider first the binary case, so $\mathcal{X} = \{0, 1\}$.

Definition 2. *For integers k,n,w,N with $1 \leq w \leq n < k$ and $1 \leq N \leq \binom{k}{w}$ we define*

$$S(k, n, w, N) = \min\{|shad_{k,n}^w(A)| : A \subset \mathcal{X}_w^k,\ |A| = N\} \tag{5}$$

where

$$shad_{k,n}^w(A) = \bigcup_{a^k \in A} shad_{k,n}^w(a^k) \tag{6}$$

and

$$shad_{k,n}^w(y^k) = \{x^n : x^n \in \mathcal{X}_w^n \text{ is an } n\text{-subword of } y^k\} \tag{7}$$

When k, n and w are specified we also use sometimes $S(N)$ for $S(k, n, w, N)$.

In this paper we solve the restricted word shadow minimization problem, namely, we determine the function $S(k, n, w, N)$ for all parameters. We also observe that our result can be easily generalized to arbitrary alphabet size.

2 The restricted word shadow problem

Let $a = a_1 a_2 \cdots a_m \in \mathcal{X}^m$ and $b = b_1 b_2 \cdots b_n \in \mathcal{X}^n$ then we denote by $ab = c = c_1 c_2 \cdots c_{m+n} \in \mathcal{X}^{m+n}$ with

$$c_1 = a_1, \cdots, c_m = a_m, c_{m+1} = b_1, \cdots, c_{m+n} = b_n.$$

For a subset $A \subset \mathcal{X}^n$ we denote by $AB = \{ab : a \in A, b \in B\}$.

It turns out to be very convenient to introduce the sets

$$A(\epsilon, \delta) = \epsilon \mathcal{X}^{k-2}_{w-\epsilon-\delta} \delta \ \text{ for } \ \epsilon, \delta \in \mathcal{X} = \{0, 1\}. \tag{8}$$

So words in $A(\epsilon, \delta)$ are from $\mathcal{X}^k_w$, start with ϵ, end with δ, and between these two letters have a word of length $k-2$ and weigth $w - \epsilon - \delta$.

Thus we have the partition

$$\mathcal{X}^k_w = \bigcup_{\epsilon,\delta \in \mathfrak{X}} A(\epsilon, \delta), \ \text{where } |A(\epsilon, \delta)| = \binom{k-2}{w-\epsilon-\delta}, \tag{9}$$

or explicitly

$$|A(1,1)| = \binom{k-2}{w-2}, |A(1,0)| = |A(0,1)| = \binom{k-2}{w-1}, |A(0,0)| = \binom{k-2}{w}.$$

Consider the following partition of $\mathcal{X}^k_w$

$$\mathcal{X}^k_w = \bigcup_s J_s \tag{10}$$

where

$$J_1 = A(1,1), J_2 = A(1,10) \cup A(01,1), J_3 = A(1,100) \cup A(01,10) \cup A(001,1),$$

and so on.

Thus, for J_{s+1} we have

$$J_{s+1} = \bigcup_{i=0}^{s} A(0^i 1, 10^{s-i}).$$

Finally, for the last partition class $J_{k-w=1}$ we have

$$J_{k-w+1} = \bigcup_{i=0}^{k-w} A(0^i 1, 10^{k-w-i}).$$

In other words, we have the following partition classes:

$$\begin{matrix} 1 & b & 1 \end{matrix} \tag{11}$$

$$\begin{matrix} 1 & b & 1 & 0 \\ 0 & 1 & b & 1 \end{matrix} \tag{12}$$

$$\begin{matrix} 1 & b & 1 & 0 & 0 \\ 0 & 1 & b & 1 & 0 \\ 0 & 0 & 1 & b & 1 \end{matrix} \tag{13}$$

for any b from $\mathcal{X}_{w-2}^{k-2}$, $\mathcal{X}_{w-2}^{k-3}$ and $\mathcal{X}_{w-2}^{k-4}$ respectively.
We continue this procedure and for any b from $\mathcal{X}_{w-2}^{k-c-1}$ we take

$$\bigcup_{i=0}^{s} 0^i 1b10^{s-i}. \tag{14}$$

For example if $s = 4$ then we take

$$\begin{matrix} 1 & b & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & b & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & b & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & b & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & b & 1 \end{matrix} \tag{15}$$

The last partition class is given by

$$\bigcup_{i=0}^{k-w} 0^i 1^w 0^{k-w-i}. \tag{16}$$

2.1 Case $n = k - 1$

Note that for the case $n = k - 1$ the shadow of $A \subset \mathcal{X}_w^k$ can be defined as $\text{shad}^w(A) = \text{shad}_L^w(A) \cup \text{shad}_R^w(A)$ where $\text{shad}_L^w(A) = \{a_2 a_3 \cdots a_k : wt(a_2 \cdots a_k) = w\}$ and $\text{shad}_R(A) = \{a_1 a_3 \cdots a_{k-1} : w(a_1 a_3 \cdots a_{k-1}) = w\}$.

In this case we have a nice graph illustration for our partition.

Consider a graph $G = (V, E)$ associated with this word-subword relation: the vertex set V is $\mathcal{X}_w^k$. Two vertices a^k and b^k form an edge $(a^k, b^k) \in E$ if and only if $\text{shad}_{k,k-1}^w(a^k) \cap \text{shad}_{k,k-1}^w(b^k) \neq \emptyset$. Note that there is one to one correspondence between edges in E and elements from $\mathcal{X}_w^{k-1}$.

It follows from the partition described above that the graph G consists of $\binom{k-2}{w-2}$ isolated vertices P_0, $\binom{k-3}{w-2}$ paths of length 1: P_1, $\binom{k-4}{w-2}$ paths of length 2: P_2 and so on.

Given integer $1 \leq N \leq \binom{n}{w}$ the restricted word shadow problem for the case $n = k-1$ is equivalent to the problem of finding N vertices of the graph G that are incident with minimal number of edges.

We order all vertices of the graph G in the following way. We start with vertices from P_0 in arbitrary order. Then we consider set P_1 from the first partition class in arbitrary order and order vertices from P_1 in compliance with (12): (first $1b10$ and then $01b1$). Then we do the same with sets $P_2, P_3, \ldots, P_{k-w+1}$. We take sets from $(s+1)$-th partition class in arbitrary order and for the set P_s order vertices from P_s in compliance with (14).

It is not hard to see now that first N vertices, in the described ordering, have minimum number of edges incident with them.This clearly dives us an optimal solution to the problem.

Hence for N $\leq \binom{k-2}{w-2}$ we have $S(k, k-1, w, N) = 0$ and for $\binom{k-2}{w-2} <$ N $\leq \binom{k-2}{w-2} + 2\binom{k-3}{w-2}$ we have

$$S(k, k-1, w, N) = \begin{cases} z & , \text{ if } N = 2z + \binom{k-2}{w-2} \\ z+1 & , \text{ if } N = 2z + 1 + \binom{k-2}{w-2}. \end{cases}$$

Now easy calculation gives us the following numerical formulation of our result.

Theorem 2. For

$$\binom{k-2}{w-2} + \ldots + (c-1)\binom{k-c}{w-2} \leq N \leq \binom{k-2}{w-2} + \ldots (c-1)\binom{k-c}{w-2} + c\binom{k-c-1}{w-2}$$

we have

$$S^c := S(\binom{k-2}{w-2} + \ldots + (c-1)\binom{k-c}{w-2}) = \binom{k-3}{w-2} + \ldots + (c-2)\binom{k-c}{w-2}$$

and

$$S(k, k-1, w, N) = \begin{cases} S^c + (c-1)z & , \text{ if } N = cz + \binom{k-2}{w-2} + \ldots + (c-1)\binom{k-c}{w-2} \\ S^c + (c-1)z + m & , \text{ if } N = cz + m + \binom{k-2}{w-2} + \ldots + (c-1)\binom{k-c}{w-2} \end{cases}$$

where $m = 1, 2, \ldots, c-1$ and $c = 2, 3, \ldots, k-w+1$.

2.2 General case: $w \le n \le k-1$

For general case we have that

$$\mathrm{shad}^{w}_{k,n} J_{s+1}(k) = J_{s-v+1}(n) \tag{17}$$

where $v = k - n$.

Thus the described above ordering of $\mathcal{X}^k_w$ also gives us an optimal solution to the problem in this general case. The set of first N vectors from $\mathcal{X}^k_w$ has the minimal possible restricted shadow and so we have

Theorem 3. For $N \le \binom{k-2}{w-2} + \ldots + v\binom{k-v-1}{w-2}$ we have $S(N) = 0$ and for

$$\binom{k-2}{w-2}+\ldots+(c-1)\binom{k-c}{w-2} \le N \le \binom{k-2}{w-2}+\ldots(c-1)\binom{k-c}{w-2}+c\binom{k-c-1}{w-2}$$

we have

$$S^c = S(\binom{k-2}{w-2}+\ldots+(c-1)\binom{k-c}{w-2}) = \binom{k-2-v}{w-2}+\ldots+(c-v-1)\binom{k-c}{w-2}$$

and

$$S(k,n,w,N) = \begin{cases} S^c + (c-1)z & \text{, if } N = cz + \binom{k-2}{w-2} + \ldots + (c-1)\binom{k-c}{w-2} \\ S^c + (c-1)z + m & \text{, if } N = cz + m + \binom{k-2}{w-2} + \ldots + (c-1)\binom{k-c}{w-2} \end{cases}$$

where $m = 1, 2, \ldots, c-1$, $v = k - m$ and $c = 2, 3, \ldots, k - w + 1$.

Remark. We note that our result can be easily extended to the q-ary case $\mathcal{X} = \{0, 1, \ldots, q-1\}$. Consider only the case $n = k - 1$.

For integers $k, w, N \in \mathbb{N}$ with $1 \le w \le k$ and $1 \le N \le \binom{k}{w}$ we define

$$S_q(N) = \min\{|\mathrm{shad}^w(A)| : A \subset \mathcal{X}^k_w, |A| = N\}$$

The proof goes along the same line as the proof of Theorem 2. We just replace a symbol 1 in (11)-(16) to any nonzero symbol from $\mathcal{X}$. So we have

Theorem 4. For

$$(q-1)^w\binom{k-2}{w-2} + \ldots + (c-1)(q-1)^w\binom{k-c}{w-2} < N \le$$

$$\leq (q-1)^w\binom{k-2}{w-2} + \ldots (c-1)(q-1)^w\binom{k-c}{w-2} + c(q-1)^w\binom{k-c-1}{w-2}$$

we have

$$S_q(\gamma) = (q-1)^w\binom{k-3}{w-2} + \ldots + (c-2)(q-1)^w\binom{k-c}{w-2},$$

where

$$\gamma := (q-1)^w\binom{k-2}{w-2} + \ldots + (c-1)(q-1)^w\binom{k-c}{w-2}$$

and

$$S_q(N) = \begin{cases} S_q(\gamma) + (c-1)z & , \text{ if } N = cz + \gamma \\ S_q(\gamma) + (c-1)z + m & , \text{ if } N = cz + m + \gamma, \end{cases}$$

where $m = 1, 2, \ldots, c-1$ and $c = 2, 3, \ldots, k-w+1$

References

[1] Ahlswede R., Lebedev V., Shadows under the word-subword relation, Twelfth International Workshop on Algebraic and Combinatorial Coding Theory, September 5-11, 2010, Akademgorodok, Novosibirsk (Russia) pp. 16-19.

[2] Ahlswede R., Lebedev V., Shadows under the word-subword relation, Problems of Information Transmission, 2012. Vol. 48 (1), pp. 30-45.

[3] Schützenberger M.P., A characteristic property of certain polynomials of E.F. Moore and C.E. Shannon, in : RLE Quarterly Progress Report No. 55, Research Laboratory of Electronics, M.I.T., 117-118, 1959.

[4] Kruskal J.B., The number of simplices in a complex, in: Mathematical optimization Techniques, Berkeley and Los Angeles, 251-278, 1963.

[5] Katona G., A theorem on finite sets, in: Theory of Graphs, Proc. Colloq. Tihany 1966, Akadémiai Kiadó, 187-207, 1968.

[6] Lindström B.A. and Zetterström H.O., A combinatorial problem in the k-adic number systems, Proc. Amer. Math. Soc., Vol. 18, 166-170, 1967.

[7] Ahlswede R. and Cai N., Shadows and isoperimetry under the sequence-subsequence relation, Combinatorica 17 (1), 11-29, 1997.

[8] Chowdhury A. and Patkos B., Shadows and intersections in vector spaces, in: J. Combin. Theory Ser. A, 117(8), 1095-1106, 2010.