AUSTRALASIAN JOURNAL OF COMBINATORICS
Volume **76(1)** (2020), Pages 208–219

Parikh-friendly permutations and uniformly Parikh-friendly words

Wen Chean Teh

School of Mathematical Sciences
Universiti Sains Malaysia
11800 USM
Malaysia
dasmenteh@usm.my

Abstract

The notion of Parikh-friendly permutations, originating from the study of Parikh matrices, was recently introduced by [Salomaa, Lec. Notes in Comp. Sci. 11011, Springer, Cham (2018), 100–112]. In this study, we show not only that every permutation is Parikh-friendly, but also that there exists a single word that witnesses the Parikh-friendliness of every permutation on a given ordered alphabet. In fact we introduce a relativized version of a Parikh-friendly permutation. As a result, words that are uniformly Parikh-friendly in a wider sense are effectively constructed.

1 Introduction

Although the Parikh vector of a word simply gives very minimal information about the word, it is useful in the study of formal languages. The classical Parikh Theorem [8] says that the set of Parikh vectors of a context-free language is a semilinear set. Parikh matrices were introduced by Mateescu, Salomaa, Salomaa, and Yu [6] as a canonical generalization of Parikh vectors. The Parikh matrix of a word carries more combinatorial information as its upper triangular entries are counts of certain subword occurrences of the word. The injectivity problem asks when two words share the same Parikh matrix. A satisfactory solution to this problem has been elusive, even for the ternary alphabet, despite various attempts over almost two decades [1–4, 9, 11, 15–18].

A subword condition was initially introduced in [13] as a conjunction of finitely many equations of the form $|w|_u = n$, where $|w|_u$ denotes the number of occurrences of u as a (scattered) subword of w. It can be extended to more complex subword conditions, for example, $|w|_a = |w|_b = |w|_c \ \wedge \ |w|_{abc} = (|w|_a)^3$, which defines the language $\{ a^n b^n c^n \mid n \in \mathbb{Z} \text{ and } n \geq 0 \}$. Languages defined by subword conditions normally cannot be described by other standard means. However, study of these languages has led to interesting results. For example, the simple subword condition

ISSN: 2202-3518

$|w|_a = |w|_u$ defines languages belonging to different classes in the Chomsky hierarchy for $u = b, ab, aba$ (see Theorem 4 in [10]).

The study of Parikh matrices naturally gives rise to languages defined by subword conditions. In 2018, Salomaa [12] introduced the notion of Parikh-friendly permutation based on Parikh matrices. By definition, a permutation is Parikh-friendly if there is a word nontrivially belonging to the corresponding language defined by some subword condition. Which permutation is Parikh-friendly was posed as an open problem. Here, we answer this question completely and it turns out unexpectedly that every permutation is Parikh-friendly. In fact, we manage to show that with respect to any ordered alphabet, a single word can be effectively constructed such that it witnesses the Parikh-friendliness of every permutation. Such a word exhibits somewhat stratified and uniform behavior with respect to subword occurrences. Furthermore, a generalization of Parikh-friendly permutations relative to a given word will be introduced. Our main result shows the existence of words that are uniformly Parikh-friendly not only with respect to every permutation but also relative to any word with a given length over the alphabet.

2 Preliminaries

2.1 Basic Notation

Suppose Σ is a (nonempty finite) alphabet. The set of words (respectively nonempty words) over Σ is denoted by Σ^* (respectively Σ^+). The empty word is denoted by λ. An *ordered alphabet* is an alphabet with a total ordering on its set of letters. Frequently, we will abuse notation and use Σ to stand for both the ordered alphabet and its *underlying alphabet*. Let $v, w \in \Sigma^*$. The concatenation of v and w is denoted by vw while the length of w is denoted by $|w|$. The *alphabet of* w, denoted $\text{alph}(w)$, is the set $\{ a \in \Sigma \mid |w|_a \geq 1 \}$. If the underlying alphabet is not explicitly mentioned, then $\text{alph}(w)$ is understood to be the minimal alphabet containing the letters of w. Suppose $\Gamma \subseteq \Sigma$. The projective morphism $\pi_\Gamma : \Sigma^* \to \Gamma^*$ is defined by

$$\pi_\Gamma(a) = \begin{cases} a, & \text{if } a \in \Gamma \\ \lambda, & \text{otherwise.} \end{cases}$$

A word u is a *(scattered) subword* of $w \in \Sigma^*$ if there exist $x_1, x_2, \ldots, x_n$, $y_0, y_1, \ldots$, $y_n \in \Sigma^*$, possibly empty, such that

$$u = x_1 x_2 \cdots x_n \text{ and } w = y_0 x_1 y_1 \cdots y_{n-1} x_n y_n.$$

The number of occurrences of u as a subword of w is denoted by $|w|_u$. For example, $|aabab|_{ab} = 5$ and $|baacbc|_{abc} = 2$. By convention, $|w|_\lambda = 1$ for all $w \in \Sigma^*$. It will be helpful later to observe that for all $w \in \{a, b, c\}^*$, we have

$$|w|_{abc} + |w|_{acb} + |w|_{bac} + |w|_{bca} + |w|_{cab} + |w|_{cba} = |w|_a |w|_b |w|_c.$$

Let $\text{Sym}(X)$ denote the symmetric group on a given set X. Suppose Σ is an alphabet. If $\sigma \in \text{Sym}(\Sigma)$, then σ naturally induces a morphism from Σ^* into Σ^*, also denoted by σ. Observe that $|\sigma(w)|_u = |w|_{\sigma^{-1}(u)}$ for all $w, u \in \Sigma^*$.

2.2 Extended Parikh Matrices

For any integer $k \geq 2$, let $\mathcal{M}_k$ denote the multiplicative monoid of $k \times k$ upper triangular matrices with nonnegative integral entries and unit diagonal.

Definition 2.1. [14] Suppose Σ is an alphabet and $u \in \Sigma^+$. Let $u = u_1 u_2 \cdots u_{|u|}$, where $u_i \in \Sigma$ for each $1 \leq i \leq |u|$. The *Parikh matrix mapping induced by* u, is the morphism $\Psi_u\colon \Sigma^* \to \mathcal{M}_{|u|+1}$ defined as follows: for each $a \in \Sigma$, if $\Psi_u(a) = (m_{i,j})_{1\leq i,j\leq |u|+1}$, then $m_{i,i} = 1$ for each $1 \leq i \leq |u|+1$, $m_{i,i+1} = \delta_{a,u_i}$ for each $1 \leq i \leq |u|$, and all other entries of the matrix $\Psi_u(a)$ are zero, where δ is the Kronecker delta function.

Matrices of the form $\Psi_u(w)$ for some $u, w \in \Sigma^*$ are called *extended Parikh matrices*. Note that Σ can be suppressed from the notation Ψ_u, not only for convenience, but also because $\Psi_u(w) = \Psi_u(\pi_{\mathrm{alph}(u)}(w))$ for all $w \in \Sigma^*$. If $\Sigma = \{a_1 < a_2 < \cdots < a_s\}$ is an ordered alphabet, then the associated standard *Parikh matrix mapping* [6] is the mapping Ψ_u induced by the word $u = a_1 a_2 \cdots a_s$.

The following characterization of extended Parikh matrices is essential throughout this work. It presents Parikh matrices as powerful tools in studying subword occurrences. Words sharing a given Parikh matrix canonically form a language defined by some subword condition.

Theorem 2.2. *[14] Suppose Σ is an alphabet and $u \in \Sigma^+$. Let $u = u_1 u_2 \cdots u_{|u|}$, where $u_i \in \Sigma$ for each $1 \leq i \leq |u|$. For every $w \in \Sigma^*$, the matrix $\Psi_u(w) = (m_{i,j})_{1\leq i,j\leq |u|+1}$ has the following properties:*

- $m_{i,i} = 1$ *for* $1 \leq i \leq |u| + 1$*;*
- $m_{i,j} = 0$ *for* $1 \leq j < i \leq |u| + 1$*;*
- $m_{i,j+1} = |w|_{u_i u_{i+1} \cdots u_j}$ *for* $1 \leq i \leq j \leq |u|$.

Example 2.3. $\Psi_{acba}(cbac) = \Psi_{acba}(c)\Psi_{acba}(b)\Psi_{acba}(a)\Psi_{acba}(c)$

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

$$= \begin{pmatrix}1&1&1&0&0\\0&1&2&1&1\\0&0&1&1&1\\0&0&0&1&1\\0&0&0&0&1\end{pmatrix} = \begin{pmatrix}1&|w|_a&|w|_{ac}&|w|_{acb}&|w|_{acba}\\0&1&|w|_c&|w|_{cb}&|w|_{cba}\\0&0&1&|w|_b&|w|_{ba}\\0&0&0&1&|w|_a\\0&0&0&0&1\end{pmatrix}.$$

2.3 t-Spectrum

The t-spectrum (also known as t-deck) of a word is defined as follows.

Definition 2.4. Suppose Σ is an alphabet. Let t be a positive integer. For every word $w \in \Sigma^*$, the *t-spectrum (over Σ)* of w is the function that sends every $u \in \Sigma^*$ with $|u| \leq t$ to $|w|_u$.

Hence, if two words $w_1, w_2 \in \Sigma^*$ have the same t-spectrum, then

$$|w_1|_u = |w_2|_u \quad \text{for all } u \in \Sigma^* \text{ such that } |u| \leq t. \qquad (\blacklozenge)$$

The following lemma is a variant of a classical fact. The reader may wish to verify it himself intuitively. This simple fact shows that no t is sufficiently large such that every word is uniquely determined by its t-spectrum [5]. Some variant of the lemma can also be found in [7].

Lemma 2.5. *Suppose $w_1, w_2, \ldots, w_n \in \Sigma^*$ have the same t-spectrum. Then for every permutation $\sigma \in \mathrm{Sym}(\{1, 2, \ldots, n\})$, the two words $w_1 w_2 \cdots w_n$ and $w_{\sigma(1)} w_{\sigma(2)} \cdots w_{\sigma(n)}$ have the same $(t+1)$-spectrum.*

3 Uniformly Parikh Friendly Words

Definition 3.1. [12] Suppose $\Sigma = \{a_1 < a_2 < \cdots < a_s\}$ is an ordered alphabet. A permutation $\sigma \in \mathrm{Sym}(\Sigma)$ is *Parikh-friendly with respect to* Σ if there exists a word $w \in \Sigma^*$ with $\mathrm{alph}(w) = \Sigma$ such that

$$\Psi_{a_1 a_2 \cdots a_s}(w) = \Psi_{\sigma(a_1 a_2 \cdots a_s)}(w).$$

We also say that w is a *Parikh-friendly witness for σ with respect to Σ.*

Example 3.2. Consider the word $w = abccba$. Let σ denote the permutation (abc) and so $\sigma^{-1} = (cba)$. Note that $\Psi_{abc}(w) = \Psi_{bca}(w) \neq \Psi_{cab}(w)$. Hence, w is a Parikh-friendly witness for σ (but not for σ^{-1}) with respect to $\{a < b < c\}$.

Example 3.3. Let $\Sigma = \{a < b < c < d < e\}$. The word $abcddcbaee$ (respectively $acddcabeeb$) is a Parikh-friendly witness for the permutation $(abcd)(e)$ (respectively $(acd)(be)$) with respect to Σ.

By constructing witnesses along the lines illustrated by Example 3.3, Salomaa arrived at the following result.

Theorem 3.4. *[12] With respect to any ordered alphabet, every product of transpositions such that no two transpositions have a common element is Parikh-friendly.*

The same construction does not work for every cyclic permutation. For example, consider the ordered alphabet $\Sigma = \{a < b < c < d\}$. The word $w = acbddbca$ is not a Parikh-friendly witness for the permutation $(acbd)$ because $|w|_{abcd} \neq |w|_{cdba}$, but $dbcaacbd$ is a Parikh-friendly witness for the same permutation. Meanwhile, $acbd$ is a Parikh-friendly witness for both $(abdc)$ and $(ac)(bd)$ with respect to Σ. Salomaa then posed the characterization of Parikh-friendly permutations as an open problem and this study was initiated with the aim of solving this problem.

As it turns out, for any ordered alphabet Σ, our result shows that every permutation on Σ is Parikh-friendly with respect to Σ. In fact, we obtain more than we expect. Since there are finitely many permutations, it is conceivable that there could exist a single word which is a common Parikh-friendly witness for every permutation on Σ, especially when the size of Σ is small. Not only is our conjecture true for any ordered alphabet, such universally Parikh-friendly words can be effectively constructed, thus bringing us to the following definition.

Definition 3.5. Suppose Σ is an alphabet and $w \in \Sigma^+$. We say that w is *uniformly Parikh-friendly with respect to* Σ if $\Psi_u(w) = \Psi_v(w)$ whenever $u, v \in \Sigma^+$ such that $|u|_a = |v|_a = 1$ for all $a \in \Sigma$.

Remark 3.6. Suppose w is uniformly Parikh-friendly with respect to Σ. Then it must be the case that $\mathrm{alph}(w) = \Sigma$ and thus the reference to Σ can be safely omitted. Also, w is a Parikh-friendly witness for every $\sigma \in \mathrm{Sym}(\Sigma)$ with respect to Σ', where Σ' is any ordered alphabet with underlying alphabet Σ. Furthermore, the reverse w^r of w is uniformly Parikh-friendly as well because $\Psi_u(w^r) = \Psi_v(w^r)$ if and only if $\Psi_{u^r}(w) = \Psi_{v^r}(w)$ whenever $u, v \in \Sigma^*$ such that $|u| = |v|$.

Theorem 3.7. *Suppose w is a uniformly Parikh-friendly word with respect to an alphabet Σ. Let $1 \le k \le |\Sigma|$ be an integer. Then $|w|_u = \frac{\alpha^k}{k!}$ for every $u \in \Sigma^*$ with $|u| = k$ such that $|u|_a \le 1$ whenever $a \in \Sigma$, where $|w|_a = \alpha$ for all $a \in \Sigma$.*

Proof. Let $\Sigma = \{a_1, a_2, \dots, a_{|\Sigma|}\}$. Fix an integer $1 \le k \le |\Sigma|$. Suppose $u \in \Sigma^*$ with $|u| = k$ such that $|u|_a \le 1$ whenever $a \in \Sigma$. Let v be any word over Σ having u as a prefix such that $|v|_a = 1$ for all $a \in \Sigma$. Since w is uniformly Parikh-friendly with respect Σ, by definition, $\Psi_v(w) = \Psi_{a_1 a_2 \cdots a_{|\Sigma|}}(w)$. By comparing the $(1, k+1)$-entries, it follows that $|w|_u = |w|_{a_1 a_2 \cdots a_k}$. Since u is arbitrary, this (when $k = 1$) implies that $|w|_a$ is equal to a constant, say α, for every $a \in \Sigma$. Now, it can be seen that

$$k!|w|_{a_1 a_2 \cdots a_k} = \sum_{\sigma \in \mathrm{Sym}(\{a_1, a_2, \dots, a_k\})} |w|_{\sigma(a_1 a_2 \cdots a_k)} = \prod_{i=1}^{k} |w|_{a_i} = \alpha^k.$$

Therefore, $|w|_u = |w|_{a_1 a_2 \cdots a_k} = \frac{\alpha^k}{k!}$. ☐

By Theorem 3.7, it follows that if a uniformly Parikh-friendly word exists, then its Parikh matrices have a special form. For example, if w is uniformly Parikh-friendly with respect to a quarternary alphabet Σ, where $|w|_a = \alpha$ for all $a \in \Sigma$, then the Parikh matrix of w is

$$\begin{pmatrix} 1 & \alpha & \frac{\alpha^2}{2!} & \frac{\alpha^3}{3!} & \frac{\alpha^4}{4!} \\ 0 & 1 & \alpha & \frac{\alpha^2}{2!} & \frac{\alpha^3}{3!} \\ 0 & 0 & 1 & \alpha & \frac{\alpha^2}{2!} \\ 0 & 0 & 0 & 1 & \alpha \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

with respect to any ordering of Σ. Hence, the notion of uniformly Parikh-friendly naturally gives rise to these languages defined by special numerical parameters. Furthermore, the existence of a uniformly Parikh-friendly word over Σ guarantees not only that each such language is nonempty but so is their intersection.

The following theorem is a consequence of Theorem 4.6 in the next section. There we generalize the notion of Parikh-friendly permutations and prove the existence of words with a much stronger Parikh-friendly property which will imply uniform Parikh-friendliness.

Theorem 3.8. *For any alphabet Σ, there exists a uniformly Parikh-friendly word with respect to Σ.*

By Remark 3.6, the next corollary follows immediately.

Corollary 3.9. *Suppose Σ is an ordered alphabet. Every permutation on Σ is Parikh-friendly with respect to Σ.*

Note that if w is a Parikh-friendly witness for a cyclic permutation on Σ, then there is a constant α such that $|w|_a = \alpha$ for all $a \in \Sigma$. Our numerical computations suggest that every cyclic permutation has a Parikh-friendly witness with α equal to one or two. Furthermore, it is plausible that through some careful analysis, a Parikh-friendly witness with minimal length can be effectively constructed for each cyclic permutation. Therefore, it leads us to the following conjecture, which is left open.

Conjecture 3.10. *Suppose Σ is an ordered alphabet. Every permutation on Σ has a Parikh-friendly witness of length at most* $2|\Sigma|$.

Example 3.11. Let $\Sigma = \{a < b < c\}$. The words abc, $abccba$, $cbaabc$, $abbac$, $bccba$, $caacb$ are Parikh-friendly witnesses for the permutations (a), (abc), (cba), (ab), (bc), and (ca) with respect to Σ respectively.

4 Generalization of Parikh Friendly Permutation

Definition 4.1. Suppose Σ is an alphabet and $u \in \Sigma^+$. A permutation $\sigma \in \mathrm{Sym}(\Sigma)$ is *u-Parikh-friendly with respect to* Σ if there exists a word w with $\mathrm{alph}(w) = \Sigma$ such that

$$\Psi_u(w) = \Psi_{\sigma(u)}(w).$$

We also say that w is a *u-Parikh-friendly witness for σ with respect to* Σ.

Note that in our definition, $\mathrm{alph}(u)$ need not be Σ.

We now introduce a stronger version of uniformly Parikh-friendliness.

Definition 4.2. Suppose Σ is an alphabet and $w \in \Sigma^+$. Let t be a positive integer. We say that w is *t-uniformly Parikh-friendly with respect to* Σ if whenever $u \in \Sigma^*$ with $|u| = t$, it holds that w is a u-Parikh-friendly witness for every $\sigma \in \mathrm{Sym}(\Sigma)$.

Example 4.3. Consider the word $w = abbabaab$. It can be verified that $|w|_a = |w|_b = 4$, $|w|_{ab} = |w|_{ba} = 8$, $|w|_{aab} = |w|_{bba} = |w|_{baa} = |w|_{abb} = 7$, and $|w|_{bab} = |w|_{aba} = 10$. It follows that $\Psi_{aab}(w) = \Psi_{bba}(w)$, $\Psi_{baa}(w) = \Psi_{abb}(w)$, $\Psi_{aba}(w) = \Psi_{bab}(w)$, and $\Psi_{aaa}(w) = \Psi_{bbb}(w)$. Therefore, w is 3-uniformly Parikh-friendly with respect to $\{a, b\}$.

Remark 4.4. If w is $|\Sigma|$-uniformly Parikh-friendly with respect to Σ, then it is uniformly Parikh-friendly with respect to Σ, but the converse is not true. None of the 66 uniformly Parikh-friendly ternary words listed in the appendix is 3-uniformly Parikh-friendly.

We will now prove the existence of t-uniformly Parikh-friendly words. The following is the key lemma.

Lemma 4.5. *Suppose* Σ *is an alphabet and* $w \in \Sigma^+$. *Let* t *be a positive integer. Suppose* $G = \{\sigma_1, \sigma_2, \dots, \sigma_n\}$ *is a subgroup of the symmetric group* $\mathrm{Sym}(\Sigma)$. *Let* v *denote the word* $\sigma_1(w)\sigma_2(w)\cdots\sigma_n(w)$. *If* $\sigma_1(w), \sigma_2(w), \dots, \sigma_n(w)$ *have the same* t*-spectrum, then* $\sigma_1(v), \sigma_2(v), \dots, \sigma_n(v)$ *have the same* $(t+1)$*-spectrum.*

Proof. Fix an integer $1 \le i \le n$. Since G is a group, it follows that

$$\{\sigma_i\sigma_1, \sigma_i\sigma_2, \dots, \sigma_i\sigma_n\} = \{\sigma_1, \sigma_2, \dots, \sigma_n\}.$$

It can be seen that

$$\sigma_i(v) = \sigma_i(\sigma_1(w))\sigma_i(\sigma_2(w))\cdots\sigma_i(\sigma_n(w)) = (\sigma_i\sigma_1)(w)(\sigma_i\sigma_2)(w)\cdots(\sigma_i\sigma_n)(w).$$

Hence, by Lemma 2.5, $\sigma_i(v)$ and v have the same $(t+1)$-spectrum. □

Theorem 4.6. *Suppose* Σ *is an alphabet. For every positive integer* t*, there exists a* t*-uniformly Parikh-friendly word with respect to* Σ.

Proof. Let $\Sigma = \{a_1, a_2, \dots, a_{|\Sigma|}\}$. Suppose $\sigma_1, \sigma_2, \dots, \sigma_n$ are all the elements of the symmetric group $\mathrm{Sym}(\Sigma)$. Let $w_1 = a_1a_2\dots a_{|\Sigma|}$. Then clearly $\sigma_1(w_1), \sigma_2(w_1), \dots, \sigma_n(w_1)$ have the same 1-spectrum. Recursively, using Lemma 4.5, for every integer $t \ge 2$, a word w_t can be effectively constructed such that $\sigma_1(w_t), \sigma_2(w_t), \dots, \sigma_n(w_t)$ have the same t-spectrum.

Fix a positive integer t. We claim that w_t is t-uniformly Parikh-friendly with respect to Σ. Suppose $u \in \Sigma^*$ with $|u| = t$. Let $\sigma \in \mathrm{Sym}(\Sigma)$ be arbitrary. We need to show that $\Psi_u(w_t) = \Psi_{\sigma(u)}(w_t)$. To see this, take an arbitrary factor v of u. By Theorem 2.2, it suffices to show that $|w_t|_v = |w_t|_{\sigma(v)}$. Let $\tau = \sigma^{-1}$. Then $|w_t|_{\sigma(v)} = |w_t|_{\tau^{-1}(v)} = |\tau(w_t)|_v$. Since $\tau(w_t)$ and w_t have the same t-spectrum and $|v| \le t$, it follows from (♦) that $|w_t|_{\sigma(v)} = |\tau(w_t)|_v = |w_t|_v$. □

By Remark 4.4, Theorem 3.8 follows immediately as a corollary.

5 Special Uniformly Parikh-friendly Words

For the quaternary alphabet, 4-uniformly Parikh-friendly words can be effectively constructed as in the proof of Theorem 4.6. However, these uniformly Parikh-friendly quaternary words have length $4(24)^3$. In this section, we present some special uniformly Parikh-friendly words over the quaternary alphabet, each with a much shorter length of 96.

By Theorem 3.7, if w is a uniformly Parikh-friendly word with respect to the ternary alphabet $\{a, b, c\}$, then $|w|_a = |w|_b = |w|_c$ must be a multiple of six and so $|w|$ must be at least 18. Computationally, we obtain a complete list of all such words with length 18, as provided in the appendix. If each of the words is split into blocks of three, it can be seen that 24 of them are quite special because each is a certain concatenation of the following words:

$$abc,\ acb,\ bac,\ bca,\ cab,\ cba.$$

Specifically, each of the ones that are underlined is one of six possible concatenations of the following three words:

$$abccba,\ bcaacb,\ cabbac.$$

Remark 5.1. Note that each letter occurs as the first letter in one of the three words. However, there is some flexibility regarding the middle four letters of each word when concatenating them to form a uniformly Parikh-friendly word. For example, the words *acbbca bcaacb cabbac* and *cabbac acbbca baccab* are both uniformly Parikh-friendly.

Inspired by this computational observation for the ternary alphabet, we discover some uniformly Parikh-friendly words with respect to the quaternary alphabet that have similar special features. Precisely, let

$$\begin{array}{llll} v_1 = abcddcba, & v_4 = bcdaadcb, & v_7 = cdbaabdc, & v_{10} = dacbbcad, \\ v_2 = dbaccabd, & v_5 = acbddbca, & v_8 = bdcaacdb, & v_{11} = cadbbdac, \\ v_3 = cbaddabc, & v_6 = dcabbacd, & v_9 = adbccbda, & v_{12} = badccdab. \end{array}$$

This time, not all possible concatenations of all the v_i's are uniformly Parikh-friendly. In fact, we found out computationally that there are exactly 432 distinct concatenations of all the v_i's that are uniformly Parikh-friendly, with 36 having v_i as the first block for each $1 \leq i \leq 12$. Specifically, let

$$V_1 = v_1 v_2 v_3, \qquad V_2 = v_6 v_5 v_4, \qquad V_3 = v_7 v_8 v_9, \qquad V_4 = v_{12} v_{11} v_{10}.$$

Then each of the 24 possible concatenations of all the V_i's is uniformly Parikh-friendly. This will be justified in our next paragraph.

Let $\Sigma = \{a, b, c, d\}$. Suppose α, β, and γ are any distinct letters of Σ. For each $1 \leq i \leq 4$, by Remark 5.1, it turns out that $\pi_{\{\alpha,\beta,\gamma\}}(V_i)$ is uniformly Parikh-friendly with respect of $\{\alpha, \beta, \gamma\}$. By Theorem 3.7, it follows that $|V_i|_u = |V_i|_v = \frac{6^k}{k!}$ for all $1 \leq i \leq 4$ whenever $u, v \in \Sigma^*$ with $|u| = |v| = k \leq 3$ such that $|u|_x \leq 1$ and $|v|_x \leq 1$ for all $x \in \Sigma$. Let $V = V_1V_2V_3V_4$. It is then straightforward to verify that $|V|_u = |V|_v$ for such u and v. For example,

$$\begin{aligned} |V|_{xy} &= \sum_{i=1}^{4} |V_i|_{xy} + |V_1|_x |V_2V_3V_4|_y + |V_2|_x |V_3V_4|_y + |V_3|_x |V_4|_y \\ &= \sum_{i=1}^{4} |V_i|_{yz} + |V_1|_y |V_2V_3V_4|_z + |V_2|_y |V_3V_4|_z + |V_3|_y |V_4|_z = |V|_{yz} \end{aligned}$$

for all distinct $x, y, z \in \Sigma$. Hence, for V to be uniformly Parikh-friendly, by definition and Theorem 2.2, it suffices to show that $|V|_u = |V|_v$ whenever $u, v \in \Sigma^*$ such that $|u|_x = |v|_x = 1$ for all $x \in \Sigma$. For such later u and v, by a similar argument, $|V|_u = |V|_v$ reduces to $\sum_{i=1}^4 |V_i|_u = \sum_{i=1}^4 |V_i|_v$, which holds because it can be verified computationally that $\sum_{i=1}^4 |V_i|_u = 216$ for every $u \in \Sigma^*$ such that $|u|_x = 1$ for all $x \in \Sigma$. Generally, this argument shows that any concatenation of all the V_i's is uniformly Parikh-friendly with respect to Σ.

6 Conclusion

The notions of uniformly Parikh-friendly and t-uniformly Parikh-friendly are new ways of defining special languages that can be described by numerical parameters. More importantly, their definitions are connected to classical objects in abstract algebra, namely, permutations. Therefore, this study opens up a possible cross-study between combinatorics on words and group theory. Since G can be any subgroup of $\mathrm{Sym}(\Sigma)$ in Lemma 4.5 (for example, the alternating group), it suggests a generalization of uniformly Parikh-friendly, not necessarily for every permutation, but only for permutations coming from G. Hence, the well-established theory of symmetric groups may prove to be useful.

In the last section, we have identified some special uniformly Parikh-friendly words. In each such word, not only that the arrangement of the blocks can be associated to some permutation, but also that the blocks can be obtained from a single block using all possible permutations of the letters. It remains to be seen whether our ad hoc construction can be generalized to higher alphabets.

Finally, we would like to bring out another potential open problem inspired by this work. In Example 4.3, although $w = abbabaab$ is 3-uniformly Parikh-friendly with respect to the binary alphabet $\{a, b\}$, it does not satisfy the following stronger property:

$$|w|_a = |w|_b,\ |w|_{ab} = |w|_{ba},\ \text{and}\ \ |w|_{abb} = |w|_{bab} = |w|_{bba} = |w|_{aab} = |w|_{aba} = |w|_{baa} \qquad (*)$$

because $|w|_{abb} \neq |w|_{aba}$. Computationally, we found out that binary words with minimal length satisfying $(*)$ have length 32 and there are exactly six of them as follows:

$$ab^5a^6ba^2ba^2bab^8a^4, \qquad a^4b^8aba^2ba^2ba^6b^5a, \qquad a^4b^8a^4b^4a^8b^4,$$
$$ba^5b^6ab^2ab^2aba^8b^4, \qquad b^4a^8bab^2ab^2ab^6a^5b, \qquad b^4a^8b^4a^4b^8a^4.$$

Therefore, for higher alphabets, we ask whether there are words satisfying the corresponding property analogous to $(*)$. Such words, if they exist, could be significantly longer and it may not be as easy to find them computationally, even for the ternary alphabet.

Appendix

The following are exactly all the 66 uniformly Parikh-friendly words over the ternary alphabet $\{a, b, c\}$ with minimal length. The ones underlined are mentioned in Section 5.

$cababcbcabaccbaacb$	$abccabbcacbabacacb$	$abccabcbbaacabccab$
$cabbcaabcbacacbcba$	$abcbcacabcbaacbbac$	$cabbacacbbcabcaacb$
$cabbacacbbcabaccab$	$\underline{cabbacabccbabcaacb}$	$cabbacabccbabaccab$
$acbbcacabbacbcaacb$	$acbbcacabbacbaccab$	$\underline{abccbacabbacbcaacb}$
$abccbacabbacbaccab$	$cabbcabaaccbcabbca$	$cabbacbcaacbacbbca$
$\underline{cabbacbcaacbabccba}$	$cabbacbaccabacbbca$	$cabbacbaccababccba$
$acbbcabcaacbcabbac$	$acbbcabaccabcabbac$	$acbbacbccaabacbbac$
$\underline{abccbabcaacbcabbac}$	$abccbabaccabcabbac$	$acbbcacbaabcbcaacb$
$acbbcacbaabcbaccab$	$acbbcabcaacbcbaabc$	$acbbcabaccabcbaabc$
$abccbacbaabcbcaacb$	$abccbacbaabcbaccab$	$abccbabcaacbcbaabc$
$abccbabaccabcbaabc$	$acbcbabacbcaabccab$	$acbbaccbabcacababc$
$bcacababcacbbaccba$	$bcaabccabacbcbabac$	$bcaacbcabbacacbbca$
$\underline{bcaacbcabbacabccba}$	$bcaacbacbbcacabbac$	$\underline{bcaacbabccbacabbac}$
$baccabcabbacacbbca$	$baccabcabbacabccba$	$baccabacbbcacabbac$
$baccababccbacabbac$	$cbaacbabbccacbaacb$	$cbaabcacbbcabcaacb$
$cbaabcacbbcabaccab$	$cbaabcabccbabcaacb$	$cbaabcabccbabaccab$
$bcaacbacbbcacbaabc$	$bcaacbabccbacbaabc$	$bcaabcaccbbabcaabc$
$baccabacbbcacbaabc$	$baccababccbacbaabc$	$cbaabcbcaacbacbbca$
$cbaabcbcaacbabccba$	$cbaabcbaccabacbbca$	$cbaabcbaccababccba$
$bcaacbcbaabcacbbca$	$bcaacbcbaabcabccba$	$baccabcbaabcacbbca$
$baccabcbaabcabccba$	$cbaacbbacabcbcacab$	$bacacbcbacabbcaabc$
$baccbacaabbcbaccba$	$cbabacacbabccabbca$	$baccbaacbcababcbca$

Acknowledgements

The author gratefully acknowledges the support for this research by Research University Grant 1001/PMATHS/8011019 of Universiti Sains Malaysia. This work was completed during his sabbatical leave from 15 Nov 2018 to 14 Aug 2019, supported by Universiti Sains Malaysia. Appreciation also goes to the referees and the editor for careful reading and suggestions for improvement of this work.

References

[1] A. Atanasiu, R. Atanasiu and I. Petre, Parikh matrices and amiable words, *Theoret. Comput. Sci.* 390(1) (2008), 102–109.

[2] A. Atanasiu, C. Martín-Vide and A. Mateescu, On the injectivity of the Parikh matrix mapping, *Fund. Inform.* 49(4) (2002), 289–299.

[3] S. Fossé and G. Richomme, Some characterizations of Parikh matrix equivalent binary words, *Inform. Process. Lett.* 92(2) (2004), 77–82.

[4] K. Mahalingam and K. G. Subramanian, Product of Parikh matrices and commutativity, *Internat. J. Found. Comput. Sci.* 23(1) (2012), 207–223.

[5] B. Manvel, A. Meyerowitz, A. Schwenk, K. Smith and P. Stockmeyer, Reconstruction of sequences, *Discrete Math.* 94(3) (1991), 209–219.

[6] A. Mateescu, A. Salomaa, K. Salomaa and S. Yu, A sharpening of the Parikh mapping, *Theor. Inform. Appl.* 35(6) (2001), 551–564.

[7] R. Milner, The Spectra of Words, In *Processes, Terms and Cycles: Steps on the Road to Infinity*, Lecture Notes in Comput. Sci. 3838, Springer, Berlin (2005), pp. 1–5.

[8] R. J. Parikh, On context-free languages, *J. Assoc. Comput. Mach.* 13 (1966), 570–581.

[9] G. Poovanandran and W. C. Teh, Elementary matrix equivalence and core transformation graphs for Parikh matrices, *Discrete Appl. Math.* 251 (2018), 276–289.

[10] A. Salomaa, On languages defined by numerical parameters, In *Formal Models, Languages and Applications*, World Scientific (2006), pp. 320–336.

[11] A. Salomaa, Criteria for the matrix equivalence of words, *Theoret. Comput. Sci.* 411 (2010), 1818–1827.

[12] A. Salomaa, Parikh Matrices: Subword Indicators and Degrees of Ambiguity, In *Adventures between lower bounds and higher altitudes*, Lec. Notes in Comp. Sci. 11011, Springer, Cham (2018), pp. 100–112.

[13] A. Salomaa and S. Yu, Subword conditions and subword histories, *Inform. and Comput.* 204(12) (2006), 1741–1755.

[14] T.-F. Şerbănuţă, Extending Parikh matrices, *Theoret. Comput. Sci.* 310 (2004), 233–246.

[15] V. N. Şerbănuţă, On Parikh matrices, ambiguity, and prints, *Internat. J. Found. Comput. Sci.* 20(1) (2009), 151–165.

[16] V. N. Şerbănuţă and T. F. Şerbănuţă, Injectivity of the Parikh matrix mappings revisited, *Fund. Inform.* 73 (2006), 265–283.

[17] W. C. Teh, Parikh matrices and Parikh rewriting systems, *Fund. Inform.* 146 (2016), 305–320.

[18] W. C. Teh, A. Atanasiu, and G. Poovanandran, On strongly *M*-unambiguous prints and Şerbănuţă's conjecture for Parikh matrices, *Theoret. Comput. Sci.* 719 (2018), 86–93.

(Received 26 June 2019; revised 5 Nov 2019)