Algebraic Aspects of Generalized Parikh Matrices on Partial Words

K. Janaki, Member, IAENG, R. Krishna Kumari, Member, IAENG, S. Marichamy, S. Felixia and R. Arulprakasam, Member, IAENG *

Abstract **In this paper, we extend the concept of a generalized Parikh vector of the partial word known as $e-$generalized Parikh vector, and its related properties are studied. We also introduce the $e-$generalized Parikh matrix of the partial word and provide its characterization theorem. Further, we discuss the algebraic properties of partial words in terms of $e-$generalized Parikh matrix. In addition, we define partial line languages and confer their properties concerning $e-$generalized Parikh vector of partial words.**

Keywords: Subword, Partial word, Parikh matrix, Generalized Parikh vector, Generalized Parikh matrix, Line language

1 Introduction

Combinatorics on words is a relatively new branch of discrete mathematics with applications in many fields. In the algebraic study of words and languages, analyzing words as numerical quantities is often convenient. In this context, the Parikh matrix mapping (Parikh matrices) [16] is now famously used to study words. The perceptible epic of the Parikh matrix mapping (Parikh matrices) begins exactly in the year 2000 by Mateescu et al. as an extension of the Parikh mapping (Parikh vector) [18]. The necessity of the Parikh matrix mapping is that it enhances the characterization of words through numerical quantities. The obstacle with the Parikh vector was that much of the information about the word was lost in the transition to the vector. In that sense, an extension to a special kind of matrix called Parikh matrix would provide more information while also remaining computationally feasible. By using matrices instead of vectors more information about the word is preserved and numerical facts such as the number of occurrences of certain subwords in a word can be elegantly computed by matrix multiplication. In general a word cannot be determined by its Parikh matrix. Hence, one of the most studied questions in this area is the injectivity problem of Parikh matrix mapping and M-equivalent classes of words, which consist of words with a common Parikh matrix. A word is M-unambiguous if and only if its M-equivalence class is a singleton. However, the injectivity problem as well as characterization of M-unambiguity have been elusive problems. Some properties related to the injectivity of the Parikh matrix over a binary alphabet are analyzed in [1, 2, 3, 4, 17]. A number of investigations on different properties related to Parikh matrices have been done extensively in [7, 8, 9, 10, 13, 14, 15, 19, 20, 21, 22, 25, 28, 29, 30, 31, 32].

The concept of generalized Parikh vector introduced by R. Siromoney and V.R.Dare [26] which gives the exact positions of the symbols in a word. The Parikh vector for a word enumerates how many times each symbol of

*Manuscript received July 31,2023; revised January 08,2024.
K. Janaki is an Assistant Professor in the Department of Mathematics, Saveetha Engineering College, Saveetha Nagar, Thandalam Chennai-602105, Tamilnadu, India E-mail: (janu89lava@gmail.com).
R. Krishna Kumari is an Assistant Professor in the Department of Career Development Centre, College of Engineering and Technology, SRM Institute of Science and Technology, Kattankulathur, Chennai-603203, Tamilnadu, India E-mail: (Corresponding author E-mail: krishrengan@gmail.com).
S. Marichamy is an Assistant Professor in the Department of Mathematics, Chennai Institute of Technology, Kundrathur, Chennai-600069, Tamilnadu, India E-mail: (marichamys@citchennai.net).
S. Felixia is an Assistant Professor in the Department of Mathematics, Panimalar Engineering College, Poonamallee, Chennai-600123, Tamilnadu, India E-mail: (felixiaraj1994@gmail.com).
R. Arulprakasam is an Assistant Professor in the Department of Mathematics, College of Engineering and Technology, SRM Institute of Science and Technology, Kattankulathur, Chennai-603203, Tamilnadu, India (Corresponding author E-mail: r.aruljeeva@gmail.com).

the alphabet occurs in it, while the generalized Parikh vector indicates its position within the word. The generalized Parikh vector was examined in [27] to decrypt public key cryptosystems based on DOL/TOL systems. Sasikala et al. intoduced a language called line language [23] and analyzed a geometrical representation for this line language with the help of generalized Parikh vector. It had been proven that the generalized Parikh vectors of the same length lie on a hyperplane. For words of the same length in a binary alphabet, the generalized Parikh vectors lie in a straight line. Several investigations to developing Parikh matrices have been shown in Figure 1.

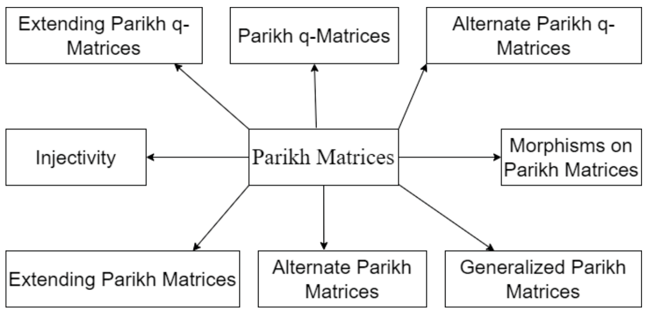

Figure 1: Studies on Parikh Matrices

As we know that genetic instructions are carried by DNA molecules. Moreover, in DNA computing, DNA strands are viewed as finite words (strings) and utilized to encode information. Some information may be missing or not visible during DNA sequencing which can be disclosed by positions representing the missing symbols in a word. As a result, in gene comparisons partial words [5] are studied instead of total words which are strings of symbols from a finite alphabet with a "don't care symbol" or "hole". Blanchet-Sadri has made a first step towards investigating languages of partial words by introducing the concept of p codes, which are sets of partial words preserving the uniqueness of factorization of partial words [6]. A partial word can also be used in a number of common and well known fields such as pattern matching and text searching. A number of studies on related to partial words have been done extensively in [11, 12]. Sasikala et al. extended the concept of line languages with respect to generalized Parikh vector to partial words where the positions representing the holes are neglected and the corresponding languages obtained are termed Partial Line languages in [24]. But words with symbols restricted to either single elements of alphabet or else alphabet itself (called a hole) have been intensively studied as partial words. This leads us to introduce $e-$generalized Parikh vector to partial words where the positions representing the holes are considered as a set of symbols over the alphabet.

The remainder of this paper is organized as follows. In Section 2, the basics are provided that are used in subsequent sections. We introduce and analyze some of the properties of $e-$generalized Parikh vectors of partial words in Section 3. Section 4 introduces the concept of $e-$generalized Parikh matrices of partial words and give a flow chart to explain the schematic moves to calculate $e-$generalized Parikh matrices of binary words defined over binary alphabets. Also we discusses the characterizations on the entries of the $e-$generalized Parikh matrices in terms of partial words. In Section 5, we extend the concept of partial line languages with respect to $e-$generalized Parikh vectors.

2 Preliminaries

In this section, we provide some basic definitions and terminologies that will be helpful for readers to understand the main section.

2.1 Words and Subwords

A finite set Σ called an alphabet whose elements are called symbols. A word x is a sequence of elements drawn from Σ. The empty word of length zero is denoted by λ. For a given Σ, let Σ^+ be the set of all possible non empty finite symbols of Σ and $\Sigma^* = \Sigma^+ \cup (\lambda)$. The length of the word x is denoted as $|x|$. If there exist words $y_1, y_2, \cdots, y_n$ and $x_0, x_1, x_2, \cdots, x_n$ over Σ such that $y = y_1 y_2 \cdots y_n$ and $x = x_0 y_1 x_1 y_2 \cdots y_n x_n$ the word $y \in \Sigma^*$ is termed as *scattered subword* of the word x. The number of occurrences of the word y in x is represented as $|x|_y$. Let a_{ij} be the word $a_i a_{i+1} \cdots a_j$ for $1 \leq i < j \leq k$ and if $i = j$ then $a_{ij} = a_i$. An ordered alphabet denoted as Σ_k is an alphabet $\Sigma = \{a_1, a_2, \cdots, a_k\}$ with the total order relation $a_1 < a_2 < \cdots < a_k$. Let a, b be two symbols in an alphabet Σ and $\delta_{a,b}$ be the *Kronecker delta* with respect to the symbols then

$$\delta_{a,b} = \begin{cases} 1 & \text{if } a = b, \\ 0 & \text{if } a \neq b. \end{cases}$$

2.2 Partial words

A partial word $u_\Diamond$ is a sequence of elements drawn from Σ which may have a number of holes. A partial word $u_\Diamond = u_\Diamond[1 \dots n]$ over Σ is a partial function

$$u : \{1, 2, \cdots, n\} \to \Sigma.$$

For $1 \le i < n$ if $u_\Diamond(i)$ is defined, then we say $i \in D(u_\Diamond)$ (the domain of $u_\Diamond$), otherwise $i \in H(u_\Diamond)$ (the set of holes). A word over Σ is a partial word over Σ with an empty set of holes. For any partial word $u_\Diamond$ over Σ, $|u|_\Diamond$ denotes its length. If $u_\Diamond$ is a partial word of length n over Σ then the companion of $u_\Diamond$ ($comp(u_\Diamond)$) is the total function $comp(u_\Diamond) : \{1, 2, \cdots, n\} \to \Sigma \cup \{\Diamond\}$ defined by

$$comp(u_\Diamond) = \begin{cases} u_\Diamond(i) & \text{if } i \in D(u) \\ \Diamond & \text{if } i \in H(u). \end{cases}$$

The symbol $\Diamond \notin \Sigma$ is viewed as a hole symbol. The set of all partial words over $\Sigma \cup \{\Diamond\} = \Sigma_\Diamond$ is denoted as $\Sigma_\Diamond^*$ and $\Sigma_\Diamond^+$ denotes the set of all partial words excluding the empty word. It is possible to identify partial words with their companions. If a partial words is complete, it does not contain $\Diamond$. Thus a partial word is any word over the alphabet $\Sigma_\Diamond$ where the symbol $\Diamond$ can be considered in a particular manner: that is the symbol $\Diamond$ can match any symbol from Σ. We note that,

1. A total word is a partial word with zero holes.
2. Empty word is not a partial word.
3. The symbol $\Diamond$ does not belong to the alphabet Σ but a standby symbol for the unknown letter.
4. The symbol $\Diamond$ is compatible to the letters of the alphabet Σ.

2.3 Parikh matrix

Let $\mathbb{N}$ be the collection of all non-negative integers and $\mathcal{M}_k$ be the collection of all $k \times k$ right triangular matrices whose elements are in $\mathbb{N}$ with unit diagonal.

Definition 2.1. *[16] The Parikh matrix mapping with respect to $\Sigma_k = \{a_1 < a_2 < \cdots < a_k\}$ denoted as ψ_k is the morphism $\psi_k : \Sigma_k^* \to \mathcal{M}_{k+1}$ defined such that for every integer $1 \le t \le k$, if $\psi_k(a_t) = (m_{ij})_{1 \le i, j \le k+1}$ then $m_{t,(t+1)} = 1$, $m_{ii} = 1$ for $1 \le i \le k+1$ and all other elements of the matrix being zero.*

Theorem 2.2. *([16]) Let x be the word over Σ_k^* then the Parikh matrix has the following characteristics*
(i) $m_{ii} = 1$ for $1 \le i \le k+1$
(ii) $m_{ij} = 0$ for each $1 \le j < i \le k+1$
(iii) $m_{i(j+1)} = |x|_{a_{ij}}$ for each $1 \le i \le j < k$.

Example 2.3. *Let $x = acbc$ be the word over Σ_3 then the Parikh matrix of x is*

$$\begin{aligned}
\psi_3(acbc) &= \psi_3(a)\psi_3(c)\psi_3(b)\psi_3(c) \\
&= \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\
&\quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & |x|_a & |x|_{ab} & |x|_{abc} \\ 0 & 1 & |x|_b & |x|_{bc} \\ 0 & 0 & 1 & |x|_c \\ 0 & 0 & 0 & 1 \end{bmatrix}.
\end{aligned}$$

Note 1. *Two words $x, y \in \Sigma_k^*$ are called* $\mathtt{M}-$equivalent *represented by $x \sim_{\mathtt{M}} y$ if and only if $\psi_k(x) = \psi_k(y)$. If there exists a word $w \neq z$ such that $z \sim_{\mathtt{M}} w$ then the word $z \in \Sigma_k^*$ is termed as* $\mathtt{M}-$ambiguous. *Otherwise z is termed as* $\mathtt{M}-$unambiguous.

Definition 2.4. *The* alternate Parikh matrix *with respect to Σ_k denoted as $\overline{\psi}_k$ is the morphism $\overline{\psi}_k : \Sigma_k^* \to \mathcal{M}_{k+1}$ defined such that*

$$\overline{\psi}_k(a_t) = (m_{ij})_{1 \le i, j \le k+1}$$

where $1 \le t \le k$ then $m_{t(t+1)} = -1$ if $t < k$, $m_{ii} = 1$ for $1 \le i \le k+1$ and all other entries of the matrix are equal to zero.

Example 2.5. *Let $x = acbc$ be the word over Σ_3 then* *the Parikh matrix of x is*

$$\begin{aligned}\overline{\psi}_3(acbc) &= \overline{\psi}_3(a)\overline{\psi}_3(c)\overline{\psi}_3(b)\overline{\psi}_3(c)\\ &= \begin{bmatrix}1 & -1 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & -1\\ 0 & 0 & 0 & 1\end{bmatrix}\\ &\quad \begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & -1 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & -1\\ 0 & 0 & 0 & 1\end{bmatrix}\\ &= \begin{bmatrix}1 & -1 & 1 & -1\\ 0 & 1 & -1 & 1\\ 0 & 0 & 1 & -2\\ 0 & 0 & 0 & 1\end{bmatrix}.\end{aligned}$$

2.4 Generalized Parikh vector and generalized Parikh matrix

Definition 2.6. *Let w be a word over Σ_k. The* generalized Parikh vector *of w is denoted as $G(w)$ and defined by*

$$\begin{aligned}G(w) &= \Big\{(g_{p_1}, g_{p_2}, \cdots, g_{p_k}) \in [0,1]^2 :\\ &\qquad g_{p_r} = \textstyle\sum_{i\in g_r} \frac{1}{2^i},\ 1 \le r \le k\Big\}\end{aligned}$$

where g_r be the set of all positions of $a_r \in \Sigma$ occurs in w.

Example 2.7. *If the word $x = ababb$ over Σ_3 then the* generalized Parikh vector *of x is*

$$\begin{aligned}G(x) &= \left(\frac{1}{2}+\frac{1}{2^3}, \frac{1}{2^2}+\frac{1}{2^4}+\frac{1}{2^5}\right)\\ &= \left(\frac{5}{2^3}, \frac{11}{2^5}\right).\end{aligned}$$

Definition 2.8. *Let w be a word of length n over Σ_k where $\Sigma_k = \{a_1, a_2, \cdots, a_k\}$ such that $w = w_1 w_2 \cdots w_t \cdots w_n$ where $w_t \in \Sigma_k$ for $1 \le t \le n$. The generalized Parikh matrix mapping denoted as ψ_{GM} is the morphism $\psi_{GM} : \Sigma_k^* \to \mathcal{M}_{k+1}$ such that $\psi_{GM}(w_t)$ is a square matrices of order $k+1$ then $\psi_{GM}(w) = \psi_{GM}(w_1)\psi_{GM}(w_2)\cdots\psi_{GM}(w_n)$ defined by the condition*

$$\psi_{GM}(w_t) = (m_{p,q})_{1\le p,q\le k+1};\ 1 \le t \le n$$

where if $w_t = a_l$ then

- *$m_{p,p} = 1$ for $1 \le p \le k+1$*
- *$m_{l,(l+1)} = \frac{1}{2^t}$ for $1 \le l \le k$*
- *remaining entries are zero*

Example 2.9. *Let $x = ababb$ be a word over Σ_2 then the generalized Parikh matrix of x is*

$$\begin{aligned}\psi_{GM}(ababb) &= \psi_{GM}(a)\psi_{GM}(b)\psi_{GM}(a)\psi_{GM}(b)\psi_{GM}(b)\\ &= \begin{bmatrix}1 & \frac{1}{2} & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & \frac{1}{2^2}\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & \frac{1}{2^3} & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\\ &\quad \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & \frac{1}{2^4}\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & \frac{1}{2^5}\\ 0 & 0 & 1\end{bmatrix}\\ &= \begin{bmatrix}1 & \frac{5}{2^3} & \frac{47}{2^8}\\ 0 & 1 & \frac{11}{2^5}\\ 0 & 0 & 1\end{bmatrix}.\end{aligned}$$

Definition 2.10. *A language $\mathbf{L}$ is said to be a line language if there exist line language l in $\mathbf{R}^2$ such that generalized Parikh vectors of $\mathbf{L}$ lie on l. The line l is called a language line. A language is said to be a finite line language if it contains only finite words.*

Throughout the paper, we consider the symbol $\Diamond$ restricted to either single elements of Σ_k or else Σ_k itself.

3 $e-$generalized Parikh vector for partial words

Dare et al. studied the concept of generalized Parikh vector to partial word where the positions associated with the holes are neglected. But words with symbols restricted to either a single element of the alphabet or the alphabet itself (called a hole) have been intensively studied as partial words. This study leads us to explore the generalization of a generalized Parikh vector to the partial word called an e-generalized Parikh vector, where the positions associated with the holes are considered a set of symbols over the alphabet. In this section, we introduce the e-generalized Parikh vector of partial words, which indicates the accurate positions of symbols in a word, including holes as a set of symbols over the alphabet.

Definition 3.1. *Consider the partial word $u_\Diamond$ over $\Sigma_\Diamond = \{a_1, a_2, \cdots, a_k\} \cup \{\Diamond\}$. Then the e-GPV of $u_\Diamond$ is denoted* *as* $G_\Diamond(u_\Diamond)$ *and defined by*

$$G_\Diamond(u_\Diamond) \quad = \quad \{(g_{p_1}, g_{p_2}, \cdots, g_{p_k}) \in [0,1]^2\}$$

where $g_{p_r} = \sum_{i \in g_r} \frac{1}{2^i} + \sum_{j \in g_\Diamond} \frac{1}{2^j},\ 1 \le r \le k.$
The set g_r *consists of all positions of* $a_r \in \Sigma$ *occurs in* $u_\Diamond$ *and the set* $g_\Diamond$ *consists of all positions of* $\Diamond$ *occurs in* $u_\Diamond$.

Example 3.2. *Let* $u_\Diamond = \Diamond ab \Diamond b$ *be the binary partial word. As "a" occurs in position* 2*, we have* $g_1 = \{2\}$ *and "*$\Diamond$*" occurs in the positions* 1 *and* 4*, we have* $g_\Diamond = \{1, 4\}$*. Therefore*

$$g_{p_1} = \left(\frac{1}{2^2}\right) + \left(\frac{1}{2} + \frac{1}{2^4}\right) = \frac{13}{2^4}.$$

Further "b" occurs in the positions 3 *and* 5*, we have* $g_2 = \{3, 5\}$*. So*

$$g_{p_2} = \left(\frac{1}{2^3} + \frac{1}{2^5}\right) + \left(\frac{1}{2} + \frac{1}{2^4}\right) = \frac{23}{2^5}.$$

The e-GPV of binary partial word $u_\Diamond$ *is*

$$G_\Diamond(u_\Diamond) = (g_{p_1}, g_{p_2}) = \left(\frac{13}{2^4}, \frac{23}{2^5}\right).$$

Theorem 3.3. *Let* $u_\Diamond, v_\Diamond \in \Sigma_\Diamond$*, then*

$$G_\Diamond(u_\Diamond v_\Diamond) = G_\Diamond(u_\Diamond) + \frac{1}{2^{|u_\Diamond|}} G_\Diamond(v_\Diamond).$$

Proof. Let the e-GPV of $u_\Diamond$ and $v_\Diamond$ be

$$\begin{aligned} G_\Diamond(u_\Diamond) &= (g_{p_1}, g_{p_2}, \cdots, g_{p_k}) \\ G_\Diamond(v_\Diamond) &= \left(g_{p'_1}, g_{p'_2}, \cdots, g_{p'_k}\right) \end{aligned}$$

where $u_\Diamond$ and $v_\Diamond$ be the partial words over $\Sigma_\Diamond$. As evidenced by the concatenation of $u_\Diamond$ and $v_\Diamond$, symbols of $v_\Diamond$ fall after the symbols of $u_\Diamond$. To remunerate this shift through the length of $u_\Diamond$, the components $g_{p'_1},\ g_{p'_2}, \cdots, g_{p'_k}$ are multiplied by $\frac{1}{2^{u_\Diamond}}$. Therefore

$$\begin{aligned} G_\Diamond(u_\Diamond v_\Diamond) &= \Bigg(g_{p_1} + \frac{1}{2^{u_\Diamond}}(g_{p'_1}), g_{p_2} + \frac{1}{2^{u_\Diamond}}(g_{p'_2}), \cdots, \\ &\qquad g_{p_k} + \frac{1}{2^{u_\Diamond}}(g_{p'_k})\Bigg) \\ &= (g_{p_1}, \cdots, g_{p_k}) + \frac{1}{2^{u_\Diamond}}\left(g_{p'_1}, \cdots, g_{p'_k}\right) \\ &= G_\Diamond(u_\Diamond) + \frac{1}{2^{|u_\Diamond|}} G_\Diamond(v_\Diamond). \end{aligned}$$

□

Theorem 3.4. *Let* $u_\Diamond \in \Sigma_\Diamond$*, then*

$$G_\Diamond(u_\Diamond^n) = G_\Diamond(u_\Diamond)\left(\frac{1 - \frac{1}{2^{n|u_\Diamond|}}}{1 - \frac{1}{2^{|u_\Diamond|}}}\right).$$

Proof. Let the e-GPV of $u_\Diamond$ be

$$G_\Diamond(u_\Diamond) \quad = \quad (g_{p_1}, g_{p_2}, \cdots, g_{p_k})$$

where $u_\Diamond$ be the partial word over $\Sigma_\Diamond$. To prove the theorem, we utilize the method of induction on the length of $u_\Diamond$. Clearly the partial word of length $n = 1$ satisfies the theorem and thus the base step holds. Now consider the induction step, let

$$\begin{aligned} G_\Diamond(u_\Diamond^2) &= G_\Diamond(u_\Diamond)\left(1 + \frac{1}{2^{|u_\Diamond|}}\right) \\ G_\Diamond(u_\Diamond^3) &= G_\Diamond(u_\Diamond)\left(1 + \left(\frac{1}{2^{|u_\Diamond|}}\right) + \left(\frac{1}{2^{|u_\Diamond|}}\right)^2\right) \\ &\ \vdots \\ G_\Diamond(u_\Diamond^n) &= G_\Diamond(u_\Diamond)\Bigg(1 + \left(\frac{1}{2^{|u_\Diamond|}}\right) + \left(\frac{1}{2^{|u_\Diamond|}}\right)^2 + \cdots \\ &\qquad + \left(\frac{1}{2^{|u_\Diamond|}}\right)^{n-1}\Bigg). \end{aligned}$$

Since by the geometric series

$$1 + x + x^2 + \cdots + x^n = \frac{1 - x^{n+1}}{1 - x}$$

and by the induction hypothesis, we have

$$G_\Diamond(u_\Diamond^n) \quad = \quad G_\Diamond(u_\Diamond)\left(\frac{1 - \frac{1}{2^{n|u_\Diamond|}}}{1 - \frac{1}{2^{|u_\Diamond|}}}\right).$$

□

Theorem 3.5. *The e-GPV of partial word is injective.*

Proof. Consider two partial words $u_\Diamond$ and $v_\Diamond$ defined over the alphabet $\Sigma_\Diamond$ such that

$$G_\Diamond(u_\Diamond) = G_\Diamond(v_\Diamond) = (g_{p_1}, g_{p_2}, \cdots, g_{p_k}).$$

Then

$$g_{p_r} = \sum_{i \in g_r} \frac{1}{2^i} + \sum_{j \in g_\Diamond} \frac{1}{2^j},\ for\ 1 \le r \le k$$

where the set g_r consists of all positions of $a_r \in \Sigma$ occurs in $u_\Diamond$ and the set $g_\Diamond$ consists of all positions of $\Diamond$ occurs in $u_\Diamond$. Thus $(g_{p_1}, g_{p_2}, \cdots, g_{p_k})$ uniquely fixes the positions of the symbols in $u_\Diamond$ and $v_\Diamond$. Therefore $u_\Diamond$ must be equal to $v_\Diamond$. $\square$

We present an algorithm for computing the e-generalized Parikh vector of partial word over $\Sigma_\Diamond$.

Algorithm 1: Algorithm for Computing $e-$generalized Parikh Vector of Partial Word

---

Data: Partial word $u_\Diamond$ over $\Sigma_\Diamond = \{a_1, a_2, \cdots, a_k\} \cup \{\Diamond\}$
Result: e-generalized Parikh vector $G_\Diamond(u_\Diamond)$

Initialize an empty set g_{pos} to keep track of positions of occurrences of symbols;
for *each symbol* a_r *in* Σ **do**
  Initialize an empty set g_r and set $g_{\text{pos}}[a_r] = g_r$;
Initialize an empty set $g_\Diamond$ and set $g_{\text{pos}}[\Diamond] = g_\Diamond$;
for *each position* i *in* $u_\Diamond$ **do**
  if $u_i = a_r$ **then**
    Add i to $g_{\text{pos}}[a_r]$;
  else if $u_i = \Diamond$ **then**
    Add i to $g_{\text{pos}}[\Diamond]$;
Compute $G_\Diamond(u_\Diamond)$ using the following condition:

$$G_\Diamond(u_\Diamond) = \left\{(g_{p_1}, g_{p_2}, \cdots, g_{p_k}) \in [0,1]^2\right\}$$

where $g_{p_r} = \sum_{i \in g_r} \frac{1}{2^i} + \sum_{j \in g_\Diamond} \frac{1}{2^j},\ 1 \le r \le k$ The resulting $G_\Diamond(u_\Diamond)$ is the e-generalized Parikh vector of partial word $u_\Diamond$.

---

4 e-generalized Parikh Matrices for Partial Words

This section introduces the e-generalized Parikh matrix (or e-GPM) of partial words and includes a flowchart to demonstrate the steps involved in calculating the e-generalized Parikh matrix for a given partial word.

Definition 4.1. *Consider $u_\Diamond$ be the partial word with length n over $\Sigma_\Diamond$ where such that $u_\Diamond = u_1 u_2 \cdots u_t \cdots u_n$ where $u_t \in \Sigma_\Diamond$ for $1 \le t \le n$. The e-GPM mapping denoted as $\psi_\Diamond$ is the morphism $\psi_\Diamond : \Sigma_\Diamond^* \to \mathcal{M}_{k+1}$ such that $\psi_\Diamond(u_t)$ is a square matrices of order $k+1$ then $\psi_\Diamond(u_\Diamond) = \psi_\Diamond(u_1)\psi_\Diamond(u_2)\cdots\psi_\Diamond(u_n)$ defined by the condition $\psi_\Diamond(u_t) = (m_{p,q})_{1 \le p,q \le k+1}$; $1 \le t \le n$ where either if $u_t = a_l$ then $m_{p,p} = 1$ for $1 \le p \le k+1$, $m_{l,(l+1)} = \frac{1}{2^t}$ for $1 \le l \le k$ and remaining entries are zero or if $u_t = \Diamond$ then $m_{p,p} = 1$ for $1 \le p \le k+1$, $m_{p,p+1} = \frac{1}{2^t}$ for $1 \le p \le k$ and remaining entries are zero.*

Example 4.2. *Let $u_\Diamond = a\Diamond bab\Diamond$ be the binary partial word. Then the e-GPM of binary partial word $u_\Diamond$ is*

$$\begin{aligned}
\psi_\Diamond(u_\Diamond) &= \psi_\Diamond(a)\psi_\Diamond(\Diamond)\psi_\Diamond(b)\psi_\Diamond(a)\psi_\Diamond(b)\psi_\Diamond(\Diamond) \\
&= \begin{bmatrix} 1 & \frac{1}{2} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & \frac{1}{2^2} & 0 \\ 0 & 1 & \frac{1}{2^2} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & \frac{1}{2^3} \\ 0 & 0 & 1 \end{bmatrix} \\
&\quad \begin{bmatrix} 1 & \frac{1}{2^4} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & \frac{1}{2^5} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & \frac{1}{2^6} & 0 \\ 0 & 1 & \frac{1}{2^6} \\ 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & \frac{53}{2^6} & \frac{263}{2^{10}} \\ 0 & 1 & \frac{27}{2^6} \\ 0 & 0 & 1 \end{bmatrix}.
\end{aligned}$$

4.1 Characterization of the Entries on the e-generalized Parikh Matrices of Partial Words

The following theorem provides a characterization of e-GPM.

Theorem 4.3. *Let $u_\Diamond$ be a partial word of length n over $\Sigma_\Diamond$ where such that $u_\Diamond = u_1 u_2 \cdots u_t \cdots u_n$ where $u_t \in \Sigma_\Diamond$ for $1 \le t \le n$. The e-GPM mapping*

$$\psi_\Diamond(u_\Diamond) = (m_{p,q})_{1 \le p,q \le k+1}$$

has the following properties:
(i) $m_{p,q} = 0$ for all $1 \le q < p \le k+1$
(ii) $m_{p,p} = 1$ for all $1 \le p \le k+1$
(iii) $m_{p,q+1} =$ e-GPV of scattered subword $a_{i,j}$ for all $1 \le i \le j \le k$.

Proof. Clearly, properties (i) and (ii) are true. Our aim is to prove the property (iii). Clearly the partial word of length $n = 1$ satisfies the theorem and thus the base step holds. We argue by induction on length of $u_\Diamond$ for proving the theorem. Let the property (iii) true for all partial words of length at most n. Assume that the partial word $u_\Diamond$ be of length $n+1$. Therefore $u_\Diamond = u'_\Diamond a_t$ where $u'_\Diamond = n$ and either $a_t \in \Sigma_k$ or $a_t = \Diamond$.
Case 1 : Suppose if $a_t \in \Sigma_k$ with $1 \le t \le k$ then

$$\psi_\Diamond(u_\Diamond) = \psi_\Diamond(u'_\Diamond a_t) = \psi_\Diamond(u'_\Diamond)\psi_\Diamond(a_t).$$

Consider that

$$\psi_\Diamond(u'_\Diamond) = \begin{bmatrix} 1 & m'_{1,2} & m'_{1,3} & m'_{1,4} & \cdots & \cdots & m'_{1,k+1} \\ 0 & 1 & m'_{2,3} & \cdots & \cdots & \cdots & m'_{2,k+1} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & \frac{1}{2^t} & \cdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & m'_{k,k+1} \\ 0 & 0 & 0 & 0 & \cdots & \cdots & 1 \end{bmatrix}$$

Assuming the induction hypothesis holds for $\psi_\Diamond(u'_\Diamond)$, it satisfies property (iii). Also from Definition 4.1, we obtain that

$$\psi_\Diamond(a_t) = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & 0 & 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & \frac{1}{2^t} & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \cdots & 1 \end{bmatrix}.$$

The matrix $\psi_\Diamond(a_t)$ contains only zeroes, except for the entries on the main diagonal, which are set to one. Additionally, the element located at position $(t, t+1)$ has a value of $\frac{1}{2^t}$. Therefore $\psi_\Diamond(u'_\Diamond) = R$.

Case 2 : Suppose if $a_t = \Diamond$ then

$$\psi_\Diamond(u_\Diamond) = \psi_\Diamond(u'_\Diamond \Diamond) = \psi_\Diamond(u'_\Diamond)\psi_\Diamond(\Diamond).$$

From Definition 4.1, we obtain that

$$\psi_\Diamond(\Diamond) = \begin{bmatrix} 1 & \frac{1}{2^t} & 0 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \frac{1}{2^t} & 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & \frac{1}{2^t} & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \cdots & 1 \end{bmatrix}.$$

The matrix $\psi_\Diamond(\Diamond)$ contains only zeroes, except for the entries on the main diagonal, which are set to one. Additionally, the elements on the super diagonal are $\frac{1}{2^t}$. Therefore

$$\psi_\Diamond(u_\Diamond) = \begin{bmatrix} 1 & \frac{1}{2^t} & 0 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \frac{1}{2^t} & 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & \frac{1}{2^t} & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \cdots & 1 \end{bmatrix}.$$

Hence $\psi_\Diamond(u'_\Diamond) = R$. Therefore the final matrix $R = (m_{r,s})_{1 \le r,s \le k+1}$ has the property that $m_{q,p+1} = m'_{q,p} + m'_{q,p+1}$ for all q, $1 \le q \le p$ and for all other entries $m_{p,q} = m'_{p,q}$. Since e-GPV of $a_{j,i} = a_j \cdots a_i$ as a scattered subword in $u_\Diamond$ equals the e-GPV of $a_{j,i}$ in $u'_\Diamond$ and e-GPV of $a_{j,i-1}$ in $u'_\Diamond$, the inductive step is complete. Thus the property (iii) is true. Hence the Theorem follows. □

We present an algorithm for computing the e-generalized Parikh matrix of a binary partial word.

Algorithm 2: Algorithm to find $\psi_\Diamond(u_\Diamond)$

Data: Partial word $u_\Diamond$ of length n over $\Sigma_\Diamond$,
Alphabet $\Sigma_\Diamond = \{a_1, a_2, \ldots, a_k\} \cup \{\Diamond\}$,
Parameter k
Result: Matrix $\psi_\Diamond(u_\Diamond)$ of order $k+1$

Initialize matrix M of order $k+1$ with all entries set to 0;
for *each symbol* u_t *in* $u_\Diamond$ **do**
  if u_t *is an element of* $\Sigma_\Diamond$ **then**
    Set $m_{p,p} = 1$ for $1 \le p \le k+1$;
    if $u_t = a_l$ **then**
      Set $m_{l,(l+1)} = \frac{1}{2^t}$ for $1 \le l \le k$;
  else if $u_t = \Diamond$ **then**
    Set $m_{p,p} = 1$ for $1 \le p \le k+1$;
    Set $m_{p,p+1} = \frac{1}{2^t}$ for $1 \le p \le k$;
The resulting matrix M is $\psi_\Diamond(u_\Diamond)$

Note 2. *We note that,*

1. *Coordinates of the e-GPV of a partial word are the elements of the super diagonal of the e-GPM.*
2. *In the binary partial word* $u_\Diamond$, *the entries* $m_{1,2}$ *and* $m_{2,2}$ *in the e-GPM represent the first and second coordinates of the e-GPV and the entry* $m_{1,3}$ *represents the e-GPV of the scattered subword ab.*

Theorem 4.4. *The e-GPM of binary partial word is injective.*

Proof. Consider two partial words $u_\Diamond$ and $v_\Diamond$ defined over the alphabet $\Sigma_\Diamond = \{a, b\} \cup \{\Diamond\}$ such that

$$\begin{aligned} \psi_\Diamond(u_\Diamond) &= \psi_\Diamond(v_\Diamond) \\ &= \begin{bmatrix} 1 & g_{p_1} & g_{p_{ab}} \\ 0 & 1 & g_{p_2} \\ 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

where g_{p_1} and g_{p_2} be the components of e-GPV and $g_{p_{ab}}$ be the e-GPV of scattered subword ab of $u_\Diamond$ and $v_\Diamond$. Then by Theorem 3.5 we have (g_{p_1}, g_{p_2}) uniquely fixes the positions of the symbols of the alphabet in the partial words $u_\Diamond$ and $v_\Diamond$. Therefore e-GPV of scattered subword ab also fixes the positions in the partial words $u_\Diamond$ and $v_\Diamond$. Hence $u_\Diamond$ must be equal to $v_\Diamond$. □

Corollary 4.5. *The e-GPM of $u_\Diamond$ over $\Sigma_\Diamond$ is injective.*

Theorem 4.6. *Let the e-GPM be* $\begin{bmatrix} 1 & g_{p_1} & g_{p_{ab}} \\ 0 & 1 & g_{p_2} \\ 0 & 0 & 1 \end{bmatrix}$ *where g_{p_1} and g_{p_2} be the components of e-GPV and $g_{p_{ab}}$ be the e-GPV of scattered subword ab such that g_{p_1}, g_{p_2} and $g_{p_{ab}}$ are in the form of $\frac{g}{2^l}$ where g, l be a positive integer with respect to a partial word $u_\Diamond$ of length n over $\Sigma_\Diamond = \{a, b\} \cup \{\Diamond\}$. Then the e-GPM satisfies the following condition:*

(i) if g_{p_1} and g_{p_2} are the components of the e-GPV then

$$g_{p_1} + g_{p_2} = \sum_{i=1}^{n} \frac{1}{2^i} + \sum_{j \in g_\Diamond}^{n} \frac{1}{2^j}$$

(ii) if $g_{p_1} = \frac{x}{2^{l_1}}$, $g_{p_2} = \frac{y}{2^{l_2}}$ and $g_{p_{ab}} = \frac{z}{2^{l_3}}$ then

$$l_3 \leq l_1 + l_2 \text{ and } z = 0 \text{ or } z < xy.$$

Proof. (i) Let $u_\Diamond$ be the partial word over $\Sigma_\Diamond = \{a, b\} \cup \{\Diamond\}$ such that the e-GPM of $u_\Diamond$ is an upper triangular square matrix whose supper diagonal elements are the components of the e-GPV. Therefore if g_1, g_2 *and* $g_\Diamond$ denote the positions of a, b *and* $\Diamond$ of the partial word of length n such that

$$\begin{aligned} g_{p_1} &= \sum_{i \in g_1} \frac{1}{2^i} + \sum_{j \in g_\Diamond} \frac{1}{2^j} \\ g_{p_2} &= \sum_{i \in g_2} \frac{1}{2^i} + \sum_{j \in g_\Diamond} \frac{1}{2^j} \text{then} \\ g_{p_1} + g_{p_2} &= \sum_{i=1}^{n} \frac{1}{2^i} + \sum_{j \in g_\Diamond} \frac{1}{2^j}. \end{aligned}$$

(ii) If g_{p_1} and g_{p_2} be the components of the e-GPV of the partial word $u_\Diamond$ such that $g_{p_1} = \frac{x}{2^{l_1}}$ and $g_{p_2} = \frac{y}{2^{l_2}}$ then the e-GPV of the scattered subword ab is

$$g_{p_{ab}} = \sum \left(\sum_{i \in g_1} \frac{1}{2^i} + \sum_{j \in g_\Diamond} \frac{1}{2^j} \right) \left(\sum_{i' \in g_2} \frac{1}{2^{i'}} + \sum_{j \in g_\Diamond} \frac{1}{2^j} \right)$$

for all $i, j < i'$ where the occurrences of a contributes $\sum_{i \in g_1} \frac{1}{2^i} + \sum_{j \in g_\Diamond} \frac{1}{2^j}$ and the occurrences of b contributes $\sum_{i' \in g_2} \frac{1}{2^{i'}} + \sum_{j \in g_\Diamond} \frac{1}{2^j}$. Therefore $l_3 \leq l_1 + l_2$. Let $g_{p_{ab}}$ be the e-GPV of the scattered subword ab of the partial word $u_\Diamond$ such that $g_{p_{ab}} = \frac{z}{2^{l_3}}$ then we have to prove that either $z = 0$ *or* $z < xy$. Now we consider two cases.

Case 1 : If the partial word $u_\Diamond$ begins with b or a sequence of b followed by $\Diamond$ or a sequence of $\Diamond$ and then followed by a or a sequence of a with no b after the occurrence of a in the partial word $u_\Diamond$ then there is no scattered subword ab in the partial word $u_\Diamond$. Therefore $z = 0$.

Case 2 : Other then *Case* 1 we have $z < xy$. □

It can be challenging to perform manual matrix multiplication when dealing with partial words of larger length. To address this issue, Figure 2 is provided to outline the process for calculating the e-GPM of a binary partial word.

4.2 e-Partial Line Languages

The concept of partial line languages in relation to e-GPVs is extended in the form of e-partial line languages with respect to e-GPVs.

Definition 4.7. *A partial language L_P is said to be an e-partial line language if there exist a language line l_{ep} in R^2 such that e-GPVs of L_P lie on l_{ep}. The line l_{ep} is called a e-partial language line. A partial language is said to be a finite e-partial line language if it contains only finite partial words.*

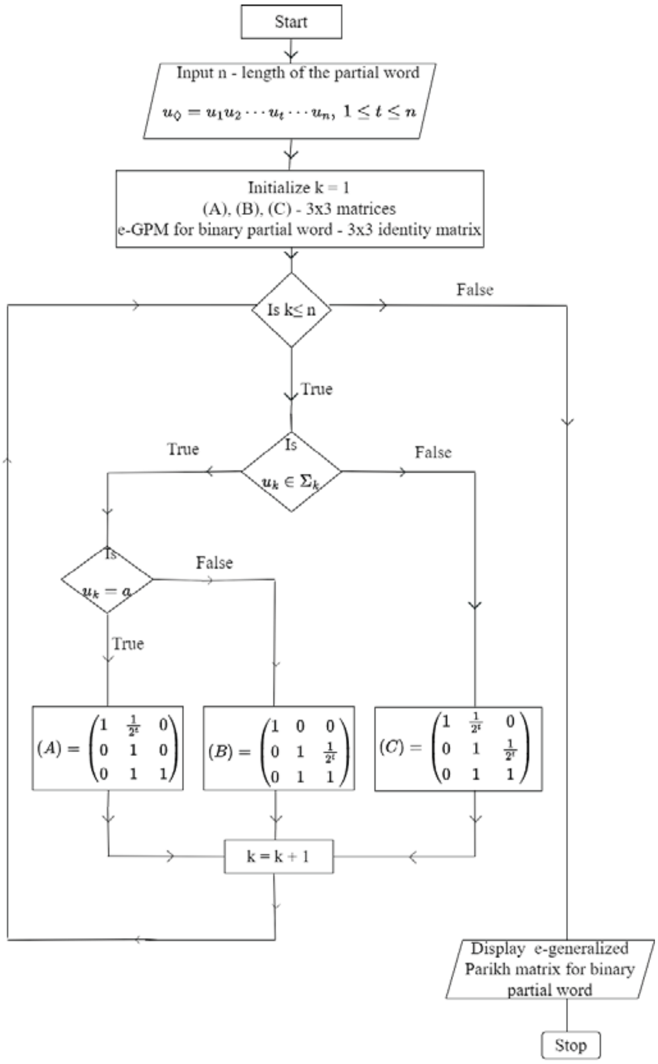

Figure 2: Flow Chart to Find e-generalized Parikh Matrix of a Binary Partial Word

Remark 4.8. *We note that,*

1. *The e-GPV of all partial words lies in the region above the lines $x = 0$, $y = 0$ and $x + y = 1$. Figure 3 illustrates this region, which is called the e-Partial Line region or $EPL-$region.*

2. *The $EPL-$region is unbounded.*

3. *There are no partial words whose e-GPVs are bounded by the lines $x = 0$, $y = 0$, and $x + y = 1$.*

4. *$L = (u)^+\Diamond$ be the partial language where $u \in \{a, b\}$ lie on the line $x + y = 1$.*

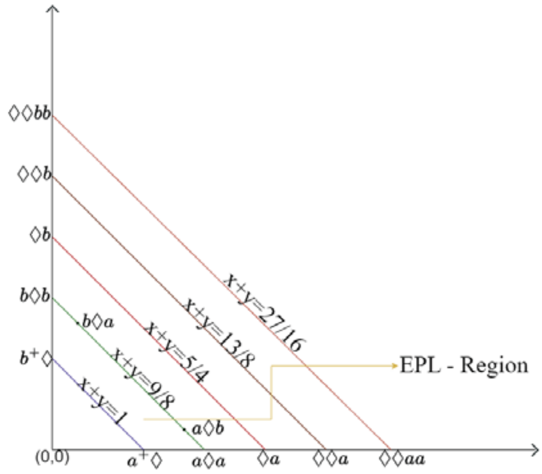

Figure 3: e-generalized Parikh Vector of Binary Partial Words

References

[1] A. Atanasiu, C. Martin-Vide and A. Mateescu, On the injectivity of the Parikh matrix mapping, *Fundamenta Informaticae*, 49 (2002), 289–299.

[2] A. Atanasiu, Binary amiable words, *International Journal of Foundations of Computer Science*, 18 (2007), 387–400.

[3] A. Atanasiu, R. Atanasiu and I. Petre, Parikh matrices and amiable words, *Theoretical Computer Science*, 390 (2008), 102–109.

[4] A. Atanasiu, G. Poovanandran, and W. C. Teh, Parikh matrices for powers of words, *Acta Informatica*, 56 (2019), 521–535.

[5] J. Berstel and L. Boasson, Partial words and a theorem of Fine and Wilf, *Theoretical Computer Science* 218(1) (1999), 135–141.

[6] F. Blanchet-Sadri, Algorithmic combinatorics on partial words, *Discrete Mathematics and its Applications, CRC Press* (2007).

[7] K. Janaki, R. Arulprakasam and V.R. Dare, 2D words and generalized Parikh matrices, *Journal of Physics: Conference Series* 1850(1) (2021).

[8] K. Janaki, and R. Arulprakasam, Tiling systems and domino systems for partial array languages, *IAENG International Journal of Applied Mathematics*, 52(2) (2022), 418-425.

[9] K. Janaki, R. Arulprakasam, P. Meenakshi and V.R. Dare, Algebraic properties of Parikh $q-$matrices on two-dimensional words, *Lecture Notes in Computer Science*, 13348 (2022).

[10] K. Janaki, R. Krishna Kumari and R. Arulprakasam, Algebraic aspects of extending parikh matrices of picture arrays, *Mathematics in Engineering, Science and Aerospace*, 14(4) (2023), 1139–1159.

[11] R. Krishna Kumari, R. Arulprakasam and V.R. Dare, Language of Lyndon partial words, *International Journal of Mathematics and Computer Science* 15(4) (2020), 1173–1177.

[12] R. Krishna Kumari, R. Arulprakasam and V.R. Dare, Combinatorial properties of Fibonacci partial words and arrays, *Journal of Discrete Mathematical Sciences and Cryptography* 24(4) (2021), 1007–1020.

[13] R. Krishna Kumari and R. Arulprakasam, Factors and subwords of rich partial words, *IAENG International Journal of Applied Mathematics* 53(2) (2023), 534–539.

[14] R. Krishna Kumari, L. Jeyanthi, K. Janaki, R. Arulprakasam, and P. Madhusoodhanan, Properties of cover and seed of partial words, *IAENG International Journal of Applied Mathematics* 53(3) (2023), 1132-1137.

[15] R. Krishna Kumari, K. Janaki and R. Arulprakasam, Block inversion on $\wp$words with zero palindromic defect, *International Journal of Applied Mathematics* 36(5) (2023), 639–649.

[16] A. Mateescu, A. Salomaa, K. Salomaa and S. Yu, A sharpening of the Parikh mapping, *Theoretical Informatics and Applications*, 35 (2001), 551–564.

[17] A. Mateescu, Algebraic aspects of Parikh matrices, *Theory is Forever*, 10 (2004), 170–180.

[18] R.J. Parikh, On context-free languages, *Journal of the ACM*, 13 (1966), 570–581.

[19] G. Poovanandran and W.C. Teh, On M-equivalence and strong M-equivalence for Parikh matrices, *International Journal of Foundations of Computer Science*, 29 (2018), 123–137.

[20] G. Poovanandran and W.C. Teh, Strong $(2, t)$ and strong $(3, t)$ transformations for strong M-equivalence, *International Journal of Foundations of Computer Science*, 30 (2019), 719–733.

[21] A. Salomaa, On the injectivity of Parikh matrix mappings, *Fundamenta Informaticae*, 64 (2005), 391-404.

[22] A. Salomaa, Parikh matrices: Subword indicators and degrees of ambiguity, *Lecture Notes in Computer Science*, 11011 (2018), 100–112.

[23] K. Sasikala, T. Kalyani, V.R. Dare and P.J. Abisha, Line languages, *Electronic Notes in Discrete Mathematics* 12 (2003), 178–185.

[24] K. Sasikala, V.R. Dare and D.G. Thomas, Learning of partial languages, *Engineering Letters* 14(2) (2007), 72–80.

[25] T.F. Serbanuta, Extending Parikh matrices, *Theoretical Computer Science*, 310 (2004), 233–246.

[26] R. Siromoney and V.R. Dare, A generalization of the Parikh vector for finite and infinite words, *International Conference on Foundations of Software Technology and Theoretical Computer Science* (1985), 290–302.

[27] G. Siromoney, R. Siromoney, K.G. Subramanian, V.R. Dare and P.J. Abisha, Generalized Parikh vector and public key cryptosystems, *World Scientific Series in Computer Science*, 16 (1989), 301–323.

[28] K.G. Subramanian, A.M. Huey and A.K. Nagar, On Parikh matrices, *International Journal of Foundations of Computer Science*, 20 (2009), 211–219.

[29] W.C. Teh and K.H. Kwa, Core words and Parikh matrices, *Theoretical Computer Science*, 582 (2015), 60–69.

[30] W.C. Teh, On core words and the Parikh matrix mapping, *International Journal of Foundations of Computer Science*, 26 (2015), 123–142.

[31] W.C. Teh, Parikh matrices and strong M-equivalence, *International Journal of Foundations of Computer Science*, 27 (2016), 545–556.

[32] W.C. Teh, Parikh-friendly permutations and uniformly Parikh-friendly words, *Australasian Journal of Combinatorics*, 76 (2020), 208–219.