International Journal of Foundations of Computer Science

Properties of Parikh Matrices of Binary Words Obtained by an Extension of a Restricted Shuffle Operator

Somnath Bera
Department of Mathematics, Indian Institute of Technology Madras, Chennai 600036 India

Kalpana Mahalingam
Department of Mathematics, Indian Institute of Technology Madras, Chennai 600036 India

K. G. Subramanian
Department of Mathematics, Madras Christian College, Tambaram, Chennai 600059 India

Received (Day Month Year)
Revised (Day Month Year)
Accepted (Day Month Year)
Communicated by (xxxxxxxxxx)

We introduce an extension of the restricted shuffle operator on binary words considered by Atanasiu and Teh (2016). We then derive properties on Parikh matrix equivalence of words over a binary alphabet based on this extended shuffle operator and a weak-ratio property of words. We also examine the recently introduced concept of core Parikh matrix equivalence of binary words in the context of the restricted shuffle operator.

Keywords: SShuffle operator; Parikh matrix; Ambiguity; Combinatorics on words

1. Introduction

The concept of Parikh matrix mapping, initiated by Mateescu et al. [8], opened up a series of investigations on different problems on words and Parikh matrices (See, for example, [1, 2, 4, 5, 12–15]). The well-known notion of Parikh vector [9] of a word which gives the number of each of the different symbols of the alphabet that define the word, has been of great significance, especially in formal language theory [11]. The Parikh matrix of a word over an ordered alphabet is a special type of a matrix whose entries are counts of certain subwords of the given word, with the subwords depending on the ordered alphabet and includes the Parikh vector in the second diagonal above the main diagonal. Parikh matrix equivalence, referred to as $M-$equivalence in the sense that several words may have the same Parikh matrix, is a problem of significant interest and has been investigated extensively (see, for example, [1, 2, 5, 12]). In the case of binary alphabet, very elegant characterizations for $M-$equivalence of words are known [5]. Very recently, Atanasiu and Teh [3] derived a number of properties, especially on $M-$equivalence of words, by developing an interesting connection between the Parikh matrix and a restricted shuffle operator on words, named $SShuffle$ in [3], initially considered in [10].

Here we introduce a special type of shuffle operator on binary words which we call as $S_{m,n}$ and which is an extension of $SShuffle$ in [3]. The shuffle operator $S_{m,n}$, $m, n \geq 1$, in operating on a pair of words (u, v) forms a word $S_{m,n}(u, v)$ by concatenating alternately consecutive factors of u and v, with the factors having length m for the former and length n for the latter. The $SShuffle$ operator is thus a special case of $S_{m,n}$ with $m = n = 1$. We derive certain properties on the $M-$equivalence of words in the binary case in the context of the extended shuffle operator $S_{m,n}$ and the weak-ratio property of words introduced in [13]. We also consider the concept of core $M-$equivalence of words introduced and investigated in [14, 15] and obtain properties again in the context of binary words formed by the restricted shuffle operator.

2. Preliminaries

We recall certain notions needed in the sequel, especially on subwords and Parikh matrices [8]. For other notions related to words and languages the reader is referred to [6, 11].

Let Σ denote a finite alphabet: a set of symbols, called letters of the alphabet. A finite word or simply a word is a finite sequence of letters from Σ. The empty word is denoted by λ. We denote by Σ^* (Σ^+) the set of all words (non empty words respectively) over Σ. For a word $w \in \Sigma^*$, the length of w, denoted by $|w|$, is the number of letters in w.

A word $u \in \Sigma^+$ is called a scattered subword or simply, a subword of w if there exist words $x_1, x_2, \cdots, x_n$ and $y_0, y_1, y_2, \cdots, y_n$ (some of them possibly empty) over Σ such that $u = x_1 x_2 \cdots x_n$ and $w = y_0 x_1 y_1 x_2 y_2 \cdots x_n y_n$. The number of occurrences of the word u as a subword of the word w is denoted by $|w|_u$.

An ordered alphabet Σ is an alphabet Σ with a linear order " $<$ " defined on it. For example, $\Sigma = \{a_1, a_2, ..., a_k\}$ with the order relation $a_1 < a_2 < \cdots < a_k$, is an ordered alphabet, which is denoted as $\Sigma_k = \{a_1 < a_2 < \cdots < a_k\}$. We denote the subword $a_i a_{i+1} ... a_j$ for $1 \leq i \leq j \leq k$ by the notation $a_{i,j}$. Through out the rest of the paper we always consider an ordered alphabet Σ_k with $k \geq 2$.

The Parikh vector [9] of a word $w \in \Sigma_k^*$, denoted by $\psi(w)$, is defined by $\psi(w) = (|w|_{a_1}, |w|_{a_2}, \cdots, |w|_{a_k})$ which counts the number of occurrences of the letters of the alphabet present in the word w.

Let $\mathbb{M}_k$ denote the set of all $k \times k$ upper triangular matrices of dimension k with entries from $\mathbb{N}$ (the set of all non-negative integers) and unit diagonal. The set $\mathbb{M}_k$ is a monoid with respect to the usual multiplication of matrices and has a unit I_k, the identity matrix of dimension k. Now we recall the notion of Parikh matrix mapping [8].

Definition 1. *[8] Let Σ_k be an ordered alphabet. The Parikh matrix mapping, denoted by ψ_{M_k}, is the morphism:*

$$\psi_{M_k} : \Sigma_k^* \longrightarrow \mathbb{M}_{k+1}$$

defined as follows: For an arbitrary $a_l \in \Sigma_k$,

$$\psi_{M_k}(\lambda) = I_{k+1}$$
$$\psi_{M_k}(a_l) = (m_{ij})_{1 \leqslant i,j \leqslant k+1},$$

where

(i) $m_{ii} = 1$ *for* $1 \leq i \leq k+1$
(ii) $m_{l(l+1)} = 1$,

and all other entries are zero.

Note that for every word $w = w_1 w_2 \cdots w_n$ with $w_i \in \Sigma_k$,
$$\psi_{M_k}(w) = \psi_{M_k}(w_1)\psi_{M_k}(w_2)\cdots\psi_{M_k}(w_n).$$
The Parikh matrix mapping is not injective. For example, consider the words $w_1 = ababa$ and $w_2 = baaab$ over the ordered binary alphabet $\Sigma_2 = \{a < b\}$. It can be verified that

$$\psi_{M_2}(ababa) = \begin{pmatrix} 1 & 3 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} = \psi_{M_2}(baaab).$$

Two words $u, v \in \Sigma_k^*$ are $M-$equivalent, if $\psi_{M_k}(u) = \psi_{M_k}(v)$ and we denote it by $u \equiv_M v$.

3. Extended shuffle operator on binary words and Parikh Matrices

In [3], a restricted s-shuffle operator on two binary words, denoted as $SShuf$, which is initially due to Păun [10], is utilized for deriving certain properties of Parikh matrices and in particular, the $M-$equivalence of words over a binary alphabet. Here we introduce an extension of the $s-$shuffle operator as follows:

Definition 2. *Let* $\Sigma_2 = \{a < b\}$ *and* u, v *be two words over* Σ_2 *such that* $u = x_1 x_2 ... x_l$ *and* $v = y_1 y_2 ... y_l$, *where* $x_i, y_i \in \Sigma_2^*$, $|x_i| = m$ *and* $|y_i| = n$, *for* $1 \leq i \leq l$ *and for two positive integers* m, n. *The extended shuffle operator, denoted as* $S_{m,n}$ *is defined on the pair* (u, v) *as follows:*

$$S_{m,n}(u, v) = x_1 y_1 x_2 y_2 ... x_l y_l.$$

In particular when $m = n = 1$, *we have* $S_{1,1}(u, v) = SShuf(u, v)$, *the restricted shuffle of binary words considered in [3].*

Remark 3. *We note that for any* $u, v \in \Sigma_2^+$, *there is at least one extended shuffle operator* $S_{m,n}$ *defining the word* $S_{m,n}(u, v)$. *In fact if* d *is a common divisor of* $|u|$ *and* $|v|$ *then for* $m = \frac{|u|}{d}, n = \frac{|v|}{d}$, $S_{m,n}(u, v)$ *can be defined.*

The problem of $M-$equivalence of binary words resulting from the extended shuffle operator $S_{m,n}$ is examined in this section. We recall first a characterization [5] for

$M-$equivalent words over a binary alphabet based on the sum of positions of letters in the words.

For a nonempty alphabet Σ_2 and $w \in \Sigma_2^*$, the sum of positions of a letter $x \in \Sigma_2$ in $w = w_1 w_2 \cdots w_n$, $w_i \in \Sigma$ is defined as follows:

$$S_x(w) = \sum_{\substack{1\le i\le n \\ w_i = x}} i = \sum_{1\le i\le n} i|w_i|_x.$$

For example, consider the ordered alphabet $\Sigma_2 = \{a < b\}$ and let $w = ab^2ab \in \Sigma_2^*$. Then $S_a(ab^2ab) = 1+4 = 5$, $S_b(ab^2ab) = 2+3+5 = 10$.

Theorem 4. *[5] Let $\Sigma_2 = \{a < b\}$ and u, v be two words over Σ_2. Then $u \equiv_M v$ if and only if $\psi(u) = \psi(v)$ and $S_b(u) = S_b(v)$. Furthermore,*

$$S_b(w) = |w|_{ab} + \frac{1}{2}|w|_b(|w|_b + 1), \text{ for all words } w \in \Sigma_2^*.$$

We now derive a formula for the sum of positions of the letter b in the extended shuffle of two words over a binary alphabet.

Theorem 5. *Let $\Sigma_2 = \{a < b\}$ and u, v be two words over Σ such that $u = x_1x_2...x_l$ and $v = y_1y_2...y_l$, where $x_i, y_i \in \Sigma_2^*$, $|x_i| = m$, $1 \le i \le l$ and $|y_i| = n$, $1 \le i \le l$. Then*

$$S_b(S_{m,n}(u,v)) = S_b(u) + S_b(v) - n|u|_b + \sum_{1\le i\le l} i[n|x_i|_b + m|y_i|_b].$$

Proof.

$$\begin{aligned}
&S_b(S_{m,n}(u,v)) \\
&= \sum_{\substack{1\le i\le m \\ u(i)=b}} i + \sum_{\substack{m+1\le i\le 2m \\ u(i)=b}} (i+n) + \cdots + \sum_{\substack{(l-1)m+1\le i\le lm \\ u(i)=b}} (i+(l-1)n) \\
&\quad + \sum_{\substack{1\le i\le n \\ v(i)=b}} (m+i) + \sum_{\substack{n+1\le i\le 2n \\ v(i)=b}} (2m+i) + \cdots + \sum_{\substack{(l-1)n+1\le i\le ln \\ v(i)=b}} (lm+i) \\
&= S_b(u) + \sum_{1\le i\le l} (i-1)n|x_i|_b + S_b(v) + \sum_{1\le i\le l} im|y_i|_b
\end{aligned}$$

which proves the result as $|u|_b = \sum_{1\le i\le l} |x_i|_b$. □

Remark 6. *When $l = 1$ in Theorem 5, we have $S_b(S_{|u|,|v|}(u,v)) = S_b(uv) = S_b(u) + S_b(v) + |u||v|_b$.*

We now consider a particular case.

Corollary 1. *When $m = 1$, we have $S_b(S_{1,n}(u,v)) = (n+1)S_b(u) - n|u|_b + S_b(v) + \sum_{1\le i\le l} i|y_i|_b$.*

A sufficient condition for the extended shuffle of a pair of words to be M-equivalent is now established, by considering the extended operator $S_{m,n}$, $m, n \geq 1$.

Theorem 7. *Let* $\Sigma_2 = \{a < b\}$ *and* $v_1, v_2, w_1, w_2 \in \Sigma_2^*$*, where* $v_1 = v_1[1]v_1[2] ...v_1[l]$*,* $v_2 = v_2[1]v_2[2]...v_2[l]$*,* $w_1 = w_1[1]w_1[2]...w_1[l]$*,* $w_2 = w_2[1]w_2[2]...w_2[l]$ *such that* $v_1 \equiv_M w_1$*,* $v_2 \equiv_M w_2$ *and* $|v_1[i]| = |w_1[i]| = m$*,* $|v_2[i]| = |w_2[i]| = n$*, for* $1 \leq i \leq l$*. If* $|v_1[i]|_b = |w_1[i]|_b$ *and* $|v_2[i]|_b = |w_2[i]|_b$*, for* $1 \leq i \leq l$*, then* $S_{m,n}(v_1, v_2) \equiv_M S_{m,n}(w_1, w_2)$ *.*

Proof. Since $v_1 \equiv_M w_1$, $v_2 \equiv_M w_2$, we have $|S_{m,n}(v_1, v_2)|_a = |S_{m,n}(w_1, w_2)|_a$ and $|S_{m,n}(v_1, v_2)|_b = |S_{m,n}(w_1, w_2)|_b$. Hence by Theorem 4, $S_{m,n}(v_1, v_2) \equiv_M S_{m,n}(w_1, w_2)$ if and only if $S_b(S_{m,n}(v_1, v_2)) = S_b(S_{m,n}(w_1, w_2))$, *i.e. if and only if*

$$S_b(v_1) + S_b(v_2) - n|v_1|_b + \sum_{1 \leq i \leq l} i[n|v_1[i]|_b + m|v_2[i]|_b] =$$

$$S_b(w_1) + S_b(w_2) - n|w_1|_b + \sum_{1 \leq i \leq l} i[n|w_1[i]|_b + m|w_2[i]|_b]$$

which is true as $v_1 \equiv_M w_1$*,* $v_2 \equiv_M w_2$ *and* $|v_1[i]|_b = |w_1[i]|_b$ *and* $|v_2[i]|_b = |w_2[i]|_b$*, for* $1 \leq i \leq l$*, by hypotheses in the Theorem.* □

We note that Theorem 7 holds even for the alphabet $\{b < a\}$. Also, the condition in Theorem 7 is only a sufficient condition but is not necessary, as seen from the following example.

Example 8. *Let* $\Sigma_2 = \{a < b\}$*. Consider the words* $v_1 = ba^3b^3a^2ba^2$ $w_1 = (ab)^3(baa)^2$*,* $v_2 = (ab)^2(ba^2)^2a^2b^2ab$ *and* $w_2 = ab^2(a^2b)^3(ab)^2$*.* $v_1 \equiv_M w_1$*,* $v_2 \equiv_M w_2$*. Also we can verify that* $S_{3,4}(v_1, v_2) \equiv_M S_{3,4}(w_1, w_2)$ *but the condition in the hypothesis of Theorem 7 is not satisfied, as* $|v_2[2]|_b = |baab|_b = 2$ *while* $|w_2[2]|_b = |abaa|_b = 1$*. (In fact* $v_2[3], v_2[4]$ *also do not have an equal number of* b*'s as in* $w_2[3], w_2[4]$ *respectively.)*

4. More Properties of Parikh Matrix of Extended Shuffle of binary words

Since the product of Parikh matrices of two words need not be commutative, the problem of obtaining conditions for Parikh matrices to commute has been studied [7]. In fact, a weak ratio property is introduced in [13] to investigate the properties of words u and v so that the words uv and vu are M-equivalent. In [7], several sufficient conditions for two words over an ordered alphabet such that their Parikh matrices commute, are provided. We now recall the definition of a weak ratio property of words.

Definition 9. *[13] Two words* $u, v \in \Sigma_k^*$ *are said to satisfy a weak ratio property, denoted by* $u \sim_{wr} v$*, if* $|u|_x = t|v|_x$ *for all* $x \in \Sigma_k$ *and for some non zero rational number* t*.*

It was shown in [13] that for any two binary words $u, v \in \Sigma_2^*$, $u \sim_{wr} v$ if and only if $\psi_M(uv) = \psi_M(vu)$.

It is natural to ask whether the extended shuffle of two words commute with respect to Parikh matrix mapping. In [3], a necessary and sufficient condition is given for the commutativity of Parikh matrices of the restricted shuffle words $SShuf(u, v)$ and $SShuf(v, u)$ of two binary words u, v having the same length. We generalize this result for all binary words such that the length of one word is k times the length of the other.

Theorem 10. *Let $\Sigma_2 = \{a < b\}$ and u, v be two words over Σ_2 such that $u = x_1x_2...x_l$ and $v = y_1y_2...y_l$, where $x_i, y_i \in \Sigma_2^+$, $|x_i| = m$, $1 \leq i \leq l$ and $|y_i| = n$, $1 \leq i \leq l$ and $|v| = k|u|$, for some non zero rational number k. Then $S_{m,n}(u, v) \equiv_M S_{n,m}(v, u)$ if and only if u and v satisfy weak ratio property with ratio constant k.*

Proof. By Theorem 4 and Theorem 5, we have

$$\begin{aligned}
&S_{m,n}(u, v) \equiv_M S_{n,m}(v, u)\\
\Leftrightarrow\ &S_b(S_{m,n}(u, v)) = S_b(S_{n,m}(v, u))\\
\Leftrightarrow\ &S_b(u) + S_b(v) - n|u|_b + \sum_{1\leq i\leq l} i[n|x_i|_b + im|y_i|_b]\\
&\qquad = S_b(v) + S_b(u) - m|v|_b + \sum_{1\leq i\leq l} i[m|y_i|_b + n|x_i|_b]\\
\Leftrightarrow\ &n|u|_b = m|v|_b\\
\Leftrightarrow\ &|v|_b = \frac{n}{m}|u|_b = k|u|_b\\
\Leftrightarrow\ &|v|_a = k|u|_a \text{ , since } |v| = k|u|.
\end{aligned}$$

□

Computation of the Parikh matrix of the restricted shuffle $SShuf(u, v)$ directly from the Parikh matrices of two binary words u and v having the same Parikh vector, was shown in [3].

Theorem 11. *[3] Let $\Sigma_2 = \{a < b\}$ and u, v be two words over Σ_2 such that $\psi(u) = \psi(v)$. Then*

$$\psi_M(SShuf(u, v)) = \begin{pmatrix} 1 & 2|u|_a & 2|u|_{ab} + 2|v|_{ab} \\ 0 & 1 & 2|u|_b \\ 0 & 0 & 1 \end{pmatrix}.$$

In the following Theorem, we generalize this result for the extended shuffle operator.

Theorem 12. *Let $\Sigma_2 = \{a < b\}$ and u, v be two words over Σ_2. Let $u = x_1x_2...x_l$ and $v = y_1y_2...y_l$, $x_i, y_i \in \Sigma_2^*$, $|x_i| = m$, $1 \leq i \leq l$ and $|y_i| = n$, $1 \leq i \leq l$ and $A = |u|_a + |v|_a$, $B = |u|_b + |v|_b$, $C = |u|_{ab} + |v|_{ab} - |u|_b|v|_b - n|u|_b + \sum_{i=1}^{l} i[n|x_i|_b + m|y_i|_b]$. Then $\psi_M(S_{m,n}(u, v)) =$*

$$\begin{pmatrix} 1 & A & C \\ 0 & 1 & B \\ 0 & 0 & 1 \end{pmatrix}.$$

Proof. We note that

$$|S_{m,n}(u,v)|_a = |u|_a + |v|_a$$
$$|S_{m,n}(u,v)|_b = |u|_b + |v|_b.$$

Now using Theorem 4 and Theorem 5, we have

$$\begin{aligned} & |S_{m,n}(u,v)|_{ab} \\ &= S_b(S_{m,n}(u,v)) - \frac{1}{2}|S_{m,n}(u,v)|_b(|S_{m,n}(u,v)|_b + 1) \\ &= S_b(u) + S_b(v) - n|u|_b + \sum_{i=1}^{l} i[n|x_i|_b + im|y_i|_b] - \frac{1}{2}(|u|_b + |v|_b)(|u|_b + |v|_b + 1) \\ &= (|u|_{ab} + \frac{1}{2}|u|_b(|u|_b + 1)) + (|v|_{ab} + \frac{1}{2}|v|_b(|v|_b + 1)) \\ &\qquad -n|u|_b + \sum_{i=1}^{l} i[n|x_i|_b + im|y_i|_b] - \frac{1}{2}(|u|_b(|u|_b + 1) + 2|u|_b|v|_b + |v|_b(|v|_b + 1)) \\ &= |u|_{ab} + |v|_{ab} - |u|_b|v|_b - n|u|_b + \sum_{i=1}^{l} i[n|x_i|_b + m|y_i|_b]. \end{aligned}$$

□

One can easily observe that by substituting $m = n = 1$ and assuming $\psi(u) = \psi(v)$ in Theorem 12, we obtain Theorem 11.

The extended shuffle operator can be considered for $n > 2$ words. We list a few main results similar to those obtained for a pair of words.

Definition 13. *Let* $\Sigma_2 = \{a < b\}$ *and* $n \geq 2$ *be an integer. The operator* $S_{m_1,m_2,\ldots,m_n}$ *is defined by*

$$S_{m_1,m_2,\ldots,m_n}(w_1, w_2, \ldots, w_n) = w_1[1]\ldots w_n[1]w_1[2]\ldots w_n[2]\ldots w_1[l]\ldots w_n[l]$$

where $w_i \in \Sigma_2^+$ *and* $w_i = w_i[1]w_i[2]\ldots w_i[l]$*, for each* $1 \leq i \leq n$ *and* $|w_i[j]| = m_i$*, for* $1 \leq j \leq l$*.*

Theorem 14. *Let* $\Sigma_2 = \{a < b\}$ *and* $n \geq 2$ *be an integer, and* $w_1, w_2, \ldots, w_n \in \Sigma_2^*$ *such that* $w_i = w_i[1]w_i[2]\ldots w_i[l]$*, for each* $1 \leq i \leq n$ *and* $|w_i[j]| = m_i$*, for* $1 \leq j \leq l$*. Then*

$$S_b(S_{m_1,m_2,\ldots,m_n}(w_1, w_2, \ldots, w_n)) = \sum_{1\leq i\leq n} S_b(w_i) + \sum_{2\leq i\leq n} \left(\sum_{1\leq k\leq l} \big(k \sum_{1\leq j\leq i-1} m_j\big)|w_i[k]|_b \right)$$
$$+ \sum_{1\leq i\leq n-1} \left(\big((k-1) \sum_{i+1\leq j\leq l} m_j\big)|w_i[k]|_b \right).$$

Proof.

$$S_{m_1,m_2,...,m_n}(w_1,w_2,...,w_n)) = (w_1[1]w_2[1]...w_i[1]...w_n[1])(w_1[2]w_2[2]...w_i[2]...w_n[2])...$$

$$(w_1[i]w_2[i]...w_i[i]...w_n[i])...(w_1[l]w_2[l]...w_i[l]...w_n[l])$$

$$\begin{aligned} & S_b(S_{m_1,m_2,...,m_n}(w_1,w_2,...,w_n)) \\ &= S_b(w_1) + \sum_{1\le k\le l}(k-1)(\sum_{2\le j\le n} m_j)|w_1[k]|_b \\ &+ S_b(w_2) + \sum_{1\le k\le l}(km_1 + (k-1)(\sum_{2\le j\le n} m_j))|w_2[k]|_b + \cdots \\ &+ S_b(w_i) + \sum_{1\le k\le l}(k\sum_{1\le j\le i-1} m_j + (k-1)(\sum_{i+1\le j\le n} m_j))|w_i[k]|_b + \cdots \\ &+ S_b(w_n) + \sum_{1\le k\le l}(k\sum_{1\le j\le n-1} m_j)|w_n[k]|_b \end{aligned}$$

which yields the result. □

Remark 15. *When* $m_i = 1$, *for* $1 \le i \le n$,

$$S_b(S_{1,1,...,1}(w_1,w_2,...,w_n)) = S_b(SShuf_n(w_1,w_2,...,w_n)).$$

where $SShuf_n(w_1,w_2,...,w_n)$ *is as in Definition* 5.1 *in [3].*

As a consequence of Theorem 14, we obtain the following result.

Theorem 16. *Let* $\Sigma_2 = \{a < b\}$ *and* $n \ge 2$ *is an integer, and* $v_i = v_i[1]v_i[2]...v_i[l]$, $w_i = w_i[1]w_i[2]...w_i[l] \in \Sigma_2^*$, $1 \le i \le n$ *such that* $v_i \equiv_M w_i$, *for* $1 \le i \le n$ *and* $|v_i[j]| = |w_i[j]| = m_i$, *for* $1 \le i \le n$. *Then,*

$$S_{m_1,m_2,...,m_n}(v_1,v_2,...,v_n) \equiv_M S_{m_1,m_2,...,m_n}(w_1,w_2,...,w_n)$$

if $|v_i[j]|_b = |w_i[j]|_b$, *for* $1 \le i \le n$ *and* $1 \le j \le l$.

5. Core M-equivalence of restricted shuffled words

In trying to capture the essential part of a word, especially in the binary case, that contributes to the Parikh matrix of the word, an interesting notion of *core* of a word and the associated concept of $core\,M-equivalence$ are introduced by Teh and Kwa in [15]. We recall here these notions.

Definition 17. *[15] Let* $\Sigma_2 = \{a < b\}$. *The core of a word* $w \in \Sigma_2^*$, *denoted as* $core(w)$, *is the unique subword* w_c *of* w *with the smallest possible length such that* $w = b^m w_c a^n$, *for some* $m, n \ge 0$.

Note that $|w|_{ab} = |core(w)|_{ab}$.

Definition 18. *[15] Let $\Sigma_2 = \{a < b\}$. Two words $u, v \in \Sigma_2^*$ are said to be core $M-$equivalent if and only if the following conditions hold:*

(i) the words u, v have the same Parikh vector and
(ii) $core(u)$ and $core(v)$ are $M-$equivalent.

For example, the words $u = b^4aba^2b^2a^3$ and $v = b^4a^2b^2aba^3$, with the same Parikh vector $(6, 7)$ are *core* $M-$*equivalent* since $core(u) = aba^2b^2$ and $core(v) = a^2b^2ab$ are $M-$equivalent.

We shall show in this paragraph that *core* $M-$*equivalence* is preserved under the restricted shuffle operator $SShuf$ when the given two words satisfy conditions analogous to the conditions on $M-$*equivalence* (Theorem 3.3 in [3]) of words obtained by the restricted shuffle operator.

Theorem 19. *Let v_1 and v_2 be two binary non-empty words over $\{a < b\}$ of equal length, respectively core $M-$equivalent to the words w_1 and w_2 where for $i = 1, 2$, $core(v_i)$ and $core(w_i)$ are nonempty. Then $SShuf(v_1, v_2)$ is core $M-$equivalent to $SShuf(w_1, w_2)$.*

Proof. By the definition of core of a word, we can write the words v_1, w_1, v_2 and w_2 as follows: $v_1 = b^p core(v_1)a^r$, $w_1 = b^p core(w_1)a^r$ and $v_2 = b^q core(v_2)a^s$ and $w_2 = b^q core(w_2)a^s$, for some $p, q, r, s \in \mathbb{N}$. Since v_1 and v_2 are respectively *core* $M-$*equivalent* to w_1 and w_2, we have $core(v_1)$ and $core(v_2)$ are respectively $M-$equivalent to $core(w_1)$ and $core(w_2)$. This implies that $|core(v_1)|_x = |core(w_1)|_x$ and $|core(v_2)|_x = |core(w_2)|_x$ for $x \in \{a, b, ab\}$.

Now there are nine cases to deal with as p can be $>$ or $=$ or $<$ q and likewise r can be $>$ or $=$ or $<$ s.
We prove the result for the case $p \geq q$ and $r \geq s$ noting that $a^0b^0 = \lambda$ and the remaining cases can be dealt with in a similar manner. We have $v_1 = b^q b^{p-q} core(v_1)a^{r-s}a^s$, $v_2 = b^q core(v_2)a^s$ and likewise $w_1 = b^q b^{p-q} core(w_1)a^{r-s}a^s$ and $w_2 = b^q core(w_2)a^s$. Now,

$$SShuf(v_1, v_2) = b^{2q} SShuf(b^{p-q} core(v_1)a^{r-s}, core(v_2))a^{2s},$$
$$SShuf(w_1, w_2) = b^{2q} SShuf(b^{p-q} core(w_1)a^{r-s}, core(w_2))a^{2s}.$$

Clearly, $SShuf(v_1, v_2)$ and $SShuf(w_1, w_2)$ have the same Parikh vector. In order to show that they are *core* $M-$*equivalent*, it is enough to show that $core(SShuf(v_1, v_2))$ and $core(SShuf(w_1, w_2))$ are $M-$equivalent, which will hold if and only if

$$core(SShuf(b^{p-q} core(v_1)a^{r-s}, core(v_2))) \equiv_M$$
$$core(SShuf(b^{p-q} core(w_1)a^{r-s}, core(w_2)))$$

if and only if

$$SShuf(core(v_2)/b, b^{p-q-1}core(v_1)a^{r-s}) \text{ and}$$
$$SShuf(core(w_2)/b, b^{p-q-1}core(w_1)a^{r-s})$$

are $M-$equivalent, which is true, where $core(v_2)/b$ is the word obtained from $core(v_2)$ by deleting the last letter and likewise for $core(w_2)/b$. Note that the $M-$equivalence of $core(v_2)$ and $core(w_2)$ implies the $M-$equivalence of $core(v_2)/b$ and $core(w_2)/b$. □

6. Conclusion

An extension of the restricted shuffle operator considered in [3] is introduced here and $M - equivalence$ and *core* $M-$equivalence of binary words under this operator are investigated. A notion of relativized core is introduced in [14]. It will be of interest to examine this notion in the context of the shuffle operator on words. It is known [3] that binary words u, v over $\{a < b\}$ with the same Parikh vector satisfy the property that $SShuf(u, v)$ and $SShuf(v, u)$ are $M-$equivalent. But words over $\Sigma_3 = \{a < b < c\}$ do not satisfy this property. For example, the words $x = abcbcacab$ and $y = bcacababc$ over Σ_3 have even the same Parikh matrix but $SShuf(x, y)$ and $SShuf(y, x)$ are not $M-$equivalent. Although the problem of the $SShuffle$ operator preserving $M-$equivalence of ternary words remains to be explored, a very simple case when this happens is the following: If u, v are two nonempty words of equal length over $\{a < b < c\}$ such that $u = \alpha_1\beta_1\alpha_2\gamma_1\alpha_3$ and $v = \alpha_1\beta_2\alpha_2\gamma_2\alpha_3$, where $\beta_1, \beta_2, \gamma_1, \gamma_2$ are words over $\{a < b\}$ with β_1, β_2 being $M-$equivalent and likewise γ_1, γ_2 being $M-$equivalent, then $SShuf(u, v)$ and $SShuf(v, u)$ are $M-$equivalent. So it may be of interest to examine $M-$equivalence of the $SShuffle$ or the extended shuffle operator on words over a ternary alphabet in the context of Parikh matrices.

Acknowledgement

The authors are very much grateful to the reviewers for their very meticulous and detailed comments on the paper. The comments have served to be very useful in correcting the errors and for improving the presentation of the paper.
This research is partially supported by the project (MAT/15-16/046/DSTX/KALP) awarded to Dr Kalpana Mahalingam by the Department of Science and Technology, India. The author K.G. Subramanian is grateful to UGC, 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. Atanasiu, Binary amiable words, International Journal of Foundations of Computer Science 18, 2007: 387-400.

[2] A. Atanasiu, C. Martin-Vide, A. Mateescu, On the injectivity of the Parikh matrix mapping, Fundamenta Informatica 46, 2001: 1-11.

[3] A. Atanasiu, W. C. Teh, A new operator over Parikh languages, International Journal of Foundations of Computer Science 27(6), 2016: 757-770.

[4] A. Atanasiu, W.C. Teh, Shuffle Operator and Parikh Languages, Multidisciplinary Creativity, Homage to Gheorghe Paun on His 65th Birthday (M. Gheorghe, I. Petre, M.J.Perez-Jimenez, G. Rozenberg, A.Salomaa Eds.) (2016) 185-191.

[5] S. Fosse, G. Richomme, Some characterizations of Parikh matrix equivalent binary words, Information Processing Letters 92, 2004: 77-82.

[6] M. Lothaire, Combinatorics on Words, Cambridge University Press, 1997.

[7] K. Mahalingam, K. G. Subramanian, Product of Parikh matrices and commutativity, International Journal of Foundations of Computer Science, 23(1), 2012: 207-223.

[8] A. Mateescu, A. Salomaa, K. Salomaa, S. Yu, A Sharpening of the Parikh Mapping, Theoretical Informatics and Applications 35, 2001: 551-564.

[9] R. J. Parikh, On context-free languages, Journal of the Association for Computing Machinery 4, 1966: 570-581.

[10] Gh. Păun, The generative mechanism of economic processes, The Technical Publishing House, Bucharest, 1980.

[11] A. Salomaa, Formal Languages, Academic Press, New York 1973.

[12] A. Salomaa, On the injectivity of Parikh matrix mappings, Fundamenta Informatica 64, 2005: 391-404.

[13] K. G. Subramanian, A. Miin Huey, A. K. Nagar, On Parikh Matrices, International Journal of Foundations of Computer Science 20(2), 2009: 211-219.

[14] W. C. Teh, On Core Words and the Parikh Matrix Mapping, International Journal of Foundations of Computer Science 26(1), 2015: 123-142.

[15] W. C. Teh, K. H. Kwa, Core words and Parikh matrices, Theoretical Computer Science 582, 2015: 60-69.