Properties of Variants of Lyndon Partial Words

R. Arulprakasam, Member, IAENG, R. Krishna Kumari, Member, IAENG,
K. Janaki, Member, IAENG, L. Jeyanthi, P. Madhusoodhanan, Member, IAENG *

Abstract—**Lyndon words have been extensively studied in different contexts of free Lie algebra and combinatorics. All Lyndon and Nyldon words are primitive and any primitive class of words contains an unique Lyndon and an unique Nyldon word. This property motivated the study of Lyndon partial word which is primitive but all primitive classes of partial words may not contain a Lyndon partial word. In this paper we introduce two variants of Lyndon partial words namely Nyldon and inverse Lyndon partial words which are constructed from the decreasing alphabetical order. We compare the properties of the variants with those of Lyndon partial words.**

Keywords:* *Partial words, Lyndon words, Nyldon words

1 Introduction

Lyndon words serve to be a useful tool for a variety of problems in combinatorics [1, 2, 9, 18]. There are many applications of Lyndon words in semigroups, pattern matching, representation theory of certain algebras and combinatorics such as they are used to describe the generators of the free Lie algebras. Lyndon words are used as a special case of Hall sets. All of these applications make use of the combinatorial properties of Lyndon words, in particular the factorization theorem. Their role in factorizing a string over an ordered alphabet was initially illustrated by Chen et.al [8]. The central result about Lyndon words is Lyndon factorization. Duval [14] presented a algorithm to derive a factorisation of strings over an ordered alphabet known as Lyndon factorization. A unique factorization theorem in terms of Lyndon trees for factoring a tree is proved in [20]. Influenced by Lyndon words, Grinberg [12] in a mathoverflow post defined Nyldon words as a variant of Lyndon words. Charlier, Philibert and Stipulanti [7] computed the Nyldon factorization of a word using an algorithm and made a comparative study between Nyldon and Lyndon words. In [5, 6] inverse Lyndon factorizations, a version of the Lyndon factorization, were introduced, and its properties were examined in order to explore its potential application in string queries. In [7] a novel proof of unique factorization into Nyldon words connected to Hall set theory was given, and Lazard process for producing Nyldon words was examined."

Partial words are nothing but words with holes over the alphabet and are considered in gene comparisons [11]. For instance, alignment of two DNA sequences which are genetic information carriers can be regarded as construction of two compatible partial words. DNA strands are viewed as finite words and are used to encode information in DNA computing. Partial words that indicate the locations of the missing symbols in a word might be used to expose information that was unseen or absent during encoding. The study of partial words was first conducted by Berstel and Boasson [3] and Blanchet Sadri [4] later expanded on their research. Both Lyndon and partial words have wide application in pattern matching.

*Manuscript received August 2, 2023; revised January 8, 2024.
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 (e-mail: r.aruljeeva@gmail.com).
R. Krishna Kumari is an Assistant Professor in Career Development Centre, College of Engineering and Technology, SRM Institute of Science and Technology, Kattankulathur, Chennai-603203, Tamilnadu, India (e-mail: krishrengan@gmail.com).
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).
L. Jeyanthi is an Assistant Professor in the Department of Mathematics, Panimalar Engineering College, Varadharajapuram, Poonamallee, Chennai-600123, Tamilnadu, India (e-mail: jeyanthi.l@gmail.com).
P. Madhusoodhanan is an Associate Professor and Head in Career Development Centre, College of Engineering and Technology, SRM Institute of Science and Technology, Kattankulathur, Chennai-603203, Tamilnadu, India (corresponding author e-mail: madhupms1973@gmail.com).

Motivated by the works on [1, 5, 6, 7, 10, 15, 16, 17, 19], here the concept of Lyndon partial words is used the two variants of Lyndon partial words namely inverse Lyndon and Nyldon partial words are established. All primitive conjugacy class of words contains exactly a single Lyndon and Nyldon word but this property is not satisfied in terms of Lyndon partial words and Nyldon partial words. There are equal count of Lyndon and Nyldon words of each length but count of Lyndon partial words and Nyldon partial words of each length are not equal. Lyndon partial words have many vital properties among which few analogues in Nyldon partial words, while many do not. Contrarily to Lyndon partial words which are constructed from the increasing alphabetical order, Nyldon partial words are constructed from the decreasing alphabetical order. Lyndon partial words are minimal in their conjugacy class but the latter are not maximal in their conjugacy class. But both Lyndon partial words and Nyldon partial words share a common property such as they have a unique factorization. In the case of variants of Lyndon partial words, both Nyldon and inverse Lyndon partial words are constructed from the decreasing alphabetical order and both are greater than any of their proper suffixes. Yet the class of Nyldon and inverse Lyndon partial words do not correspond to each other since they show major differences such as inverse Lyndon partial words are prefix-closed while the former are not. Also factorization is not unique in inverse Lyndon partial words. The paper has the following organization. We recall some basics in Section 2 and in Section 3 we introduce Nyldon partial words and study their properties. A comparison between the Lyndon and Nyldon partial words is established. We also introduce inverse Lyndon partial words and study their properties. Finally we conclude the paper in Section 4.

2 Preliminaries

Definition 2.1. *[18] The* alphabet Σ *is a finite set of symbols (or letters) and is non-empty.*

Definition 2.2. *[18] A* total word *or* string *is a series of letters over* Σ. *The set of all total words from* Σ *is denoted by Sigma∗ and* $\Sigma^+ = \Sigma^* \setminus \lambda$ *where* λ *is the identity.*

Definition 2.3. *[18] A language* L *is a subset of* Σ^*.

Definition 2.4. *[18] The total word* p *is a* sub word (or factor) *of* q *if* $q = xpy$ *where* x *and* y *are total words.* p *exists as a proper subword of* q *if* $xy \neq \epsilon$. *If* $x = \epsilon$ *(resp.* $y = \epsilon$,*), then* p *is a prefix (resp. suffix) of* q.

Definition 2.5. *[18] A rotation of the total word* $p^r = yx$ *of* $p = xy$ *is said to be non-trivial if both* x *and* y *are nonempty.*

Definition 2.6. *[14] An* ordered alphabet *is an alphabet with a total order so that comparisons of any two symbols from the alphabet can be computed in constant time. The* alphabetical order (lexicographical order) $\prec$ *on* $(\Sigma^*, \prec)$ *is defined by setting* $u \prec v$ *if any of the following conditions are met:*

1. u *is a proper prefix of* v,
2. *there exists words* x, y, z *(possibly empty) and elements* a *and* b *of* Σ *such that* $u = xay, v = xbz, a \prec b$.

The Table 1 *illustrates examples of alphabetical order of words over a given alphabet.*

Table 1: Illustration of alphabetical order of words over binary and tertiary alphabet

Alphabet	*Order*	*Alphabetical Order*
a, b	$a \prec b$	$aa \prec aab \prec aba \prec b$
a, b, c	$a \prec b \prec c$	$ab \prec abb \prec abbc \prec abc \prec acbc$

Definition 2.7. *[6] A* finite Lyndon word l *is non-empty and primitive such that it is less than all its rotations in the alphabetical order.*

Definition 2.8. *[6] A word* u *over* Σ^+ *is an inverse Lyndon word if* $u \succ q$ *for each non-empty proper suffix* q *of* u.

Definition 2.9. *[7] A* Nyldon word *is recursively defined as any primitive word of length at least three in the set* N *over* $\Sigma = \{a, b\}$ *with order* $\{a \prec b\}$ *such that it cannot factorized into a alphabetically non-decreasing sequence of shorter words of* N. N *is called as the finite Nyldon words set over* Σ.

Definition 2.10. *[4] The sequence containing a number of do not know symbols or holes denoted as* $\diamond$ *is termed as a* partial word. *A total word is a partial word with zero* *holes. Empty word is not a partial word. The symbol* $\Diamond$ *is not an element of* Σ *but a back-up symbol for the unknown letter.* $\Diamond$ *alone of any length cannot exist as a word. A partial word* p *of length* n *over* Σ *is a partial function* $p : \{0, 1, 2 \cdots, n-1\} \to \Sigma$*. For* $0 \leq i < n$*, if* $p(i)$ *is defined, then we say* $i \in D(p)$ *(the domain of* p*), otherwise* $i \in H(p)$ *(the set of holes). The positions of the holes in the partial words are represented using the following definition. The companion of* p*, denoted by* $p_\Diamond$ *is the total function* $p_\Diamond : \{0, 1, 2 \cdots, n-1\} \to \Sigma_\Diamond = \Sigma \cup \{\Diamond\}$ *defined by*

$$p_\Diamond(i) = \begin{cases} p(i) & \text{if } i \in D(p) \\ \diamond & \text{if } i \in H(p). \end{cases}$$

Definition 2.11. *[4] A partial word* $p_\Diamond$ *over* $\Sigma_\Diamond$ *is* primitive *if there does not exist any partial word* $q_\Diamond$ *over* $\Sigma_\Diamond$ *such that* $p_\Diamond = q_\Diamond^i$ *with* $i \geq 2$.

Definition 2.12. *[4] If* $p_\Diamond = xy$ *for some* x *and* y *over* $\Sigma_\Diamond$*, then* yx *is said to be the a rotation (or a conjugate) of* $p_\Diamond$.

Definition 2.13. *[4] A nonempty partial word* $p_\Diamond$ *over* $\Sigma_\Diamond$ *is unbordered if no nonempty partial words* $x_\Diamond, q_\Diamond, r_\Diamond$ *over* $\Sigma_\Diamond$ *exist such that* $p_\Diamond \subset x_\Diamond q_\Diamond$ *and* $p_\Diamond \subset r_\Diamond x_\Diamond$*. Unbordered partial words are primitive.*

3 Characteristics and Applications of Lyndon partial words

Lyndon partial words, also known as Lyndon partial words with gaps, represent a specialized and invaluable concept within the realm of DNA sequence analysis. They serve as a powerful tool for modeling and articulating DNA sequences, particularly in scenarios involving incomplete or uncertain data. The primary utility of Lyndon partial words comes to the forefront when dealing with DNA sequences that exhibit gaps or regions of missing information.

These partial words are crafted by transforming a given DNA sequence into a structured representation. This representation consists of a sequence of standard DNA base letters, namely A (adenine), C (cytosine), G (guanine), and T (thymine). Importantly, the unique characteristic of Lyndon partial words lies in their capacity to incorporate gaps or placeholder symbols, which stand in for segments of the sequence where information is either absent or uncertain. Lyndon partial words are particularly instrumental in the field of DNA sequencing due to their ability to capture and convey the nuances of real-world genetic data. When DNA sequencing is conducted, it is common to encounter regions where the sequence quality is compromised, leading to uncertainties or outright gaps in the data. Lyndon partial words elegantly address this challenge by providing a structured framework to account for these irregularities. By including gap symbols, often represented as '-' or $\Diamond$ or other designated symbols, researchers and bioinformaticians can meticulously document and manipulate sequences that may have segments with unclear or incomplete information. In essence, Lyndon partial words serve as a bridge between the idealized, complete DNA sequences and the complex, often imperfect data obtained through sequencing processes. They allow for a more faithful representation of biological reality, enabling researchers to work with and analyze genetic information in a manner that acknowledges and accommodates the uncertainties and gaps that are inherent to DNA sequencing.

A C G - T C - G A

Figure 1: Lyndon Partial Word

The following are key characteristics and applications of Lyndon partial words:

1. **Representation of Uncertainty**: In DNA sequencing, it's common to encounter regions with uncertain bases due to sequencing errors or low-quality data. Lyndon partial words allow for the representation of these uncertainties by using gap symbols in the sequence.

2. **Capturing Partial Information**: Lyndon partial words can capture partial information about a DNA sequence. Instead of completely discarding regions with gaps, they preserve the available data while indicating where gaps exist.

3. **Sequence Alignment**: In bioinformatics, Lyndon partial words can be used in sequence alignment algorithms, such as pairwise or multiple sequence alignment, where gaps need to be introduced to align sequences with

insertions or deletions.

4. **Error-Tolerant Matching**: When comparing sequences with errors, Lyndon partial words are helpful in performing error-tolerant matching or searching, as they consider gaps and uncertainties in the comparison.

5. **Phylogenetic Analysis**: Lyndon partial words can be applied in phylogenetic analysis to account for missing or uncertain data in DNA sequences when constructing phylogenetic trees.

6. **Fragment Assembly**: In DNA sequence assembly, which involves piecing together shorter DNA reads to reconstruct the full genome, Lyndon partial words can be used to represent partial sequences when dealing with gaps or ambiguous data.

These partial words are a valuable tool for DNA sequence analysis, especially in scenarios where data quality is a concern, or when dealing with ancient or degraded DNA samples where uncertainty is high. They provide a way to incorporate and work with incomplete or uncertain information within the context of DNA sequences.

Algorithm 1 Generate Lyndon Partial Words

```
function GenerateLyndonPartialWords(length, alphabet, partial_word)
    if length = 0 then
        word ← concatenate symbols in partial_word
        if IsLyndon(word) then
            Print word
        end ifreturn
    end if
    for symbol in alphabet do
        GenerateLyndonPartialWords(length − 1, alphabet, partial_word + [symbol])
        GenerateLyndonPartialWords(length − 1, alphabet, partial_word + ['*'])
    end for
end function
function IsLyndon(word)
    n ← length of word
    for i in 1 to n − 1 do
        if word[: i] > word[−i :] then
            return False
        end if
    end for
    return True
end function
```

4 Finite Nyldon and Inverse Lyndon Partial Words

Here we introduce and study the properties of finite Nyldon partial words with respect to the properties of Lyndon partial words. We take partial words over $\Sigma_\Diamond$ with a single hole into consideration throughout the work. The symbol $\Diamond$ is not an element of the ordered alphabet Σ but a back-up symbol for the unknown letter and is compatible or matches to any of the symbol in Σ.

In [1], the authors have defined that a primitive partial word is a partial Lyndon word iff it is minimal in its conjugate class with respect to alphabetical order by assuming the order of $\Diamond$ as $\{a \prec b \prec ....\Diamond\}$. The order of $\Diamond$ does not play a special role in the definition for studying properties of partial Lyndon words since the $\Diamond$ is considered as a letter with highest order which makes the definition of partial Lyndon words similar to that of Lyndon words. In the following definition of Lyndon partial word, the order of $\Diamond$ plays a special role in studying properties. Here we assume the order of $\Diamond$ as $a_1 \preceq \Diamond$ and $\Diamond \preceq a_k$ such that $\Diamond$ is compatible with all the elements of $\Sigma_k = \{a_1 \prec a_2 \prec ... \prec a_k\}, k \succ 1$.

Definition 4.1. *A finite Lyndon partial word $l_\Diamond$ belonging to the set of all Lyndon partial words $L_\Diamond$ over the ordered alphabet $\Sigma_\Diamond = \Sigma_k \bigcup \{\Diamond\} = \{a_1 \prec a_2 \prec ... \prec a_k\} \bigcup \{\Diamond\}, k > 1$ is a non-empty primitive partial word which is less than all its conjugates (rotations) with respect to the alphabetical order.*

In other words, a finite primitive partial word $l_\Diamond$ of length at least two is recursively defined as a Lyndon partial word if it cannot be factorized into an alphabetically non-increasing sequence of shorter Lyndon partial words.

Example 4.2. Consider the ordered alphabet $\Sigma_\Diamond = \{a \prec b\} \bigcup \{\Diamond\}$ with order $\{a \preceq \Diamond$ and $\Diamond \preceq b\}$. The finite Lyndon partial words set of length ≤ 4 are

$$\{a\Diamond, \Diamond b, aa\Diamond, a\Diamond b, \Diamond bb, aaa\Diamond, aa\Diamond b, a\Diamond bb, \Diamond bbb\}.$$

Remark 4.3. *Any Lyndon partial word is primitive but the converse may not be true. For instance $\Diamond abb$ is a primitive partial word but its conjugacy class does not contain a Lyndon partial word. This shows that the lexi-*

cographical order relation among Lyndon partial words is not always a total order relation due to the existence of hole.

Definition 4.4. *A* Nyldon partial word *over the ordered alphabet $\Sigma_\Diamond$ is recursively defined as any non-empty primitive word of length at least two in the set of words N such that it cannot be factorized into an alphabetically non-decreasing sequence of shorter words of $N_\Diamond$. Here $N_\Diamond$ is called as the finite set of Nyldon partial words over $\Sigma_\Diamond$.*

Any factorization $(n^1_\Diamond, ..., n^r_\Diamond)$ of $n_\Diamond$ into Nyldon partial words such that $n^1_\Diamond \preceq ...... \preceq n^r_\Diamond$ is called Nyldon factorization of $n_\Diamond \in N_\Diamond$

Table 2 shows the collection of Lyndon words along with Lyndon partial words and Nyldon words with Nyldon partial words of length atmost five over the ordered alphabet $\Sigma_\Diamond = \{a \prec b\} \bigcup \{\Diamond\}$.

Table 2: Illustration of Lyndon words along with Lyndon partial words and Nyldon words with Nyldon partial words

Length	*Lyndon words and Lyndon partial words*	*Nyldon words and Nyldon partial words*
1	a, b	a, b
2	$a\Diamond, ab, \Diamond b$	$b\Diamond, ba, \Diamond a$
3	$aa\Diamond, aab, a\Diamond b, abb, \Diamond bb$	$b\Diamond b, b\Diamond a, bab, baa$
4	$aaa\Diamond, aaab, aa\Diamond b, aabb, a\Diamond ab, a\Diamond bb, abbb, ab\Diamond b, \Diamond bbb$	$b\Diamond bb, b\Diamond ba, b\Diamond ab, b\Diamond aa, ba\Diamond a, babb, baab, baaa$
5	$aaaa\Diamond, aaaab, aaa\Diamond b, aabab, aab\Diamond b, aabbb, ababb, abbbb, \Diamond bbbb, aaabb, aa\Diamond ab, a\Diamond abb, aa\Diamond bb, a\Diamond bbb, ab\Diamond bb$	$baaaa, baaab, baaba, ba\Diamond ab, b\Diamond abb, babbb, b\Diamond bab, b\Diamond aab, b\Diamond aaa, baabb, babba, ba\Diamond aa, b\Diamond bbb, bab\Diamond a, b\Diamond bba$

Remark 4.5. *All Nyldon partial words are primitive but the converse may not be true.*

Example 4.6. *Let $y_\Diamond = bba\Diamond b$ be a primitive partial word over $\Sigma_\Diamond$. The conjugacy class of $y_\Diamond$ contains $\{bba\Diamond b, ba\Diamond bb, a\Diamond bbb, \Diamond bbba, bbba\Diamond\}$. All the words contained in the conjugacy class are primitive but none of them is a Nyldon partial word.*

Remark 4.7. *Nyldon partial words are not always alphabetically extremal among their rotation.*

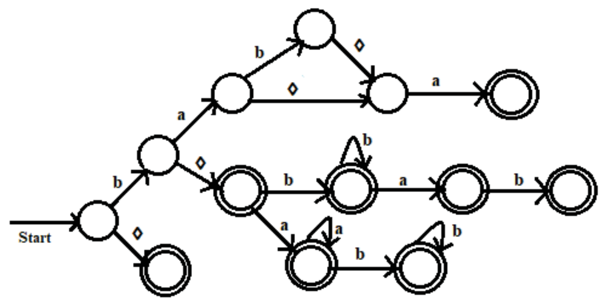

Figure 2: Estimated automaton for some Nyldon Partial Words in Table 2

Example 4.8. *Let the partial word $v_\Diamond = b\Diamond ab \in N_\Diamond$ over $\Sigma_\Diamond$. The conjugacy class of $v_\Diamond$ contains $\{b\Diamond ab, \Diamond abb, abb\Diamond, bb\Diamond a\}$. Among the conjugacy class $bb\Diamond a$ is the maximal but it is not a Nyldon partial word.*

Proposition 4.9. *Each Lyndon partial word $u_\Diamond$ over $\Sigma_\Diamond$ is unbordered but the converse is not true.*

Proof. Assume that $u_\Diamond$ has a non-overlapping border x. Then $u_\Diamond = xu_\Diamond x$. Let $igeq0$ be maximal such that $u_\Diamond = x^i u^1_\Diamond$. Then $u_\Diamond = x^{i+1}u^1_\Diamond x$. Then $x^{i+2}u^1_\Diamond$ is lexicographically smaller than $x^{i+1}u^1_\Diamond$, a contradiction with $u_\Diamond = x^{i+1}u^1_\Diamond x$ being Lyndon partial word. The following example illustrates that the converse is not true. □

Example 4.10. *For instance consider the unbordered partial word $u_\Diamond = bb\Diamond aa$ over $\Sigma_\Diamond$ which is also primitive. But $u_\Diamond$ is not a Lyndon partial word.*

Proposition 4.11. *Each Nyldon partial word $v_\Diamond$ is not unbordered.*

For instance $ba\Diamond b$ is a Nyldon partial word which is bordered.

Proposition 4.12. *For any partial word $v_\Diamond$ over $\Sigma^+_\Diamond$, a factorization $(n^1_\Diamond, ..., n^r_\Diamond)$ of $v_\Diamond$ as $N_\Diamond$ over the alphabet $\Sigma^+_\Diamond$ exists such that $n^1_\Diamond \preceq ...... \preceq n^r_\Diamond$.*

Theorem 4.13. *Each Nyldon partial word of length at least 2 over $\Sigma_\Diamond$ starts with ba or $b\Diamond$.*

Proof. We prove by contradiction. Consider $n_\Diamond = baq_\Diamond$ with $q_\Diamond \in \Sigma^+_\Diamond$. Let $(n^1_\Diamond, ..., n^r_\Diamond)$ be Nyldon partial factorization of $aq_\Diamond$. Then $n^1_\Diamond$ begins with a. Since b is a Nyldon word, $b \preceq a \preceq n_1$ and $rgeq1$. By Proposition

4.12, we get the Nyldon factorization of $n_\Diamond$ of length at least2 in the form $(b, n_\Diamond^1, ..., n_\Diamond^r)$. Hence $n_\Diamond$ is not Nyldon. □

Theorem 4.14. *[15] A proper subword cannot exist as prefix as well as suffix of a Lyndon partial word.*

Theorem 4.15. *A proper subword cannot exist as prefix as well as suffix of a Nyldon partial word.*

Proof. Consider $v_\Diamond$ to be a partial word over $\Sigma_\Diamond^+$. Assume a proper subword of $v_\Diamond$ say $p_\Diamond$ with $p_\Diamond$ exists as both prefix and suffix of $v_\Diamond$. The partial word $v_\Diamond = p_\Diamond q_\Diamond^i$ and $v_\Diamond = q_\Diamond^j p_\Diamond$ for some $q_\Diamond^i, q_\Diamond^j \in \Sigma_\Diamond^+$. Consider $v_\Diamond \in N_\Diamond$ over $\Sigma_\Diamond^+$. Then according to the notion of finite Nyldon partial words, $v_\Diamond \succ q_\Diamond^i p_\Diamond$ and $v_\Diamond \succ p_\Diamond q_\Diamond^j$. Then $q_\Diamond^j p_\Diamond \succ q_\Diamond^i p_\Diamond$ and $p_\Diamond q_\Diamond^i \succ p_\Diamond q_\Diamond^j$. This implies that $q_\Diamond^j \succ q_\Diamond^i$ and $q_\Diamond^i \succ q_\Diamond^j$ which is impossible. Thus a Nyldon partial word does not have a proper subword as both prefix and suffix. □

Theorem 4.16. *[15] Any partial word $u_\Diamond \in L_\Diamond$ if and only if $u_\Diamond \prec q_\Diamond$ for each proper suffix $q_\Diamond$ of $u_\Diamond$.*

Example 4.17. *Consider a Lyndon partial word $u_\Diamond = aaa\Diamond b = p_\Diamond q_\Diamond$ where $p_\Diamond = a$ and $q_\Diamond = aa\Diamond b$. Here $u_\Diamond = aaa\Diamond b \prec aa\Diamond b = q_\Diamond$ and $q_\Diamond$ is also a proper suffix of $u_\Diamond$.*

Theorem 4.18. *Any partial word $v_\Diamond \in N_\Diamond$ if $v_\Diamond \succ q_\Diamond$ for each proper suffix $q_\Diamond$ of $v_\Diamond$ but the converse is not true.*

Proof. Consider $v_\Diamond = p_\Diamond q_\Diamond$ to be a finite Nyldon partial word such that $p_\Diamond$ and $q_\Diamond$ are non-empty. Let $q_\Diamond$ be a proper suffix of $v_\Diamond = p_\Diamond q_\Diamond$. Then $q_\Diamond p_\Diamond \prec p_\Diamond q_\Diamond$. By Theorem 4.15, $q_\Diamond$ is not a proper prefix of $p_\Diamond q_\Diamond$. Thus we get $q_\Diamond \prec v_\Diamond$ for each proper suffix $q_\Diamond$. The converse is not true in the case of Nyldon partial words. For instance consider the partial word $u_\Diamond = ba\Diamond bab$. The Nyldon proper suffixes of u is (b, bab) but u is not a Nyldon partial word. □

Theorem 4.19. *Any partial word $v_\Diamond$ over $\Sigma_\Diamond^+$ such that $v_\Diamond = n_\Diamond^1 .....n_\Diamond^r$ and $n_\Diamond^1, ..., n_\Diamond^r \in N_\Diamond$ has*
(i) $n_\Diamond^r$ as the maximum suffix of $v_\Diamond$,
(ii) $n_\Diamond^r$ as the longest suffix of $v_\Diamond$.

Proof. Consider the partial word $v_\Diamond$ over $\Sigma_\Diamond^+$
(i) Let $q_\Diamond$ be a suffix of $v_\Diamond$ in the form

$$q_\Diamond = n_\Diamond^j n_\Diamond^{i+1} ......n_\Diamond^r$$

such that $n_\Diamond^j$ is non-empty and it is the suffix of $n_\Diamond^i$ with $igeqr$. Then we have $n_\Diamond^j \preceq n_\Diamond^i$. This implies $n_\Diamond^j \preceq n_\Diamond^i ...... \preceq n_\Diamond^r$. Therefore $n_\Diamond^r$ as the maximum suffix of $v_\Diamond$.
(ii) Assume that $q_\Diamond$ is the longest suffix of $v_\Diamond \in N_\Diamond$. Then $i > m$ and $n_\Diamond^j \succ q_\Diamond$. This shows $n_\Diamond^r \succ q_\Diamond$ and $q_\Diamond \notin N_\Diamond$. Therefore $n_\Diamond^r$ as the longest suffix of $v_\Diamond$. □

Theorem 4.20. *A Nyldon factorization of a Nyldon partial word always exists.*

Proof. Let $n^\Diamond$ be a Nyldon partial word over the ordered alphabet $\Sigma^\Diamond$. We want to show that there exists a Nyldon factorization of $n^\Diamond$, i.e., a factorization $(n_1^\Diamond, n_2^\Diamond, \ldots, n_r^\Diamond)$ such that $n_1^\Diamond \preceq n_2^\Diamond \preceq \ldots \preceq n_r^\Diamond$.

Base Case ($|n^\Diamond| = 2$: If $n^\Diamond$ is a primitive word of length two, it is, by definition, a Nyldon partial word. The Nyldon factorization is trivially $(n^\Diamond)$ itself, and the theorem holds.

Inductive Step: Assume the theorem holds for all Nyldon partial words of length less than k. Now, consider a Nyldon partial word $n^\Diamond$ of length $k \geq 3$. Since $n^\Diamond$ is a Nyldon partial word, it cannot be factorized into an alphabetically non-decreasing sequence of shorter words of $N^\Diamond$. Therefore, $n^\Diamond$ itself is a Nyldon partial word.

Consider the factorization $(n_1^\Diamond, n_2^\Diamond, \ldots, n_r^\Diamond)$ where $n_i^\Diamond$ is the maximal Nyldon partial word starting at the i-th position in $n^\Diamond$. By the induction hypothesis, each $n_i^\Diamond$ has a Nyldon factorization. Since $n_i^\Diamond$ is maximal, it cannot be further factorized into shorter Nyldon partial words. Therefore, we have a Nyldon factorization of $n^\Diamond$ as desired.

□

Theorem 4.21. *Every factor of a Nyldon partial word is also a Nyldon partial word.*

Proof. Let $n^\Diamond$ be a Nyldon partial word. We want to show that every factor of $n^\Diamond$ is also a Nyldon partial word.

We will prove this by contradiction. Assume there exists a factor $x^\Diamond$ of $n^\Diamond$ such that $x^\Diamond$ is not a Nyldon partial word. Then, $x^\Diamond$ can be factorized into an alphabetically non-decreasing sequence of shorter words of $N^\Diamond$. However, $x^\Diamond$ is also a factor of $n^\Diamond$, which is a Nyldon partial word, and it cannot be factorized in such a way. This contradiction implies that our assumption was incorrect. Therefore, every factor of $n^\Diamond$ is also a Nyldon partial word.

□

Theorem 4.22. ***Factorization Theorem*** *Any partial word $u_\Diamond$ over $\Sigma_\Diamond^+$ can be uniquely represented as $u_\Diamond = l_\Diamond^1.....l_\Diamond^r$ with $l_\Diamond^1 \succeq ...... \succeq l_\Diamond^r$ where $l_\Diamond^1, ..., l_\Diamond^r \in L_\Diamond$.*

Theorem 4.23. *Any partial word $v_\Diamond$ over $\Sigma_\Diamond^+$ can be uniquely represented as $v_\Diamond = n_\Diamond^1.....n_\Diamond^r$ with $n_\Diamond^1 \preceq ...... \preceq n_\Diamond^r$ where $n_\Diamond^1, ..., n_\Diamond^r \in N_\Diamond$.*

Proof. Proposition 4.12 shows the existence of factorization. We prove the uniqueness of factorization by induction. Consider $|v_\Diamond|$ of minimal length 2. All finite words shorter than $v_\Diamond$ admits a unique Nyldon factorization. Assume $v_\Diamond$ as a non-Nyldon partial word. Let $(n_\Diamond^1, ..., n_\Diamond^r)$ be the factorization of $v_\Diamond$. By Theorem 4.19, $n_\Diamond^r$ is the longest suffix of the Nyldon partial word $v_\Diamond$ such that by induction $v_\Diamond$ determines the subwords $n_\Diamond^1, ..., n_\Diamond^r$. Thus there does not exist another factorization of $v_\Diamond$ other than $n_\Diamond^1, ..., n_\Diamond^r$. □

Definition 4.24. *A partial word $u_\Diamond$ over $\Sigma_\Diamond$ is an inverse Lyndon partial word if $u_\Diamond \succ q_\Diamond$ for each non-empty proper suffix $q_\Diamond$ of $u_\Diamond$.*

Example 4.25. *The partial words $bb\Diamond a, baaa\Diamond, bb\Diamond ba$ are inverse Lyndon partial words on $\Sigma_\Diamond$.*

Theorem 4.26. *Every proper suffix of an inverse Lyndon partial word is also an inverse Lyndon partial word.*

Proof. Let $u\Diamond$ be an inverse Lyndon partial word. We need to show that every proper suffix of $u\Diamond$ is also an inverse Lyndon partial word.

By the definition of an inverse Lyndon partial word, $u\Diamond \succ q\diamond$ for each non-empty proper suffix $q\Diamond$ of $u\Diamond$. Since $q\Diamond$ is a proper suffix of $u\Diamond$, the last symbol of $q\Diamond$ is the last symbol of $u\Diamond$. Now, consider any proper suffix $q\Diamond$ of $u\Diamond$. Since $u\Diamond$ is an inverse Lyndon partial word, $u\Diamond \succ q\Diamond$, and this relationship holds for every proper suffix $q\Diamond$. Therefore, every proper suffix of $u\Diamond$ is also an inverse Lyndon partial word.

□

Algorithm 2 Generate Inverse Lyndon Partial Words

```
function GenerateInverseLyndonPartialWords(length, alphabet, partial_word)
    if length = 0 then
        word ← concatenate symbols in partial_word
        if IsLyndon(word) then
            Print word^-1    ▷ Print the inverse of the Lyndon word
        end ifreturn
    end if
    for symbol in alphabet do
        GenerateInverseLyndonPartialWords(length − 1, alphabet, partial_word + [symbol])
        GenerateInverseLyndonPartialWords(length − 1, alphabet, partial_word + ['*'])
    end for
end function
function IsLyndon(word)
    n ← length of word
    for i in 1 to n − 1 do
        if word[: i] > word[−i :] then
            return False
        end if
    end for
    return True
end function
```

Theorem 4.27. *A concatenation of inverse Lyndon partial words is an inverse Lyndon partial word if and only if the last symbol of each component is greater than the first symbol of the next component.*

Proof. Let $u_1\Diamond, u_2\Diamond, \ldots, u_k\Diamond$ be inverse Lyndon partial words such that $u_i\Diamond \succ u_{i+1}\Diamond$ for $1 \leq i < k$.

Assume that $u = u_1\Diamond u_2\Diamond \ldots u_k\Diamond$ is an inverse Lyndon partial word. We need to show that $u_i\Diamond \succ q\Diamond$ for every non-empty proper suffix $q\Diamond$ of $u_i\Diamond$ and every i (where $1 \leq i \leq k$). Consider $i = 1$. For any proper suffix $q\Diamond$ of $u_1\Diamond$, $q\Diamond$ is also a proper suffix of u since $u_1\Diamond$ is the first component of u. Therefore, $u \succ q\Diamond$. Now, consider $i > 1$. Since $u_{i-1}\Diamond \succ u_i\diamond$, any proper suffix $q\Diamond$ of $u_i\Diamond$ is also a proper suffix of $u_{i-1}\Diamond u_i\Diamond \ldots u_k\Diamond = u$. Therefore, $u \succ q\Diamond$.

Conversely, assume that $u_i\Diamond \succ q\Diamond$ for every non-empty proper suffix $q\Diamond$ of $u_i\Diamond$ and every i (where $1 \le i \le k$). We need to show that $u = u_1\Diamond u_2\Diamond \ldots u_k\Diamond$ is an inverse Lyndon partial word. Consider any proper suffix $q\Diamond$ of u. It must be a proper suffix of one of the components $u_i\Diamond$. Since $u_i\Diamond \succ q\Diamond$, it follows that $u \succ q\Diamond$. Therefore, u is an inverse Lyndon partial word.

□

Proposition 4.28. *If $u_\Diamond$ over $\Sigma_\Diamond^+$ is not an inverse Lyndon partial word, then a non-empty proper suffix $q_\Diamond$ of $u_\Diamond$ exists such that $u_\Diamond$ is not a proper prefix of $q_\Diamond(u_\Diamond \prec\prec q_\Diamond)$.*

Proof. Let $u_\Diamond$ be a non-inverse Lyndon partial word over the alphabet Σ^+, and let $v_\Diamond$ be the longest proper prefix of $u_\Diamond$ that is also a Lyndon partial word. Since $u_\Diamond$ is not an inverse Lyndon partial word, $v_\Diamond$ is a proper prefix of $u_\Diamond$.

Now, let $q_\Diamond$ be the shortest non-empty proper suffix of $u_\Diamond$ such that $u_\Diamond$ is a proper prefix of $q_\Diamond(u_\Diamond \prec\prec q_\Diamond$. Since $v_\Diamond$ is the longest proper prefix of $u_\Diamond$ that is also a Lyndon partial word, $q_\Diamond$ cannot be equal to $u_\Diamond$, and therefore, $q_\Diamond$ is a non-empty proper suffix of $u^\diamond$. □

Theorem 4.29. *Any non-empty prefix of an inverse Lyndon partial word is an inverse Lyndon partial word.*

Proof. Let us prove by contradiction. Assume $u_\Diamond$ over $\Sigma_\Diamond^+$ to be an inverse Lyndon partial word of the form $u_\Diamond = p_\Diamond q_\Diamond$ where $p_\Diamond$ is a non-empty proper prefix of $u_\Diamond$ such that it is not an inverse Lyndon partial word. By Proposition 4.28, a non-empty proper suffix of $p_\Diamond$ say $r_\Diamond$ exists such that $p_\Diamond \prec\prec r_\Diamond$. For $q_\Diamond \in \Sigma_\Diamond^+$, if $p_\Diamond$ such that $p_\Diamond \prec\prec r_\Diamond$ then $p_\Diamond q_\Diamond \prec\prec r_\Diamond q_\Diamond$. Thus $u_\Diamond \prec\prec r_\Diamond q_\Diamond$. This shows that $u_\Diamond$ is smaller than its non-empty proper suffix $r_\Diamond q_\Diamond$ which contradicts the assumption. □

Theorem 4.30. *For any partial word $u_\Diamond$ over $\Sigma_\Diamond^+$, there exists a non-unique factorization $(s_\Diamond^1, \ldots\ldots, s_\Diamond^r)$ of $u_\Diamond$ into inverse Lyndon partial words such that $u_\Diamond = (s_\Diamond^1 \ldots\ldots s_\Diamond^r)$ and $s_\Diamond^1 \prec\prec \ldots. \prec\prec s_\Diamond^r$.*

Proof. Unlike Lyndon partial words and Nyldon partial words, factorization of inverse Lyndon partial words is not unique. For instance, consider the inverse Lyndon partial word $u_\Diamond = baabb\Diamond abbbaa$ over $\Sigma_\Diamond$.We get three sequences $(baa, bb\Diamond a, bbbaa)$, $(baabb\Diamond a, bbbaa)$ and $(baab, b\Diamond abbbaa)$ as factorizations of the inverse Lyndon partial word $u_\Diamond$. Thus $u_\Diamond = (baa, bb\Diamond a, bbbaa) = (baabb\Diamond a, bbbaa) = (baab, b\Diamond abbbaa)$. Here baa,$bb\Diamond a$, $bbbaa$, $baabb\Diamond a$, $baab$, $b\Diamond abbbaa$ are all inverse Lyndon partial words. Further $baa \prec\prec bb\Diamond a \prec\prec bbbaa$, $baabb\Diamond a \prec\prec bbbaa$ and $baab \prec\prec b\Diamond abbbaa$. □

Table 3: Characteristics of Inverse Lyndon Words

Characteristic	Description
Lexicographic Order	Inverse Lyndon words are arranged in non-increasing lexicographic order.
Factorization	Every factor of an inverse Lyndon word is also an inverse Lyndon word.
Length Property	An inverse Lyndon word cannot be the prefix of a longer inverse Lyndon word.
Conjugacy Class	Inverse Lyndon words in the same conjugacy class are cyclic shifts of each other.
Lyndon Factorization	Every inverse Lyndon word has a Lyndon factorization.
Generation	Inverse Lyndon words can be generated iteratively using a specific algorithm.
Uniqueness	Inverse Lyndon words are unique representatives of their conjugacy classes.
Subword Complexity	The subword complexity of inverse Lyndon words is known and can be calculated efficiently.
Applications	Used in various applications, such as data compression, genomics, and coding theory.

5 Conclusion and Future Work

In this paper, we studied properties of two distinct variants of Lyndon partial words defined over a binary alphabet. These variants are carefully examined and their properties are established. Moreover, as we look ahead to future research, we turn our attention to the study of free monoid morphisms that possess the remarkable quality of preserving finite Lyndon partial words. This investigation promises to shed light on the profound interplay between structural transformations and Lyndon partial words, a relationship that holds significant implications in various areas of computer science and mathematical formalism.

A key aspect of the future exploration involves

- Study free monoid morphisms that preserve finite Lyndon partial words.
- Investigate whether these morphisms maintain or disrupt the lexicographic order.
- Explore the interplay between structural transformations and Lyndon partial words.
- Examine the impact of these investigations on computer science and mathematical formalism.
- Extend our understanding of Lyndon partial words and their applications in diverse domains.

References

[1] Nayak A.C. and Kapoor K.,(2015), On the Language of Primitive Partial Words, International Conference on Language and Automata Theory and Applications, LNCS, 8977, pp. 436–455.

[2] Barcelo H.,(1990), On the action of the symmetric group on the free Lie algebra and the partition lattice, Journal of Combinatorial Theory, Series A, 55, pp. 93–129.

[3] Berstel J. and Boasson L.,(1999), Partial Words and a Theorem of Fine and Wilf, Theoret. Comput.Sci., 218, pp. 135-141.

[4] Blanchet-Sadri F., (2005), Primitive Partial Words, Discrete Applied Mathematics, 148, pp. 195–213.

[5] Bonizzoni P.,De Felicem C.,Zaccagnino R. and Zizza R., (2018), Inverse Lyndon Words and Inverse Lyndon Factorizations of Words, Adv. Appl. Math.101, pp. 281–319.

[6] Bonizzoni P., De Felice C., Zaccagnino R. and Zizza R.,(2020), Lyndon Words versus Inverse Lyndon Words: Queries on Suffixes and Bordered Words, A. Leporati et al. (Eds.): LATA 2020, LNCS 12038, pp. 385–396.

[7] Charlier E., Philibert M., and Stipulanti M., Nyldon Words,(2019), J. Combin. Theory, Ser. A, 167, pp. 60–90.

[8] Chen K.T., Fox R.H., Lyndon R.C.,(1958), Free differential calculus, IV. The quotient groups of the lower central series, Ann. of Math., 68 (1), pp. 81–95.

[9] Christophe Reutenauer,(1993), Free Lie Algebras, London Mathematical Society Monographs. New Series. The Clarendon Press, Oxford University Press, New York, 7.

[10] Garg S.,(2021), New Results on Nyldon Words Derived Using an Algorithm from Hall Set Theory, Advances in Applied Mathematics, 131(102249), pp. 1–31.

[11] Fischer M.J. and Paterson M.S.,(1974), String Matching and other Products, In: Karp, R.M. (ed.) Complexity of Computation. SIAM-AMS Proceedings, pp. 113–125.

[12] Grinberg D.,(2014), Nyldon words: Understanding a Class of Words Factorizing the Free Monoid Increasingly, https://mathoverflow.net/questions/187451/.

[13] Janaki K and Arulprakasam R.,(2022), Tiling Systems and Domino Systems for Partial Array Languages, IAENG International Journal of Applied Mathematics, 52(2) , pp. 418–425.

[14] Jean-Pierre Duval,(1983), Factorizing Words over an Ordered Alphabet, J. Algorithms, 4(4), pp. 363–381.

[15] Krishna Kumari R., Arulprakasam R. and Dare V.R.,(2020), Language of Lyndon Partial Words, International Journal of Mathematics and Computer Science, 15, pp. 1173–1177.

[16] Krishna Kumari R and Arulprakasam R.,(2023), Factors and Subwords of Rich Partial Words, IAENG International Journal of Applied Mathematics, 53(2) , pp. 534–539.

[17] Krishna Kumari R, Jeyanthi L, Arulprakasam R, Janaki K and Madhusoodhanan P.,(2023), Properties of Cover and Seed of Partial Words, IAENG International Journal of Applied Mathematics, 53(3), pp. 1132–1137.

[18] Lothaire M.,(1983), Combinatorics on Words, Cambridge University Press.

[19] Sasikala K, Rajkumar Dare V and Thomas D.G., (2007), Learning of Partial Languages, Engineering Letters, vol.14(2), pp.72–80.

[20] Subramanian K.G., Rani Siromoney and Lisa Mathew,(1992), Lyndon Trees, Theoret. Comput. Sci., 106, pp. 373–383.