Journal of Mathematics and Informatics
Vol. 8, 2017, 95-101
ISSN: 2349-0632 (P), 2349-0640 (online)
Published 24 September 2017
www.researchmathsci.org
DOI: http://dx.doi.org/10.22457/jmi.v8a10

Journal of
Mathematics and
Informatics

Precedence Matrices of Binary and Ternary Words and their Algebraic Properties

V. Nithya Vani[1], R. Stella Maragatham[2], S. Sriram[3] and K.G. Subramanian[4]

[1,2]Department of Mathematics, Queen Mary's College, Chennai 600004 India
Email: [1]nithyavaniv@gmail.com; [2]rstellamar@yahoo.com
[3]Department of Science and Humanities (Mathematics)
Saveetha School of Engineering, Saveetha University, Chennai 602105, India
Email: sriram.discrete@gmail.com
[4]Department of Mathematics, Madras Christian College
Tambaram, Chennai 600059, India
[4]Corresponding author. Email: kgsmani1948@gmail.com

Received 1 September 2017; accepted 22 September 2017

Abstract. A word, mathematically expressed, is a sequence of symbols in a finite set, called an alphabet. Parikh matrix is an ingenious tool providing information on certain subsequences of a word, referred to as subwords. On the other hand, based on subwords of a word, the notion of precedence matrix or p-matrix of a word has been introduced in studying a property, known as fair words. In this paper we consider p-matrix for words especially over binary and ternary alphabets and obtain several algebraic properties of the p-matrix.

Keywords: Combinatorics on words, subwords, precedence matrix

AMS Mathematics Subject Classification (2010): 68R15

1. Introduction

The theory of formal languages [6] is one of the fundamental areas of theoretical computer science. Combinatorics on words [2] is one of the topics of study and research (see, for example, [3, 7]) in the theory of formal languages but is a comparatively new area of research in Discrete Mathematics, with applications in many fields. The concept of Parikh vector [6], which gives counts of the symbols in a word, has been an important notion in the theory of formal languages. Extending this concept Mateescu et al. [5] introduced the notion of Parikh matrix of a word which gives numerical information about certain subwords of the word, including the information given by the Parikh vector of the word. Cerny [1] introduced another notion called precedence matrix or p-matrix of a word which is motivated by the notion of a fair word. Here we consider p-matrices and derive certain algebraic properties of binary and ternary words.

2. Preliminaries

A word is a finite sequence of symbols taken from a finite set called an alphabet. For example the word *abaabb* is over the binary alphabet $\{a, b\}$. An ordered alphabet is an alphabet with an ordering on its elements, denoted by the symbol <. For example, the ternary alphabet {*a*, *b*, *c*} with an ordering $a < b < c$ is an ordered alphabet, denoted as $\{a < b < c\}$. For a word *w* , the mirror image or reversal of $w = a_1 a_2 \cdots a_{n-1} a_n$, $n \geq 1$, is the word $mi(w) = a_n\, a_{n-1} \cdots a_2 a_1$ where each a_i is a symbol in an alphabet. A subword *u* of a given word *w* is a subsequence of *w*. We denote the number of such subwords *u* in a given word *w* by $|w|_u$. For example, if the word is *w* = *abaabb* over $\{a < b\}$, the number of subwords *ab* in *w* is $|w|_{ab} = 7$. The Parikh vector [6] of a word *w* gives the number of occurrences of each of the symbols in the word. For example, (3,4,2) is the Parikh vector of the word *babcbaacb* over the ternary alphabet {*a* ,*b*, *c*}. An extension of the notion of Parikh vector is the *Parikh matrix* [5] of a word. For a word *w* over an ordered alphabet Σ, the Parikh matrix *M(w)* of *w* is a triangular matrix, with 1*'s* on the main diagonal and 0*'s* below it but the entries above the main diagonal provide information on the number of certain subwords in *w*. For a binary word *u* over the ordered binary alphabet $\{a < b\}$, the Parikh matrix is

$$M(u) = \begin{pmatrix} 1 & |u|_a & |u|_{ab} \\ 0 & 1 & |u|_b \\ 0 & 0 & 1 \end{pmatrix}.$$

The notion of precedence matrix or p-matrix of a word over an alphabet has been introduced in [1]. Given the square matrices *A*, *B* of the same order and with integer entries, the matrix $A \circ B$ is defined as follows: the (*i, j*)*th* entry of $A \circ B$ is given by

$$(A \circ B)_{ij} = \begin{cases} A_{ii} + B_{ii} & if\ i = j \\ A_{ij} + B_{ij} + A_{ii} B_{jj} & if\ i \neq j \end{cases},$$

where A_{ij}, B_{ij} are the (*i, j*)*th* entries of *A*, *B* respectively.

Definition 2.1. [1] Let $\Sigma = \{a_1, a_2, \cdots, a_k\}$ be an alphabet. For a symbol $a_s \in \Sigma$ for $1 \leq s \leq k$, let E_{as} be the k×k matrix defined as $(E_{a_s})_{i,j} = 1$, if $i = j = s$ and $(E_{a_s})_{i,j} = 0$, otherwise. The precedence morphism or p-morphism on Σ is the morphism φ_k given by $\varphi_k(a_s) = E_{as}$. For a word $w = a_{i1} a_{i2} \ldots a_{im}$, $a_{ij}\ \varepsilon\ \Sigma$ for $1 \leq j \leq m$, we have $\varphi_k(w) = \varphi_k(a_{i1}) \circ \varphi_k(a_{i2}) \circ \ldots \circ \varphi_k(a_{im})$. In other words $\varphi_k(w)$ is computed by the operation $\circ$ on matrices as defined earlier. The resulting matrix $\varphi_k(w)$ is called the precedence matrix or p-matrix of *w*.

As an illustration, let $\Sigma = \{a, b, c\}$ with *a*<*b*<*c*, so that k = 3. Then

$$\varphi_3(a) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \varphi_3(b) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \varphi_3(c) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

$$\varphi_3(abcb) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \circ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \circ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \circ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 2 & 1 \\ 0 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

In fact the p-matrix φ_3 for a ternary word w over $\{a, b, c\}$ is given by

$$\varphi_3(w) = \begin{pmatrix} |w|_a & |w|_{ab} & |w|_{ac} \\ |w|_{ba} & |w|_b & |w|_{bc} \\ |w|_{ca} & |w|_{cb} & |w|_c \end{pmatrix}$$

while the p-matrix $\varphi_2(w)$ for a binary word w over $\{a, b\}$ is given by

$$\varphi_2(w) = \begin{pmatrix} |w|_a & |w|_{ab} \\ |w|_{ba} & |w|_b \end{pmatrix}.$$

3. **Properties precedence matrices of binary and ternary words**

We mainly consider only binary and ternary words and derive properties of p-matrices of these words.

For a binary word w over $\{a < b\}$, using the well-known identity [4], namely, $|w|_{ab} + |w|_{ba} = |w|_a \times |w|_b$, we note that the p-matrix of w can be formed, if the Parikh matrix of w is known. This remark cannot be extended to ternary words.

Analogous to the notion of M-ambiguity [7] of a word defined in terms of Parikh matrix, we can define p-matrix ambiguity of a word.

Definition 3.1. Let $\Sigma = \{a, b, c\}$. A ternary word w over Σ is said to be p-matrix ambiguous if there exists another ternary word v over Σ such that $\varphi_3(v) = \varphi_3(w)$. Otherwise, w is said to be p-matrix unambiguous.

For a binary word, p-matrix ambiguity can be similarly defined.

As an illustration, the ternary word u = *acbaabcc* is p-matrix ambiguous since both the words u, v = *aabccbac* have the same p-matrix $\begin{pmatrix} 3 & 4 & 4 \\ 2 & 2 & 7 \\ 2 & 2 & 3 \end{pmatrix}$.

Extending an observation in [4], we consider the following rules *A, B, C, D, E, F* and obtain conditions for p-matrix ambiguity of a ternary word. The rules *A* to *F* are given below: $(A): ab \to ba$ $(B): ba \to ab$; $(C): bc \to cb$ $(D): cb \to bc$; $(E): ac \to ca$ $(F): ca \to ac$

Theorem 3.1. Let $\Sigma = \{a, b, c\}$ with $a<b<c$. If any of the following sets (1) to (3) of rules is applicable to a ternary word w with the p-matrix M, then the application of

the rules in (1), (2) or (3) to the word w yields another ternary word with the same p-matrix M.

(1) *Rule A followed by B or B followed by A* ;

(2) *Rule C followed by D or D followed by C* ;

(3) *Rule E followed by F or F followed by E* ;

As a consequence the word w is p-matrix ambiguous.

Proof: If rule (A) is applied to the word w, then in the resulting word, the number of subword ab is decreased by 1 while if rule (B) is applied to the word w, then the number of subword ab is increased by 1 so that application of the rules as in (1), on yielding a new word does not change the number of subword ab and also the number of subword ba. Similar arguments can be made for (2) and (3).

A notion of weak-ratio property is considered in [8], which we recall here.

Two ternary words u, v over $\Sigma=\{a,b,c\}$ are said to satisfy weak-ratio property and we write $u \sim_{wr} v$, , if $|v|_x = k|u|_x$, for $x \in \{a,b,c\}$ and for some rational constant $k > 0$.

In [3], conditions are derived for the equality of the Parikh matrices of the words uv and vu , where the words u, v are over $\Sigma=\{a,b,c\}$ while in [8], given two words u, v over $\Sigma=\{a,b,c\}$, conditions are obtained for the equality of p-matrices of the words uv and vu. The following Theorem 3.3 has been established in [8].

Theorem 3.2. Let $\Sigma=\{a,b,c\}$ and let w_1, w_2 be ternary words over Σ such that $w_1 \sim_{wr} w_2$. Then $\varphi_3(w_1w_2) = \varphi_3(w_2w_1)$.

As a consequence of Theorem 3.3, we have the following result.

Theorem 3.3. Let $\Sigma=\{a,b,c\}$,

(i) For any ternary word w over Σ, $\varphi_3(w\,mi(w))=\varphi_3(mi(w)w)$.

(ii) For any two ternary words w_1, w_2 over Σ having the same Parikh vector, $\varphi_3(w_1 w_2)=\varphi_3(w_2 w_1)$.

Proof: Statements (i) and (ii) follow from Theorem 3.3, since

$|mi(w)|_x = |w|_x$, $x \in \{a,b,c\}$ so that $mi(w) \sim_{wr} w$ and $|w_1|_x = |w_2|_x$, $x \in \{a,b,c\}$ so that again $w_1 \sim_{wr} w_2$.

Theorem 3.4. Let $\Sigma=\{a,b,c\}$ and let w_1, w_2 be ternary words over Σ such that $w_1 \sim_{wr} w_2$. Then $\varphi_3(\alpha w_1 w_2 \beta w_2 w_1 \gamma) = \varphi_3(\alpha w_2 w_1 \beta w_1 w_2 \gamma)$.

Proof: We have by Theorem 3.3, $\varphi_3(w_1w_2) = \varphi_3(w_2w_1)$. Hence

$$\begin{aligned}\varphi_3(\alpha w_1 w_2 \beta w_2 w_1 \gamma) &= \varphi_3(\alpha)\varphi_3(w_1 w_2)\varphi_3(\beta)\varphi_3(w_2 w_1)\varphi_3(\gamma)\\ &= \varphi_3(\alpha)\varphi_3(w_2 w_1)\varphi_3(\beta)\varphi_3(w_1 w_2)\varphi_3(\gamma)\\ &= \varphi_3(\alpha w_2 w_1 \beta w_1 w_2 \gamma).\end{aligned}$$

Motivated by Theorem 3.3, we obtain conditions for the equality of p-matrix of uv and the transpose of the p-matrix of vu.

Theorem 3.5. Let $\Sigma = \{a, b, c\}$ and let w_1, w_2 be ternary words over Σ such that (i) $|w_1|_{ab} + |w_2|_{ab} = |w_1|_{ba} + |w_2|_{ba}$,

(ii) $|w_1|_{ac} + |w_2|_{ac} = |w_1|_{ca} + |w_2|_{ca}$ and

(iii) $|w_1|_{bc} + |w_2|_{bc} = |w_1|_{cb} + |w_2|_{cb}$.

Then $\varphi_3(w_1 w_2) = [\varphi_3(w_1 w_2)]^t$ where M^t is the transpose of the matrix *M*.

Proof: For *i* = 1,2, let $|w_i|_a = p_i$, $|w_i|_b = q_i$, $|w_i|_c = r_i$,

$|w_i|_{ab} = s_i$, $|w_i|_{ba} = t_i$, $|w_i|_{ac} = x_i$, $|w_i|_{ca} = y_i$, $|w_i|_{bc} = h_i$, $|w_i|_{cb} = k_i$.

Then

$$\varphi_3(w_i) = \begin{pmatrix} p_i & s_i & x_i \\ t_i & q_i & h_i \\ y_i & k_i & r_i \end{pmatrix}, \ i = 1,2 .$$

Now $\varphi_3(w_1 w_2) = \begin{pmatrix} p_1 + p_2 & s_1 + s_2 + p_1 q_2 & x_1 + x_2 + p_1 r_2 \\ t_1 + t_2 + q_1 p_2 & q_1 + q_2 & h_1 + h_2 + q_1 r_2 \\ y_1 + y_2 + r_1 p_2 & k_1 + k_2 + r_1 q_2 & r_1 + r_2 \end{pmatrix}$

By hypothesis, $s_1 + s_2 = t_1 + t_2$, $x_1 + x_2 = y_1 + y_2$, $h_1 + h_2 = k_1 + k_2$. On using these relations, we have

$$[\varphi_3(w_2w_1)]^t = \begin{pmatrix} p_2+p_1 & s_2+s_1+p_2q_1 & x_2+x_1+p_2r_1 \\ t_2+t_1+q_2p_1 & q_2+q_1 & h_2+h_1+q_2r_1 \\ y_2+y_1+r_2p_1 & k_2+k_1+r_2q_1 & r_2+r_1 \end{pmatrix}^t$$

$$= \begin{pmatrix} p_1+p_2 & t_1+t_2+p_1q_2 & y_1+y_2+p_1r_2 \\ s_1+s_2+q_1p_2 & q_1+q_2 & k_1+k_2+q_1r_2 \\ x_1+x_2+r_1p_2 & h_1+h_2+r_1q_2 & r_1+r_2 \end{pmatrix}$$

$$= \begin{pmatrix} p_1+p_2 & s_1+s_2+p_1q_2 & x_1+x_2+p_1r_2 \\ t_1+t_2+q_1p_2 & q_1+q_2 & h_1+h_2+q_1r_2 \\ y_1+y_2+r_1p_2 & k_1+k_2+r_1q_2 & r_1+r_2 \end{pmatrix} = \varphi_3(w_1w_{2.}).$$

4. Conclusion

Properties of precedence matrices of binary and ternary words have been studied in this paper. It will be interesting to make a study of such matrices in the context of picture arrays taking motivation from corresponding studies of Parikh matrices of picture arrays [11].

Acknowledgement. The authors thank the reviewers for useful comments. The first author K.G. Subramanian is grateful to University Grants Commission, India, for the award of Emeritus Fellowship (No.F.6-6/2016-17/EMERITUS-2015-17-GEN-5933 / (SA-II)) to him to execute his work in the Department of Mathematics, Madras Christian College.

REFERENCES

1. A.Cerny, On fair words, *Journal of Automata, Languages and Combinatorics*, 14 (2009) 163-174.
2. M.Lothaire, *Combinatorics on Words*, Cambridge Mathematical Library, Cambridge university Press, (1997.
3. L.Kari and K.Mahalingam, Involutively bordered words, *International Journal of Foundations of Computer Science*, 18 (2007) 1089-1106.
4. K.Mahalingam and K.G.Subramanian, Product of Parikh Matrices and Commutativity, *International Journal of Foundations of Computer Science*, 23 (2012) 207-223.
5. A.Mateescu and A.Salomaa, Matrix indicators for subword occurrences and ambiguity, *International Journal of Foundations of Computer Science*, 17 (2004) 277-292.
6. A.Mateescu, A.Salomaa, K.Salomaa and S.Yu, A Sharpening of the Parikh Mapping. *RAIRO-Theoretical Informatics and Applications*, 35 (2001) 551-564.
7. A.Priya Darshini, V.R.Dare, I.Venkat and K.G.Subramanian, Factors of words under an involution, *Journal of Mathematics and Informatics*, 1 (2013-14) 52-59.
8. G.Rozenberg and A.Salomaa, *Handbook of Formal Languages*, Springer-Verlag New York, Inc., 1997, Vols. 1-3.
9. A.Salomaa, Characteristic words for Parikh matrices, *Automata, Formal languages and Related topics*, (2009)117-127.

10. K.G.Subramanian, A.M.Huey and A.K.Nagar, On parikh matrices, *International Journal of Foundations of Computer Science*, 20 (2009) 211-219.
11. K.G.Subramanian, K.Mahalingam, R.Abdullah and A.K.Nagar, Two-dimensional digitized picture arrays and parikh matrices, *International Journal of Foundations of Computer Science*, 24 (2013) 393-408.