Factors and Subwords of Rich Partial Words

R. Krishna Kumari Member, IAENG and R. Arulprakasam Member, IAENG *

Abstract—Many classes of finite words have noticeable properties with reference to their palindromic factors and one among them are the words having zero palindromic defect i.e., words rich in palindromes. In this paper we introduce rich partial word and discuss its combinatorial properties. We show that the palindromic richness of a partial word can be studied by including the positions of the missing symbols in that word. The significant difference between rich and rich partial word is that a rich word of length \(n\) contains exactly \(n+1\) distinct palindromic factors whereas a rich partial word of length \(n\) contains at least \(n+1\) distinct palindromic factors. These factors differ from the classical palindromes due to the presence of holes.

Keywords: palindromes, rich words, factors, partial words, primitivity.

1 Introduction

In the study of the various properties of words with finite length [8] such as structural and combinatorial properties, palindromes are natural objects which play a vital role in word combinatorics, automata theory and formal languages. Palindromes often occur in DNA and are extensively present in human cancer cells [11]. In biological context, complement DNA characters are considered by palindromes. By identification of these segments of DNAs, the instability of genomes could be understood. Biologists believe that palindromes play an major role in cell processes and other regulation gene activity because these are frequently noticed near introns, promoters and specific untranslated regions. So, locating palindromic factors in any genome sequence is vital. Also for comparison study, locating common palindromes in two genome sequences can be a major criterion. A palindromic word is a word when taken in reverse order gives the same word. Many classes of words have prominent properties with regard to their palindromic factors [7]. Algorithmic and combinatorial studies of palindromes are considered as a favorable tool to construct linear-time recognizable languages [3, 15].

In the study of palindromes, one of the recent topics of interest concerns an interesting class of finite words termed as rich words. Words comprising the greatest number of distinct and palindromic factors are rich words and are called as words with zero palindromic defect \([2, 12, 13, 14, 17]\). In [18], X. Droubay et al. showed that a finite word \(x\) of length \(|x|\) has maximum \(|x|\) distinct palindromic factors, excluding the empty word. Characterized by this palindromic richness property in [3], the authors launched a unified study of words with finite and infinite length. Accordingly we say that a finite word \(x\) is rich if and only if it has exactly \(|x|+1\) distinct palindromic factors. In various contexts, rich words have appeared such as complementation-symmetric sequences, episturmian words and a certain class of words associated with \(\beta\)-expansions where \(\beta\) represents a simple Parry number. The number of rich binary words of length \(n\) can be referred in https://oeis.org/A216264.

Partial words are nothing but words with holes and are considered in gene comparisons [1, 9, 16]. For instance, alignment of two DNA sequences which are genetic information carriers can be regarded as formation of two compatible partial words. The DNA sequence is treated as a finite word in DNA computations, and is used to encode information. When encoding information, some parts of the information may be hidden or not visible, which are revealed by using a partial word which represents the position of a missing symbol in a word. Initial research on partial words was initiated by Berstel and Boasson [10] and later expanded by Blanchet-Sadri [4, 5, 6]. Partial words and palindromicity of words are classical topics in molecular biology and language theory which inspired and initiated a unified study of rich words and partial words. The hole(s) present in partial words is not a character of the alphabet but survives as a back-up symbol for the unknown letter. Since it is compatible to any of the letter(s) in the alphabet, if a hole in a rich partial word over the alphabet is replaced by a letter in the alphabet, the rich partial word turns out to be a rich word. On the other hand, since holes do not belong to the alphabet, we study the palindromic richness of a partial word by including the positions of the missing symbols in that word. This paper introduces rich partial words and study some combinatorial properties. We initially recall in Section 2 the fundamental notions and properties. We define rich partial words and discuss some properties based on
their palindromic richness in Section 3. In Section 4 we discuss the relation between partial palindromic perfect factors and partial palindromic perfect subwords of rich partial words followed by conclusion in Section 5.

2 Preliminaries

Let the finite alphabet \( A \) represent a non-empty set of symbols (or letters). A total word (or string) is a sequence of letters over \( A \). The length (or size) of a total word \( x = x[1\ldots n] \) is \( n \). The length of a total word \( x \) is denoted by \( |x| \). \( \text{Alph}(x) \) denotes the set of all elements in \( x \). \( \lambda \) denotes the empty word. \( A^* \) denotes the set of all total words from \( A \) including \( \lambda \) and \( A^+ \) denotes the set of all total words from \( A \) excluding \( \lambda \). A language \( L \) is a subset of \( A^* \).

The total word \( x \) is a subword (or factor) of \( y \) if the total words \( u \) and \( v \) exist such that \( y = uve \). If \( u, v \neq \lambda \) then \( x \) is a proper subword of \( y \). If \( u = \lambda \) then \( x \) is a prefix of \( y \). If \( v = \lambda \) then \( x \) is a suffix of \( y \). A finite total word \( x \) is called a palindrome if \( x = x^R \) where \( x^R \) is the reversal (mirror image) of \( x \). A total word \( x \) is rich if it has exactly \( |x| + 1 \) distinct factors that are palindromic including the empty word \( \lambda \). A non-empty factor \( x \) of a finite word \( u \) is unioccurent in \( y \) if \( x \) has exactly one occurrence in \( y \). If \( x \) has more than one occurrence in \( y \), then there exists a factor \( z \) of \( y \) having exactly two distinct occurrences of \( x \), one as a prefix and other as a suffix. Such a factor \( z \) is called a complete return to \( x \) in \( y \). For example, \( bbaacab \) is a complete return to \( bb \) in the rich word \( bbcaacbba \).

The sequence of symbols that contains a number of “do not know symbols” or “holes” denoted as \( \diamond \) is termed as a finite partial word (or partial word).

The partial word of \( u \) denoted by \( u_\diamond \) is the total function \( u_\diamond : \{1,2,\ldots,n\} \to A_\diamond = A \cup \{\diamond\} \) defined by

\[
\forall i \in \text{dom}(u_\diamond) \quad u_\diamond(i) = \begin{cases} u(i) & \text{if } i \in \text{dom}(u) \\ \diamond & \text{if } i \in \text{dom}(H(u)) \end{cases}
\]

where \( \text{dom}(u) \) represents the domain set and \( \text{dom}(H(u)) \) denotes the set of holes in \( u \). The set of all partial words over \( A_\diamond \) is denoted as \( A_\diamond^* \). \( A^*_\diamond \) denotes the set of all partial words excluding the empty word. A partial language \( L_\diamond \subseteq A_\diamond^* \) is a set of all partial words over \( A_\diamond \).

We note that,

(i) A total word is a partial word with zero holes and the empty word is not a partial word.

(ii) The symbol \( \diamond \) does not belong to the alphabet \( A \) but a standby symbol for the unknown letter.

(iii) The symbol \( \diamond \) is compatible to the letters of the alphabet \( A \).

(iv) The symbol \( \diamond \) alone of any length cannot exist as a word. In other words, the hole of any length is neither a total word nor a partial word.

A partial word \( u_\diamond = u_\diamond[1\ldots n] \) is primitive (non-periodic) if no word \( v \) exists such that \( u_\diamond \subset v^i \) with \( i \geq 2 \). Partial words that are not primitive are said to be periodic partial words. If \( u_\diamond \) and \( v_\diamond \) are two partial words of equal length and if all the elements in domain of \( u_\diamond \) are also in domain of \( v_\diamond \) with \( u_\diamond(i) = v_\diamond(i) \) for all \( i \in \text{dom}(u_\diamond) \), then \( u_\diamond \) is contained in \( v_\diamond \) and is denoted by \( u_\diamond \subset v_\diamond \). Two partial words \( u_\diamond \) and \( v_\diamond \) are compatible, denoted by \( u_\diamond \uparrow v_\diamond \) if \( u_\diamond(i) = v_\diamond(i) \) for all \( i \in \text{dom}(u_\diamond) \cap \text{dom}(v_\diamond) \). Equivalently, the partial words \( u_\diamond \) and \( v_\diamond \) are compatible if a partial word (or a total word) \( u_\diamond \) exists such that \( u_\diamond \subset v_\diamond \) and \( v_\diamond \subset u_\diamond \). A finite partial word \( u_\diamond \) is a palindrome if \( u_\diamond \) is compatible with its reversal (denoted by \( u_\diamond \uparrow v_\diamond^R \)). For instance \( u_\diamond = \diamond ab\diamond \diamond ab \) is a palindrome.

3 Rich Partial Words

This section defines rich partial words in view of their palindromic richness and discusses their combinatorial properties. The empty word \( \lambda \) is regarded as a palindrome.

**Definition 1.** A factor \( p_\diamond \) of a partial word \( u_\diamond \) over \( A_\diamond \) is called a partial palindromic proper factor if \( p_\diamond \) is compatible with its reversal (denoted by \( p_\diamond \uparrow p_\diamond^R \)). The set of all non-empty partial palindromic proper factors of \( u_\diamond \) is denoted by \( \text{PPPF}(u_\diamond) \).

**Example 1.** Consider a partial word \( u_\diamond = babaab \) over \( A_\diamond = \{a,b\} \cup \{\diamond\} \). The palindromic factors of \( u_\diamond \) are

\[
\{\lambda, a, b, aa, b\diamond, b, ab, b\diamond b, baab\}.
\]

Here the factors \( \{b\diamond, b, ab, b\diamond b\} \) are termed as partial palindromic factors.

**Definition 2.** Any partial word over \( A_\diamond \) with length \( n \) is a rich partial word if it has at least \( n \) distinct partial palindromic proper factors.

**Example 2.** Consider a partial word \( u_\diamond = ba\diamond aba \) over \( A_\diamond = \{a,b\} \cup \{\diamond\} \) with \( |u_\diamond| = 6 \). The set of all distinct partial palindromic proper factors of \( u_\diamond \) are

\[
\{\lambda, a, b, a\diamond, a\diamond a, ba\diamond, a\diamond ab, aba, ba\diamond ab, a\diamond aba\}.
\]

Among the above set, the set of all distinct partial palindromic factors of \( u_\diamond \) are

\[
\{a\diamond, a\diamond a, ba\diamond, a\diamond ab, ba\diamond ab, a\diamond aba\}.
\]

Here the number of distinct partial palindromic proper factors is equal to \( |u_\diamond| + 1 \). Hence \( u_\diamond \) is a rich partial word.

**Example 3.** Consider a partial word \( v_\diamond = \diamond abab \) with length \( |v_\diamond| = 6 \) over \( A_\diamond = \{a,b\} \cup \{\diamond\} \). Then the partial palindromic proper factors of \( v_\diamond \) are

\[
v_\diamond = \{\diamond a, \diamond ab, \diamond abab\}.
\]
The set of all suffixes of $p$ are uniquely represented as a pair $(u, v)$ such that $u$ and $v$ are the prefix and suffix of $p$ with maximum length and not factors of one another.

To prove the uniqueness, for any finite rich partial word $u_0$ with factors $v_0$ and $r_0$ having the same palindromic factor $s$ and same palindromic factor $t$ with maximum length. We have to show that $v_0 = r_0$. Let us prove by contradiction. Suppose $v_0 \neq r_0$ such that both $v_0$ and $r_0$ are not palindromes. Then $v_0$ and $r_0$ are not factors of one another and $p_0$ and $q_0$ are uniuoccurrent in each of $v_0$ and $r_0$. Let $k_0$ be a factor of $u_0$ of least length. Let us assume that the factor $v_0$ is the prefix and the factor $r_0$ is the suffix of $k_0$. Then $p_0$ (resp.$q_0$) occurs twice in $k_0$ as a prefix (resp. suffix) of each of $v_0$ and $r_0$ respectively. Since $p_0$ and $q_0$ are uniuoccurrent in $v_0$ and $r_0$ respectively, we conclude that a factor say $l_0$ has a proper prefix (resp. suffix) starting with $v_0$ (resp.$r_0$) and concluding with $r_0$ (resp.$v_0$) which is a contradiction for the minimality of $k_0$. Hence $v_0 = r_0$. Conversely, to prove $u_0$ is a rich partial word, we have to verify that each prefix of $u_0$ has a uniuoccurrent palindromic factor. Consider the prefix of $u_0$ as $v_0$ and the palindromic prefix of $u_0$ with maximum length as $q_0$. Suppose $v_0$ is palindromic then $v_0 = q_0$ and thus $q_0$ is uniuoccurrent in $v_0$. Suppose $v_0$ is not palindromic, then let $p_0$ be the palindromic factor of $v_0$ with maximum length. If $q_0$ is not uniuoccurrent in $v_0$ then $v_0$ has a proper factor $r_0$ starting with $p_0$ and ending with $q_0$ where $p_0$ and $q_0$ are not factors of one another. Then $p_0$ is the palindromic factor of $r_0$ with maximum length. Similarly we can show that $q_0$ is the palindromic factor of $r_0$ with maximum length which contradicts our assumption. Hence $q_0$ is uniuoccurrent in $v_0$.

Theorem 1. If $u_0$ is a rich partial word over $\mathcal{A}_0$ and $u_0r_0$ has a uniuoccurrent palindromic factor $s_0$ such that $r_0 \in \mathcal{A}_0$ and $2|q_0| \geq |u_0r_0|$ then $u_0r_0$ is a rich partial word.

Proof. Let us assume that $q_0$ is the palindromic factor of $u_0r_0$ with maximum length. Suppose $q_0$ is not uniuoccurrent in $u_0r_0$ such as if $q_0$ has another occurrence in $u_0r_0$, then as $2|q_0| + 1 \geq |u_0r_0|$, the two occurrences overlap each other or separated from each other by maximum of one letter of $\mathcal{A}_0$. Thus both the occurrences form a palindromic factor of $u_0r_0$ such that they are strictly longer than $q_0$ which is a contradiction. Therefore $q_0$ is the uniuoccurrent palindromic factor of $u_0r_0$ such that $u_0r_0$ is rich. Hence the proof.

Theorem 2. If the rich partial word $u_0$ over $\mathcal{A}_0$ is the product of two rich palindromic factors $p_0$ and $q_0$ and satisfies the conditions:

\[(i)\] $|u_0| - 4 \leq 2|q_0|$

\[(ii)\] $|u_0| - 4 \leq 2|p_0|$

then the products $p_0q_0p_0$ and $q_0p_0q_0$ are also rich partial words.
Proof. Let us use contradiction. Consider the rich partial word $u_0 = p_0q_0$ satisfying the condition $|u_0| - 4 \leq 2|q_0|$ such that $p_0q_0p_0$ is not rich. Let $r_0 \in \mathbb{A}_0$, $s_0 \in \{\text{Alph}(u_0)\}$ with $r_0s_0\space as \space the \space prefix \space of \space p_0$ of minimum length such that $p_0q_0r_0s_0$ is not rich. Let $k_0$ be the pal$_s$ of $p_0q_0r_0s_0$ with maximum length. Then as $s_0r_0^nq_0r_0s_0$ is the suffix of $p_0q_0r_0s_0$, we have $|q_0| + 2r_0 + 2 \leq |k_0|$ which further infers that $|u_0| \leq |u_0| + 4|r_0| \leq 2|q_0| + 4|r_0| + 4 \leq |k_0|$. Thus by Theorem 2 we get $p_0q_0r_0s_0$ to be a rich partial word which contradicts our assumption. Therefore $p_0q_0p_0$ is a rich partial word only if $u_0 = p_0q_0$ is rich and $|u_0| - 4 \leq 2|q_0|$. Similarly we can prove that $q_0p_0q_0$ is a rich partial word.

Example 5. Let $u_0 = p_0q_0$ over $\mathbb{A}_0 = \{a,b\} \cup \{\varnothing\}$ be a rich partial word with rich palindromic factors $p_0 = ab\varnothing ba$ and $q_0 = b$. Also

\begin{align*}
(i) |u_0| - 4 &= 2|q_0| \\
(ii) |u_0| - 4 &= 2 < 2|p_0|
\end{align*}

Then the products $p_0q_0p_0 = ab\varnothing bab\varnothing ba$ and $q_0p_0q_0 = bab\varnothing bab$ are also rich partial words.

Theorem 4. For any non-empty rich partial word $u_0$ over $\mathbb{A}$, if $u_0u_0 \uparrow v_0u_0w_0$ for some rich partial words $v_0, w_0$ such that $v_0 = \lambda$ or $w_0 = \lambda$ then $u_0$ is primitive.

Proof. Let us assume that $u_0u_0 \uparrow v_0u_0w_0$ such that $v_0 = \lambda$ or $w_0 = \lambda$. Suppose to the contrary that $u_0$ is not primitive then a non-empty rich word $x$ exists such that $u_0 \subset x^m$ where $m \geq 2$ is an integer. But then $v_0u_0 \uparrow x^{m-1}u_0x$, and using our assumption we get $x^{m-1} = \lambda$ or $x = \lambda$, a contradiction. Therefore $u_0$ is a primitive rich partial word.

Example 6. Assume the rich partial words $u_0, v_0$ and $w_0$ over $\mathbb{A}_0 = \{a,b,c\} \cup \{\varnothing\}$ such that $u_0 = ac\varnothing cb$, $v_0 = \lambda$ and $w_0 = acc\varnothing cb$. Then $u_0$ is primitive since

$u_0w_0 = ac\varnothing cbac\varnothing cb \uparrow acc\varnothing cbacc\varnothing cb = xuy.$

Theorem 5. Let $u_0$ and $v_0$ be non-empty rich partial words. If $u_0$ and $v_0$ are conjugate, then a rich partial word $w_0$ exists such that $u_0w_0 \uparrow w_0u_0$. Also there exist rich partial words $p_0, q_0$ such that $u_0 \subset p_0q_0$, $v_0 \subset q_0p_0$ and $w_0 \subset p_0(q_0p_0)^m$ for some $m \geq 1$.

Proof. Let $u_0$ and $v_0$ be non-empty rich partial words. Suppose that $u_0$ and $v_0$ are conjugate and let $p_0q_0$ be rich partial words such that $u_0 \subset p_0q_0$ and $v_0 \subset q_0p_0$.

Then $u_0w_0 \subset p_0q_0w_0$ and $p_0v_0 \subset p_0q_0v_0$ and so for $w_0 = p_0$ we have $u_0w_0 \uparrow w_0u_0$.

Theorem 6. Let $u_0$ be a rich partial word over $\mathbb{A}_0$ and let $x$ and $y$ be two rich words over $\mathbb{A}$. If $u_0 \subset xy$ and $u_0 \subset xy$ then $xy = yx$.

Proof. To prove the theorem, we consider $|x| \leq |y|$. Let $y = x'y'$ such that $|x'| = |x|$ where $x'$ and $y'$ are rich words. Also let $u_0 = v_0w_0$ with $|v_0| = |w_0|$ where $v_0, w_0$ are rich partial words. Since $u_0 \subset xy$, we have $v_0w_0 \subset xy$ such that we get $v_0 \subset x$ and $w_0 \subset y$. Likewise $u_0 \subset xy$ implies that $v_0w_0 \subset yx$ which further implies that $v_0w_0 \subset x'y'y'$ such that we get $v_0 \subset x'$ and $w_0 \subset y'$. Since $u_0 = v_0w_0$ is a rich partial word, it has exactly one hole. Then the following two cases arise:

Case (i): If $v_0$ is a rich partial word with zero hole and $w_0$ is a rich partial word with one hole. Then $v_0 = x' = x$ and $w_0 \subset y'x$. Thus by induction process, $xy = y'x$ which follows that $xy = yx$.

Case (ii): If $v_0$ is a rich partial word with one hole and $w_0$ is a rich partial word with zero hole. Then $v_0 \subset x' = x$ and $w_0 = y'x = x'y' = y$. Then there exists two rich words $p$ and $q$ such that $x = pq$ and $x' = qp$ and $y' = (qp)mq$ for $m \geq 0$ where $x$ and $y'$ are conjugates to each other. Hence by induction process, $pq = qp$ which follows that $xy = yx$ since $v_0 \subset pq$ and $v_0 \subset qp$.

3.1 Rich Palindromic Partial Words

A rich partial word is closed by factors and also under the operations of reversal and palindromic closures. Palindromic partial words help in encoding and decoding the information contained in DNA strands. The palindromic defect of rich partial words is zero: Most of the rich partial words are also palindromic which is not a necessary condition. In this section, the rich palindromic partial words are to be analyzed and examined to find the periodicity of possible elements in the $\varnothing$ positions of the partial word sequence.

Definition 3. Rich partial words that are also palindromic are termed as rich palindromic partial words.

Example 7. Assume a partial word $u_0 = ab\varnothing aba$ with $|u_0| = 5$ over $\mathbb{A}_0 = \{a,b\} \cup \{\varnothing\}$. Here $u_0 \uparrow u_0^R$ and so $u_0$ is a palindromic partial word. The set of all distinct partial palindromic proper factors of $u_0$ are

\{ba\varnothing ab, a\varnothing aba, aba\varnothing a, a\varnothing a, \varnothing a, a, \varnothing ab\}.

Here the number of distinct partial palindromic proper factors are more than $|u_0|$. Hence $u_0$ is a rich palindromic partial word.

Theorem 7. Let $u_0$ be a rich palindromic partial word. Then $u_0u_0^R$ and $u_0^Ru_0^R$ are periodic partial words.

Proof. Let the rich partial word $u_0 \in \mathbb{A}_0^*$ be a palindromic. We have $u_0 \uparrow u_0^R$. By the notion of compatibility, a total word $x$ exists such that $u_0 \subset x$ and $u_0^R \subset x$. Hence by the law of multiplication, $u_0u_0^R \subset x \cdot x = x^2$. Thus $u_0u_0^R$ is periodic. Similarly it is easy to follow that $u_0^Ru_0^R$ is a periodic partial word.
Theorem 8. For a rich partial word \( u_\circ \in \mathcal{A}_\circ \), if \( u_\circ^m \) is a rich palindromic partial word for \( m > 0 \) then \( u_\circ \) is a rich palindromic partial word.

Proof. We prove it by induction hypothesis. For \( m = 1 \), the assertion is accurate. Let us assume that it is true for all \( l < m \) that is if \( u_\circ \) is a palindrome for all \( l \leq m - 1 \), then \( u_\circ \) is a palindrome. Now to prove it for \( m \), assume that \( u_\circ^m \) is a palindrome. We can write

\[
u_\circ^m = u_\circ^{m-1} u_\circ = u_\circ^* u_\circ^{m-1}.
\]

Now

\[
u_\circ^m = u_\circ u_\circ^{m-1} \uparrow (u_\circ^m)^R = (u_\circ^R)^m = (u_\circ^R)^m u_\circ^{m-1}.
\]

As \( |u_\circ| = |u_\circ^R| \) and \( u_\circ^{m-1} \uparrow (u_\circ^m)^R \), then by simplification we have \( u_\circ \uparrow u_\circ^m \). Hence \( u_\circ \) is a palindrome. \( \square \)

4 Partial Palindromic Proper Subwords of Rich Partial Words

The study of subsequences (or subwords) in partial words involves a number of combinatorial complexities. One of them is the detection of palindromes, or subwords that are symmetric when reversed, in partial words. Since the last two decades, there has been interest in researching the characteristics of palindromes. In this section, we prove that the maximum number of partial palindromic perfect subwords in a partial word relies on both the length and the number of distinct letters in the partial word.

Definition 4. A partial palindromic proper subword (or partial palindromic scattered proper subword) of a partial word \( u_\circ \) over the alphabet \( \mathcal{A}_\circ \) is a sequence that can be derived by deleting zero or more letters from it without altering the order of the remaining letters. The set of all non-empty partial palindromic proper subwords of \( u_\circ \) is denoted by \( \text{PPPS}(u_\circ) \).

Example 8. Consider a partial word \( u_\circ = aab \uparrow ba \) over \( \mathcal{A}_\circ = \{a,b\} \cup \{\Diamond\} \). The set of all distinct palindromic proper subwords of \( u_\circ \) are

\[
\{aa, bb, a\Diamond, \Diamond a, b\Diamond, \Diamond b, aab, a\Diamond a, ab\Diamond, b\Diamond b, \Diamond ba, aab\Diamond a, \}
\]

\[
\text{abba, a\Diamond ba, ab\Diamond ba, aab\Diamond a}\}.
\]

Among the above set, the set of all distinct partial palindromic proper subwords of \( u_\circ \) are

\[
\{a\Diamond, \Diamond a, b\Diamond, \Diamond b, aab, a\Diamond a, ab\Diamond, b\Diamond b, \Diamond ba, aab\Diamond a, a\Diamond ba, ab\Diamond ba, aab\Diamond a}\}.
\]

Theorem 9. Any rich partial word of length \( n \) with no three consecutive similar letters over \( \mathcal{A}_\circ = \{a,b\} \cup \{\Diamond\} \) has a partial palindromic proper subword of length at least \( \frac{2}{3} (n - 2) \).

Proof. Consider a rich partial word \( u_\circ = u_\circ[1 \ldots t] \) with no three consecutive similar letters over \( \mathcal{A}_\circ = \{a,b\} \cup \{\Diamond\} \). Let each \( u_\circ[i] \) be made up of only \( a \)s or only \( b \)s and let two consecutive partial words \( u_\circ[j] \) and \( u_\circ[j+1] \) consist of different letters. Then we have length of each \( u_\circ[j] \) as at most \( 2 \). Now assume \( t \) to be even. Then at least one letter from each pair \( u_\circ[j], u_\circ[t-j+1] \) together with all the letters from \( u_\circ[\frac{t+1}{2}] \) results in a partial palindromic subword. Thus we get a partial palindromic subword of length at least \( \frac{2}{3} (n - 2) \). Hence the proof. \( \square \)

Theorem 10. For a given rich partial word \( u_\circ \), \( |u_\circ| \leq |\text{PPPF}(u_\circ)| \leq |\text{PPPS}(u_\circ)| \).

Proof. It is clear from the notion of rich partial words that \( |u_\circ| \leq |\text{PPPF}(u_\circ)| \). Let \( u_\circ = u_\circ[1 \ldots n] \) be a rich partial word where \( u_\circ[i] \in \mathcal{A}_\circ \). Let the prefix of length \( t \) of \( u_\circ \) be \( v_\circ = u_\circ[1 \ldots t] \). We observe that on the concatenation of each \( u_\circ[i] \) to \( v_\circ[i-1] \), an additional partial palindromic perfect subword \( u_\circ^*_i \), where \( s = |v_\circ[i]| |u_\circ[i]| \) is formed. Hence, at least one partial palindromic perfect subword is always formed on the concatenation of each letter of \( u_\circ \). Thus \( |\text{PPPS}(u_\circ)| \geq |u_\circ| \). Also the set of all partial palindromic perfect factors is a subset of the set of all partial palindromic perfect subwords of \( u_\circ \). Therefore \( |\text{PPPF}(u_\circ)| \leq |\text{PPPS}(u_\circ)| \). Hence the result. \( \square \)

5 Conclusion

In this paper, we introduced rich partial words and studied the combinatorial properties. We also discussed the relation between partial palindromic perfect factors and partial palindromic perfect subwords of rich partial words.

References


