Partial words and the critical factorization theorem

F. Blanchet-Sadri*, S. Duncan

Department of Mathematical Sciences, University of North Carolina, P.O. Box 26170, Greensboro, NC 27402-6170, USA

Received 29 July 2004
Communicated by Andrew M. Odlyzko
Available online 30 December 2004

Abstract

The study of combinatorics on words, or finite sequences of symbols from a finite alphabet, finds applications in several areas of biology, computer science, mathematics, and physics. Molecular biology, in particular, has stimulated considerable interest in the study of combinatorics on partial words that are sequences that may have a number of “do not know” symbols also called “holes”. This paper is devoted to a fundamental result on periods of words, the critical factorization theorem, which states that the period of a word is always locally detectable in at least one position of the word resulting in a corresponding critical factorization. Here, we describe precisely the class of partial words \( w \) with one hole for which the weak period is locally detectable in at least one position of \( w \). Our proof provides an algorithm which computes a critical factorization when one exists. A World Wide Web server interface at http://www.uncg.edu/mat/cft/ has been established for automated use of the program.

© 2004 Elsevier Inc. All rights reserved.

Keywords: Word; Partial word; Period; Weak period; Local period

1. Introduction

Words, or finite sequences of symbols from a finite alphabet, are natural objects in several areas of biology, computer science, mathematics, and physics. Molecular biology, in particular, has stimulated considerable interest in the study of the combinatorial problems
on sequences that may have a number of “do not know” symbols. Such sequences are referred to as partial words and appear, for instance, when genes or proteins are compared. Another area of current interest for the study of the combinatorics on partial words is data communication where some information may be missing, lost, or unknown. While a word can be described by a total function, a partial word can be described by a partial function. More precisely, a partial word of length  \( n \) over a finite alphabet \( A \) is a partial function from \( \{0, \ldots, n - 1\} \) into \( A \). Elements of \( \{0, \ldots, n - 1\} \) without an image are called holes. A word is just a partial word without holes. We refer the reader to [1,3–9] for recent results in combinatorics on partial words.

The basic concept of this paper is that of period of a word. In the case of partial words, there are two notions: one is that of period, the other is that of weak period. There are many fundamental periodicity results on words. Among them are:

- The well-known result of Fine and Wilf [22] which intuitively determines how far two periodic events have to match in order to guarantee a common period. This result was extended to partial words with one hole by Berstel and Boasson [1], to partial words with two or three holes by Blanchet-Sadri and Hegstrom [8], and to partial words with an arbitrary number of holes by Blanchet-Sadri [5].
- The well known and unexpected result of Guibas and Odlyzko [24] which states that the set of all periods of a word is independent of the alphabet size (a new proof of this result appears in [26]). Guibas and Odlyzko’s result states that for every word \( u \), there exists a word \( v \) over the alphabet \( \{0, 1\} \) such that \( u \) and \( v \) have the same length and the same set of periods. In [7], Blanchet-Sadri and Chriscoe extend Guibas and Odlyzko’s result to partial words with one hole. There, the authors show that for every partial word \( u \) with one hole, there exists a partial word \( v \) over the alphabet \( \{0, 1\} \) with at most one hole such that \( u \) and \( v \) have the same length, the same set of periods, and the same set of weak periods.
- The well-known critical factorization theorem, of which several versions exist [13,14,19,29,30], states that the period of a word \( u \) is always locally detectable in at least one position of \( u \) resulting in a corresponding critical factorization. In this paper, we extend the critical factorization theorem to partial words with one hole.

The study of periods in words finds applications in different areas such as data compression [15,35,35], theory of codes [2], computational biology [11,25,31], string searching algorithm design [10,16,18,23,28], and serial data communication systems [12]. The critical factorization theorem, in particular, has found several applications which include the design of efficient string matching algorithms [16] and the design of efficient approximation algorithms for the shortest superstring problem [11]. Other important applications of the critical factorization theorem are found in [27,29]. A periodicity theorem on words that has strong analogies with the critical factorization theorem and some applications are proved in [33].

The contents of our paper are summarized as follows. In Section 2, we fix our terminology. In particular, we discuss compatibility and conjugacy on partial words. In Section 3, we define two total orderings \( \preceq_l \) and \( \preceq_r \) on partial words and state some of their properties. The ordering \( \preceq_l \) is simply the lexicographic ordering related to a fixed total ordering on the alphabet, and \( \preceq_r \) is obtained from \( \preceq_l \) by reversing the order of letters in the alphabet.
Section 4 is devoted to the critical factorization theorem on words which relates \textit{periods} and \textit{local periods}. Informally, a local period at a given position of a word \( u \) is the length of a square around that position. It turns out that a \textit{critical factorization} of \( u \) (for which the minimal local period is equal to the minimal period of \( u \)) can be computed very efficiently in time linear in the length of \( u \) [16]. Indeed, a critical factorization can be obtained from the computation of the maximal suffixes of \( u \) with respect to \( \leq_l \) and \( \leq_r \). Sections 5 and 6 are devoted to our critical factorization theorem on partial words with one hole which relates \textit{weak periods} and \textit{local periods}. First, in Section 5, we note that a class of partial words with one hole do not admit any critical factorization (a factorization is critical when the minimal local period is equal to the minimal weak period of the partial word). We then give our first version of the critical factorization theorem for the class of the so-called \textit{nonspecial} partial words (see Theorem 2). Finally, in Section 6, we refine Theorem 2 (see Theorem 3). There, we describe precisely the class of partial words with one hole that do not admit critical factorizations. Our proof leads to an efficient algorithm which, given a partial word with one hole, outputs a critical factorization when one exists or outputs “no such factorization exists”. Our algorithm refines the method based on the maximal suffixes with respect to the orderings \( \leq_l \) and \( \leq_r \).

2. Preliminaries

In this section, we recall some basic notions on partial words that are used in this paper.

Let \( A \) be a finite \textit{alphabet}. Elements of \( A \) are called \textit{letters}, and finite sequences of letters of \( A \) are called \textit{words} over \( A \). In particular, the \textit{empty word}, which is denoted by \( \varepsilon \), is the sequence with no letters. The set of all words (respectively, all nonempty words) over \( A \) is denoted by \( A^* \) (respectively, \( A^+ \)). It is a monoid (respectively, semigroup) under the operation of \textit{concatenation} of words. Moreover, each word has a unique representation as a concatenation of letters, so that \( A^* \) and \( A^+ \) are \textit{free}, referred to as the \textit{free monoid} and \textit{free semigroup} generated by \( A \).

A word of length \( n \) over \( A \) can be defined by a total function \( u : \{0, \ldots, n − 1\} \rightarrow A \) and is usually represented as \( u = a_0a_1 \ldots a_{n−1} \) with \( a_i \in A \). A \textit{partial word} of length \( n \) over \( A \) is a partial function \( u : \{0, \ldots, n − 1\} \rightarrow A \). For \( 0 \leq i \leq n \), if \( u(i) \) is defined, then we say that \( i \) belongs to the \textit{domain} of \( u \) (denoted by \( i \in D(u) \)), otherwise we say that \( i \) belongs to the \textit{set of holes} of \( u \) (denoted by \( i \in H(u) \)). A word over \( A \) is a partial word over \( A \) with an empty set of holes (we sometimes refer to words as \textit{full} words). For any partial word \( u \) over \( A \), \( |u| \) denotes its length. Clearly, \( |\varepsilon| = 0 \). We denote by \( W_0(A) \) the set \( A^* \), and for \( i \geq 1 \), by \( W_i(A) \) the set of partial words over \( A \) with at most \( i \) holes. We put \( W(A) = \bigcup_{i \geq 0} W_i(A) \), the set of all partial words over \( A \) with an arbitrary number of holes.

If \( u \) is a partial word of length \( n \) over \( A \), then the \textit{companion} of \( u \) (denoted by \( u_\diamond \)) is the total function \( u_\diamond : \{0, \ldots, n − 1\} \rightarrow A \cup \{\diamond\} \) defined by

\[
u_\diamond(i) = \begin{cases} 
u(i) & \text{if } i \in D(u), \\ \diamond & \text{otherwise.} \end{cases}
\]

The symbol \( \diamond \notin A \) is viewed as a “do not know” symbol. The word \( u_\diamond = abbc\diamond bc\diamond ac \) is the companion of the partial word \( u \) of length 10 where \( D(u) = \{0, 1, 2, 3, 5, 6, 8, 9\} \) and
$H(u) = \{4, 7\}$. The bijectivity of the map $u \mapsto u_\diamond$ allows us to define for partial words concepts such as concatenation and powers in a trivial way. More specifically, for partial words $u, v$, the concatenation of $u$ and $v$ is defined by $(uv)_\diamond = u_\diamond v_\diamond$, and the $i$-power of $u$ is defined by $(u^i)_\diamond = (u_\diamond)^i$ where $(u_\diamond)^0 = \varepsilon$ and $(u_\diamond)^i = u_\diamond (u_\diamond)^{i-1}$. The set $W(A)$ is a monoid under the concatenation of partial words ($\varepsilon$ serves as identity). For a subset $X$ of $W(A)$, we use the notation $|X|$ for the cardinality of $X$.

For a pair $u, v$ of partial words, we define: $u$ is a prefix of $v$, if there exists a partial word $x$ such that $v = ux$; $u$ is a suffix of $v$, if there exists a partial word $x$ such that $v = xu$; and $u$ is a factor of $v$, if there exist partial words $x$ and $y$ such that $v = xy$. When the trivial cases $u = \varepsilon$ or $u = v$ are excluded, the factor $u$ is called proper. Two partial words $u$ and $v$ always have a unique maximal common prefix denoted by $u \wedge v$. For a subset $X$ of $W(A)$, we denote by $F(X)$ the set of factors of elements in $X$. More specifically,

$$F(X) = \{ u \mid u \in W(A) \text{ and there exist } x, y \in W(A) \text{ such that } xuy \in X \}.$$ 

We similarly denote by $P(X)$ (respectively, $S(X)$) the set of prefixes (respectively, suffixes) of elements in $X$. In case $X$ is the singleton $\{v\}$, we abbreviate $F(X)$, $P(X)$ and $S(X)$ respectively by $F(v)$, $P(v)$ and $S(v)$.

A period of a partial word $u$ is a positive integer $p$ such that $u(i) = u(j)$ whenever $i, j \in D(u)$ and $i \equiv j \mod p$. In such a case, we call $u$ $p$-periodic. Similarly, a weak period of $u$ is a positive integer $p$ such that $u(i) = u(i + p)$ whenever $i, i + p \in D(u)$. In such a case, we call $u$ weakly $p$-periodic. The partial word with companion $abba\diamond bacb$ is weakly 3-periodic but is not 3-periodic. The latter shows a difference between partial words and words since every weakly $p$-periodic word is $p$-periodic. Another difference worth noting is the fact that even if the length of a partial word $u$ is a multiple of a weak period of $u$, then $u$ is not necessarily a power of a shorter partial word. The smallest period of $u$ is called the minimal period of $u$, and is denoted by $p(u)$, and the smallest weak period of $u$ is called the minimal weak period of $u$, and is denoted by $p'(u)$. The set of all periods of $u$ will be denoted by $\mathcal{P}(u)$ and the set of all weak periods of $u$ by $\mathcal{P}'(u)$. Note that, for any partial word $u$, $\mathcal{P}(u) \neq \emptyset$, since $|u| \in \mathcal{P}(u)$ (a similar statement holds for $\mathcal{P}'(u)$).

In the sequel, for convenience, we will consider a partial word over $A$ as a word over the enlarged alphabet $A \cup \{\diamond\}$, where the additional symbol $\diamond$ plays a special role. Thus, we say for instance “the partial word $\diamond ab \diamond baa$” instead of “the partial word with companion $\diamond ab \diamond baa$”.

2.1. Compatibility

In this section, we discuss compatibility on partial words.

If $u$ and $v$ are two partial words of equal length, then $u$ is said to be contained in $v$, denoted by $u \subseteq v$, if all elements in $D(u)$ are in $D(v)$ and $u(i) = v(i)$ for all $i \in D(u)$. We sometimes write $u \subseteq v$ if $u \subseteq v$ but $u \neq v$.

The partial words $u$ and $v$ are called compatible, denoted by $u \uparrow v$, if there exists a partial word $w$ such that $u \subseteq w$ and $v \subseteq w$. We denote by $u \vee v$ the least upper bound of $u$ and $v$ (in other words, $u \subseteq u \vee v$ and $v \subseteq u \vee v$ and $D(u \vee v) = D(u) \cup D(v)$). As an example, $u = aba\diamond a$ and $v = a\diamond b\diamond a$ are compatible and $u \vee v = abab\diamond a$. For a subset $X$ of
Lemma 2. Let \( u \) be a word and \( z \) a partial word. Assume that \( u \) and \( z \) are full. Then one of the following holds:

1. There exist partial words \( x, y, x_1, x_2 \) such that \( u = x_1 y, v = yx_2, x \subset x_1, x \subset x_2, \) and \( z = (x_1 y)^m x (yx_2)^n \) for some integers \( m, n \geq 0 \).
2. There exist partial words \( x, y, y_1, y_2 \) such that \( u = xy_1, v = y_2 x, y \subset y_1, y \subset y_2, \) and \( z = (xy_1)^m y (xy_2)^n x \) for some integers \( m, n \geq 0 \).

Proof. Assume that \( uz \uparrow vz \) with \( \|H(z)\| = 1 \) (the case where \( z \) is full comes from Lemma 2). Let \( m \) be such that \( m|u| > |z| \geq (m - 1)|u| \). Put \( u = x_1 y_1 \) and \( v = y_2 x_2 \) where \( |x_1| = W(A) \), we denote by \( C(X) \) the set of all partial words compatible with elements of \( X \). More specifically,

\[
C(X) = \{ u \mid u \in W(A) \text{ and there exists } v \in X \text{ such that } u \uparrow v \}.
\]

If \( X = \{u\} \), then we denote \( C(\{u\}) \) simply by \( C(u) \).

The following rules are useful for computing with partial words.

- **Multiplication**: If \( u \uparrow v \) and \( x \uparrow y \), then \( ux \uparrow vy \).
- **Simplification**: If \( ux \uparrow vy \) and \( |u| = |v| \), then \( u \uparrow v \) and \( x \uparrow y \).
- **Weakening**: If \( u \uparrow v \) and \( w \subset u \), then \( w \uparrow v \).

The following result extends to partial words the so-called lemma of Lévi.

**Lemma 1** (Berstel and Boasson [1]). Let \( u, v, x, y \in W(A) \) be such that \( ux \uparrow vy \).

- If \( |u| \geq |v| \), then there exist \( w, z \in W(A) \) such that \( u = wz, v \uparrow w, \) and \( y \uparrow zx \).
- If \( |u| \leq |v| \), then there exist \( w, z \in W(A) \) such that \( v = wz, u \uparrow w, \) and \( x \uparrow zy \).

2.2. Conjugacy

In this section, we discuss conjugacy on partial words. Two partial words \( u \) and \( v \) are called **conjugate** if there exist partial words \( x \) and \( y \) such that \( u \subset xy \) and \( v \subset yx \) [9]. The following lemma extends to partial words a fundamental property of words.

**Lemma 2** (Blanchet-Sadri and Luhmann [9]). Let \( u, v \in W(A) \setminus \{z\} \) and let \( z \in A^a \). If \( uz \uparrow zv \), then there exist words \( x, y \) such that \( u \subset xy, v \subset yx, \) and \( z \subset (xy)^n x \) for some integer \( n \geq 0 \).

Lemma 2 does not necessarily hold if \( z \) is not full even if \( u, v \) are full. The partial words \( u = a, v = b, \) and \( z = \#bb \) provide a counterexample. However, the following holds.

**Lemma 3** (Blanchet-Sadri and Luhmann [9]). Let \( u, v \in W(A) \setminus \{z\} \) and let \( z \in W(A) \). If \( uz \uparrow zv \) and \( uz \lor zv \) is \( |u| \)-periodic, then there exist words \( x, y \) such that \( u \subset xy, v \subset yx, \) and \( z \subset (xy)^n x \) for some integer \( n \geq 0 \).

The following lemma will be useful for our purposes in this paper. It is used extensively to prove our main theorems (Theorems 2 and 3).

**Lemma 4.** Let \( u, v \in A^+ \) and let \( z \in W_1(A) \). If \( uz \uparrow zv \), then one of the following holds:

1. There exist partial words \( x, y, x_1, x_2 \) such that \( u = x_1 y, v = yx_2, x \subset x_1, x \subset x_2, \) and \( z = (x_1 y)^m x (yx_2)^n \) for some integers \( m, n \geq 0 \).
2. There exist partial words \( x, y, y_1, y_2 \) such that \( u = xy_1, v = y_2 x, y \subset y_1, y \subset y_2, \) and \( z = (xy_1)^m y (xy_2)^n x \) for some integers \( m, n \geq 0 \).
|z_2| = |z| - (m - 1)|u| and |y_1| = |y_2| (here |u| = |v|). Put \( z = y_1y'_1x_2y_2' \ldots x_{m-1}'y_{m-1}'x_m' \) where \( |x_i'| = \ldots = |x_{m-1}'| = |x_m'| = |x_1| = |x_2| \) and \( |y_i'| = \ldots = |y_{m-1}'| = |y_1| = |y_2| \).

Since \( uz \uparrow zv \), we get

\[
\begin{align*}
&x_1 y_1' x_1' y_1' x_2' y_2' \ldots x_{m-2}' y_{m-2}' x_{m-1}' y_{m-1}' x_m' \\
&\uparrow \\
&x_1' y_1' x_2' y_2' x_3' y_3' \ldots x_{m-1}' y_{m-1}' x_m' y_2 \quad x_2.
\end{align*}
\]

If the hole is in \( x_m' \), then \( y_1 = y_1' = y_2' = \ldots = y_{m-1}' = y_2, x_m' \subset x_2, \) and \( x_m' \subset x_{m-1} = \ldots = x_1' = x_1 \). Here, \( u = x_1y_1, v = y_1x_2, z = (x_1y_1)^{m-1}x_m' \). Now, if the hole is in \( x_i' \) for some \( 1 \leq i < m \), then \( y_1 = y_1' = y_2' = \ldots = y_{m-1}' = y_2, x_i' \subset x_{i+1}' = \ldots = x_m' = x_2, \) and \( x_i' \subset x_{i-1}' = \ldots = x_1' = x_1 \). Here, \( u = x_1y_1, v = y_1x_2, z = (x_1y_1)^{m-1}x_i'y_i'(x_1y_2)^{m-i-1}x_1 \) and Statement 1 holds.

If the hole is in \( y_i' \) for some \( 1 \leq i < m \), then \( x_1 = x_1' = x_2' = \ldots = x_m' = x_2, \)
\( y_i' \subset y_{i+1}' = \ldots = y_{m-1}' = y_2, \) and \( y_i' \subset y_{i-1}' = \ldots = y_1' = y_1 \). Here, \( u = x_1y_1, v = y_2x_1, z = (x_1y_1)^{m-1}x_1y_1'(x_1y_2)^{m-i-1}x_1 \) and Statement 2 holds.

The following lemma is used in particular to prove Theorem 4.

**Lemma 5.** Let \( u, v \in A^+ \).

1. Let \( z \in W_1(A) \setminus A^+ \) and let \( z' \in A^+ \). If \( z \uparrow z' \) and \( uz \uparrow z'v \), then one of the following holds:
   (a) There exist partial words \( x, y, x_1, x_2 \) such that \( u = x_1y, v = yx_2, x \subset x_1, x \subset x_2, \)
   \( z = (x_1y)^m x(yx_2)^n, \) and \( z' = (x_1y)^m x_1(yx_2)^n \) for some integers \( m, n \geq 0 \).
   (b) There exist partial words \( x, y, y_1, y_2 \) such that \( u = xy_1, v = y_2x, y \subset y_1, y \subset y_2, \)
   \( z = (x_1y)^m yx(yx_2)^n x, \) and \( z' = (x_1y)^m yx(yx_2)^n x \) for some integers \( m, n \geq 0 \).

2. Let \( z \in A^+ \) and let \( z' \in W_1(A) \setminus A^+ \). If \( z \uparrow z' \) and \( uz \uparrow z'v \), then one of the following holds:
   (a) There exist partial words \( x, y, x_1, x_2 \) such that \( u = x_1y, v = yx_2, x \subset x_1, x \subset x_2, \)
   \( z = (x_1y)^m x_2(yx_2)^n, \) and \( z' = (x_1y)^m x_2(yx_2)^n \) for some integers \( m, n \geq 0 \).
   (b) There exist partial words \( x, y, y_1, y_2 \) such that \( u = xy_1, v = y_2x, y \subset y_1, y \subset y_2, \)
   \( z = (x_1y)^m yx_1(yx_2)^n x, \) and \( z' = (x_1y)^m yx_1(yx_2)^n x \) for some integers \( m, n \geq 0 \).

**Proof.** Let us prove Statement 1 (Statement 2 is proved similarly). By weakening, \( uz \uparrow zv \).

If Lemma 4(1) holds, then there exist partial words \( x, y, x_1, x_2 \) such that \( u = x_1y, v = yx_2, x \subset x_1, x \subset x_2, \) and \( z = (x_1y)^m x(yx_2)^n \) for some integers \( m, n \geq 0 \). Since \( u, v \) are full, we have \( y, x_1, x_2 \) full and thus, \( ||H(x)|| = 1 \). Since \( z \uparrow z' \), there exists a word \( x' \) such that \( x \subseteq x' \) and \( z' = (x_1y)^m x' (yx_2)^n \).

Now, \( uz \uparrow z'v \) implies \( (x_1y)^m x(yx_2)^n \uparrow (x_1y)^m x'(yx_2)^n+1 \)
and by simplification, \( x_1yx \uparrow x'yx_2 \). Thus, \( x_1 \uparrow x' \). The latter along with the fact that both \( x_1 \) and \( x' \) are full lead to \( x' = x_1 \), and Lemma 5(1)(a) holds in this case. If Lemma 4(2) holds, then Lemma 5(1)(b) follows. \( \square \)

Throughout the rest of this paper, \( A \) denotes a fixed alphabet.
3. Orderings

In this section, we define some total orderings of $W(A)$ that will be used in the sequel.

First, assume that $A$ is totally ordered by $<$ and assume that $\diamond < a$ for all $a \in A$. The first total ordering of $W(A)$, denoted by $\preceq_l$, is defined as follows: $u \preceq_l v$ if either $u$ is a proper prefix of $v$, or $u = (u \wedge v)a_k$, $v = (u \wedge v)b_k$ by with $a, b \in A \cup \{\diamond\}$ satisfying $a \preceq_l b$. The second total ordering of $W(A)$, denoted by $\preceq_r$, is obtained from $\preceq_l$ by reversing the order of letters, that is, for $a, b \in A$, $a \preceq_l b$ if and only if $b \preceq_r a$. For all $u, v, x, y \in W(A)$, the following hold:

$$
\begin{align*}
& u \preceq_l v \quad \text{if and only if} \quad xu \preceq_l xv, \\
& u \preceq_r v \quad \text{if and only if} \quad xu \preceq_r xv, \\
& u \preceq_l v \quad \text{and} \quad u \not\in P(v) \quad \text{imply} \quad ux \preceq_l vy, \\
& u \preceq_r v \quad \text{and} \quad u \not\in P(v) \quad \text{imply} \quad ux \preceq_r vy.
\end{align*}
$$

Now, for a partial word $u$ and $0 \leq i < j < |u|$, $u[i..j]$ denotes the factor of $u$ satisfying $(u[i..j])_0 = u_0(i) \ldots u_0(j - 1)$. The maximal suffix of $u$ with respect to $\preceq_l$ (respectively, $\preceq_r$) is defined as $u[i..|u|]$ where $0 \leq i < |u|$ and where $u[j..|u|] \preceq_l u[i..|u|]$ (respectively, $u[j..|u|] \preceq_r u[i..|u|]$) for all $0 \leq j < |u|$. For example, if $a \preceq_l b$, then the maximal suffix of $ba \diamond bbaab$ with respect to $\preceq_l$ is $bbaab$, and with respect to $\preceq_r$ is $aab$.

To prove our main results, we need three auxiliary results. The first two hold for the orderings $\preceq_l$ and $\preceq_r$.

**Lemma 6.** Let $u, v, w \in W(A)$.

1. If $v$ is the maximal suffix of $w = uv$ with respect to $\preceq_l$, then no nonempty partial words $x, y$ are such that $y \subset x, u = rx$ and $v = ys$ for some $r, s \in W(A)$.
2. If $v$ is the maximal suffix of $w = uv$ with respect to $\preceq_r$, then no nonempty partial words $x, y$ are such that $y \subset x, u = rx$ and $v = ys$ for some $r, s \in W(A)$.

**Proof.** We prove Statement 1 (Statement 2 is similar). Let $x, y$ be nonempty partial words satisfying $y \subset x, u = rx$ and $v = ys$ for some $r, s \in W(A)$. Since $w = uv = rxv = rxys$, by the maximality of $v$, we have $xv \preceq_l v$ and $s \preceq_l v$. Since $v = ys$, these inequalities can be rewritten as $xys \preceq_l ys$ and $s \preceq_l ys$. Now, from the former inequality we obtain that $yys \preceq_l ys$ since $y \subset x$. We then obtain that $ys \preceq_l s$, which together with $s \preceq_l ys$ imply that $s = ys$. Therefore, $y = \varepsilon$ and $x = \varepsilon$ leading to a contradiction. □

**Lemma 7.** Let $u, v, w \in W(A)$.

1. If $v$ is the maximal suffix of $w = uv$ with respect to $\preceq_l$, then no nonempty partial words $x, y, s$ are such that $y \subset x, u = rx$ and $y = vs$ for some $r \in W(A)$.
2. If $v$ is the maximal suffix of $w = uv$ with respect to $\preceq_r$, then no nonempty partial words $x, y, s$ are such that $y \subset x, u = rx$ and $y = vs$ for some $r \in W(A)$.
Proof. We prove Statement 1 (Statement 2 is similar). Let $x, y, s$ be nonempty partial words satisfying $y \subset x$, $u = rx$ and $y = vs$ for some $r \in W(A)$. Here $w = uv = rxv$, and since $v$ is the maximal suffix with respect to $\leq_l$, we get $xv \leq_l v$. Since $y \subset x$, we get $yv \leq_l v$. Replacing $y$ by $vs$ in the latter inequality yields $vsv \leq_l v$, leading to a contradiction. □

The third one, which comes from the definitions, claims that under some restrictions on $u, v$, the relations $u \leq_l v$ and $u \leq_r v$ together define “$u$ is a prefix of $v$”.

Lemma 8. Let $u \in A^+$ and $v \in W_1(A) \setminus \{\varepsilon\}$. Then both $u \leq_l v$ and $u \leq_r v$ if and only if $u \in P(v)$.

Proof. The result being trivial for $v \in A^+$, assume that $\|H(v)\| = 1$. If $u \in P(v)$, then both $u \leq_l v$ and $u \leq_r v$. Conversely, if both $u \leq_l v$ and $u \leq_r v$, then either $u$ is a prefix of $v$, or $u = (u \land v)ax$, $v = (u \land v)by$ with $a, b \in A \cup \{\phi\}$ satisfying $a \leq_l b$ and $a \leq_r b$. The latter possibility leads to $a = \phi$, contradicting the fact that $u$ is full. □

4. Critical factorization theorem on words

In this section, we discuss the version of the critical factorization theorem on words which appears in [14]. Intuitively, the theorem states that the minimal period $p(w)$ of a word $w$ of length at least two can be locally determined in at least one position of $w$. This means that there exists a factorization $(u, v)$ of $w$ with $u, v \neq \varepsilon$ such that $p(w)$ is the minimal local period of $w$ at position $|u| - 1$ (a factorization of a word $w$ is any tuple $(u, v)$ of words such that $w = uv$). A local period at a given position is defined as follows.

Definition 1. Let $w \in A^+$. A positive integer $p$ is called a local period of $w$ at position $i$ if there exist $u, v \in A^+$ and $x \in A^*$ such that $w = uv$, $|u| = i + 1$, $|x| = p$, and such that one of the following conditions holds for some words $r, s$:

1. $u = rx$ and $v = xs$ (internal square),
2. $x = ru$ and $v = xs$ (left external square if $r \neq \varepsilon$),
3. $u = rx$ and $x = vs$ (right external square if $s \neq \varepsilon$),
4. $x = ru$ and $x = vs$ (left and right external square if $r, s \neq \varepsilon$).

The minimal local period of $w$ at position $i$, denoted by $p(w, i)$, is defined as the smallest local period of $w$ at position $i$.

Intuitively, around position $i$, there exists a factor of $w$ having as its minimal period this minimal local period. A factorization $(u, v)$ of $w$ is called critical when $u, v \neq \varepsilon$ and $p(w) = p(w, |u| - 1)$. Clearly, $1 \leq p(w, i) \leq p(w) \leq |w|$. For more on the local periodic structure of a word, we refer the reader to [20, 21].

As an example, consider the word $w = babbaab$ with minimal period 6. The minimal local periods of $w$ are: $p(w, 0) = 2$, $p(w, 1) = 3$, $p(w, 2) = 1$, $p(w, 3) = 6$, $p(w, 4) = 1$, and $p(w, 5) = 3$. Here, $p(w) = p(w, 3)$ which means that the factorization $(babb, aab)$ is critical.
The following theorem states that each word of length at least two has at least one critical factorization.

**Theorem 1.** Let \( w \) be a word such that \(|w| \geq 2\). Then \( w \) has at least one critical factorization \((u,v)\) with \( u, v \neq \varepsilon \) and \( p(w) = p(w, |u| - 1) \).

According to Theorem 1, the minimal period of a word \( w \) is simply the maximum among all minimal local periods. Theorem 1’s proof not only shows the existence of a critical factorization, but also shows that such a factorization can be found by computing two maximal suffixes of \( w \) with respect to \( \preceq_l \) and \( \preceq_r \). If \( v \) denotes the maximal suffix of \( w \) with respect to \( \preceq_l \) and \( v' \) the maximal suffix of \( w \) with respect to \( \preceq_r \), then let \( u, u' \) be such that \( w = uv = u'v' \). The factorization \((u,v)\) is critical when \(|v| \leq |v'|\), and the factorization \((u',v')\) is critical when \(|v| > |v'|\). There exist linear time (in the length of \( w \)) algorithms for such computations [16,17,32] (the latter two use the suffix tree construction).

### 5. Critical factorization theorem on partial words with one hole

In this section, we discuss our first version of the critical factorization theorem on partial words with one hole. Intuitively, our theorem states that the minimal weak period \( p'(w) \) of a partial word \( w \) with one hole of length at least two can be locally determined in at least one position of \( w \) provided \( w \) is not special as stated in Definition 3. This means that there exists a factorization \((u,v)\) of \( w \) with \( u, v \neq \varepsilon \) such that \( p'(w) \) is the minimal local period of \( w \) at position \(|u| - 1\). The following definition states precisely what we mean by a local period at a given position.

**Definition 2.** Let \( w \in W(A) \setminus \{\varepsilon\} \). A positive integer \( p \) is called a local period of \( w \) at position \( i \) if there exist \( u, v \in W(A) \setminus \{\varepsilon\} \) and \( x, y \in W(A) \) such that \( w = uv, |u| = i + 1, |x| = p, x \uparrow y \), and such that one of the following conditions holds for some partial words \( r, s \):

1. \( u = rx \) and \( v = ys \),
2. \( x = ru \) and \( v = ys \),
3. \( u = rx \) and \( y = vs \),
4. \( x = ru \) and \( y = vs \).

The minimal local period of \( w \) at position \( i \), denoted by \( p(w, i) \), is defined as the smallest local period of \( w \) at position \( i \).

Intuitively, around position \( i \), there exists a factor of \( w \) having as its minimal weak period this minimal local period. A factorization \((u,v)\) of \( w \) where \( u, v \neq \varepsilon \) is called critical when \( p'(w) = p(w, |u| - 1) \). Clearly, \( 1 \leq p(w, i) \leq p'(w) \leq |w| \).

In order to distinguish partial words from special partial words, the following restrictions are placed on the above conditions.

**Definition 3.** Let \( w \) be a partial word with one hole satisfying \( p'(w) > 1 \). Let \( v \) (respectively, \( v' \)) be the maximal suffix with respect to \( \preceq_l \) (respectively, \( \preceq_r \)). Let \( u, u' \) be partial words such that \( w = uv = u'v' \).
• If $|v| \leq |v'|$, then let $t$ be the partial word such that $u = u't$ and $v' = tv$. Then $w$ is called special if one of the following holds:

1. Definition 2(1) is satisfied with $r \neq \epsilon$, $\|H(x)\| = 1$, and either Condition (a) or (b) holds:
   (a) $|t| > |x|$ and $(r \notin C(S(u))$ or $s \notin C(P(v))$).
   (b) $|t| \leq |x|$, $\|H(t)\| \neq 1$, and $r \notin C(S(u))$.
2. Definition 2(2) is satisfied with $r \neq \epsilon$, $s \notin C(P(v))$, and either Condition (c) or (d) holds:
   (c) $\|H(s)\| = 1$.
   (d) $\|H(t')\| = 1$ where $t'$ is the suffix of length $|t|$ of $y$.
3. Definition 2(3) is satisfied with $r \notin C(S(u))$, $s \neq \epsilon$, and $\|H(\gamma)\| = 1$ (where $\gamma$ is the partial word satisfying $x = \gamma s$).

• If $|v| \geq |v'|$, then let $t$ be the partial word such that $u' = ut$ and $v = tv'$. Then $w$ is called special if one of the above holds when referring to Definition 2 where $u$ is replaced by $u'$ and $v$ by $v'$.

The partial word $w$ is called nonspecial otherwise.

Referring to Definition 3, the following table illustrates the (underlined) factor of the special partial word $w = uv = u'tv = u'v'$ that contains the hole:

<table>
<thead>
<tr>
<th>Definition 3</th>
<th>$u$</th>
<th>$v$</th>
<th>$x$</th>
<th>$y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1(a)</td>
<td>$r \notin C(S(u))$</td>
<td>$t &gt;</td>
<td>x</td>
<td>$</td>
</tr>
<tr>
<td></td>
<td>$r \neq \epsilon$</td>
<td>$s \notin C(P(v))$</td>
<td>$t &gt;</td>
<td>x</td>
</tr>
<tr>
<td>1(b)</td>
<td>$r \notin C(S(u))$</td>
<td>$t \leq</td>
<td>x</td>
<td>$</td>
</tr>
<tr>
<td>2(c)</td>
<td>$r \neq \epsilon$</td>
<td>$s \notin C(P(v))$</td>
<td>$y$</td>
<td>$r$</td>
</tr>
<tr>
<td>2(d)</td>
<td>$r \neq \epsilon$</td>
<td>$s \notin C(P(v))$</td>
<td>$y$</td>
<td>$r$</td>
</tr>
<tr>
<td>3</td>
<td>$r \notin C(S(u))$</td>
<td>$s \neq \epsilon$</td>
<td>$x$</td>
<td>$y$</td>
</tr>
</tbody>
</table>

The following theorem claims that each nonspecial partial word with one hole of length at least two has at least one critical factorization. Our proof not only shows the existence of a critical factorization, but also gives an algorithm to compute such a factorization explicitly.

**Theorem 2.** If $w$ is a nonspecial partial word with one hole such that $|w| > 2$, then $w$ has at least one critical factorization. More specifically, if $p'(w) > 1$, let $v$ denote the maximal suffix of $w$ with respect to $\preceq_l$ and $v'$ the maximal suffix of $w$ with respect to $\preceq_r$. Let $u$, $u'$ be partial words such that $w = uv = u'v'$. Then the factorization $(u, v)$ is critical when $|v| \leq |v'|$, and the factorization $(u', v')$ is critical when $|v| > |v'|$. Moreover, if $|v| \leq |v'|$ and the factorization $(u, v)$ is critical, then $w$ is nonspecial, and if $|v| > |v'|$ and the factorization $(u', v')$ is critical, then $w$ is nonspecial.

**Proof.** If $p'(w) = 1$, then $w = a^m \diamond b^n$ for some $a, b \in A$ and integers $m, n \geq 0$. The result trivially holds in this case. We now assume that $p'(w) > 1$ and that $|v| \leq |v'|$. In the case
where \( p'(w) > 1 \) and \(|v| > |v'| \) is proved similarly but requires that the orderings \( \leq_l \) and \( \leq_r \) be interchanged. Put \( u = u't \) and \( v' = tv \) for some \( t \). We first show that \( u \neq v \).

Assume to the contrary that \( u = v \), and thus \( w = v \). Since \(|v| \leq |v'| \), we also have \( w = v' \). In the case where \( w = az = v = v' \) for some \( a \in A \) and \( z \in W(A) \setminus A^* \), we argue as follows. If \( b \in A \) is a letter in \( z \), then \( b \leq_l a \) and \( b \leq_r a \). Thus \( b = a \) and \( w \) is unary. We get \( p'(w) = 1 \), contradicting our assumption.

Now, let us denote \( p(w, |u| - 1) \) by \( p \). We consider the following cases:

**Case 1:** \( p \geq |u| \) and \( p \geq |v| \)

If \( p \geq |u| \) and \( p \geq |v| \), then Condition (4) is satisfied. There exist \( x, y, r, s \in W(A) \) such that \(|x| = p, x \uparrow y, x = ru, \) and \( y = vs \).

First, assume that \(|r| > |v| \). We may choose \( r, s \in W(A) \setminus \{e\} \) such that \( r = vz \) and \( zu = s \). Thus, \( p = |x| = |ru| = |vzu| > |uv| = |w| \geq p'(w) \), which leads to a contradiction.

Now, assume that \(|r| \leq |v| \). We may choose \( r, s \in W(A) \) such that \( v = rz, \) \( u = zs, \) and \( z \in z' \). If \( \|H(u)\| = 1 \) or \( \|H(r)\| = 1 \), then \( z' \subset z \). Thus, \( vu \subset zsrz \) showing that \( p = |zsr| \) is a weak period of \( uv \). If \( \|H(z)\| = 0 \), then \( z \subset z' \). Thus, \( uv \subset zsrz' \) showing that \( p = |z'sr| \) is a weak period of \( uv \). In any case, \( p'(w) \leq p \). On the other hand, \( p'(w) \geq p \). Therefore, \( p'(w) = p \) which shows that the factorization \((u, v)\) is critical.

**Case 2:** \( p < |u| \) and \( p > |v| \)

If \( p < |u| \) and \( p > |v| \), then Condition (3) is satisfied. There exist \( x, y, r, s, \gamma \in W(A) \) such that \(|x| = p, \gamma \uparrow v, u = rx = r\gamma s, \) and \( y = vs \).

If \( v \subset \gamma \), then \( y \subset x, \) and \( v \) being the maximal suffix of \( w \) with respect to \( \leq_l \), we get a contradiction with Lemma 7.

If \( \gamma \subset v \), then we consider whether or not \( r \in C(S(u)) \). If \( r \notin C(S(u)) \), then \( w \) is special by Definition 3(3). If \( r \in C(S(u)) \), then \( x'r \uparrow rx \) for some \( x' \). Since \( r \) is full, by Lemma 2, there exist words \( t_1, t_2 \) such that \( x' \subset t_2t_1, x \subset t_2t_1, \) and \( r \subset (t_1t_2)^k \) for some integer \( k \geq 0 \). Thus, \( uv = rxyv = r\gamma s \subset (t_1t_2)^k t_1 \gamma s v \) with \( s \subset t_2t_1 \) and \( \gamma \subset v \) showing that \( p = |x| = |t_1t_2| \) is a weak period of \( uv \).

**Case 3:** \( p < |u| \) and \( p \leq |v| \)

If \( p < |u| \) and \( p \leq |v| \), then Condition (1) is satisfied. There exist \( x, y, r, s \in W(A) \) such that \(|x| = p, x \uparrow y, u = rx, \) and \( v = ys \).

If \( y \subset x \), then \( v \) being the maximal suffix of \( w \) with respect to \( \leq_l \), we get a contradiction with Lemma 6.

If \( x \subset y \), then we argue as follows. We first show that if \( r \in C(S(u)) \) and \( s \in P(v) \), then \( p \) is a weak period of \( w \). If the two conditions hold, then \( x'r \uparrow rx \) and \( ys = sy' \) for some \( x', y' \). Since \( r \) and \( s \) are full, by Lemma 2, there exist words \( t_1, t_2 \) and an integer \( k \geq 0 \) as in Case 2, and words \( t_3, t_4 \) such that \( y \subset t_3t_4, y' \subset t_4t_3, \) and \( s \subset (t_3t_4)^k t_3 \) for some integer \( \ell \geq 0 \). We get \( x \subset t_2t_1, x \subset y \subset t_3t_4, uv = rxyv \subset (t_1t_2)^k t_1x(t_3t_4)^{\ell+1} t_3, \) and \( p = |x| = |t_2t_1| = |t_3t_4| \) is a weak period of \( w \). Now, \( u = rx = u't \), and we consider the case where \( |r| \leq |x| \) and then the case where \( |r| > |x| \).

- If \( |r| \leq |x| \), then put \( x = zt \) for some \( z \). Since \( x \uparrow y, \) put \( y = z't' \) where \( z \uparrow z' \) and \( t \uparrow t' \).
  - If \( \|H(t)\| = 1 \), then \( z = z' \) and \( t \subset t' \). Since \( v' = tzt's \) is maximal with respect to \( \leq_r \), the
The latter implies that $\|H(z)\| = 1$, then $z \sqsubseteq z'$ and $t = t'$. Since $v' = t'z'ts$ is maximal with respect to $\leq_r$, the suffix $ts$ of $v'$ satisfies $ts \preceq_r v' = tv$, which implies $s \preceq_r v$. Since $v = z'ts$ is maximal with respect to $\preceq$, we have $s \preceq_t v$, and $s \in P(v)$ by Lemma 8. If $r \in C(S(u))$, then $p$ is a weak period of $w$. Otherwise, $w$ is special by Definition 3(1)(b).

- If $|t| > |v|$, then by Definition 3(1)(a), $w$ is special unless $r \in C(S(u))$ and $s \in P(v)$.

  If the latter conditions hold, then $p$ is a weak period of $w$.

**Case 4:** $p \geq |u|$ and $p < |v|$

If $p \geq |u|$ and $p < |v|$, then Condition (2) is satisfied. There exist $x, y, r, s, \gamma, t' \in W(A)$ such that $|x| = p, u' \uparrow \gamma, t \uparrow t', x = ru = ru't$, and $v = ys = r\gamma't's$. Note that since $v' = tv = tr\gamma't's$ is maximal with respect to $\leq_r$, the suffix $t's$ of $v'$ satisfies $t's \preceq_r v' = tv$. The latter implies that $t' \subset t$. Note also that if $r = e$ and $\gamma \subset u'$, then $y \subset x = u$ leading to a contradiction with Lemma 6.

We first assume that $t' \subset t$, and thus $w$ is special by Definition 3(2)(d) unless $s \in C(P(v))$. If $s \in C(P(v))$, then $ys \uparrow sy'$ for some $y'$, and since $s$ is full, by Lemma 2, there exist words $t_3, t_4$ and an integer $\ell \geq 0$ as in Case 3. We get $ruv = xys \subset xy(t_3t_4)\ell t_3$ with $y \subset x$ and $y \subset t_3t_4$. So $p = |x| = |t_3t_4|$ is a weak period of $uv$.

We now assume that $t' = t$. We can rewrite $v, v'$ as $v' = r\gamma' ts$ and $v' = tr\gamma' ts$. We consider whether or not $s$ is full.

- If $s$ is full, then since $v'$ is maximal with respect to $\leq_r$, the suffix $ts$ of $v'$ satisfies $ts \preceq_r v' = tv$, which implies that $s \preceq_r v$. Since $v$ is maximal with respect to $\preceq$, we have $s \preceq_t v$. Since both $s \preceq_t v$ and $s \preceq_r v$, we get $s \in P(v)$ by Lemma 8. Put $v = ys = sy'$ for some $y'$. There exist words $t_3, t_4$ and an integer $\ell \geq 0$ as above. If $\gamma \subset u'$, then as above $p$ is a weak period of $uv$. If $u' \subset \gamma$, then $x \subset y \subset t_3t_4$, $ruv = xys \subset (t_3t_4)^{\ell+2}t_3$ and $p$ is a weak period of $uv$.

- If $s$ is not full, then $w$ is special by Definition 3(2)(c) unless $s \in C(P(v))$. If $s \in C(P(v))$, then $ys \uparrow sy'$ for some word $y'$. By Lemma 4, two possibilities can occur:

  **Possibility 1:** there exist partial words $t_1, t_2, t_3, t_4$ such that $y = t_3t_2, y' = t_2t_4, t_1 \subset t_3, t_1 \subset t_4$, and $s = (t_3t_2)^{k_1}(t_2t_4)\ell$ for some integers $k, \ell \geq 0$. We get $ruv = xys = yys = (t_3t_2)^{k+2}t_1(t_2t_4)\ell$ and so $p = |x| = |t_3t_2| = |t_1t_2| = |t_4t_2|$ is a weak period of $uv$.

  **Possibility 2:** there exist partial words $t_1, t_2, t_3, t_4$ such that $y = t_1t_3, y' = t_4t_1, t_2 \subset t_3, t_2 \subset t_4$, and $s = (t_1t_3)^{k_1}t_1t_2(t_1t_4)\ell$ for some integers $k, \ell \geq 0$. We get $ruv = xys = yys = (t_1t_3)^{k+2}t_1(t_2t_4)\ell t_1$ and so $p = |x| = |t_1t_3| = |t_1t_2| = |t_1t_4|$ is a weak period of $uv$. $\Box$

We now give examples showing why our result excludes the special partial words. It is assumed that $a \prec_l b$.

**Example 1.** The partial word $w = baabab$ has $v = bb$ as maximal suffix with respect to $\preceq_l$ and $v' = ababab$ as maximal suffix with respect to $\leq_r$. We can compute $p'(w) = 5$ and $p(w, 3) = 1$. Referring to Definition 3(1)(a), $(baa)\diamond (b)(b) = rxyz = uyz = uv, \|H(x)\| = 1, |t| = |aab| > |x|, s \in P(v)$ but $r \notin C(S(u))$. No position $i$ exists such that $p'(w) = p(w, i)$.
Algorithm 1. Compute the maximal suffix $v$ of $w$ with respect to $\preceq_l$ (respectively, $\preceq_r$). We can compute $p'(w) = 7$ and $p(w, 5) = 3$. Referring to Definition 3(1)(a). $(aa)(aa)(baa)(a) = rxy = uys = uv$, $\|H(x)\| = 1$, $|r| = |aa\circ ab| > |x|$, $r \in C(S(u))$ but $s \not\in P(v)$. There exists no position $i$ such that $p'(w) = p(w, i)$.

- If $w = ab\circ ab\circ ba$, then $v = bba$ and $v' = abba$. Here, $p'(w) = 4$ and $p(w, 4) = 3$. Referring to Definition 3(1)(b). $(ab)(ba)(bba)(bbab) = uys = uv$, $\|H(s)\| = 1$, and $s \not\in C(P(v))$. There does not exist any position $i$ such that $p'(w) = p(w, i)$.

- If $w = babbabbab\circ b$, then the maximal suffixes with respect to $\preceq_l$ and $\preceq_r$ are $v = bbabbbab\circ b$ and $v' = abbbabbab\circ b$. Here, $p'(w) = 8$ and $p(w, 1) = 3$. Referring to Definition 3(2)(c). $(ba)(ba)(bbab) = uys = uv$, $\|H(t')\| = 1$, and $s \not\in C(P(v))$. There does not exist any position $i$ satisfying $p'(w) = p(w, i)$.

- Consider the partial word $w = bab\circ ab$ where $v = bb\circ ab$ and $v' = ab\circ ab$. Computations show that $p'(w) = 4$ and $p(w, 1) = 3$. Referring to Definition 3(2)(d). $(ba)(b)(b)(ab) = ur\gamma t's = uv$, $\|H(t')\| = 1$, and $s \not\in C(P(v))$. No position $i$ satisfies $p'(w) = p(w, i)$.

- Consider the partial word $w = ab\circ aab$. Computations show that $v = ba$, $v' = aaba$, $p'(w) = 4$, and $p(w, 4) = 3$. Referring to Definition 3(3). $(ab)(a)(a)(ba) = r\gamma sv = uv$, $\|H(\gamma)\| = 1$, and $r \not\in C(S(u))$. No position $i$ satisfies $p'(w) = p(w, i)$.

From the proof of Theorem 2, we obtain an algorithm that outputs a critical factorization for a given partial word $w$ with $p'(w) > 1$ and with one hole of length at least two when $w$ is nonspecial, and that outputs “special” otherwise.

Algorithm 1. Compute the maximal suffix $v$ of $w$ with respect to $\preceq_l$ and the maximal suffix $v'$ of $w$ with respect to $\preceq_r$. Find partial words $u, u'$ such that $w = uv = u'v'$. Then do one of the following steps:

Step 1: If $|v| \leq |v'|$, then compute $p = p(w, |u| - 1)$ and find the partial word $t$ such that $u = u't$ and $v' = tv$. Then do one of the following:

1. If $p \geq |u|$ and $p \geq |v|$, then output $(u, v)$.
2. If $p < |u|$ and $p \leq |v|$, then find partial words $x, y, r, s$ satisfying Definition 2(1). If $\|H(x)\| = 1$ and $|r| > |x|$ and $r \not\in C(S(u))$ or $s \not\in P(v)$, then output “special Definition 3(1)(a)”. If $\|H(x)\| = 1$ and $|r| \leq |x|$ and $\|H(s)\| \neq 1$ and $r \not\in C(S(u))$, then output “special Definition 3(1)(b)”. Otherwise, output $(u, v)$.
3. If $p \geq |u|$ and $p < |v|$, then find partial words $x, y, r, s$ satisfying Definition 2(2). If $r \neq \varepsilon$ and $s \not\in C(P(v))$ and $\|H(s)\| = 1$, then output “special Definition 3(2)(c)”.
4. If $p < |u|$ and $p > |v|$, then find partial words $x, y, r, s$ satisfying Definition 2(3). Find the prefix $\gamma$ of $x$ of length $|v|$. If $\|H(\gamma)\| = 1$ and $r \not\in C(S(u))$, then output “special Definition 3(3)”.

Step 2: If $|v| > |v'|$, then compute $p = p(w, |u'| - 1)$ and find the partial word $t$ such that $u' = ut$ and $v = tv'$. Then do one of the above where $u$ is replaced by $u'$ and $v$ by $v'$.
6. A class of special partial words

Consider the ordering \( \preceq_l \) where \( \diamond < a < b < c \). The nonempty suffixes of \( w = abca\diamond cab \) are ordered as follows:

\[
v_7 \prec_l v_6 \prec_l v_5 \prec_l v_4 \prec_l v_3 \prec_l v_2 \prec_l v_1 \prec_l v_0,
\]

where \( v_7 = \diamond cab, v_6 = a\diamond cab, v_5 = ab, v_4 = abca\diamond cab, v_3 = b, v_2 = bca\diamond cab, v_1 = ca\diamond cab, \) and \( v_0 = cab \). Now, consider the ordering \( \preceq_r \) where \( \diamond < c < b < a \). The nonempty suffixes of \( w \) are ordered as:

\[
v_7' \prec_r v_6' \prec_r v_5' \prec_r v_4' \prec_r v_3' \prec_r v_2' \prec_r v_1' \prec_r v_0',
\]

where \( v_7' = \diamond cab, v_6' = ca\diamond cab, v_5' = cab, v_4' = bca\diamond cab, v_3' = ab, \) and \( v_0' = abca\diamond cab \). We get the factorizations \( (u_i, v_i) \) and \( (u_j', v_j') \) for \( i, j = 0, 1, \ldots, 7 \). In this example, the factorization \( (u_0, v_0) \) is not critical, nor the factorization \( (u_0', v_0') \). However, the factorization \( (u_1, v_1) \) is critical.

In the following, the nonempty suffixes of a given partial word \( w \) of length \( n \) are ordered as follows according to \( \preceq_l \):

\[
v_{n-1} \prec_l v_{n-2} \prec_l \cdots \prec_l v_0.
\]

There result the factorizations of \( w \) called \( (u_0, v_0), (u_1, v_1), \ldots \). We will abbreviate \( p(w, |u_i| - 1) \) by \( p_i \) for all \( i = 0, 1, \ldots, n - 1 \). Similarly, the nonempty suffixes of \( w \) are ordered as follows according to \( \preceq_r \):

\[
v_{n-1}' \prec_r v_{n-2}' \prec_r \cdots \prec_r v_0'.
\]

There result the factorizations of \( w \) called \( (u_0', v_0'), (u_1', v_1'), \ldots \). We will abbreviate \( p(w, |u_j'| - 1) \) by \( p_j' \) for all \( j = 0, 1, \ldots, n - 1 \).

Let us look at the following examples of special partial words where \( a \prec_l b \):

**Example 2.** The partial word \( w = aab\diamond babbabba \) has \( v_0 = bbabba \) as maximal suffix with respect to \( \preceq_l \) and \( v_0' = aab\diamond babbabba \) as maximal suffix with respect to \( \preceq_r \). Computations give \( p'(w) = 11 \) and \( p(w, 5) = 3 \). Referring to Definition 3(1)(a), \( (aab)(\diamond ba)(bba)(bba) = rxys = uys = uv, \|H(x)\| = 1, |t| = |aabba| > |x|, s \in P(v) \) but \( r \notin C(S(u)) \). None of the factorizations \( (u_0, v_0) \) or \( (u_0', v_0') \) is critical. However, there exists a position \( i \) such that \( p'(w) = p(w, i) \). Indeed, \( (u_5, v_5) = (aa, b\diamond babbabba) \) is critical.

- The partial word \( w = aabba\diamond babbabba \) has \( v_0 = bbababa \) and \( v_0' = aabba\diamond babbabba \) as maximal suffixes with respect to \( \preceq_l \) and \( \preceq_r \). We can compute \( p'(w) = 13 \) and \( p(w, 6) = 5 \). Referring to Definition 3(1)(a), \( (aa)(bba\diamond a)(bba)(ba) = rxys = uys = uv, \|H(x)\| = 1, |t| = |aabba\diamond a| > |x|, r \in C(S(u)) \) but \( s \notin P(v) \). None of the factorizations \( (u_0, v_0) \) or \( (u_0', v_0') \) is critical. However, \( (u_1', v_1') = (aabba\diamond abb, ababa) \) is critical.

- If \( w = b\diamond baabbaab \), then \( v_0 = bbaab \) and \( v_0' = aabbaab \). Here, \( p'(w) = 9 \) and \( p(w, 4) = 4 \). Referring to Definition 3(1)(b), \( (b)(\diamond ba)(bbaab) = rxv = uv, \|H(x)\| = \)

...
1, \(|t| = |aa| \leq |x|, \|H(t)\| \neq 1\), and \(r \notin C(S(u))\). The factorization \((u_0, v_0)\) is not critical, however \((u_0, v_0')\) is critical although \(|v_0| < |v_0'|\).

- If \(w = baababaob\), then the maximal suffixes with respect to \(\preceq_l\) and \(\preceq_r\) are \(v_0 = babaaob\) and \(v_0' = aababaoob\). Here, \(p'(w) = 9\) and \(p(w, 2) = 5\). Referring to Definition 3(2)(c), \((baa)(babaa)(ob) = uys = uv, \|H(s)\| = 1\), and \(s \notin C(P(v))\). None of the factorizations \((u_0, v_0)\) or \((u_0', v_0')\) is critical. There exists however a position \(i\) satisfying \(p'(w) = p(w, i)\). Indeed, the factorization \((u_1', v_1') = (baabab, aaob)\) is critical.

- Consider the partial word \(w = baabbaobaa\) where \(v_0 = bbbaobaa\) and \(v_0' = aabbbaoobaa\). Computations show that \(p'(w) = 8\) and \(p(w, 2) = 5\). Referring to Definition 3(2)(d), \((baa)(bb)(b)(aob)(baa) = ur^3t's = uv, \|H(t')\| = 1\), and \(s \notin C(P(v))\). None of the factorizations \((u_0, v_0)\) or \((u_0', v_0')\) is critical. However, there exists a position \(i\) satisfying \(p(w, i) = 8\). Indeed, \(i = 5\) shows that \((u_8, v_8) = (baabbb, aoboob)\) is critical.

- Consider the partial word \(w = aaabobabb\). Computations show that \(v_0 = bb, v_0' = aaabobobb, p'(w) = 9\), and \(p(w, 6) = 3\). Referring to Definition 3(3), \((aaab)(ob)(a)(bb) = r^3sv = uv, \|H(y)\| = 1\), and \(r \notin C(S(u))\). Position 2 satisfies \(p'(w) = p(w, 2)\) showing that \((u_2, v_2) = (aaa, bobbabb)\) is critical, although none of the factorizations \((u_0, v_0)\) or \((u_0', v_0')\) is critical.

The above examples lead us to refine Theorem 2. First, we define the concept of an \((i, j)\)-special partial word.

**Definition 4.** Let \(w\) be a partial word with one hole satisfying \(p'(w) > 1\). Let \(i, j\) be a pair of nonnegative integers where \(v_i\) (respectively, \(v_j'\)) is a nonempty suffix in the ordering \(\preceq_l\) (respectively, \(\preceq_r\)). Let \(u_i, u_j'\) be partial words such that \(w = u_i v_i = u_j' v_j'\).

- If \(|v_i| \leq |v_j'|\), then let \(t\) be the partial word such that \(u_i = u_j't\) and \(v_j' = tv_j'\). Then \(w\) is called \((i, j)\)-special if one of the following holds:

1. Definition 2(1) is satisfied with \(r \neq \varepsilon\) and Condition (a) or (b) or (c) or (d) holds:
   - (a) \((i, j) = (0, 0), \|H(x)\| = 1, \|t\| > |x|,\text{ and } (r \notin C(S(u_i)))\text{ or } s \notin C(P(v_i)))\).
   - (b) \((i, j) = (0, 0), \|H(x)\| = 1, \|H(t)\| \neq 1, \|t\| \leq |x|,\text{ and } r \notin C(S(u_i)))\).
   - (c) \(i = 0, j > 0, \|H(x)\| = 1,\text{ and } (r \notin C(S(u_i))\text{ or } s \notin C(P(v_i)))\).
   - (d) \(i > 0, \|H(y)\| \neq 1,\text{ and } (r \notin C(S(u_i))\text{ or } s \notin C(P(v_i)))\).

The following table illustrates the (underlined) factor of \(w\) that contains the hole:

<p>| | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(u_i)</td>
<td>(v_i)</td>
<td>(x)</td>
<td>(i)</td>
<td>(j)</td>
<td></td>
</tr>
<tr>
<td>1(a)</td>
<td>(r \notin C(S(u_i)))</td>
<td>(t \geq</td>
<td>x</td>
<td>)</td>
<td>(\varepsilon)</td>
<td>(y)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(r \neq \varepsilon)</td>
<td>(s \notin C(P(v_i)))</td>
<td>(t \geq</td>
<td>x</td>
<td>)</td>
<td>(\varepsilon)</td>
<td>(y)</td>
</tr>
<tr>
<td>1(b)</td>
<td>(r \notin C(S(u_i)))</td>
<td>(t \leq</td>
<td>x</td>
<td>)</td>
<td>(\varepsilon)</td>
<td>(y)</td>
<td>(t)</td>
</tr>
<tr>
<td>1(c)</td>
<td>(r \neq \varepsilon)</td>
<td>(s \notin C(P(v_i)))</td>
<td>(\varepsilon)</td>
<td>(y)</td>
<td>(t)</td>
<td>0</td>
<td>+</td>
</tr>
<tr>
<td>1(d)</td>
<td>(r \neq \varepsilon)</td>
<td>(s \notin C(P(v_i)))</td>
<td>(\varepsilon)</td>
<td>(y)</td>
<td>(t)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>(r \neq \varepsilon)</td>
<td>(s \notin C(P(v_i)))</td>
<td>(\varepsilon)</td>
<td>(y)</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>
2. Definition 2(2) is satisfied with \( s \not\in C(P(v_i)) \) and either Condition (e) or (f) or (g) or (h) or (i) holds:

- (e) \( \|H(s)\| = 1 \) and \( r \neq \varepsilon \) or \( i > 0 \).
- (f) \( \|H(t')\| = 1 \) (where \( t' \) is the suffix of length \( |t| \) of \( y \)) and \( r \neq \varepsilon \) or \( i > 0 \).
- (g) \( \|H(t)\| = 1 \) and \( j > 0 \).
- (h) \( \|H(u'_j)\| = 1 \) and \( (i, j) \neq (0, 0) \).
- (i) \( \|H(r\gamma)\| = 1 \) (where \( r\gamma \) is the prefix of length \( ru'_j \) of \( y \)) and \( i > 0 \) or \( (r \neq \varepsilon \) and \( j > 0 \)).

In the following table, the + indicates a positive integer:

<table>
<thead>
<tr>
<th></th>
<th>( u_i )</th>
<th>( v_i )</th>
<th>( x )</th>
<th>( y )</th>
<th>( i )</th>
<th>( j )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2(e)</td>
<td>( r \neq \varepsilon )</td>
<td>( s \not\in C(P(v_i)) )</td>
<td>( y_\leq )</td>
<td>( ru_i )</td>
<td>( 0 )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( s \not\in C(P(v_i)) )</td>
<td>( y_\leq )</td>
<td>( ru_i )</td>
<td></td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>2(f)</td>
<td>( r \neq \varepsilon )</td>
<td>( s \not\in C(P(v_i)) )</td>
<td>( y_\leq )</td>
<td>( ru_i )</td>
<td>( r\gamma' )</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>( s \not\in C(P(v_i)) )</td>
<td>( y_\leq )</td>
<td>( ru_i )</td>
<td>( r\gamma' )</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>2(g)</td>
<td>( s \not\in C(P(v_i)) )</td>
<td>( u'_j )</td>
<td>( y_\leq )</td>
<td>( ru_i )</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( s \not\in C(P(v_i)) )</td>
<td>( u'_j )</td>
<td>( y_\leq )</td>
<td>( ru_i )</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>2(h)</td>
<td>( s \not\in C(P(v_i)) )</td>
<td>( u'_j )</td>
<td>( y_\leq )</td>
<td>( ru_i )</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( s \not\in C(P(v_i)) )</td>
<td>( u'_j )</td>
<td>( y_\leq )</td>
<td>( ru_i )</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>2(i)</td>
<td>( r \neq \varepsilon )</td>
<td>( s \not\in C(P(v_i)) )</td>
<td>( y_\leq )</td>
<td>( ru_i )</td>
<td>( r\gamma' )</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>( s \not\in C(P(v_i)) )</td>
<td>( y_\leq )</td>
<td>( ru_i )</td>
<td>( r\gamma' )</td>
<td></td>
<td>+</td>
</tr>
</tbody>
</table>

3. Definition 2(3) is satisfied with \( r \not\in C(S(u_i)) \), \( s \neq \varepsilon \), and either Condition (j) or (k) holds:

- (j) \( i = 0 \) and \( \|H(\gamma)\| = 1 \) (where \( \gamma \) is the partial word satisfying \( x = \gamma s \)).
- (k) \( i > 0 \) and \( \|H(u_i)\| = 1 \).

<table>
<thead>
<tr>
<th></th>
<th>( u_i )</th>
<th>( x )</th>
<th>( y )</th>
<th>( i )</th>
</tr>
</thead>
<tbody>
<tr>
<td>3(j)</td>
<td>( r \not\in C(S(u_i)) )</td>
<td>( s \neq \varepsilon )</td>
<td>( rx )</td>
<td>( \gamma s )</td>
</tr>
<tr>
<td>3(k)</td>
<td>( r \not\in C(S(u_i)) )</td>
<td>( s \neq \varepsilon )</td>
<td>( rx )</td>
<td>( v_is )</td>
</tr>
</tbody>
</table>

- If \( |v_i| \geq |v'_j| \), then let \( t \) be the partial word such that \( u'_j = u_it \) and \( v_i = tv'_j \). Then \( w \) is called \((i, j)\)-special if one of the above holds when referring to Definition 2 where \( u_i \) is replaced by \( u'_j \) and \( v_i \) by \( v'_j \).

The partial word \( w \) is called \((i, j)\)-nonspecial otherwise.

Note that the concept of special in Definition 3 is equivalent to the concept of \((0,0)\)-special in Definition 4.

We now describe our algorithm (based on Theorem 3) that outputs a critical factorization for a given partial word \( w \) with \( p'(w) > 1 \) and with one hole of length at least two when such a factorization exists, and that outputs “no critical factorization exists” otherwise.

**Algorithm 2.**

1. **Step 1:** Compute the nonempty suffixes of \( w \) with respect to \( \leq_l \) (say \( v_{|w|-1} <_l \cdots <_l v_0 \)) and the nonempty suffixes of \( w \) with respect to \( \leq_r \) (say \( v'_{|w|-1} <_r \cdots <_r v'_0 \)).

2. **Step 2:** Set \( i = 0 \) and \( j = 0 \), and \( mw_p = 0 \).
Step 3: If $i \geq |w| - 1$ or $j \geq |w| - 1$, then output “no critical factorization exists”.

Step 4: If $v_i < v_i^{j'}$, then update $j$ with $j + 1$ and go to Step 3. If $v_i^{j'} < v_i$, then update $i$ with $i + 1$ and go to Step 3.

Step 5: If $i > 0$ and $v_i^{j'} = w$, then update $j$ with $j + 1$ and go to Step 3. If $j > 0$ and $v_i = w$, then update $i$ with $i + 1$ and go to Step 3.

Step 6: Find partial words $u_i, u_i'$ such that $w = u_i v_i = u_i' v_i'$.

Step 7: If $|v_i| \leq |v_i^{j'}|$, then compute $p_i = p(w, |u_i| - 1)$. If $p_i \leq mw_p$, then move up which means to update $i$ with $i + 1$ and to go to Step 3. If $p_i > mw_p$, then update $mw_p$ with $p_i$ and find the partial word $t$ such that $u_i = u_i't$ and $v_i' = tv_i$. Then do one of the following:

1. If $p_i \geq |u_i|$ and $p_i \geq |v_i|$, then output $(u_i, v_i)$.

2. If $p_i < |u_i|$ and $p_i \leq |v_i|$, then find partial words $x, y, r, s$ satisfying Definition 2(1). Do one of the following:
   
   (a) If $(i, j) = (0, 0)$, then do one of the following: If $|H(x)| = 1$ and $|t| > |x|$ and $(r \notin C(S(u_i))$ or $s \notin C(P(v_i)))$, then $w$ is $(i, j)$-special according to Definition 4(1)(a) and move up. If $|H(x)| = 1$ and $|H(r)| \neq 1$ and $|t| \leq |x|$ and $r \notin C(S(u_i))$, then $w$ is $(i, j)$-special according to Definition 4(1)(b) and move up. Otherwise, output $(u_i, v_i)$.
   
   (b) If $(i, j) \neq (0, 0)$, then do one of the following: If $i = 0$ and $|H(y)| = 1$ and $(r \notin C(S(u_i))$ or $s \notin C(P(v_i)))$, then move up since $w$ is $(i, j)$-special according to Definition 4(1)(c). If $i > 0$ and $|H(y)| \neq 1$ and $(r \notin C(S(u_i))$ or $s \notin C(P(v_i)))$, then move up since $w$ is $(i, j)$-special according to Definition 4(1)(d). Otherwise, output $(u_i, v_i)$.

3. If $p_i \geq |u_i|$ and $p_i < |v_i|$, then find partial words $x, y, r, s$ satisfying Definition 2(2). If $(r \neq \varepsilon$ or $i > 0$ and $|H(s)| = 1$ and $s \notin C(P(v_i)))$, then $w$ is $(i, j)$-special according to Definition 4(2)(e) and move up. If $(r \neq \varepsilon$ or $i > 0$ and $|H(t')| = 1$ (where $t'$ is the suffix of length $|t|$ of $y$) and $s \notin C(P(v_i)))$, then $w$ is $(i, j)$-special according to Definition 4(2)(f) and move up. If $j > 0$ and $|H(r)| = 1$ and $s \notin C(P(v_i)))$, then $w$ is $(i, j)$-special according to Definition 4(2)(g) and move up. If $(i, j) \neq (0, 0)$ and $|H(u_i')| = 1$ and $s \notin C(P(v_i)))$, then $w$ is $(i, j)$-special according to Definition 4(2)(h) and move up. If $(i > 0$ or $(r \neq \varepsilon$ and $j > 0$) and $|H(r'y)| = 1$ and $s \notin C(P(v_i)))$, then $w$ is $(i, j)$-special according to Definition 4(2)(i) and move up. Otherwise, output $(u_i, v_i)$.

4. If $p_i < |u_i|$ and $p_i > |v_i|$, then find partial words $x, y, r, s$ satisfying Definition 2(3). Do one of the following:
   
   (a) If $i = 0$, then find the prefix $y$ of $x$ of length $|v_i|$. If $|H(y)| = 1$ and $r \notin C(S(u_i)))$, then move up since $w$ is $(i, j)$-special according to Definition 4(3)(j). Otherwise, output $(u_i, v_i)$.
   
   (b) If $i > 0$, then do one of the following: If $|H(u_i)| = 1$ and $r \notin C(S(u_i)))$, then move up since $w$ is $(i, j)$-special according to Definition 4(3)(k). Otherwise, output $(u_i, v_i)$.

Step 8: If $|v_i| > |v_i^{j'}|$, then do the work of Step 7 with $p_i^{j'}, u_i^{j'}$ and $v_i^{j'}$ instead of $p_i, u_i$ and $v_i$. Move up here means to update $j$ with $j + 1$ and to go to Step 3. The algorithm may produce $(u_i^{j'}, v_i^{j'})$ unless $w$ is $(i, j)$-special. In those cases, move up. The partial word $t$ here satisfies $u_i^{j'} = u_i t$ and $v_i = tv_i^{j'}$.
We illustrate Algorithm 2 with the following examples.

**Example 3.** Referring to Example 1 where \( w = aaa \triangleleft aabaaa \), the nonempty suffixes of \( w \) are ordered as follows (where \( \triangleleft \prec_l a \prec_l b \) and \( \triangleleft \prec_r b \prec_r a \)):

<table>
<thead>
<tr>
<th>( i )</th>
<th>( p_i )</th>
<th>( v_i )</th>
<th>( v'_i )</th>
<th>( p'_i )</th>
<th>( j )</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>∆aabaaa</td>
<td>∆aabaaa</td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>a</td>
<td>baaa</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>1 a∆aabaaa</td>
<td>a</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1 aa</td>
<td>a∆aabaaa</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1 a∆aabaaa</td>
<td>aaba</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>aaa</td>
<td>aa</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>a∆aabaaa</td>
<td>a∆aabaaa</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1 abaaaa</td>
<td>abaaaa</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1 abaaa</td>
<td>aaa</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>3 baaa</td>
<td>aaa∆aabaaa</td>
<td>0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Algorithm 2 starts with the pair \((v_0, v'_0)\) and selects the shortest component, that is, \( v_0 = baaa \). In Step 7, \( p_0 \) is computed as 3 and the algorithm discovers that \( w \) is \((0,0)\)-special according to Definition 4(1)(a). The variable \( i \) is then updated with 1 and the pair \((v_1, v'_0)\) is considered. In Step 5, \( j \) is updated with 1 since \( i = 1 > 0 \) and \( v'_j = v'_0 = w \). Now, the pair \((v_1, v'_1)\) is treated and the shortest component \( v'_1 = aaa \) gets selected with \( p'_1 = 4 \) calculated in Step 8. But \( w \) turns out to be \((1,1)\)-special according to Definition 4(3)(k) and no position leads to an improvement of the number 4.

Referring to Example 2 where \( w = aaab\triangleleft babb \), the nonempty suffixes of \( w \) are ordered as follows (where \( \triangleleft \prec_l a \prec_l b \) and \( \triangleleft \prec_r b \prec_r a \)):

<table>
<thead>
<tr>
<th>( i )</th>
<th>( p_i )</th>
<th>( v_i )</th>
<th>( v'_i )</th>
<th>( p'_i )</th>
<th>( j )</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>∆babbb</td>
<td>∆babbb</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>aab\triangleleft babbb</td>
<td>b</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>aab\triangleleft babbb</td>
<td>b\triangleleft babbb</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>ab\triangleleft babbb</td>
<td>bb</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>abb</td>
<td>babbb</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>b</td>
<td>ab\triangleleft babbb</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>9 b\triangleleft babbb</td>
<td>abb</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1 babbb</td>
<td>aab\triangleleft babbb</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>3 bb</td>
<td>aab\triangleleft babbb</td>
<td>0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Algorithm 2 starts with the pair \((v_0, v'_0)\) and selects the shortest component, that is, \( v_0 = bb \). In Step 7, \( p_0 \) is computed as 3 and the algorithm discovers that \( w \) is \((0,0)\)-special according
to Definition 4(3)(j). The variable \( i \) is then updated with 1 and the pair \((v_1, v'_0)\) is considered. In Step 5, \( j \) is updated with 1 since \( i = 1 > 0 \) and \( v'_j = v'_0 = w \). Now, the pair \((v_1, v'_1)\) is treated and the shortest component \( v_1 = \text{babb} \) gets selected with \( p_1 = 1 \) calculated in Step 7. Since \( p_1 < p_0, i \) is updated with 1 and the pair \((v_2, v'_1)\) is considered. In Step 7, the algorithm outputs \((u_2, v_2)\) with \( p_2 = 9 = p'(w) \) (here \( w \) is (2, 1)-nonspecial).

We now prove Theorem 3.

**Theorem 3.** Let \((i, j)\) be a pair of nonnegative integers being considered at Step 7 (when \( p_i > mw_p \) or at Step 8 (when \( p'_j > mw_p \)). If \( w \) is an \((i, j)\)-nonspecial partial word with one hole such that \(|w| \geq 2 \) and \( p'(w) > 1 \), then \( w \) has at least one critical factorization. More specifically, the factorization \((u_i, v_i)\) is critical when \(|v_i| \leq |v'_j|\), and the factorization \((u'_j, v'_j)\) is critical when \(|v_i| > |v'_j|\). Moreover, if \(|v_i| \leq |v'_j|\) and the factorization \((u_i, v_i)\) is critical, then \( w \) is \((i, j)\)-nonspecial, and if \(|v_i| > |v'_j|\) and the factorization \((u'_j, v'_j)\) is critical, then \( w \) is \((i, j)\)-nonspecial.

**Proof.** The pair \((i, j) = (0, 0)\) was treated in Theorem 2. So, we may assume that \((i, j) \neq (0, 0)\). We treat the case where \(|v_i| \leq |v'_j|\) (the case where \(|v_i| > |v'_j|\) is treated similarly but requires that the orderings \( \leq_i \) and \( \leq_r \) be interchanged). Put \( u_i = u'_j t \) and \( v'_j = tv_i \) for some \( t \). Here, \( u_i \neq \varepsilon \) unless \( v_i = v'_j = w \). In such case where \( v_i = v'_j = w \), if \( w \) begins with \( \omega \), then \( i = |w| - 1 \) and \( j = |w| - 1 \) and the algorithm discovers in Step 3 that \( w \) has no critical factorization, and if \( w = az = v_i = v'_j \) for some \( a \in A \), then \( i < |w| - 1 \) and \( j < |w| - 1 \). In the latter case, Steps 4 or 5 will increase \( i \) or \( j \) by 1 resulting in the pair \((i + 1, j)\) or \((i, j + 1)\) being treated.

We now consider the following cases where \( p_i \) denotes \( p(w, |u_i| - 1) \).

**Case 1:** \( p_i > |u_i| \) and \( p_i > |v_i| \)

If \( p_i > |u_i| \) and \( p_i > |v_i| \), then Condition (4) is satisfied. There exist \( x, y, r, s \in W(A) \) such that \(|x| = p_i, x \uparrow y, x = ru_i \) and \( y = v_i s \).

First, assume that \(|r| > |v_i|\). By Lemma 1, there exist \( r', z \in W(A) \) such that \( r = r'z, r' \uparrow v_i, \) and \( zu_i \uparrow s \). Thus, \( p_i = |x| = |ru_i| = |r'zu_i| > |u_i v_i| = |w| \geq p'(w) \), which leads to a contradiction.

Now, assume that \(|r| \leq |v_i|\). By Lemma 1, there exist \( r', z \in W(A) \) such that \( v_i = r'z, r \uparrow r', \) and \( u_i \uparrow zs \). If \( \|H(u_i)\| = 1 \) or \( \|H(r')\| = 1 \), then \( u_i \subset zs \). Thus, \( u_i v_i \subset zsr'z \) showing that \( p_i = |zsr'| \) is a weak period of \( u_i v_i \). If \( \|H(z)\| = 1 \), then \( zs \subset u_i \). Put \( u_i = z's \) with \( z \subset z' \). Thus, \( u_i v_i = z'sr'z \subset z'sr'z' \) showing that \( p_i = |z'sr'| \) is a weak period of \( u_i v_i \). In any case, \( p'(w) \leq p_i \). On the other hand, \( p'(w) \geq p_i \). Therefore, \( p'(w) = p_i \) which shows that the factorization \((u_i, v_i)\) is critical.

**Case 2:** \( p_i < |u_i| \) and \( p_i > |v_i| \)

If \( p_i < |u_i| \) and \( p_i > |v_i| \), then Condition (3) is satisfied. There exist \( x, y, r, s, \gamma \in W(A) \) such that \(|x| = p_i, \gamma \uparrow v_i, u_i = rx = r's, \) and \( y = v_i s \). Note that if \( i = 0 \) and \( v_i \subset \gamma \), then \( y \subset x \) and we get a contradiction with Lemma 7.

We first assume that \( v_i \subset \gamma \), and thus we may assume that \( i > 0 \). Here \( v_i <_l \gamma s v_i = xv_i \) and the suffix \( xv_i \) was considered before. Let \( q \) denote the minimal local period of \( w \) at
position $|r| - 1$ (here $q < p_i$). If $p_i$ is a weak period of $w$, then there exist a word $z$ of length $q$ and words $x, \beta$ such that $xz = z\beta = x$ and either $xz$ is a suffix of $r$ or $r$ is a suffix of $xz$. In such case, since $z$ is full, by Lemma 2, there exist words $t_1, t_2$ such that $x \subseteq t_1 t_2, \beta \subset t_2 t_1$, and $z \subset (t_1 t_2)^m t_1$ for some integer $m \geq 0$. Let us first assume that $t_1 \neq e$. In this case, note that $m$ needs to be zero since $q$ is minimal. Thus, $y \subset z\beta \subset t_1 t_2 t_1$. Consequently, we obtain $|t_1 t_2 t_1| = p_i \leq |t_1|$ leading to a contradiction. Now, if $t_1 = e$, then $m$ needs to be one since $q$ is minimal. Thus, $y \subset x = z\beta \subset t_2 t_1$. Consequently, we obtain $|t_2 t_1| = p_i \leq |t_2|$ leading to a contradiction.

We now assume that $\gamma \subset v_i$. If $r \notin C(S(u_i))$, then $w$ is $(i, j)$-special by Definition 4(3) (note that in the case where $i = 0$, we have $\gamma \subset v_i$). If $r \in C(S(u_i))$, there then exist words $t' \in W(x)$ such that $x' r \uparrow x$. If $r$ is full, then the result follows as in Case 2 of Theorem 2. If $r$ has a hole, then by Lemma 4, two possibilities can occur. **Possibility 1:** there exist partial words $t_1, t_2, t_3, t_4$ such that $x' \equiv t_3 t_2, x = t_2 t_4, t_1 \subset t_3, t_1 \subset t_4$, and $r = (t_3 t_2)^m t_1 (t_2 t_4)^n$ for some integers $m, n \geq 0$. We get $u_i v_i = r x v_i = (t_3 t_2)^m t_1 (t_2 t_4)^n v_i$ with $v_i \in P(t_2 t_4)$. Here $p_i = |x| = |t_3 t_2| = |t_1 t_2| = |t_4 t_2|$ is a weak period of $u_i v_i$. **Possibility 2:** there exist partial words $t_1, t_2, t_3, t_4$ such that $x' \equiv t_1 t_3, x = t_4 t_1, t_2 \subset t_3, t_2 \subset t_4$, and $r = (t_1 t_3)^m t_1 t_2 (t_1 t_4)^n$ for some integers $m, n \geq 0$. We get $u_i v_i = r x v_i = (t_1 t_3)^m t_1 t_2 (t_1 t_4)^n v_i$ with $v_i \in P(t_4 t_1)$. Here $p_i = |x| = |t_1 t_3| = |t_1 t_2| = |t_1 t_4|$ is a weak period of $u_i v_i$.

**Case 3:** $p_i < |u_i|$ and $p_i \leq |v_i|$. If $p_i < |u_i|$ and $p_i \leq |v_i|$, then Condition (1) is satisfied. There exist $x, y, r, s \in W(A)$ such that $|x| = p_i, x \uparrow y, u_i = r x$, and $v_i = y s$. Note that if $i = 0$ and $y \subset x$, then we get a contradiction with Lemma 6.

We first assume that $y \subset x$, and thus we may assume that $i > 0$. Here $v_i = y s \prec_l y x s = x v_i$, and the suffix $x v_i$ was considered before. We can argue similarly as in Case 2 where $v_i \subset \gamma$.

We now assume that $x \subset y$, and thus $y$ is full. Here $w$ is $(i, j)$-special by Definition 4(1) unless $r \in C(S(u_i))$ and $s \in C(P(v_i))$. If the two conditions hold, then $x' r \uparrow x r$ and $y s \uparrow y s'$ for some $x', y'$. We consider whether or not $r, s$ are full.

- If $r$ and $s$ are full, then by Lemma 2, there exist words $t_1, t_2$ such that $x' \subset t_1 t_2, x \subset t_2 t_1$, and $r \subset (t_1 t_2)^m t_1$ for some integer $m \geq 0$, and words $t_3, t_4$ such that $y \subset t_3 t_4, y' \subset t_4 t_3$, and $s \subset (t_3 t_4)^n t_3$ for some integer $n \geq 0$. The result follows as in Case 3 of Theorem 2.

- If $r$ has a hole, then $s$ is full and words $t_3, t_4$ exist as above. By Lemma 4, two possibilities can occur. **Possibility 1:** there exist partial words $t'_1, t'_2, t'_3, t'_4$ such that $x' = t'_4 t'_2, x = t'_2 t'_4$, $t'_1 \subset t'_3, t'_1 \subset t'_4$, and $r = (t_3 t_2)^m t'_1 (t'_2 t'_4)^n$ for some integers $m', n' \geq 0$. We get $u_i v_i = r x y s = (t_3 t_2)^m t'_1 (t'_2 t'_4)^n t'_1 (t_3 t_4)^n t_3$ with $t'_2 t'_4 = x = y = t_3 t_4$. Here, $p_i = |x| = |t'_3 t'_2| = |t'_1 t'_2| = |t'_4 t'_2|$ is a weak period of $u_i v_i$. **Possibility 2:** there exist partial words $t'_1, t'_2, t'_3, t'_4$ such that $x' = t'_4 t'_3, x = t'_4 t'_1, t'_2 \subset t'_3, t'_2 \subset t'_4$, and $r = (t_1 t_3)^m t'_1 t'_2 (t_1 t_4)^n t'_1$ for some integers $m', n' \geq 0$. We get $u_i v_i = r x y s = (t_1 t_3)^m t'_1 t'_2 (t_1 t_4)^n t'_1 t_3$ with $t'_4 t'_1 = x = y = t_3 t_4$. Here $p_i = |x| = |t'_4 t'_3| = |t'_1 t'_2| = |t'_4 t'_1|$ is a weak period of $u_i v_i$.

- If $s$ has a hole, then $r$ is full and words $t_1, t_2$ exist as above. This case follows also from Lemma 4.

**Case 4:** $p_i \geq |u_i|$ and $p_i < |v_i|$
If \( p_i \geq |u_i| \) and \( p_i < |v_i| \), then Condition (2) is satisfied. There exist \( x, y, r, s \in W(A) \) such that \( |x| = p_i, x \uparrow y, x = ru_i = ru_j \) and \( v_i = y \). Since \( ru_i \uparrow y \), there exist \( y, t' \in W(A) \) such that \( y = r\gamma t', u' \uparrow \gamma \) and \( t' = \gamma t' \). We can rewrite \( v_i, v'_j \) as \( v_i = r\gamma t' s \) and \( v'_j = tu_i = tr\gamma t' s \). Note that if \( i = 0 \) and \( r = -e \) and \( y \subset x \), then we get a contradiction with Lemma 6.

First, assume that \( t \sqsubset t' \). The case where \( j = 0 \) leads to a contradiction, and the case where \( j > 0 \) leads to \( s \leq l_i v_i \). To see this, if \( j = 0 \), then since \( v'_j \) is maximal with respect to \( \leq_r \), the suffix \( t's \) of \( v'_j \) satisfies \( t's \leq_r v'_j \). This implies \( t's \leq_r tv_i \), which yields a contradiction. If \( j > 0 \), then assume first that \( i = 0 \). In this case, the relationship holds since \( v_i \) is maximal with respect to \( \leq_l \) and \( s \) is a suffix of \( v_i \). If \( i > 0 \), then assume to the contrary that \( v_i \leq_l s \). The suffix \( s \) was considered before, and let \( q \) denote the minimal local period of \( w \) at position \( |u_i| - 1 \) (here \( q < p_i \)). If \( p_i \) is a weak period of \( w \), then there exists a word \( z \) of length \( q \) and words \( z', \beta \) such that \( y = zz = z\beta \) and either \( z\beta \) is a prefix of \( s \) or \( s \) is a prefix of \( z\beta \). Here, we can argue similarly as in Case 2.

Now, since \( s \leq q \leq v_i \), we have \( s \in P(v_i) \) or \( v_i <_r s \). If \( v_i <_r s \), then \( w \) is \((i, j)\)-special by Definition 4.2. If \( s \in P(v_i) \), then \( ys = a_j \) for some \( a_j \). By Lemma 2, there exist words \( t_1, t_2 \) such that \( y = t_1 t_2, y' = t_2 t_1 \), and \( s = (t_1 t_2)^m t_1 t_2 \) for some integer \( m \geq 0 \). Let \( t_3 \) be the suffix of length \( \gamma t' \) of \( t_1 t_2 \). We have \( u_i v_i \gamma t' s \subset t_3(t_1 t_2)^m t_1 t_2 \) and \( p_i = |t_1 t_2| \) is a weak period of \( u_i v_i \).

Second, assume that \( t' \subset t \). Here, \( w \) is \((i, j)\)-special by Definition 4.2 unless \( s \in C(P(v_i)) \). If \( s \in C(P(v_i)) \), then \( ys \uparrow a_j \) for some \( a_j \).

- If \( s \) is full, then by Lemma 2, there exist words \( t_1, t_2 \) such that \( y \subset t_1 t_2, y' \subset t_2 t_1 \), and \( s \subset (t_1 t_2)^k t_1 \) for some integer \( k \geq 0 \). If \( u'_j \subset \gamma \), then \( x \subset y \subset t_1 t_2, ru_i v_i = xys \subset (t_1 t_2)^{k+2} t_1 \) and \( p_i = |x| = |t_1 t_2| \) is a weak period of \( u_i v_i \). If \( \gamma \subset u'_j \), then \( ru_i v_i = xys \subset xy(t_1 t_2)^k t_1 \) with \( y \subset x \) and \( y \subset t_1 t_2 \), and \( p_i = |x| = |t_1 t_2| \) is a weak period of \( u_i v_i \).
- If \( s \) has a hole, then two possibilities can occur as in Case 4 of Theorem 2 where \( t' = t \) and \( s \) is not full.

We end this section with a characterization of the special partial words that admit a critical factorization. If \( w \) is such a special partial word with one hole satisfying \( |w_0| \leq |v'_0| \), then the minimal local period of \( w \) at position \( |u_0| - 1 \), or \( p_0 \), is smaller than the desired number \( p'(w) \). The following theorem gives a bound of how far \( p_0 \) is from \( p'(w) \) (in some instances \( p_0 < p'(w) \leq 2p_0 \) say) and gives an explanation why Algorithm 2 is faster in average than a trivial algorithm where every position would be tested for critical factorization.

**Theorem 4.** Let \( w \) be a special partial word with one hole that admits a critical factorization, and let \( v_0 \) (respectively, \( v'_0 \)) be the maximal suffix with respect to \( \leq_l \) (respectively, \( \leq_r \)). Let \( u_0, u'_0 \) be partial words such that \( w = u_0 v_0 = u'_0 v'_0 \).

- If \( |v_0| \leq |v'_0| \), then the following hold:
  1. If \( w \) is special according to Definition 3(1), then one of the following holds:
2. If \( w \) is special according to Definition 3(2), then one of the following holds where \( r \) is the nonempty prefix of length \( p_0 - |u_0| \) of \( v_0 \):

(a) If \( w \) is special according to Definition 3(2)(c), then there exist partial words \( t_1, t_2 \) and a nonnegative integer \( n \) such that

\[
\begin{align*}
  u_0 &= (t_1 t_2)^n t_1, \\
  v_0 &= P((r(t_3 t_2)^m t_1 (t_2 t_4)^n) \omega), \\
  p_0 &= |(r(t_3 t_2)^m t_4 (t_2 t_4)^n)\omega| < |(r(t_3 t_2)^m t_4 (t_2 t_4)^n)\omega| = p'(w), \quad (m = 0 \text{ or } n > 0).
\end{align*}
\]

(b) If \( w \) is special according to Definition 3(2)(d), then one of the following holds:

(i) There exist partial words \( t_1, t_2, t_3, t_4, t_5 \) and nonnegative integers \( m, n \) such that \( t_1 \sqsubset t_3, t_1 \sqsubset t_4, t_1 \sqsubset t_5 \),

\[
\begin{align*}
  u_0 &= (t_3 t_2)^m t_4 (t_2 t_4)^n, \\
  v_0 &= P((r(t_3 t_2)^m t_1 (t_2 t_4)^n)\omega (r(t_3 t_2)^m t_5 (t_2 t_4)^n)\omega), \\
  p_0 &= |(r(t_3 t_2)^m t_4 (t_2 t_4)^n)\omega| < |(r(t_3 t_2)^m t_4 (t_2 t_4)^n)\omega| = p'(w).
\end{align*}
\]

(ii) There exist partial words \( t_1, t_2, t_3, t_4, t_5 \) and nonnegative integers \( m, n \) such that \( t_2 \sqsubset t_3, t_2 \sqsubset t_4, t_2 \sqsubset t_5 \),

\[
\begin{align*}
  u_0 &= (t_1 t_3)^m (t_1 t_4)^n+1 t_1, \\
  v_0 &= P((r(t_1 t_3)^m t_1 t_2 (t_1 t_4)^n+1 t_1 (r(t_1 t_3)^m t_5 (t_1 t_4)^n+1 t_1)\omega), \\
  p_0 &= |(r(t_1 t_3)^m (t_1 t_4)^n+1 t_1)\omega| < |(r(t_1 t_3)^m (t_1 t_2 (t_1 t_4)^n+1 t_1)\omega| = p'(w) < 2p_0.
\end{align*}
\]
3. If \( w \) is special according to Definition 3(3), then one of the following holds where \( s \) is the nonempty suffix of length \( p_0 - |v_0| \) of \( u_0 = \omega \):

(a) There exist partial words \( t_1, t_2, t_3, t_4, t_5 \) and nonnegative integers \( m, n \) such that
\[
\begin{align*}
u_0 &\in S(((t_3t_2)^m t_5(t_2t_4)^n s)\omega(t_3t_2)^m t_1(t_2t_4)^n s), \\
v_0 &= (t_3t_2)^m t_3(t_2t_4)^n,
\end{align*}
\]
and
\[
p_0 = |(t_3t_2)^m t_1(t_2t_4)^n s| < |(t_3t_2)^m t_1(t_2t_4)^n s| = p'(w).
\]

(b) There exist partial words \( t_1, t_2, t_3, t_4, t_5 \) and nonnegative integers \( m, n \) such that
\[
\begin{align*}
u_0 &\in S(((t_3t_2)^m t_5(t_2t_4)^n s)\omega(t_3t_2)^m t_1 t_2(t_1t_4)^n t_1 s), \\
v_0 &= (t_3t_2)^m t_1(t_1t_4)^n t_1,
\end{align*}
\]
and
\[
p_0 = |(t_3t_2)^m t_1 t_2(t_1t_4)^n t_1 s| < |(t_3t_2)^m t_1 t_2(t_1t_4)^n t_1 s| = p'(w) < 2p_0.
\]

- If \( |v_0| \geq |v'_0| \), then the above hold when replacing \( u_0, v_0, p_0 \) by \( u'_0, v'_0, p'_0 \), respectively.

**Proof.** Let us first prove Statement 1. Let \( x, y, r, s \in W(A) \) be such that \( |x| = p_0, x \uparrow y, u_0 = rx, \) and \( v_0 = ys \). Since \( w \) admits a critical factorization, there exists \( (i, j) \neq (0, 0) \) such that \( w \) is \( (i, j) \)-nonspecial and either \((u_i, v_i)\) (if \( |v_i| \leq |v'_j| \)) or \((u'_j, v'_j)\) (if \( |v_i| > |v'_j| \)) is critical with minimal local period \( q \) (here \( p_0 < q = p'(w) \)). Let \( \alpha, \beta \in A^+ \) be such that \( \alpha x \uparrow y \beta, |\alpha x| = |y \beta| = q \), either \( u_0 \) is a suffix of \( \alpha x \) or \( \alpha x \) is a suffix of \( u_0 \), and either \( y \beta \) is a prefix of \( v_0 \) or \( v_0 \) is a prefix of \( y \beta \). Since \( ||H(x)|| = 1 \), by Lemma 5(1), one of the following holds:

**Possibility 1:** There exist partial words \( t_1, t_2, t_3, t_4 \) such that \( x = t_3t_2, \beta = t_2t_4, t_1 \sqsubset t_3, t_1 \sqsubset t_4, x = (t_3t_2)^m t_1(t_2t_4)^n, \) and \( y = (t_3t_2)^m t_3(t_2t_4)^n \) for some integers \( m, n \geq 0 \). Since \( q \) is the minimal weak period of \( w \), a word \( t_5 \) exists as desired and the result follows. Since \( p_0 \) is minimal, we have \( m = 0 \) or \( n > 0 \). **Possibility 2:** There exist partial words \( t_1, t_2, t_3, t_4 \) such that \( x = t_1t_3, \beta = t_4t_1, t_2 \sqsubset t_3, t_2 \sqsubset t_4, x = (t_1t_3)^m t_1t_2(t_1t_4)^n t_1, \) and \( y = (t_1t_3)^m t_1 t_2(t_1t_4)^n t_1 \) for some integers \( m, n \geq 0 \). Here \( t_1 = \varepsilon \) and \( (m = 0 \text{ or } n > 0) \) since \( p_0 \) is minimal, and the result follows.

Now, let us prove Statement 3. Let \( x, y, r, s \in W(A) \) be such that \( |x| = p_0, x \uparrow y, u_0 = rx = r'ys, \) and \( y = v_0s \). Since \( w \) admits a critical factorization, there exist \( (i, j) \) and \( q \) as above. Let \( \alpha, \beta \in A^+ \) be such that \( \alpha x = \omega \beta, |\alpha x| = |\omega \beta| = |y \beta| = q \), and \( u_0 \) is a suffix of \( \alpha x \) or \( \alpha x \) is a suffix of \( u_0 \). By simplification, \( \alpha x \uparrow v_0 \beta \), and we also have \( \gamma \uparrow v_0 \). The result follows similarly as above since \( ||H(\gamma)|| = 1 \).

Finally, let us prove Statement 2. Let \( x, y, r, s \in W(A) \) be such that \( |x| = p_0, x \uparrow y, x = ru_0 = ru'_{0'}t, v_0 = ys = r'_{0'}t' \) where \( t' \) is the suffix of length \( |t| \) of \( y \), and \( (i, j) \) and
For Statement 2(a), let $x, \beta$ be such that $ru_0t \uparrow r\gamma t'\beta$, $|zx| = |ru_0t| = |r\gamma t'\beta| = |y\beta| = q$, and either $y\beta$ is a prefix of $v_0$ or $v_0$ is a prefix of $y\beta$. By simplification, $zu_0t \uparrow \gamma t'\beta$, and we also have $u_0t = \gamma t' \in A^*$ since $\|H(s)\| = 1$. The result follows using Lemma 2. For Statement 2(b), $zu_0t \uparrow \gamma t'\beta$ with $x, \beta \in A^+$, and $u_0t \uparrow \gamma t'$ with $\|H(t')\| = 1$. The result follows from Lemma 5(2). □

7. Conclusion

In this paper, we considered one of the most fundamental results on periodicity of words, namely the critical factorization theorem. Given a word $w$ and nonempty words $u, v$ satisfying $w = uv$, the minimal local period associated to the factorization $(u,v)$ is the length of the shortest square at position $|u| - 1$. It is easy to see that no minimal local period is longer than the minimal period of the word. The critical factorization theorem on words shows that critical factorizations are unavoidable. Indeed, for any word, there is always a factorization whose minimal local period is equal to the minimal period of the word. A critical factorization can be found efficiently from the computation of the maximal suffixes of the word with respect to two total orderings on words: the lexicographic ordering related to a fixed total ordering on the alphabet, and the lexicographic ordering obtained by reversing the order of letters in the alphabet.

Our goal was to extend the critical factorization theorem to partial words with one hole (such sequences contain a do not know symbol). In this case, we called a factorization critical if its minimal local period is equal to the minimal weak period of the partial word. It turned out that for partial words, critical factorizations may be avoidable. We characterized the class of the so-called special partial words with one hole that possibly avoid critical factorizations. We first gave a version of the critical factorization theorem for the nonspecial partial words with one hole. Then, by refining the method based on the maximal suffixes with respect to the lexicographic/reverse lexicographic orderings, we gave a version of the critical factorization theorem for the so-called $(i,j)$-nonspecial partial words with one hole. Our proof led to an efficient algorithm which, given a partial word with one hole, outputs a critical factorization when one exists or outputs “no such factorization exists”.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. CCF-0207673. We thank Ajay Chriscoe for very valuable comments and suggestions, and for implementing Algorithm 2 and creating a World Wide Web site for this research. We also thank the referee of a preliminary version of this paper for his/her very valuable comments and suggestions.

References