COUNTING SCATTERED PALINDROMES IN FINITE WORDS

PALAK PANDOH

Assistant Professor
School of Mathematics,
SMVD university,
JK, India.

Jun 21, 2021

OUTLINE

- Introduction and Motivation
- Basic Terminologies
- Scattered palindromes in 1D words
- Conclusions and Future directions.

Introduction and Motivation

- Palindromes, i.e., words symmetric under reversal, have been the subject of many works in the past especially in the study of Sturmian words.
- There has been an interest in the identification of palindromes in the areas of mathematics, theoretical computer science, and biology and have applications to sting processing, bio-informatics, error correcting codes, etc.
- Bounds on the number of palindromic factors in words have been extensively studied. (Droubay, X. et al[1], M.C., Anisiu, et.al [2], Fici, Zamboni [3], etc.).

---

[1] Droubay, X., Justin, J. and Pirillo, G. (2001).Episturmian words and some constructions of de Luca and Rauzy. Theoretical Computer Science.

[2] Anisiu, Mira Cristiana and Anisiu, Valeriu and Kasa, Zoltan (2010). Total palindromic complexity of finite words. Discrete Mathematics.

[3] Gabriele, F. and Zamboni, L. Q. (2013). On the least number of palindromes contained in an infinite word. Theoretical Computer Science.

Basic Terminologies (One-Dimensional Words)

- Alphabet Σ: A finite non-empty set of symbols
- Word u : $u_1 u_2 \cdots u_n; u_i \in \Sigma$
- Factor of u : $u_i u_{i+1} \cdots u_j; 1 \leq i \leq j \leq n$
- $v = \alpha_1 \cdots \alpha_n$ is a scattered subword of $u = \beta_0 \alpha_1 \beta_1 \cdots \alpha_n \beta_n$
- $|u|$: Length of a finite word u is the number of letters in u

Example

Let $\Sigma = \{a, b, c\}$. Then $u = abcaba$ is a word over Σ.

- $|u| = 6$.
- bca is a factor of u.
- $abba$ is a scattered subword of u.

BASIC TERMINOLOGIES (ONE-DIMENSIONAL WORDS)

- Σ^n: Set of all words of length n over Σ
- Σ^* : Set of all possible words over Σ
- $Alph(u)$: Set of all sub-strings of u of length 1
- u is said to be primitive if $u = v^n$ for $n > 0$, implies $n = 1$
- For $u = u_1 u_2 \cdots u_n$, u^R= $u_n \cdots u_2 u_1$
- u is a palindrome if $u = u^R$

Example

The word $abccba$ is a palindrome over $\Sigma = \{a, b, c\}$.

Scattered Palindromes in 1D words

- Palindromic subsequences, i.e., scattered palindromic subwords, in words are studied in Holub, S. et al.[4], Mullner, C. et al.[5] Ago, K. et al.[6].
- The authors in Holub, S. et al. and Ago, K. et al. studied some properties of the scattered palindromic subwords of binary and ternary words.
- Mullner, C. et al. provided bounds on the minimum length of the longest scattered palindromic subword in some restricted classes of words of a given length.
- We enumerate scattered palindromic subwords in one-dimensional words.

---

[4] Holub, S. and Saari, K. (2009). On highly palindromic words. Discrete Applied Mathematics.

[5] Mullner, C. Ryzhikov, A.(2019). Palindromic Subsequences in Finite Words. In. Language and automata theory and applications, Lecture Notes in Computer Science, Springer.

[6] Ago, K. and Basic, B. (2020).On highly palindromic words: The ternary case. Discrete Applied Mathematics.

TERMINOLOGIES

- $PAL(u)$: Set of all non-empty palindromic factors of u
- $P(u) = |PAL(u)|$
- $SPAL(u)$: Set of all non-empty scattered palindromic subwords of u
- $SP(u) = |SPAL(u)|$

Example

Let $u = abaa$. Then

$$PAL(u) = \{a, b, aba, aa\} \implies P(u) = 4$$

$$SPAL(u) = \{a, b, aba, aa, aaa\} \implies SP(u) = 5.$$

Scattered palindromes in 1D words

Droubay, X. et al.[7] proved that at most one new palindromic factor can be created on the concatenation of a letter to a 1D word, thus $P(u) \leq |u|$.

Theorem (Mahalingam , Pandoh)

For a given finite word u, $P(u) \leq |u| \leq SP(u)$.

Block word u: $u = v_1 v_2 \cdots v_r$; $v_i = a_i^{n_i}$ such that $a_i \in \Sigma$ are distinct.

Theorem (Mahalingam , Pandoh)

For a finite word u, $SP(u) = |u|$ *if and only if* u *is a block word.*

[7] Droubay, X., Justin, J. and Pirillo, G. (2001).Episturmian words and some constructions of de Luca and Rauzy. Theoretical Computer Science.

SCATTERED PALINDROMES IN 1D WORDS

Lemma (Mahalingam , Pandoh)

At most F_n scattered palindromic subwords can be created on the concatenation of a letter to a word of length $n-1$.

- Let $u = u'x$ be a word of length n for $x \in \Sigma$ and $u' \in \Sigma^*$
- Let n_1 be the number of scattered palindromic subwords added by the concatenation of x to the word u'
- If $|u'|_x = 0$, then $n_1 = 1 \leq F_n$
- If $|u'|_x \neq 0$, then $u' = \alpha x \beta$ such that $|\alpha|_x = 0$ and $\beta \in \Sigma^*$
- $n_1 \leq SP(\beta) + 1$
- Since $|\alpha x \beta| = n - 1$ and $|\beta| \leq n - 2$, by induction, $n_1 \leq F_n - 1 + 1 = F_n$

Theorem (Mahalingam , Pandoh)

A word of length n has at most $F_{n+2} - 1$ scattered palindromic subwords.

Scattered Palindromes of a given length

We investigate the maximum number of scattered palindromic subwords of length i in a word u, denoted by $SP_i(u)$, of length n.

- $SP_1(u) \leq n$.
- $SP_2(u) \leq \lfloor \frac{n}{2} \rfloor$.
- $SP_3(u) \leq \begin{cases} \frac{n^2-2n}{4}, & if\ n\ is\ even, \\ (\frac{n-1}{2})^2, & if\ n\ is\ odd. \end{cases}$
- $SP_4(u) \leq \begin{cases} \frac{n^2-2n}{4}, & if\ n\ is\ even, \\ (\frac{n-1}{2})^2, & if\ n\ is\ odd. \end{cases}$
- $SP_{n-2}(u) \leq \lceil \frac{n}{2} \rceil$, for $n \geq 6$.
- $SP_{n-1}(u) \leq \begin{cases} 1, & if\ n\ is\ odd, \\ 2, & if\ n\ is\ even. \end{cases}$
- $SP_n(u) \leq 1$.

Scattered palindromes in 1D words

We deduce the following:

Proposition

The number of scattered palindromic subwords in a word of length $n \geq 7$*, with alphabet size* q *is bounded above by*

$$\frac{n^2 + 2n + 4}{2} + \sum_{i=5}^{n-3} q^{\lceil \frac{i}{2} \rceil}.$$

We show that the maximum number of scattered palindromic subwords in a word depends on both the length and the number of distinct letters in the word.

SCATTERED PALINDROMES IN 1D WORDS

Let $a_i \in \Sigma$ be distinct and

$SP^{n,q} \;=\; \max\{SP(w) \mid w \in \Sigma^n$ and $|Alph(w)| = q\}$.

q	$SP^{4,q}$	Words
1	4	$a_1a_1a_1a_1$
2	6	$a_1a_2a_1a_2, a_1a_2a_2a_1$
3	6	$a_1a_2a_3a_1$
4	4	$a_1a_2a_3a_4$

q	$SP^{5,q}$	Words
1	5	$a_1a_1a_1a_1a_1$
2	9	$a_1a_2a_1a_2a_1$
3	10	$a_1a_2a_3a_2a_1$
4	8	$a_1a_2a_3a_4a_1$
5	5	$a_1a_2a_3a_4a_5$

Theorem (Mahalingam , Pandoh)

Let $SP^{n,q} \;=\; Max\{SP(u) \mid u \in \Sigma^n$ *and* $|Alph(u)| = q\}$*, then for* $q \geq \frac{n}{2}$*,*

$$SP^{n,q} = 2^{n-q}(2q - n + 2) - 2.$$

If $q \geq \frac{n}{2}$, the maximum number of scattered palindromic subwords in a word of length n is

$$\begin{cases} 3(2^{\lceil \frac{n}{2} \rceil - 1}) - 2, & if\ n\ is\ odd, \\ 2^{\frac{n}{2}+1} - 2, & if\ n\ is\ even \end{cases}$$

and this bound is achieved when $q = \begin{cases} \lceil \frac{n}{2} \rceil, & if\ n\ is\ odd \\ \frac{n}{2} \text{ and } \frac{n}{2}+1, & if\ n\ is\ even. \end{cases}$

Scattered palindromes in 1D words

Theorem (Pandoh)

Let w be a word of length n with $|Alph(w)| = q < \lceil \frac{n}{2} \rceil$. Then there exists a word w' of length n with $|Alph(w')| = q+1$ such that $SP(w) < SP(w')$.

For the alphabet size $q \leq \lceil \frac{n}{2} \rceil$, the maximum number of scattered palindromic subwords in a word of length n is achieved only when $q = \lceil \frac{n}{2} \rceil$.

Theorem (Pandoh)

The maximum number of scattered palindromic subwords in a word of length n

$$= \begin{cases} 3(2^{\lceil \frac{n}{2} \rceil - 1}) - 2, & if\ n\ is\ odd, \\ 2^{\frac{n}{2}+1} - 2, & if\ n\ is\ even \end{cases}$$

and this bound is achieved only when the alphabet size

$$q = \begin{cases} \lceil \frac{n}{2} \rceil, & if\ n\ is\ odd, \\ \frac{n}{2} \text{ and } \frac{n}{2}+1, & if\ n\ is\ even. \end{cases}$$

Scattered palindromes in a finite word

A word w of length n is called scattered palindromic rich if

$$SP(w) = \begin{cases} 3(2^{\lceil \frac{n}{2} \rceil - 1}) - 2, & if\ n\ is\ odd, \\ 2^{\frac{n}{2}+1} - 2, & if\ n\ is\ even. \end{cases}$$

Theorem (Pandoh)

A word w of length n is scattered palindromic rich if and only if it is one of the following forms:

$$\begin{cases} a_1a_2a_2a_1,\ a_1a_2a_1a_2,\ a_1a_2a_3a_1; & for\ n = 4, \\ a_1 \cdots a_{\frac{n}{2}} a_{\frac{n}{2}} \cdots a_1,\ a_1 \cdots a_{\frac{n}{2}} a_{\frac{n}{2}+1} a_{\frac{n}{2}-1} \cdots a_1; & \text{for } n \geq 6 \text{ and even,} \\ a_1a_2 \cdots a_{\frac{n-1}{2}-1} a_{\frac{n+1}{2}} a_{\frac{n-1}{2}-1} \cdots a_1; & \text{for } n \text{ odd,} \end{cases}$$

where $a_i \in \Sigma$ are distinct.

Scattered palindromes in a finite word

$n\backslash q$	1	2	3	4	5	6	7	8	9	10
1	1	–	–	–	–	–	–	–	–	–
2	2	2	–	–	–	–	–	–	–	–
3	3	4	3	–	–	–	–	–	–	–
4	4	6	6	4	–	–	–	–	–	–
5	5	9	10	8	5	–	–	–	–	–
6	6	12	14	14	10	6	–	–	–	–
7	7	17	21	22	18	12	7	–	–	–
8	8	22	28	30	30	22	14	8	–	–
9	9	30	41	45	46	38	26	16	9	–
10	10	38	54	60	62	62	46	30	18	10

Table: Values of $SP^{n,q}$ for $1 \leq n, q \leq 10$

SCATTERED PALINDROMES IN A FINITE WORD

A word w of length n with $|Alph(w)| = q$ is called scattered q-ary palindromic rich if $SP(w) = SP^{n,q}$.

Mahalingam, Pandoh proved:

- A word of length n with $|Alph(w)| = q$ such that $q \geq \frac{n}{2}$ is scattered q-ary palindromic rich if and only if $SP(w) = 2^{n-q}(2q - n + 2) - 2$.
- A word of length n is scattered 2-ary palindromic rich if and only if

$$\begin{cases} F_{\frac{n+9}{2}} - 4, & \text{if } n \text{ is odd,} \\ 2F_{\frac{n}{2}+3} - 4, & \text{if } n \text{ is even.} \end{cases}$$

  where F_n is the n-th Fibonacci number.

Classification of scattered q-ary palindromic rich words:

- Pandoh: A word w of length n with $|Alph(w)| = q$ such that $q > \frac{n}{2}$ is scattered q-ary palindromic rich if and only if w is of the form

$$a_1 a_2 \cdots a_q a_{n-q} a_{n-q-1} \cdots a_1,$$

where $a_i \in \Sigma$ are distinct.

- Mahalingam, Pandoh: A word w of length n is scattered palindromic 2-ary rich if and only if

$$w = \begin{cases} (xy)^{\frac{n-1}{2}} x, & \text{if } n \text{ is odd,} \\ (xy)^{\frac{n}{2}},\ vv^R \ \text{ for } \ v = (xy)^{(\frac{n}{4})} & \text{if } n \text{ is even;} \end{cases}$$

.

Conclusions and Future Directions

- We have studied scattered palindromic subwords in a word of length n.
- We have obtained lower and upper bounds for the number of scattered palindromic subwords in any word.
- We gave the value of $SP^{n,q}$ for $q \geq \frac{n}{2}$ and $q = 2$.
- Finding a tight bound of $SP^{n,q}$ for all values of q is one of our future work.
- Characterizing scattered palindromic rich words is another future direction.

Thank you!