

---

This item was submitted to [Loughborough's Research Repository](#) by the author.  
Items in Figshare are protected by copyright, with all rights reserved, unless otherwise indicated.

## Weakly unambiguous morphisms

PLEASE CITE THE PUBLISHED VERSION

<http://dx.doi.org/10.1016/j.tcs.2012.04.045>

PUBLISHER

© Elsevier

VERSION

AM (Accepted Manuscript)

LICENCE

CC BY-NC-ND 4.0

REPOSITORY RECORD

Freydenberger, Dominik D., Hossein Nevisi, and Daniel Reidenbach. 2019. "Weakly Unambiguous Morphisms". figshare. <https://hdl.handle.net/2134/10120>.

This item was submitted to Loughborough's Institutional Repository (<https://dspace.lboro.ac.uk/>) by the author and is made available under the following Creative Commons Licence conditions.



For the full text of this licence, please go to:  
<http://creativecommons.org/licenses/by-nc-nd/2.5/>

# Weakly Unambiguous Morphisms<sup>☆</sup>

Dominik D. Freydenberger<sup>a</sup>, Hossein Nevisi<sup>b,\*</sup>, Daniel Reidenbach<sup>b</sup>

<sup>a</sup>*Institut für Informatik, Goethe-Universität, Frankfurt am Main, Germany*

<sup>b</sup>*Department of Computer Science, Loughborough University, Loughborough, Leicestershire, LE11 3TU, United Kingdom*

---

## Abstract

A nonerasing morphism  $\sigma$  is said to be weakly unambiguous with respect to a word  $s$  if  $\sigma$  is the only nonerasing morphism that can map  $s$  to  $\sigma(s)$ , i. e., there does not exist any other nonerasing morphism  $\tau$  satisfying  $\tau(s) = \sigma(s)$ . In the present paper, we wish to characterise those words with respect to which there exists such a morphism. This question is nontrivial if we consider so-called length-increasing morphisms, which map a word to an image that is strictly longer than the word. Our main result is a compact characterisation that holds for all morphisms with ternary or larger target alphabets. We also comprehensively describe those words that have a weakly unambiguous length-increasing morphism with a unary target alphabet, but we have to leave the problem open for binary alphabets, where we can merely give some non-characteristic conditions.

*Keywords:*

Nonerasing morphisms, ambiguity

---

## 1. Introduction

For any alphabets  $\mathcal{A}$  and  $\mathcal{B}$ , a morphism  $\sigma : \mathcal{A}^* \rightarrow \mathcal{B}^*$  is said to be *ambiguous* with respect to a word  $s$  if there exists a second morphism  $\tau : \mathcal{A}^* \rightarrow \mathcal{B}^*$  mapping  $s$  to the same image as  $\sigma$ ; if such a morphism  $\tau$  does not exist, then  $\sigma$  is called *unambiguous* (with respect to  $s$ ). For example, if we consider  $\mathcal{A} := \{A, B, C\}$ ,  $\mathcal{B} := \{a, b\}$  and  $s := A B B C A C$ , then the morphism  $\sigma$ , defined by  $\sigma(A) := abb$ ,  $\sigma(B) := abbb$ ,  $\sigma(C) := abbbb$ , is ambiguous with respect to  $s$ , since there exists a different morphism  $\tau$ , given by  $\tau(A) := abbab$ ,  $\tau(B) := bbab$ ,  $\tau(C) := bbb$ , satisfying  $\tau(s) = \sigma(s)$ :

$$\begin{array}{cccccc} \sigma(A) & \sigma(B) & \sigma(B) & \sigma(C) & \sigma(A) & \sigma(C) \\ \underbrace{a b b}_{\tau(A)} & \underbrace{a b b b}_{\tau(B)} & \underbrace{a b b b}_{\tau(B)} & \underbrace{a b b b b}_{\tau(C)} & \underbrace{a b b}_{\tau(A)} & \underbrace{a b b b b}_{\tau(C)} \\ \end{array} .$$

---

<sup>☆</sup>A preliminary version [2] of this paper was presented at the conference STACS 2011.

\*Corresponding author.

*Email addresses:* [freydenberger@em.uni-frankfurt.de](mailto:freydenberger@em.uni-frankfurt.de) (Dominik D. Freydenberger),  
[H.Nevisi@lboro.ac.uk](mailto:H.Nevisi@lboro.ac.uk) (Hossein Nevisi), [D.Reidenbach@lboro.ac.uk](mailto:D.Reidenbach@lboro.ac.uk) (Daniel Reidenbach)

In contrast to this, as can be verified with little effort, the morphism  $\sigma' : \mathcal{A}^* \rightarrow \mathcal{B}^*$ , defined by  $\sigma'(A) := \sigma'(C) := a$  and  $\sigma'(B) := b$ , is unambiguous with respect to  $s$ .

The potential ambiguity of morphisms is not only a fundamental phenomenon in combinatorics on words, but it also shows connections to various concepts in computer science. This particularly holds for *equality sets* (and, hence, the *Post Correspondence Problem*, see Harju and Karhumäki [6]) and *pattern languages* (see Mateescu and Salomaa [8]). Regarding the latter topic, insights into the ambiguity of morphisms have been used to solve a number of prominent problems (see, e. g., Reidenbach [9, 10, 11]), revealing that unambiguous morphisms, in a setting where various morphisms are applied to the same word, have the ability to optimally encode information about the structure of the word. This shows an interesting contrast to the foundations of coding theory (see Berstel and Perrin [1]), which is based on *injective* morphisms.

Since unambiguity can, thus, be seen as a desirable property of morphisms, the initial work on this topic by Freydenberger, Reidenbach and Schneider [4] and most of the subsequent papers have focused on the following question:

**Problem 1.** Let  $s$  be a word over an arbitrary alphabet. Does there *exist* a morphism (preferably with a finite target alphabet comprising at least two letters) that is *unambiguous* with respect to  $s$ ?

In order to further qualify this problem, [4] introduces two types of unambiguity: The first type follows our intuitive definition given above; more precisely, a morphism  $\sigma$  is called *strongly unambiguous* with respect to a word  $s$  if there does not exist a morphism  $\tau$  satisfying  $\tau(s) = \sigma(s)$  and, for a symbol  $x$  occurring in  $s$ ,  $\tau(x) \neq \sigma(x)$ . The second type slightly relaxes this requirement by calling  $\sigma$  *weakly unambiguous* with respect to  $s$  if there is no *nonerasing* morphism  $\tau$  (which means that  $\tau$  must not map any symbol to the empty word) showing the above properties. Thus, e. g., our initial example morphism  $\sigma$  is weakly unambiguous with respect to  $s' := A A B$ , but it is not strongly unambiguous, since the morphism  $\tau$ , given by  $\tau(A) := \varepsilon$  and  $\tau(B) := \sigma(s')$  (where  $\varepsilon$  stands for the empty word), satisfies  $\tau(s') = \sigma(s')$ . By definition, every strongly unambiguous *nonerasing* morphism is also weakly unambiguous, but – as shown by this example – the converse does not necessarily hold.

Apart from some very basic considerations, previous research has focussed on *strongly* unambiguous morphisms, partly giving comprehensive results on their existence; positive results along this line then automatically also hold for weak unambiguity. Freydenberger et al. [4] characterise those words with respect to which there exist strongly unambiguous nonerasing morphisms, and their characteristic criterion reveals that the existence of such morphisms is equivalent to a number of other vital properties of words, such as being a fixed point of a nontrivial morphism (see, e. g., Hamm and Shallit [5]) or being a shortest generator of a terminal-free E-pattern language. Freydenberger and Reidenbach [3], among other results, improve and deepen the techniques used in [4]. Schneider [14] studies the more general problem of the existence of arbitrary (i. e., possibly erasing) strongly unambiguous morphisms. While [14] provides a characterisation of those words that have a strongly unambiguous erasing morphism with an infinite target alphabet, a comprehensive result on finite target alphabets is still open. It is known, however, that a distinct characteristic criterion

is required for every alphabet size (unlike the restricted problem for strongly unambiguous nonerasing morphisms, the existence of which can be characterised for all non-unary alphabets identically), and that each of these criteria is NP-hard. Reidenbach and Schneider [13] continue this strand of research, demonstrating that the existence of strongly unambiguous erasing morphisms is closely related to decision problems for multi-pattern languages, and they show that the same criterion that characterises the existence of such morphisms for infinite target alphabets also, for all binary or larger alphabets, characterises the existence of erasing morphisms with a strongly restricted ambiguity.

In the present paper, we wish to investigate the existence of *weakly* unambiguous non-erasing morphisms; in other words, we initiate the research on the ambiguity of morphisms in free semigroups without empty word. When considering this problem as indicated above, we can already refer to a strong yet trivial insight mentioned by Freydenberger et al. [4], stating that there indeed is a weakly unambiguous morphism with respect to every word. More precisely, it directly follows from the definitions that every 1-uniform morphism (i. e., a morphism that maps each variable in the pattern to a word of length 1) is weakly unambiguous with respect to every word. Despite this immediate and unexciting observation, weak unambiguity deserves further research, since there are major fields of study that are exclusively based on nonerasing morphisms; this particularly holds for pattern languages, where so-called *nonerasing* (or *NE* for short) pattern languages have been intensively investigated. We therefore exclude the 1-uniform morphisms from our considerations and study *length-increasing* nonerasing morphisms instead, i. e., we deal with morphisms  $\sigma$  that, for the word  $s$  they are applied to, satisfy  $|\sigma(s)| > |s|$ . Hence, we wish to examine the following problem:

**Problem 2.** Let  $s$  be a word over an arbitrary alphabet. Does there exist a *length-increasing* nonerasing morphism that is *weakly* unambiguous with respect to  $s$ ?

Our results in the present paper shall provide a nearly comprehensive answer to this question, demonstrating that a combinatorially rich theory results from it. In particular, we show that the existence of weakly unambiguous length-increasing morphisms depends on the size of the target alphabet considered. However, unlike the above-mentioned result by Schneider [14] on the existence of strongly unambiguous erasing morphisms, we can give a compact and efficiently decidable characteristic condition on Problem 2, which holds for *all* target alphabets that consist of at least three letters and which describes a type of words we believe has not been discussed in the literature so far. Interestingly, this characterisation does not hold for binary target alphabets. In this case, we can give a number of strong conditions, but still do not even know whether Problem 2 is decidable. In contrast to this phenomenon, it is of course not surprising that for unary target alphabets again a different approach is required. Regarding this specification of Problem 2, we shall give a characteristic condition.

## 2. Definitions

For notations not explained explicitly, we refer the reader to Freydenberger et al. [4].

An *alphabet*  $\mathcal{A}$  is a nonempty set of symbols, and a *word* (over  $\mathcal{A}$ ) is a finite sequence of symbols taken from  $\mathcal{A}$ . We denote the empty word by  $\varepsilon$ . The notation  $\mathcal{A}^*$  refers to the set of all (empty and non-empty) words over  $\mathcal{A}$ , and  $\mathcal{A}^+ := \mathcal{A}^* \setminus \{\varepsilon\}$ . For the *concatenation* of two words  $w_1, w_2$ , we write  $w_1 \cdot w_2$  or simply  $w_1 w_2$ . The word that results from  $n$ -fold concatenation of a word  $w$  is denoted by  $w^n$ . The notation  $|x|$  stands for the size of a set  $x$  or the length of a word  $x$ . We call a word  $v \in \mathcal{A}^*$  a *factor* of a word  $w \in \mathcal{A}^*$  if, for some  $u_1, u_2 \in \mathcal{A}^*$ ,  $w = u_1 v u_2$ ; moreover, if  $v$  is a factor of  $w$  then we say that  $w$  *contains*  $v$  and denote this by  $v \sqsubseteq w$ . If  $v \neq w$ , then we say that  $v$  is a *proper* factor of  $w$  and denote this by  $v \sqsubset w$ . If  $u_1 = \varepsilon$ , then  $v$  is a prefix of  $w$ , and if  $u_2 = \varepsilon$ , then  $v$  is a suffix of  $w$ . For any words  $v, w \in \mathcal{A}^*$ ,  $|w|_v$  stands for the number of (possibly overlapping) occurrences of  $v$  in  $w$ . The symbol  $[\dots]$  is used to omit some canonically defined parts of a given word, e. g.,  $\alpha = 1 \cdot 2 \cdot [\dots] \cdot 5$  stands for  $\alpha = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5$ .

Let  $\mathbb{N} := \{1, 2, 3, \dots\}$  and  $\Sigma$  be alphabets. We call any symbol in  $\mathbb{N}$  a variable and any symbol in  $\Sigma$  a letter. In order to distinguish between a word over  $\mathbb{N}$  and a word over  $\Sigma$ , we call the former a *pattern*. We name patterns with lower case letters from the beginning of the Greek alphabet such as  $\alpha, \beta, \gamma$ . With regard to an arbitrary pattern  $\alpha$ ,  $\text{var}(\alpha)$  denotes the set of all variables occurring in  $\alpha$ .

A *morphism* is a mapping that is compatible with concatenation, i. e.,  $\sigma : \mathbb{N}^* \rightarrow \Sigma^*$  is a morphism if it satisfies  $\sigma(\alpha \cdot \beta) = \sigma(\alpha) \cdot \sigma(\beta)$  for all patterns  $\alpha, \beta \in \mathbb{N}^*$ . A morphism  $\sigma : \mathbb{N}^* \rightarrow \Sigma^*$  is called *nonerasing* provided that, for every  $i \in \mathbb{N}$ ,  $\sigma(i) \neq \varepsilon$ . If  $\sigma$  is nonerasing, then we often indicate this by writing  $\sigma : \mathbb{N}^+ \rightarrow \Sigma^+$ . A morphism  $\sigma$  is *length-increasing* (for  $\alpha$ ) if  $|\sigma(\alpha)| > |\alpha|$ , and it is called *1-uniform* if, for every  $i \in \mathbb{N}$ ,  $|\sigma(i)| = 1$ .

For any alphabet  $\Sigma$ , for any morphism  $\sigma : \mathbb{N}^+ \rightarrow \Sigma^+$  and for any pattern  $\alpha \in \mathbb{N}^+$ , we call  $\sigma$  *weakly unambiguous* with respect to  $\alpha$  if there is no morphism  $\tau : \mathbb{N}^+ \rightarrow \Sigma^+$  with  $\tau(\alpha) = \sigma(\alpha)$  and, for some variable  $q \in \text{var}(\alpha)$ ,  $\tau(q) \neq \sigma(q)$ . Moreover, for any morphism  $\sigma : \mathbb{N}^* \rightarrow \Sigma^*$ ,  $\sigma$  is said to be *strongly unambiguous* with respect to  $\alpha$ , if there is no morphism  $\tau : \mathbb{N}^* \rightarrow \Sigma^*$  with  $\tau(\alpha) = \sigma(\alpha)$  and, for some variable  $q \in \text{var}(\alpha)$ ,  $\tau(q) \neq \sigma(q)$ . On the other hand,  $\sigma$  is *ambiguous* with respect to  $\alpha$ , if there is a morphism  $\tau : \mathbb{N}^+ \rightarrow \Sigma^+$  with  $\tau(\alpha) = \sigma(\alpha)$  and, for some variable  $q \in \text{var}(\alpha)$ ,  $\tau(q) \neq \sigma(q)$ .

We now introduce some terminology that is helpful when comparing two morphisms that are applied to the same pattern, in terms of the positions of the letters in their images: Let  $\alpha := x_1 \cdot x_2 \cdot [\dots] \cdot x_n$ ,  $x_k \in \mathbb{N}$ ,  $1 \leq k \leq n$ , and let  $\sigma : \text{var}(\alpha)^+ \rightarrow \Sigma^+$  and  $\tau : \text{var}(\alpha)^+ \rightarrow \Sigma^+$  be morphisms. Assume that we are comparing  $\sigma(\alpha)$  with  $\tau(\alpha)$ . We say that  $\tau(x_i)$  is *located at the position of  $\sigma(x_i)$  in  $\sigma(\alpha)$*  if and only if

$$|\sigma(x_1 \cdot x_2 \cdot [\dots] \cdot x_{i-1})| < |\tau(x_1 \cdot x_2 \cdot [\dots] \cdot x_i)| \leq |\sigma(x_1 \cdot x_2 \cdot [\dots] \cdot x_i)|, \text{ and}$$

$$|\tau(x_1 \cdot x_2 \cdot [\dots] \cdot x_{i-1})| \geq |\sigma(x_1 \cdot x_2 \cdot [\dots] \cdot x_{i-1})|.$$

The following example illustrates this definition: Let  $\alpha := 1 \cdot 2 \cdot 3$ , and let the morphism  $\sigma : \mathbb{N}^+ \rightarrow \{a, b\}^+$  be given by  $\sigma(1) := a$ ,  $\sigma(2) := b$  and  $\sigma(3) := ab$ . Furthermore, let the morphism  $\tau : \mathbb{N}^+ \rightarrow \{a, b\}^+$  be defined by  $\tau(1) := ba$ ,  $\tau(2) := b$  and  $\tau(3) := b$ . Using the above terminology, we can say that  $\tau(3)$  is located at the position of  $\sigma(3)$ . However,  $\tau(1)$  and  $\tau(2)$  are not located at the positions of  $\sigma(1)$  and  $\sigma(2)$ .

The following concept is known to be vital for the research on the unambiguity of morphisms (see, e. g., Theorem 8 below): We call any  $\alpha \in \mathbb{N}^+$  *prolix* if and only if, there exists a factorisation  $\alpha = \beta_0 \gamma_1 \beta_1 \gamma_2 \beta_2 [\dots] \gamma_n \beta_n$  with  $n \geq 1$ ,  $\beta_k \in \mathbb{N}^*$  and  $\gamma_k \in \mathbb{N}^*$ ,  $k \leq n$ , such that

1. for every  $k$ ,  $1 \leq k \leq n$ ,  $|\gamma_k| \geq 2$ ,
2. for every  $k$ ,  $1 \leq k \leq n$  and, for every  $k'$ ,  $0 \leq k' \leq n$ ,  $\text{var}(\gamma_k) \cap \text{var}(\beta_{k'}) = \emptyset$ ,
3. for every  $k$ ,  $1 \leq k \leq n$ , there exists a variable  $i_k \in \text{var}(\gamma_k)$  such that  $|\gamma_k|_{i_k} = 1$  and, for every  $k'$ ,  $1 \leq k' \leq n$ , if  $i_k \in \text{var}(\gamma_{k'})$  then  $\gamma_k = \gamma_{k'}$ .

We call  $\alpha \in \mathbb{N}^+$  *succinct* if and only if it is not prolix. Thus, for example, the pattern  $1 \cdot 2 \cdot 3 \cdot 2 \cdot 4 \cdot 2 \cdot 1 \cdot 5 \cdot 5 \cdot 4 \cdot 2 \cdot 1 \cdot 1 \cdot 2 \cdot 3 \cdot 2$  is prolix (with  $\beta_0 := \varepsilon$ ,  $\gamma_1 := 1 \cdot 2 \cdot 3 \cdot 2$ ,  $\beta_1 := \varepsilon$ ,  $\gamma_2 := 4 \cdot 2 \cdot 1$ ,  $\beta_2 := 5 \cdot 5$ ,  $\gamma_3 := 4 \cdot 2 \cdot 1$ ,  $\beta_3 := \varepsilon$ ,  $\gamma_4 := 1 \cdot 2 \cdot 3 \cdot 2$ ,  $\beta_4 := \varepsilon$ ), whereas  $1 \cdot 2 \cdot 3 \cdot 3 \cdot 4 \cdot 2 \cdot 4 \cdot 2 \cdot 1$  is succinct.

Note that the set of succinct patterns is equivalent to the set of words that are not a *fixed point* of a nontrivial morphism. Furthermore, it corresponds to the set of *morphically primitive* words and the set of shortest generators of *terminal-free E-pattern languages*. These aspects are discussed by Reidenbach and Schneider [12] in more details.

### 3. Loyal neighbours

Before we begin our examination of Problem 2, we introduce some notions on structural properties of variables in patterns that shall be used in the subsequent sections.

In our first definition, we introduce a concept that collects the neighbours of a variable in a pattern.

**Definition 3.** Let  $\alpha \in \mathbb{N}^+$ . For every  $j \in \text{var}(\alpha)$ , we define the following sets:

$$\begin{aligned} L_j &:= \{k \in \text{var}(\alpha) \mid \alpha = \dots \cdot k \cdot j \cdot \dots\}, \\ R_j &:= \{k \in \text{var}(\alpha) \mid \alpha = \dots \cdot j \cdot k \cdot \dots\}. \end{aligned}$$

Moreover, if  $\alpha = j \dots$ , then  $\varepsilon \in L_j$ , and if  $\alpha = \dots j$ , then  $\varepsilon \in R_j$ .

Thus, the notation  $L_j$  refers to all left neighbours of variable  $j$  and  $R_j$  to all right neighbours of  $j$ . To illustrate these notions, we give an example.

**Example 4.** We consider  $\alpha := 1 \cdot 2 \cdot 3 \cdot 1 \cdot 4 \cdot 5 \cdot 6 \cdot 1 \cdot 4 \cdot 7 \cdot 8$ . For the variable 1, we have  $L_1 = \{\varepsilon, 3, 6\}$  and  $R_1 = \{2, 4\}$ .

We now introduce the concept of loyalty of neighbouring variables, which is vital for the examination of weakly unambiguous morphisms.

**Definition 5.** Let  $\alpha \in \mathbb{N}^+$ . A variable  $i \in \text{var}(\alpha)$  has *loyal neighbours (in  $\alpha$ )* if and only if at least one of the following cases is satisfied:

1.  $\varepsilon \notin L_i$  and, for every  $j \in L_i$ ,  $R_j = \{i\}$ , or
2.  $\varepsilon \notin R_i$  and, for every  $j \in R_i$ ,  $L_j = \{i\}$ .

Using the above definition, we can divide the variables of any pattern into two sets.

**Definition 6.** For any pattern  $\alpha \in \mathbb{N}^+$ ,  $|\alpha| \geq 2$ , let  $S_\alpha$  be the set of variables that have loyal neighbours and  $E_\alpha$  be the set of variables that do not have loyal neighbours in  $\alpha$ .

Note that in Definition 6 the notations  $S_\alpha$  and  $E_\alpha$  are short for “stable” and “(possibly) expanding”, respectively. These terms refer to the length of the morphic images of the variables in these sets under potentially unambiguous morphisms and, hence, anticipate some of the main results of the present paper (such as Theorem 13 and Corollary 19 below).

The following example clarifies these definitions.

**Example 7.** Let  $\alpha := 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 4 \cdot 3 \cdot 7 \cdot 8$ . Definition 3 implies that

$$\begin{aligned} L_1 &= \{\varepsilon\}, & L_2 &= \{1\}, & L_3 &= \{2, 4\}, & L_4 &= \{3, 6\}, \\ L_5 &= \{4\}, & L_6 &= \{5\}, & L_7 &= \{3\}, & L_8 &= \{7\}, \\ R_1 &= \{2\}, & R_2 &= \{3\}, & R_3 &= \{4, 7\}, & R_4 &= \{5, 3\}, \\ R_5 &= \{6\}, & R_6 &= \{7\}, & R_7 &= \{8\}, & R_8 &= \{\varepsilon\}. \end{aligned}$$

According to Definition 5, the variables 3 and 4 do not have loyal neighbours. Thus, due to Definition 6,  $S_\alpha = \{1, 2, 5, 6, 7, 8\}$  and  $E_\alpha = \{3, 4\}$ .

Freydenberger et al. [4] demonstrate that the partition of the set of all patterns into succinct and prolix ones is characteristic for the existence of strongly unambiguous nonerasing morphisms:

**Theorem 8** (Freydenberger et al. [4]). *Let  $\alpha \in \mathbb{N}^*$ , let  $\Sigma$  be an alphabet,  $|\Sigma| \geq 2$ . There exists a strongly unambiguous nonerasing morphism  $\sigma : \mathbb{N}^* \rightarrow \Sigma^*$  with respect to  $\alpha$  if and only if  $\alpha$  is succinct.*

Our subsequent remark shows that having a variable with loyal neighbours is a sufficient, but not a necessary condition for a pattern being prolix.

**Proposition 9.** *Let  $\alpha \in \mathbb{N}^+$ . If  $S_\alpha \neq \emptyset$ , then  $\alpha$  is prolix. In general, the converse of this statement does not hold true.*

*Proof.* Let  $i \in S_\alpha$ . According to Definition 5, one of the following cases is satisfied:

1.  $\varepsilon \notin L_i$  and, for every  $j \in L_i$ ,  $R_j = \{i\}$ , or
2.  $\varepsilon \notin R_i$  and, for every  $j \in R_i$ ,  $L_j = \{i\}$ .

Let  $\Sigma$  be an alphabet. For every nonerasing morphism  $\sigma : \mathbb{N}^* \rightarrow \Sigma^*$  over  $\alpha$ , we define a morphism  $\tau : \mathbb{N}^* \rightarrow \Sigma^*$  by, for every  $x \in \text{var}(\alpha)$ ,

$$\tau(x) := \begin{cases} \varepsilon, & x = i, \\ \sigma(x)\sigma(i), & \text{Case 1 is satisfied and } x \in L_i, \\ \sigma(i)\sigma(x), & \text{Case 1 is not satisfied, Case 2 is satisfied and } x \in R_i, \\ \sigma(x), & \text{else .} \end{cases}$$



It is easily verified that  $\tau(\alpha) = \sigma(\alpha)$ . Consequently, there is no strongly unambiguous nonerasing morphism  $\sigma$  with respect to  $\alpha$ . So, according to Theorem 8,  $\alpha$  is prolix.

For the second statement of Proposition 9, let  $\alpha := 1 \cdot 2 \cdot 2$ . It can be verified with little effort that  $\alpha$  is prolix, and  $S_\alpha = \emptyset$ .  $\square$

#### 4. Weakly unambiguous morphisms with $|\Sigma| \geq 3$

We now make use of the concepts introduced in the previous section to comprehensively solve Problem 2 for all but unary and binary target alphabets of the morphisms.

We start this section by giving some lemmata that are required when proving the main results of this paper. The first lemma is a general combinatorial insight that can be used in the proof of Lemma 11 – which, in turn, is a fundamental lemma in this paper.

**Lemma 10.** *Let  $v$  be a word and  $n$  be a natural number. If, for a word  $w$ ,  $w^n$  is a proper factor of  $v^n$ , then  $w$  is a proper factor of  $v$ .*

*Proof.* Let  $v^n := v_1 \cdot v_2 \cdot [\dots] \cdot v_n$  with, for every  $j$ ,  $1 \leq j \leq n$ ,  $v_j = v$ , and let  $w^n := w_1 \cdot w_2 \cdot [\dots] \cdot w_n$  with, for every  $k$ ,  $1 \leq k \leq n$ ,  $w_k = w$ . Moreover, assume that for every  $j$ ,  $1 \leq j \leq n$ ,  $v_j = p_j \cdot s_j$  such that  $p_j$  is an arbitrary nonempty prefix of  $v_j$  and,  $s_j$  is an arbitrary nonempty suffix of  $v_j$ . We assume to the contrary that  $w$  is not a proper factor of  $v$ . Consequently, for every  $j$ ,  $1 \leq j \leq n$ , and for every  $k$ ,  $1 \leq k \leq n$ ,  $w_k \not\sqsubseteq v_j$ . So, we can assume that  $w^n$  starts from the position of the first letter of  $s_q$ ,  $1 \leq q \leq n$ . Since  $w_1 \not\sqsubseteq v_q$ ,  $w_1 = s_q \cdot p_{q+1}$ . Then, due to  $w_2 \not\sqsubseteq v_{q+1}$ ,  $(q+1) \leq n$ , and  $w^n$  being a proper factor of  $v^n$ ,  $w_2 = s_{q+1} \cdot p_{q+2}$ ,  $(q+2) \leq n$ . If we continue the above reasoning, then  $w^{(n-q)}$  with

$$w^{n-q} = s_q \cdot p_{q+1} \cdot s_{q+1} \cdot p_{q+2} \cdot s_{q+2} \cdot p_{q+3} \cdot [\dots] \cdot s_{n-1} \cdot p_n$$

is a proper factor of  $v^n$ . Since  $p_n$  is a prefix of  $v_n$ , and  $w^n$  is a proper factor of  $v^n$ ,  $w_{n-q+1}w_{n-q+2}w_{n-q+3}[\dots]w_n$  must be a factor of  $s_n$ . Consequently,  $w^q$  must be a proper factor of  $v_n$ , and as a result  $w$  must be a proper factor of  $v$ , which is a contradiction.  $\square$

We continue our studies with the following lemma, which is a vital tool for the proof of many statements of this paper. It features an important property of two different morphisms that map a pattern to the same image.

**Lemma 11.** *Let  $\alpha \in \mathbb{N}^+$ ,  $|\alpha| \geq 2$ , and let  $\Sigma$  be an alphabet. Assume that  $\sigma : \mathbb{N}^+ \rightarrow \Sigma^+$  is a morphism such that, for a variable  $i \in \text{var}(\alpha)$ ,  $|\sigma(i)| \geq 2$  and, for every  $x \in \text{var}(\alpha) \setminus \{i\}$ ,  $|\sigma(x)| = 1$ . Moreover, assume that  $\tau$  is a nonerasing morphism satisfying  $\tau(\alpha) = \sigma(\alpha)$ . If there exists a variable  $j \in \text{var}(\alpha)$  with  $\tau(j) \neq \sigma(j)$ , then  $\tau(i) \sqsubset \sigma(i)$ .*

*Proof.* Assume to the contrary that there exists a variable  $j \in \text{var}(\alpha)$  with  $\tau(j) \neq \sigma(j)$ , and  $\tau(i) \not\sqsubset \sigma(i)$ . We now consider the following cases:

- $\tau(i) = \sigma(i)$

According to the assumption of Lemma 11, there exists a variable  $j \in \text{var}(\alpha)$  with  $\tau(j) \neq \sigma(j)$ ; hence,  $j \neq i$ . Since  $\sigma$  maps all variables except  $i$  to a word of length 1

and  $|\sigma(\alpha)| = |\tau(\alpha)|$ , if  $|\tau(j)| > 1$ , then we must have a variable  $x$  in  $\alpha$  with  $\tau(x) = \varepsilon$ . This is a contradiction to the fact that morphism  $\tau$  is nonerasing. If  $|\tau(j)| = 1$ , then this contradicts  $\sigma(\alpha) = \tau(\alpha)$ , since  $\tau(j) \neq \sigma(j)$ .

- $|\tau(i)| > |\sigma(i)|$

Since  $\sigma$  maps all variables except  $i$  to a word of length 1 and due to the fact that  $\tau$  is nonerasing,  $|\tau(\alpha)| > |\sigma(\alpha)|$ . Hence, necessarily,  $\tau(\alpha) \neq \sigma(\alpha)$ , which contradicts the assumption of Lemma 11.

- $|\tau(i)| \leq |\sigma(i)|$  and  $\tau(i) \neq \sigma(i)$

Assume that  $\alpha = \alpha_1 \cdot i_1^{p_1} \cdot \alpha_2 \cdot i_2^{p_2} \cdot [\dots] \cdot \alpha_n \cdot i_n^{p_n} \cdot \alpha_{n+1}$  where,  $\alpha_2, \alpha_3, \dots, \alpha_n \in \mathbb{N}^+$ ,  $\alpha_1, \alpha_{n+1} \in \mathbb{N}^*$  and, for every  $k$ ,  $1 \leq k \leq n$ ,  $i_k = i$ ,  $p_k \in \mathbb{N}$ ,  $i \not\sqsubseteq \alpha_k, \alpha_{n+1}$ . It follows from  $\tau$  being nonerasing and, for every  $q$ ,  $1 \leq q \leq n+1$ ,  $|\sigma(\alpha_q)| = |\alpha_q|$  that  $|\tau(\alpha_q)| \geq |\sigma(\alpha_q)|$ . As a result,  $|\tau(\alpha_1)| \geq |\sigma(\alpha_1)|$ . Now, assume that  $|\tau(\alpha_1 \cdot i_1^{p_1})| \leq |\sigma(\alpha_1 \cdot i_1^{p_1})|$ ; thus, due to  $\tau(\alpha) = \sigma(\alpha)$ ,  $\tau(\alpha_1 \cdot i_1^{p_1}) \sqsubseteq \sigma(\alpha_1 \cdot i_1^{p_1})$ . Since  $|\tau(\alpha_1)| \geq |\sigma(\alpha_1)|$ , this implies that  $\tau(i_1)^{p_1} \sqsubseteq \sigma(i_1)^{p_1}$ . Moreover, according to the assumption of this case,  $\tau(i) \neq \sigma(i)$ . These results satisfy the conditions of Lemma 10, and therefore  $\tau(i_1) \sqsubset \sigma(i_1)$ . However, this contradicts  $\tau(i) \not\sqsubseteq \sigma(i)$ . Consequently, we must have  $|\tau(\alpha_1 \cdot i_1^{p_1})| > |\sigma(\alpha_1 \cdot i_1^{p_1})|$ . Since  $|\tau(\alpha_2)| \geq |\sigma(\alpha_2)|$ , we can conclude  $|\tau(\alpha_1 \cdot i_1^{p_1} \cdot \alpha_2)| > |\sigma(\alpha_1 \cdot i_1^{p_1} \cdot \alpha_2)|$ . Using the same reasoning as above, we can show that  $|\tau(\alpha_1 \cdot i_1^{p_1} \cdot \alpha_2 \cdot i_2^{p_2})| > |\sigma(\alpha_1 \cdot i_1^{p_1} \cdot \alpha_2 \cdot i_2^{p_2})|$ . By extending this argument,

$$|\tau(\alpha_1 \cdot i_1^{p_1} \cdot \alpha_2 \cdot i_2^{p_2} \cdot [\dots] \cdot \alpha_n \cdot i_n^{p_n})| > |\sigma(\alpha_1 \cdot i_1^{p_1} \cdot \alpha_2 \cdot i_2^{p_2} \cdot [\dots] \cdot \alpha_n \cdot i_n^{p_n})|$$

Due to  $|\tau(\alpha_{n+1})| \geq |\sigma(\alpha_{n+1})|$ , we can conclude that  $|\tau(\alpha)| > |\sigma(\alpha)|$ , which contradicts  $\tau(\alpha) = \sigma(\alpha)$ .

Consequently, in all cases, our assumption leads to a contradiction. Hence,  $\tau(i) \sqsubset \sigma(i)$ .  $\square$

The next lemma, which directly results from Definition 5 and shall support the proof of the main result in the present section, discusses those patterns that have at least one square; more precisely, there exists a variable  $i \in \mathbb{N}$  with  $i^2 \sqsubset \alpha$ .

**Lemma 12.** *Let  $\alpha \in \mathbb{N}^+$ . If, for an  $i \in \mathbb{N}$ ,  $i^2 \sqsubseteq \alpha$ , then  $i \in E_\alpha$ .*

*Proof.* Assume that  $i^2 \sqsubseteq \alpha$ . If there exists a variable  $x_1 \in \text{var}(\alpha) \setminus \{i\}$  satisfying  $x_1 \cdot i \sqsubset \alpha$ , then  $\{i, x_1\} \subseteq L_i$ ; otherwise,  $L_i = \{i, \varepsilon\}$ . Moreover, if there exists a variable  $x_2 \in \text{var}(\alpha) \setminus \{i\}$  satisfying  $i \cdot x_2 \sqsubset \alpha$ , then  $\{i, x_2\} \subseteq R_i$ ; otherwise,  $R_i = \{i, \varepsilon\}$ . We assume to the contrary that  $i \notin E_\alpha$ . This means that  $i$  has loyal neighbours in  $\alpha$ . Hence, due to Definition 5, we need to consider two cases. If  $\varepsilon \notin L_i$  and, for every  $j \in L_i$ , we have  $R_j = \{i\}$ , then  $i \in L_i$  and  $R_i \neq \{i\}$ , which is a contradiction. If  $\varepsilon \notin R_i$  and, for every  $j \in R_i$ ,  $L_j = \{i\}$ , then  $i \in R_i$  and  $L_i \neq \{i\}$ , and this is again a contradiction.  $\square$

The subsequent characterisation of those patterns that have a weakly unambiguous length-increasing morphism with ternary or larger target alphabets is the main result of

this paper. It yields a novel partition of the set of all patterns over any sub-alphabet of  $\mathbb{N}$ . This partition is different from the partition into prolix and succinct patterns, which characterises the existence of strongly unambiguous nonerasing morphisms (see Theorem 8 and Proposition 9).

**Theorem 13.** *Let  $\alpha \in \mathbb{N}^+$  with  $|\alpha| \geq 2$ , and let  $\Sigma$  be an alphabet,  $|\Sigma| \geq 3$ . There is a weakly unambiguous length-increasing morphism  $\sigma : \mathbb{N}^+ \rightarrow \Sigma^+$  with respect to  $\alpha$  if and only if  $E_\alpha$  is not empty.*

*Proof.* Let  $\{a, b, c\} \subseteq \Sigma$ .

We begin with the *if* direction. Assume that  $E_\alpha$  is not empty. This means that there is at least one variable  $i \in \text{var}(\alpha)$  that does not have loyal neighbours, i. e.,  $i \in E_\alpha$ . Due to Definition 5 and Lemma 12, one of the following cases is satisfied:

*Case 1:*  $i^2 \sqsubseteq \alpha$ .

We define a morphism  $\sigma$  by  $\sigma(x) := bc$  if  $x = i$  and  $\sigma(x) := a$  if  $x \neq i$ . So,  $\sigma(i^2) = bc bc$ . Assume to the contrary that there is a morphism  $\tau : \mathbb{N}^+ \rightarrow \Sigma^+$  with  $\tau(\alpha) = \sigma(\alpha)$  and, for some  $q \in \text{var}(\alpha)$ ,  $\tau(q) \neq \sigma(q)$ . According to Lemma 11,  $\tau(i) \neq \sigma(i)$  must be satisfied, and this means that  $\tau(i)$  needs to be a proper factor of  $\sigma(i)$ . This implies that  $\tau(i) = b$  or  $\tau(i) = c$  and, as a result,  $\tau(i^2) = bb$  or  $\tau(i^2) = cc$ . Since  $\sigma(\alpha)$  does not contain the factors  $bb$  and  $cc$ , we can conclude that  $\tau(\alpha) \neq \sigma(\alpha)$ , which is a contradiction. Consequently,  $\sigma$  is weakly unambiguous with respect to  $\alpha$ .

*Case 2:*  $i^2 \not\sqsubseteq \alpha$ , and one of the following cases is satisfied:

*Case 2.1:* If  $\varepsilon \notin L_i$ , then there exists a variable  $j \in L_i$  such that  $R_j \neq \{i\}$ , and if  $\varepsilon \notin R_i$ , then there exists a variable  $j' \in R_i$  such that  $L_{j'} \neq \{i\}$ .

*Case 2.2:*  $\varepsilon \in L_i$  and  $\varepsilon \in R_i$ .

Let  $\sigma : \mathbb{N}^+ \rightarrow \{a, b, c\}^+$  be the morphism defined in Case 1. We assume to the contrary that there is a morphism  $\tau : \mathbb{N}^+ \rightarrow \Sigma^+$  with  $\tau(\alpha) = \sigma(\alpha)$  and, for some  $q \in \text{var}(\alpha)$ ,  $\tau(q) \neq \sigma(q)$ . Lemma 11 again implies that  $\tau(i) \sqsubset \sigma(i)$  must be satisfied. Thus,  $\tau(i) = b$  or  $\tau(i) = c$ .

With regard to Case 2.1, we first consider  $\tau(i) = c$  and  $\varepsilon \notin L_i$ . Due to the number of occurrences of  $c$  in  $\sigma(\alpha)$ , which equals the number of occurrences of  $i$  in  $\alpha$ , and also due to  $\sigma(i) = bc$ , the positions of  $c$  of  $\tau(i)$  must be at the same positions as  $c$  of  $\sigma(i)$  in  $\sigma(\alpha)$ . Therefore, the condition  $\tau(\alpha) = \sigma(\alpha)$  implies that, for every  $l \in L_i$ ,  $b$  is a suffix of  $\tau(l)$ , which means that  $b$  is suffix of  $\tau(j)$ . However, since  $R_j \neq \{i\}$ , the number of occurrences of  $b$  in  $\tau(\alpha)$  is greater than the number of occurrences of  $b$  in  $\sigma(\alpha)$ . Hence,  $\tau(\alpha) \neq \sigma(\alpha)$ , which is a contradiction.

We now consider  $\tau(i) = b$  and  $\varepsilon \notin R_i$ . Due to the number of occurrences of  $b$  in  $\sigma(\alpha)$ , which equals the number of occurrences of  $i$  in  $\alpha$ , and also due to  $\sigma(i) = bc$ , the positions of  $b$  of  $\tau(i)$  are at the same positions as  $b$  of  $\sigma(i)$  in  $\sigma(\alpha)$ . Hence, since  $\tau(\alpha) = \sigma(\alpha)$ , for every  $r \in R_i$ ,  $c$  is a prefix of  $\tau(r)$  and, consequently,  $c$  is prefix of  $\tau(j')$ . However, because of  $L_{j'} \neq \{i\}$ , the number of occurrences of  $c$  in  $\tau(\alpha)$  is greater than the number of occurrences of  $c$  in  $\sigma(\alpha)$ . This again implies  $\tau(\alpha) \neq \sigma(\alpha)$ .

Case 2.2 means that  $\alpha = i \cdot \alpha' \cdot i$ ,  $\alpha' \in \mathbb{N}^*$ . So,  $\sigma(\alpha) = bc \cdot \sigma(\alpha') \cdot bc$ . As mentioned above, due to Lemma 11,  $\tau(i) = b$  or  $\tau(i) = c$ . This implies that  $\tau(\alpha)$  starts with  $b$  and ends with

$b$ , or it starts with  $c$  and ends with  $c$ . Thus,  $\tau(\alpha) \neq \sigma(\alpha)$ . Hence, we can conclude that if  $E_\alpha \neq \emptyset$ , then there is a weakly unambiguous length-increasing morphism with respect to  $\alpha$ .

We now prove the *only if* direction. Hence, we shall demonstrate that if there is a weakly unambiguous length-increasing morphism  $\sigma : \mathbb{N}^+ \rightarrow \Sigma^+$  with respect to  $\alpha$ , then  $E_\alpha$  is not empty. Since  $\sigma$  is length-increasing, there exists a variable  $i$  that is mapped by  $\sigma$  to a word of length more than 1. Let  $\sigma(i) := a_1 a_2 [\dots] a_n$  with  $n \geq 2$  and, for every  $k$ ,  $1 \leq k \leq n$ ,  $a_k \in \Sigma$ . Assume to the contrary that  $E_\alpha$  is empty. Thus, due to Lemma 12,  $i^2 \not\sqsubseteq \alpha$ . According to Definition 5, one of the following cases is satisfied:

*Case 1:*  $\varepsilon \notin L_i$  and, for every  $j \in L_i$ ,  $R_j = \{i\}$ .

From this condition, we can directly conclude that

$$\alpha := \alpha_1 \cdot l_1 \cdot i \cdot \alpha_2 \cdot l_2 \cdot i \cdot [\dots] \cdot \alpha_m \cdot l_m \cdot i \cdot \alpha_{m+1},$$

with  $|\alpha|_i = m$  and, for every  $k$ ,  $1 \leq k \leq m$  and, for every  $k'$ ,  $1 \leq k' \leq m+1$ ,  $l_k \in L_i$ ,  $\alpha_{k'} \in \mathbb{N}^*$ ,  $i \neq l_k$  and,  $i, l_k \notin \text{var}(\alpha_{k'})$ . Thus,

$$\begin{aligned} \sigma(\alpha) &= \sigma(\alpha_1) \sigma(l_1) a_1 a_2 [\dots] a_n \cdot \sigma(\alpha_2) \sigma(l_2) a_1 a_2 [\dots] a_n \\ &\quad \cdot [\dots] \cdot \sigma(\alpha_m) \sigma(l_m) a_1 a_2 [\dots] a_n \cdot \sigma(\alpha_{m+1}). \end{aligned}$$

We now define a nonerasing morphism  $\tau$  such that, for every  $k$ ,  $1 \leq k \leq m$ ,  $\tau(l_k) := \sigma(l_k) a_1$ ,  $\tau(i) := a_2 a_3 [\dots] a_n$  and, for all other variables in  $\alpha$ ,  $\tau$  is identical to  $\sigma$ . Due to the fact that, for every  $k$ ,  $1 \leq k \leq m$ ,  $R_{l_k} = \{i\}$ , we can conclude that  $\tau(\alpha) = \sigma(\alpha)$ . Since  $\tau$  is nonerasing,  $\sigma$  is not weakly unambiguous, which is a contradiction.

*Case 2:*  $\varepsilon \notin R_i$  and, for every  $j \in R_i$ ,  $L_j = \{i\}$ .

We can directly conclude that

$$\alpha := \alpha_1 \cdot i \cdot r_1 \cdot \alpha_2 \cdot i \cdot r_2 \cdot [\dots] \cdot \alpha_m \cdot i \cdot r_m \cdot \alpha_{m+1}$$

with  $|\alpha|_i = m$  and, for every  $k$ ,  $1 \leq k \leq m$  and, for every  $k'$ ,  $1 \leq k' \leq m+1$ ,  $r_k \in R_i$ ,  $\alpha_{k'} \in \mathbb{N}^*$ ,  $i \neq r_k$ , and  $i, r_k \notin \text{var}(\alpha_{k'})$ . So,

$$\begin{aligned} \sigma(\alpha) &= \sigma(\alpha_1) a_1 a_2 [\dots] a_n \sigma(r_1) \cdot \sigma(\alpha_2) a_1 a_2 [\dots] a_n \sigma(r_2) \\ &\quad \cdot [\dots] \cdot \sigma(\alpha_m) a_1 a_2 [\dots] a_n \sigma(r_m) \cdot \sigma(\alpha_{m+1}). \end{aligned}$$

If we consider the nonerasing morphism  $\tau$  that satisfies, for every  $k$ ,  $1 \leq k \leq m$ ,  $\tau(r_k) := a_n \sigma(r_k)$  and  $\tau(i) := a_1 a_2 [\dots] a_{n-1}$  and that is identical to  $\sigma$  for all other variables in  $\alpha$ , then we can conclude that  $\tau(\alpha) = \sigma(\alpha)$ . Since  $\tau$  is nonerasing,  $\sigma$  is not weakly unambiguous. Hence,  $E_\alpha = \emptyset$  implies that  $\sigma$  is not weakly unambiguous, which contradicts the assumption. Consequently,  $E_\alpha$  is not empty.  $\square$

In order to illustrate Theorem 13 and its proof, we give two examples:

**Example 14.** Let  $\alpha := 1 \cdot 2 \cdot 3 \cdot 4 \cdot 1 \cdot 2 \cdot 3$ . According to Definition 6,  $S_\alpha = \{1, 2, 3\}$  and  $E_\alpha = \{4\}$ . In other words, the variable 4 does not have loyal neighbours. We define a morphism  $\sigma$  by  $\sigma(4) := bc$  and, for every other variable  $j \in \text{var}(\alpha)$ ,  $\sigma(j) := a$ . Due to Lemma 11, any morphism  $\tau$  with  $\tau(\alpha) = \sigma(\alpha)$  and, for a variable  $k \in \text{var}(\alpha)$ ,  $\tau(k) \neq \sigma(k)$  needs to split the factor  $bc$ . Hence,  $\tau(1)$  needs to contain  $c$ , or  $\tau(3)$  needs to contain  $b$ . However, since  $|\alpha|_1 = 2$  and  $|\alpha|_3 = 2$ ,  $|\tau(\alpha)|_c > |\sigma(\alpha)|_c$ , or  $|\tau(\alpha)|_b > |\sigma(\alpha)|_b$ . Consequently,  $\tau(\alpha) \neq \sigma(\alpha)$  and as a result,  $\sigma$  is weakly unambiguous with respect to  $\alpha$ .

**Example 15.** Let  $\alpha := 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 4 \cdot 7 \cdot 8 \cdot 3$ . According to Definition 5, all variables have loyal neighbours; in other words,  $E_\alpha = \emptyset$ . Hence, it follows from Theorem 13 that there is no weakly unambiguous length-increasing morphism  $\sigma : \mathbb{N}^+ \rightarrow \Sigma^+$ ,  $|\Sigma| \geq 3$ , with respect to  $\alpha$ .

We now give an alternative version of Theorem 13 that is based on regular expressions.

**Corollary 16.** Let  $\alpha \in \mathbb{N}^+$ , and let  $\Sigma$  be an alphabet,  $|\Sigma| \geq 3$ . There is no weakly unambiguous length-increasing morphism  $\sigma : \mathbb{N}^+ \rightarrow \Sigma^+$  with respect to  $\alpha$  if and only if, for every  $i \in \text{var}(\alpha)$ , at least one of the following statements is satisfied:

- there exists a partition  $L, N, \{i\}$  of  $\text{var}(\alpha)$  such that  $\alpha \in (N^*Li)^+N^*$ ,
- there exists a partition  $R, N, \{i\}$  of  $\text{var}(\alpha)$  such that  $\alpha \in (N^*iR)^+N^*$ .

*Proof.* According to the definition of loyal neighbours, it is easily verified that the first statement of Corollary 16 is equivalent to the first case of Definition 5, and the second one is equivalent to the second case of Definition 5. More precisely, the first statement is equivalent to, for every  $x \in L$ ,  $R_x = \{i\}$ , and the second one is equivalent to, for every  $x \in R$ ,  $L_x = \{i\}$ . Consequently, for every  $i \in \text{var}(\alpha)$ , one of the above statements being satisfied is equivalent to  $E_\alpha = \emptyset$ . Hence, Corollary 16 directly follows from Theorem 13.  $\square$

We conclude this section by determining the complexity of the decision problem resulting from Theorem 13.

**Theorem 17.** Let  $\alpha \in \mathbb{N}^+$  with  $|\alpha| \geq 2$ , and let  $\Sigma$  be an alphabet,  $|\Sigma| \geq 3$ . The problem of whether there is a length-increasing morphism  $\sigma : \mathbb{N}^+ \rightarrow \Sigma^+$  that is weakly unambiguous with respect to  $\alpha$  is decidable in polynomial time.

*Proof.* According to Theorem 13, a procedure deciding on the problem in Theorem 17 needs to test whether  $E_\alpha$  is empty. This can be accomplished by first producing the sets  $L_i$  and  $R_i$  for all  $i \in \text{var}(\alpha)$  and then scanning these sets for a variable that does not have loyal neighbours. The former task can be completed in time  $O(|\alpha|)$ , and the latter task requires  $O(|\text{var}(\alpha)|^2)$  steps.  $\square$

Hence, the complexity of Problem 2 is comparable to that of the equivalent problem for strongly unambiguous nonerasing morphisms (this is a consequence of the characterisation by Freydenberger et al. [4] and the complexity consideration by Holub [7]). In contrast to this, deciding on the existence of strongly unambiguous erasing morphisms is NP-hard (according to Schneider [14]).

## 5. Weakly unambiguous morphisms with $|\Sigma| = 2$

As we shall demonstrate below, our characterisation in Theorem 13 does not hold for binary target alphabets  $\Sigma$  (see Corollary 29). Hence, we have to study this case separately. We cannot give a characteristic condition on the existence of weakly unambiguous length-increasing morphisms with  $|\Sigma| = 2$ . Instead we shall present two criteria, namely Theorems 20 and 30, that can be interpreted as sufficient conditions on the existence of such morphisms, and one criterion, namely Theorem 27, that is a sufficient condition on the their non-existence. A comparison of these criteria, which shall be supported by a number of examples, then facilitates insights into the rather specific type of patterns that we cannot classify in this respect. The main result of this section is Theorem 20, which requires an extensive reasoning that is based on Lemmata 21, 22, 23, and 24, and on Proposition 25. However, before we study the technical details of our considerations on morphisms with binary target alphabets, we shall briefly discuss some basic yet vital observations that directly result from our work in Section 4.

Despite being restricted to ternary or larger alphabets, Theorem 13 and its proof have two important implications that also hold for unary and binary alphabets. The first of them shows that  $E_\alpha$  being empty for any given pattern  $\alpha$  is a sufficient condition for  $\alpha$  not having any weakly unambiguous length-increasing morphism:

**Corollary 18.** *Let  $\alpha \in \mathbb{N}^+$ , and let  $\Sigma$  be any alphabet. If  $E_\alpha = \emptyset$ , then there is no weakly unambiguous length-increasing morphism  $\sigma : \mathbb{N}^+ \rightarrow \Sigma^+$  with respect to  $\alpha$ . In general, the converse of this statement does not hold true.*

*Proof.* The first statement of Corollary 18 directly follows from the proof of *only if* direction of Theorem 13.

For the second statement of Corollary 18, we refer to the pattern  $\alpha := 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 4 \cdot 3 \cdot 7 \cdot 8$ . It can be verified with little effort that the variables 3 and 4 do not have loyal neighbours in  $\alpha$ . In Theorem 27, we demonstrate that, nevertheless, every length-increasing morphism  $\sigma : \mathbb{N}^+ \rightarrow \{a, b\}^+$  is ambiguous with respect to  $\alpha$ .  $\square$

Hence, if we wish to characterise those patterns with respect to which there is a weakly unambiguous morphism  $\sigma : \mathbb{N}^+ \rightarrow \Sigma^+$ ,  $|\Sigma| \leq 2$ , then we can safely restrict our considerations to those patterns  $\alpha$  where  $E_\alpha$  is a nonempty set.

The second implication of Theorem 13 demonstrates that any length-increasing morphism that is weakly unambiguous with respect to a pattern  $\alpha$  must have a particular, and very simple, shape for all variables in  $S_\alpha$ :

**Corollary 19.** *Let  $\alpha \in \mathbb{N}^+$ , let  $\Sigma$  be any alphabet, and let  $\sigma : \mathbb{N}^+ \rightarrow \Sigma^+$  be a length-increasing morphism that is weakly unambiguous with respect to  $\alpha$ . Then, for every  $i \in S_\alpha$ ,  $|\sigma(i)| = 1$ .*

*Proof.* Corollary 19 directly follows from the proof of the *only if* direction of Theorem 13.  $\square$

Thus, any weakly unambiguous length-increasing morphism with respect to a pattern  $\alpha$  must not be length-increasing for the variables in  $S_\alpha$ . This insight is very useful when searching for morphisms that might be weakly unambiguous with respect to a given pattern.

As shown by Corollary 18, if  $E_\alpha$  is empty, then there is no weakly unambiguous length-increasing morphism  $\sigma : \mathbb{N}^+ \rightarrow \Sigma^+$  with respect to  $\alpha$ . In the next step, we give a strong *necessary* condition on the structure of those patterns  $\alpha$  that satisfy  $E_\alpha \neq \emptyset$ , but nevertheless do *not* have a weakly unambiguous morphism  $\sigma : \mathbb{N}^+ \rightarrow \Sigma^+$ ,  $|\Sigma| = 2$ .

**Theorem 20.** *Let  $\alpha \in \mathbb{N}^+$  such that  $E_\alpha$  is nonempty. Let  $\Sigma$  be an alphabet,  $|\Sigma| = 2$ . If there is no weakly unambiguous length-increasing morphism  $\sigma : \mathbb{N}^+ \rightarrow \Sigma^+$  with respect to  $\alpha$ , then for every  $e \in E_\alpha$  there exists an  $e' \in E_\alpha$ ,  $e' \neq e$ , such that  $e \cdot e'$  and  $e' \cdot e$  are factors of  $\alpha$ .*

Before we can prove Theorem 20, we first need to introduce some technical lemmata. Referring to Section 4, if  $i^2 \sqsubseteq \alpha$ ,  $i \in \text{var}(\alpha)$ , then there is a weakly unambiguous length-increasing morphism  $\sigma : \mathbb{N}^+ \rightarrow \Sigma^+$ ,  $|\Sigma| \geq 3$ , with respect to  $\alpha$ ; this is a direct consequence of Lemma 12 and Theorem 13. We now investigate this case for  $|\Sigma| = 2$ .

**Lemma 21.** *Let  $\alpha \in \mathbb{N}^+$  such that, for an  $i \in \mathbb{N}$ ,  $i^2 \sqsubseteq \alpha$ . Let  $\Sigma$  be an alphabet,  $|\Sigma| = 2$ . There is a weakly unambiguous length-increasing morphism  $\sigma : \mathbb{N}^+ \rightarrow \Sigma^+$  with respect to  $\alpha$  that maps  $i$  to an image of length more than 1 and every variable in  $\text{var}(\alpha) \setminus \{i\}$  to images of length 1 if*

- (I) *for every occurrence of  $i$  in  $\alpha$ , the right or left neighbour of  $i$  is  $i$ , or*
- (II) *for every  $(i' \cdot i) \sqsubseteq \alpha$  with  $i' \in \text{var}(\alpha) \setminus \{i\}$ ,  $(i \cdot i') \not\sqsubseteq \alpha$ .*

*Proof.* Let  $\Sigma := \{a, b\}$ .

We first prove that Condition (I) implies the existence of a weakly unambiguous length-increasing morphism with respect to  $\alpha$ . Let

$$\alpha := \alpha_1 \cdot i^{p_1} \cdot \alpha_2 \cdot i^{p_2} \cdot [\dots] \cdot \alpha_n \cdot i^{p_n} \cdot \alpha_{n+1},$$

with  $n \in \mathbb{N}$ ,  $\alpha_2, \alpha_3, \dots, \alpha_n \in (\mathbb{N} \setminus \{i\})^+$ ,  $\alpha_1, \alpha_{n+1} \in (\mathbb{N} \setminus \{i\})^*$  and, for every  $j$ ,  $1 \leq j \leq n$ ,  $p_j \in \mathbb{N}$ . It follows from Condition (I) that, for every  $j$ ,  $p_j \geq 2$ . We define a morphism  $\sigma : \mathbb{N}^+ \rightarrow \Sigma^+$  by, for every  $x \in \mathbb{N}$ ,

$$\sigma(x) := \begin{cases} ab, & x = i, \\ b, & x \neq i. \end{cases}$$

Thus,  $\sigma(\alpha) = b \cdot b \cdot [\dots] \cdot b \cdot (ab)^{p_1} \cdot b \cdot b \cdot [\dots] \cdot b \cdot (ab)^{p_2} \cdot [\dots] \cdot b \cdot b \cdot [\dots] \cdot b \cdot (ab)^{p_n} \cdot b \cdot b \cdot [\dots] \cdot b$ .

We now assume to the contrary that  $\sigma$  is not weakly unambiguous with respect to  $\alpha$ . Hence, there is a morphism  $\tau : \mathbb{N}^+ \rightarrow \Sigma^+$  such that  $\tau(\alpha) = \sigma(\alpha)$  and, for some  $q \in \text{var}(\alpha)$ ,  $\tau(q) \neq \sigma(q)$ . According to Lemma 11, it is required to split the factor  $ab$  when defining  $\tau(i)$ . If we consider  $\tau(i) = a$ , then, due to the fact that there is no factor  $a^k$ ,  $k \geq 2$ , in  $\sigma(\alpha)$ ,  $\tau(\alpha) \neq \sigma(\alpha)$ . Thus,  $\tau(i) = b$ . As a result,  $\tau(\alpha) = \tau(\alpha_1) \cdot b^{p_1} \cdot \tau(\alpha_2) \cdot b^{p_2} \cdot [\dots] \cdot \tau(\alpha_n) \cdot b^{p_n} \cdot \tau(\alpha_{n+1})$ . Due to  $\tau(\alpha) = \sigma(\alpha)$ , one of the following cases is satisfied:

- $|\tau(\alpha_1)| < |\sigma(\alpha_1)|$ .

This means that there exists a variable  $z \in \text{var}(\alpha_1)$  with  $\tau(z) = \varepsilon$ ; however, this contradicts the fact that  $\tau$  is nonerasing.

- $|\tau(\alpha_1)| > |\sigma(\alpha_1)|$ .

Since  $\sigma(i^{p_1})$  has no factor  $b^k$ ,  $k > 1$ ,  $|\tau(\alpha_1 \cdot i^{p_1})| > |\sigma(\alpha_1 \cdot i^{p_1})|$ . This implies that  $\tau(i^{p_2})$  cannot be located to the left of the position of  $\sigma(i^{p_2})$  in  $\sigma(\alpha)$ ; otherwise, for some  $z \in \text{var}(\alpha_2)$ ,  $\tau(z) = \varepsilon$ . Thus,  $|\tau(\alpha_1 \cdot i^{p_1} \cdot \alpha_2 \cdot i^{p_2})| > |\sigma(\alpha_1 \cdot i^{p_1} \cdot \alpha_2 \cdot i^{p_2})|$ . Consequently, if we continue our above reasoning, this finally implies that

$$|\tau(\alpha_1 \cdot i^{p_1} \cdot \alpha_2 \cdot i^{p_2} \cdot [\dots] \cdot \alpha_n \cdot i^{p_n})| > |\sigma(\alpha_1 \cdot i^{p_1} \cdot \alpha_2 \cdot i^{p_2} \cdot [\dots] \cdot \alpha_n \cdot i^{p_n})|$$

and there exist some variable  $z \in \text{var}(\alpha_{n+1})$  such that  $\tau(z) = \varepsilon$ . However, this contradicts the fact that  $\tau$  is nonerasing.

It follows from our reasoning on the above cases that the morphism  $\tau$  does not exist. Hence, if Condition (I) is satisfied, then  $\sigma$  is weakly unambiguous with respect to  $\alpha$ .

We now prove that Condition (II) also implies the existence of a weakly unambiguous length-increasing morphism with respect to  $\alpha$ . According to Condition (II),  $(R_i \cap L_i) \setminus \{i\} = \emptyset$ . So, by considering Condition (II), we can define a morphism  $\sigma : \mathbb{N}^+ \rightarrow \Sigma^+$  with

$$\sigma(x) = \begin{cases} ab, & x = i, \\ b, & x \in L_i, \\ a, & x \in R_i, \\ b, & \text{else.} \end{cases}$$

Without loss of generality, we can assume that Condition (I) is not satisfied. So, any two consecutive occurrences of  $i$ , which are denoted by  $i_1$  and  $i_2$ , can occur in  $\alpha$  according to one of the following cases:

1.  $\alpha = \alpha_1 \cdot l_1 \cdot i_1 \cdot r_1 \cdot \alpha_2 \cdot l_2 \cdot i_2 \cdot r_2 \cdot \alpha_3$ ,
2.  $\alpha = \alpha_1 \cdot l_1 \cdot i_1 \cdot r_1 \cdot \alpha_2 \cdot l_2 \cdot i_2^{p_2} \cdot r_2 \cdot \alpha_3$ ,
3.  $\alpha = \alpha_1 \cdot l_1 \cdot i_1^{p_1} \cdot r_1 \cdot \alpha_2 \cdot l_2 \cdot i_2 \cdot r_2 \cdot \alpha_3$ ,
4.  $\alpha = \alpha_1 \cdot l_1 \cdot i_1^{p_1} \cdot r_1 \cdot \alpha_2 \cdot l_2 \cdot i_2^{p_2} \cdot r_2 \cdot \alpha_3$ ,

where  $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{N}^*$ ,  $l_1, r_1, l_2, r_2 \in \text{var}(\alpha) \setminus \{i\}$ ,  $i_1 = i_2 = i$ ,  $i \not\sqsubseteq \alpha_2$ , and  $p_1, p_2 > 1$ .

We assume to the contrary that  $\sigma$  is not weakly unambiguous with respect to  $\alpha$ . Hence, there is a morphism  $\tau : \mathbb{N}^+ \rightarrow \Sigma^+$  satisfying  $\tau(\alpha) = \sigma(\alpha)$  and for some  $q \in \text{var}(\alpha)$ ,  $\tau(q) \neq \sigma(q)$ . According to Lemma 11, it is required to split the factor  $ab$  when defining  $\tau(i)$ . This means that  $\tau(i) = a$  or  $\tau(i) = b$ . Furthermore, for all of the above-mentioned cases, we assume that

$$|\tau(\alpha_1 \cdot l_1)| \geq |\sigma(\alpha_1 \cdot l_1)|. \quad (1)$$

Referring to this assumption, we now compare the position of  $\tau(i)$  to that of  $\sigma(i)$  in  $\sigma(\alpha)$  for the above four cases. Our corresponding insights shall be applied further below.



In Case 1,  $\sigma(\alpha) = \sigma(\alpha_1) \cdot b \cdot ab \cdot a \cdot \sigma(\alpha_2) \cdot b \cdot ab \cdot a \cdot \sigma(\alpha_3)$ . We assume that  $\tau(i_1) = a$  such that  $a$  is located at the same position as  $a$  of  $\sigma(i_1)$  in  $\sigma(\alpha)$ . Since  $\tau$  is nonerasing and  $\sigma(l_2) = b$ ,  $a$  of  $\tau(i_2)$  is located at the same position as  $a$  of  $\sigma(i_2)$  in  $\sigma(\alpha)$  or it is located to the right of that position; otherwise, there should be a variable  $z \in (\text{var}(\alpha_2) \cup \{r_1, l_2\})$  with  $\tau(z) = \varepsilon$ . If the letter  $a$  of  $\tau(i_1) = a$  is located to the right of the position of the letter  $a$  of  $\sigma(i_1)$  in  $\sigma(\alpha)$ , due to  $\tau$  being nonerasing, the letter  $a$  of  $\tau(i_2)$  is located to the right of the position of the letter  $a$  of  $\sigma(i_2)$  in  $\sigma(\alpha)$ . We can apply the same reasoning to  $\tau(i_1) = b$ .

In Case 2,  $\sigma(\alpha) = \sigma(\alpha_1) \cdot b \cdot ab \cdot a \cdot \sigma(\alpha_2) \cdot b \cdot (ab)^{p_2} \cdot a \cdot \sigma(\alpha_3)$ . We assume that  $\tau(i_1) = a$  such that  $a$  is located at the same position as  $a$  of  $\sigma(i_1)$  in  $\sigma(\alpha)$ . So,  $\tau(i_2^{p_2}) = a^{p_2}$ . Since  $\sigma(l_2 \cdot i_2^{p_2}) = b \cdot (ab)^{p_2}$ ,  $a^{p_2}$  of  $\tau(i_2^{p_2})$  must be located to the left or to the right of  $\sigma(l_2 \cdot i_2^{p_2})$  in  $\sigma(\alpha)$ . However, it cannot be located to the left of this factor, since  $\tau$  is nonerasing. If  $\tau(i_1) = a$  and  $a$  is located to the right of the position of the letter  $a$  of  $\sigma(i_1)$  in  $\sigma(\alpha)$ , then  $\tau(i_2^{p_2})$  must be located to the right of  $\sigma(l_2 \cdot i_2^{p_2})$  using the same reasoning. An analogous reasoning can also be used for  $\tau(i_1) = b$ .

In Case 3,  $\sigma(\alpha) = \sigma(\alpha_1) \cdot b \cdot (ab)^{p_1} \cdot a \cdot \sigma(\alpha_2) \cdot b \cdot ab \cdot a \cdot \sigma(\alpha_3)$ . We assume that  $\tau(i_1) = a$ . Since  $a^{p_1} \not\sqsubseteq \sigma(i_1^{p_1})$ , and due to Relation (1), the factor  $\tau(i_1^{p_1})$  must be located to the right position of  $\sigma(i_1^{p_1})$  in  $\sigma(\alpha)$ . This implies that, since  $|\tau(i_1^{p_1})| \geq 2$  and  $\tau$  is nonerasing,  $a$  of  $\tau(i_2)$  must be located to the right of the position of the letter  $a$  of  $\sigma(i_2)$  in  $\sigma(\alpha)$ . This reasoning is also valid if  $\tau(i_1) = b$ .

In Case 4,  $\sigma(\alpha) = \sigma(\alpha_1) \cdot b \cdot (ab)^{p_1} \cdot a \cdot \sigma(\alpha_2) \cdot b \cdot (ab)^{p_2} \cdot a \cdot \sigma(\alpha_3)$ . We assume that  $\tau(i_1) = a$ . Since  $a^{p_1} \not\sqsubseteq \sigma(i_1^{p_1})$ , and due to Relation (1), the factor  $\tau(i_1^{p_1})$  must be located to the right of the position of  $\sigma(i_1^{p_1})$  in  $\sigma(\alpha)$ . This implies that, since  $\tau$  is nonerasing and there is no factor  $a^{p_2}$  in  $\sigma(i_2^{p_2})$ , the factor  $a^{p_2}$  of  $\tau(i_2)$  must be located to the right of the factor  $(ab)^{p_2}$  of  $\sigma(i_2^{p_2})$  in  $\sigma(\alpha)$ . The same reasoning applies to  $\tau(i_1) = b$ .

Now, let  $\alpha := \alpha' \cdot i \cdot \alpha''$ ,  $i \not\sqsubseteq \alpha'$ . Since  $\tau$  is nonerasing and  $\sigma$  maps every variable of  $\alpha'$  to words of length 1,  $|\tau(\alpha')| \geq |\sigma(\alpha')|$ . This result satisfies Relation (1). Hence, we can consider one of the above cases to investigate  $\tau$  when applied to the first occurrence of  $i$  in  $\alpha$ . This means  $i \not\sqsubseteq \alpha_1$ . All cases lead to the fact that  $\tau(i_2)$  or  $\tau(i_2^{p_2})$  cannot be located to the left of the positions of  $\sigma(i_2)$  or  $\sigma(i_2^{p_2})$ , respectively, in  $\sigma(\alpha)$ . Consequently,

$$\begin{aligned} |\tau(\alpha_1 \cdot l_1 \cdot i_1 \cdot r_1 \cdot \alpha_2 \cdot l_2)| &\geq |\sigma(\alpha_1 \cdot l_1 \cdot i_1 \cdot r_1 \cdot \alpha_2 \cdot l_2)| \text{ or} \\ |\tau(\alpha_1 \cdot l_1 \cdot i_1^{p_1} \cdot r_1 \cdot \alpha_2 \cdot l_2)| &\geq |\sigma(\alpha_1 \cdot l_1 \cdot i_1^{p_1} \cdot r_1 \cdot \alpha_2 \cdot l_2)|. \end{aligned} \quad (2)$$

In the next step, if we consider  $i_2$  or  $i_2^{p_2}$  as  $i_1$  or  $i_1^{p_1}$ , respectively, and the next occurrence of  $i$  or  $i^k$ ,  $k > 1$ , as  $i_2$  or  $i_2^{p_2}$ , respectively, due to Relation (2), Relation (1) of our cases is satisfied again. Consequently, we can extend this result to the last occurrence of  $i$ .

We now consider Cases 2, 3, and 4. In these cases, the factor  $\tau(i_2)$  is not located to the left or even at the same position as  $\sigma(i_2)$  in  $\sigma(\alpha)$ . Moreover, as mentioned in Case 1, if the letter  $a$  of  $\tau(i_1) = a$  is located to the right of the position of the letter  $a$  of  $\sigma(i_1)$  in  $\sigma(\alpha)$ , the letter  $a$  of  $\tau(i_2)$  is located to the right of the position of the letter  $a$  of  $\sigma(i_2)$  in  $\sigma(\alpha)$  – the same happens if  $\tau(i_1) = b$ . Hence, since there is at least one  $i^k$ ,  $k \geq 2$ , in  $\alpha$ , by considering Cases 1, 2, 3, and 4, which can be extended over the other occurrences of  $i$ , and due to  $\tau$

being nonerasing,  $|\tau(\alpha)| > |\sigma(\alpha)|$ . Thus, the morphism  $\tau$  does not exist. This implies that  $\sigma$  is weakly unambiguous with respect to  $\alpha$ .  $\square$

In the following lemma, we introduce a special pattern with respect to which there is a weakly unambiguous length-increasing morphism  $\sigma : \mathbb{N}^+ \rightarrow \{a, b\}^+$ .

**Lemma 22.** *Let  $\alpha := \alpha_1 \cdot e \cdot \alpha_2 \cdot e \cdot [\dots] \cdot \alpha_{n-1} \cdot e \cdot \alpha_n$  with  $e \in E_\alpha$ ,  $\alpha_1, \alpha_n \in \mathbb{N}^*$ ,  $\alpha_2, \alpha_3, \dots, \alpha_{n-1} \in \mathbb{N}^+$  and, for every  $j$ ,  $1 \leq j \leq n$ ,  $e \not\sqsubseteq \alpha_j$ . Suppose that there exists a factor  $l \cdot e \cdot r \sqsubseteq \alpha$ ,  $l, r \in \text{var}(\alpha)$ , such that  $l$  and  $r$  satisfy the following conditions:*

- *there exists an occurrence of  $l$  in  $\alpha$  such that the right neighbour of this occurrence is not  $e$  and the left neighbour of this occurrence is not  $e$ , and*
- *there exists an occurrence of  $r$  in  $\alpha$  such that the right neighbour of this occurrence is not  $e$  and the left neighbour of this occurrence is not  $e$ .*

*If  $\sigma : \mathbb{N}^+ \rightarrow \{a, b\}^+$  is a nonerasing morphism with  $\sigma(e) = bb$  and, for every  $x \in \text{var}(\alpha) \setminus \{e\}$ ,  $\sigma(x) = a$ , then  $\sigma$  is weakly unambiguous with respect to  $\alpha$ .*

*Proof.* Let  $\alpha := \alpha_1 \cdot e_1 \cdot \alpha_2 \cdot e_2 \cdot [\dots] \cdot \alpha_{n-1} \cdot e_{n-1} \cdot \alpha_n$  with, for every  $k$ ,  $1 \leq k \leq n-1$ ,  $e_k = e$ . Also, let  $\sigma(e) := b_1 b_2$  with  $b_1 = b_2 = b$ . Assume to the contrary that  $\sigma$  is not weakly unambiguous with respect to  $\alpha$ . So, there exists a morphism  $\tau$  satisfying  $\tau(\alpha) = \sigma(\alpha)$  and, for some  $q \in \text{var}(\alpha)$ ,  $\tau(q) \neq \sigma(q)$ . Lemma 11 implies that  $\tau(e) = b$ .

We claim that, for every  $k$ ,  $1 \leq k \leq n-1$ ,  $\tau(e_k)$  is located at the same position as the first or second  $b$  of  $\sigma(e_k)$  in  $\sigma(\alpha)$ . To prove this claim, we assume to the contrary that there exists a  $j$ ,  $1 \leq j \leq n-1$ , such that  $\tau(e_j)$  is not at the position of the first or second  $b$  of  $\sigma(e_j)$  in  $\sigma(\alpha)$ . Thus, the following cases need to be considered:

- $\tau(e_j)$  is located to the left of the position of  $\sigma(e_j)$  in  $\sigma(\alpha)$ .

If there is no occurrence of  $e$  to the left of  $e_j$  in  $\alpha$ , then  $\tau(\alpha) \neq \sigma(\alpha)$ . So, assume that there is an occurrence of  $e_{j-1}$  to the left of  $e_j$ . Since  $\tau(e_j)$  is located to the left of the position of  $\sigma(e_j)$ , it must be located at the position of the first  $b$  or the second  $b$  of  $\sigma(e_{j-1})$ , or it is located to the left of the position of the first  $b$  of  $\sigma(e_{j-1})$  in  $\sigma(\alpha)$ . In both cases, due to the facts that  $\tau$  is nonerasing and there exists at least one variable between  $e_{j-1}$  and  $e_j$ ,  $\tau(e_{j-1})$  must be located to the left of the position of  $\sigma(e_{j-1})$ . Now, if we continue the above reasoning for  $\tau(e_{j-1}), \tau(e_{j-2}), \dots, \tau(e_1)$ , the factor  $\tau(e_1)$  must be located to the left of the position of  $\sigma(e_1)$  in  $\sigma(\alpha)$ ; however, since there is no occurrence of  $e$  to the left of  $e_1$  in  $\alpha$ ,  $\tau(\alpha) \neq \sigma(\alpha)$ .

- $\tau(e_j)$  is located to the right of the position of  $\sigma(e_j)$  in  $\sigma(\alpha)$ .

In this case, an analogous reasoning to that in the previous case leads to the insight that  $\tau(e_{n-1})$  must be located to the right of the position of  $\sigma(e_{n-1})$  in  $\sigma(\alpha)$ , which again is a contradiction.

Hence, for every  $k$ ,  $1 \leq k \leq n-1$ ,  $\tau(e_k)$  is located at the same position as the first or second  $b$  of  $\sigma(e_k)$  in  $\sigma(\alpha)$ . This insight has two implications. The first one is that, due to  $\tau$  being nonerasing and  $l \cdot e \cdot r$  being a factor of  $\alpha$ ,

$$\begin{aligned}\tau(l) &= v \cdot b_1, v \in \{a, b\}^* \text{ or} \\ \tau(r) &= b_2 \cdot v, v \in \{a, b\}^*.\end{aligned}\tag{3}$$

The second implication is that, since for any two consecutive occurrences of  $e$  in  $\alpha$ , the word  $e \cdot z_1 \cdot z_2 \cdot [\dots] \cdot z_{n-1} \cdot z_n \cdot e$ ,  $z_j \in \text{var}(\alpha) \setminus \{e\}$ ,  $1 \leq j \leq n$ , is a factor of  $\alpha$ ,  $\tau(z_j)$  must satisfy the following conditions:

$$\tau(z_j) = \begin{cases} \begin{cases} b_2 \text{ or } b_2 \cdot \sigma(z_j) \text{ or } b_2 \cdot \sigma(z_j) \cdot \sigma(z_{j+1}) \text{ or} \\ \sigma(z_j) \text{ or } \sigma(z_j) \cdot \sigma(z_{j+1}), \end{cases} & \text{if } j = 1, \\ \begin{cases} b_1 \text{ or } \sigma(z_j) \cdot b_1 \text{ or } \sigma(z_{j-1}) \cdot \sigma(z_j) \cdot b_1 \text{ or} \\ \sigma(z_j) \text{ or } \sigma(z_{j-1}) \cdot \sigma(z_j), \end{cases} & \text{if } j = n, \\ \begin{cases} \sigma(z_j) \text{ or } \sigma(z_{j+1}) \text{ or } \sigma(z_{j-1}) \text{ or } \sigma(z_j) \cdot \sigma(z_{j+1}) \\ \text{or } \sigma(z_{j-1}) \cdot \sigma(z_j) \text{ or } \sigma(z_{j-1}) \cdot \sigma(z_j) \cdot \sigma(z_{j+1}), \end{cases} & \text{if } 1 < j < n. \end{cases}\tag{4}$$

According to the assumption of Lemma 22, there exist an occurrence of  $l$  and an occurrence of  $r$  in  $\alpha$  such that the right neighbour and the left neighbour of these occurrences are not  $e$ . So, by considering Condition (4),  $\tau(l)$  and  $\tau(r)$  cannot contain any factor  $b$ . This contradicts Condition (3). Hence,  $\sigma$  is weakly unambiguous with respect to  $\alpha$ .  $\square$

Before we continue with the next two lemmata that are required to prove Theorem 20, we wish to briefly clarify their subject in an informal manner: Let  $\alpha \in \mathbb{N}^+$ ,  $|\alpha| \geq 2$ , and let  $\sigma : \mathbb{N}^+ \rightarrow \Sigma^+$  be a nonerasing morphism satisfying for a variable  $e \in \text{var}(\alpha)$ ,  $|\sigma(e)| > 1$  and, for every  $i \in \text{var}(\alpha) \setminus \{e\}$ ,  $|\sigma(i)| = 1$ . Moreover, assume that  $\tau$  is a nonerasing morphism satisfying  $\tau(\alpha) = \sigma(\alpha)$ . According to Lemma 11, if there exists a variable  $j \in \text{var}(\alpha)$  with  $\tau(j) \neq \sigma(j)$ , then  $\tau(e) \sqsubset \sigma(e)$ . In the following lemmata, we examine the position of  $\tau(e)$  in comparison with the position of  $\sigma(e)$  in  $\sigma(\alpha)$ .

**Lemma 23.** *Let  $\alpha \in \mathbb{N}^+$  such that  $E_\alpha \neq \emptyset$ . Let  $e \in E_\alpha$  with  $L_e \cap R_e = \emptyset$ . Let  $\alpha = \alpha_1 \cdot e_1 \cdot \alpha_2 \cdot e_2 \cdot [\dots] \cdot \alpha_{n-1} \cdot e_{n-1} \cdot \alpha_n$  with  $\alpha_1, \alpha_n \in \mathbb{N}^*$  and, for every  $k$ ,  $2 \leq k \leq n-1$ ,  $\alpha_k \in \mathbb{N}^+$ ,  $|\alpha_k| \geq 2$ , and, for every  $j$ ,  $1 \leq j \leq n-1$ ,  $e_j = e$  and,  $e \not\sqsubseteq \alpha_j, \alpha_n$ . Let  $\sigma : \mathbb{N}^+ \rightarrow \{a, b\}^+$  be any morphism satisfying*

$$\sigma(x) = \begin{cases} ab, & x = e, \\ b, & x \in L_e, \\ a, & x \in R_e, \end{cases}$$

and  $|\sigma(x)| = 1$  for every  $x \in \text{var}(\alpha) \setminus (\{e\} \cup L_e \cup R_e)$ . Assume that there exists a nonerasing morphism  $\tau$  with  $\tau(\alpha) = \sigma(\alpha)$  and, for some  $j \in \text{var}(\alpha)$ ,  $\tau(j) \neq \sigma(j)$ . Then, for every occurrence of  $e_i$ ,  $1 \leq i \leq n-1$ , one of the following cases is satisfied:

(I)  $\tau(e_i) = a$ , and this letter is located at the same position in  $\sigma(\alpha)$  as the letter  $a$  of  $\sigma(e_i)$ ,  
or

(II)  $\tau(e_i) = b$ , and this letter is located at the same position in  $\sigma(\alpha)$  as the letter  $b$  of  $\sigma(e_i)$ .

*Proof.* For every  $i$ ,  $1 \leq i \leq n-1$ , let  $\sigma(e_i) := a_i b_i$ ,  $a_i = a$ ,  $b_i = b$ . Also, for every  $j$ ,  $1 \leq j \leq n$ , let  $\alpha_j := l_j \cdot \alpha'_j \cdot r_j$ ,  $\alpha'_j \in (\text{var}(\alpha) \setminus \{e\})^*$ ,  $l_j, r_j \in \text{var}(\alpha) \setminus \{e\}$ . Thus,

$$\begin{aligned} \sigma(\alpha) &= \sigma(l_1) \cdot \sigma(\alpha'_1) \cdot b \cdot a_1 b_1 \cdot a \cdot \sigma(\alpha'_2) \cdot b \cdot a_2 b_2 \cdot [\dots] \cdot \\ &\quad \sigma(\alpha'_{n-1}) \cdot b \cdot a_{n-1} b_{n-1} \cdot a \cdot \sigma(\alpha'_n) \cdot \sigma(r_n). \end{aligned}$$

According to Lemma 11,  $\tau(e) = a$  or  $\tau(e) = b$ . In order to prove Case (I), assume to the contrary that there exists a  $k$ ,  $1 \leq k \leq n-1$ , with  $\tau(e_k) = a$ , but this  $a$  is not located at the same position as the letter  $a_k$  in  $\sigma(\alpha)$ . This leads to the following cases:

- The letter  $a$  of  $\tau(e_k)$  is located to the left of the position of the letter  $a_k$  in  $\sigma(\alpha)$ .

If there is no occurrence of  $e$  to the left of  $e_k$ , then  $\tau(\alpha) = \sigma(\alpha)$  implies for some variables  $z \in \alpha_k$ ,  $\tau(z) = \varepsilon$ . However, this contradicts  $\tau$  being nonerasing.

Assume that there is an occurrence of  $e$  to the left of  $e_k$ . Due to the fact that there is an occurrence of  $b$  as a left neighbour of  $a_k$  in  $\sigma(\alpha)$ , the difference of the position of the nearest occurrence of  $a$  to the position of  $a_k$  in  $\sigma(\alpha)$  is at least 2. If  $\tau(e_{k-1})$  is located at the position of  $a_{k-1}$  in  $\sigma(\alpha)$ , or it is located at any of the positions of  $\sigma(\alpha_k)$ , then this leads to  $|\tau(\alpha_k)| \leq (|\alpha_k| - 2) + 1$  – note that “+1” results from  $b_{k-1} \sqsubseteq \tau(\alpha_k)$  if  $\tau(e_{k-1})$  is located at the position of  $a_{k-1}$ . This means that, for some variables  $z \in \alpha_k$ ,  $\tau(z) = \varepsilon$ , which contradicts  $\tau$  being nonerasing. However, if  $a$  of  $\tau(e_{k-1})$  is located to the left of the position of  $a_{k-1}$ , then we continue our above reasoning. This argument finally leads to  $\tau(e_1)$  being located to the left of  $a_1$  in  $\sigma(\alpha)$ ; however, this means that, for some  $z \in \text{var}(\alpha_1)$ ,  $\tau(z) = \varepsilon$ , which again contradicts the fact that  $\tau$  is nonerasing.

- The letter  $a$  of  $\tau(e_k)$  is located to the right of the position of the letter  $a_k$  in  $\sigma(\alpha)$ .

In this case, an analogous reasoning to that in the previous case – now considering  $a_k, a_{k+1}, \dots, a_{n-1}$  instead of  $a_k, a_{k-1}, \dots, a_1$  – leads to an equivalent contradiction.

To prove Case (II), assume to the contrary that there exists a  $k$ ,  $1 \leq k \leq n-1$ , with  $\tau(e_k) = b$ ; however,  $b$  is not at the position of  $b_k$  in  $\sigma(\alpha)$ . Then we can use an analogous reasoning to that on Case (I).  $\square$

Lemma 23 and its proof enable us in the following lemma to investigate the morphism  $\tau$ , which is defined in Lemma 23, for the variables occurring between two consecutive occurrences of  $e$ .

**Lemma 24.** *Let  $\alpha \in \mathbb{N}^+$  such that  $E_\alpha \neq \emptyset$ . Let  $e \in E_\alpha$  with  $L_e \cap R_e = \emptyset$ . Let  $\alpha := \alpha_1 \cdot e_1 \cdot x_1 \cdot x_2 [\dots] \cdot x_n \cdot e_2 \cdot \alpha_2$ ,  $\alpha_1, \alpha_2 \in \mathbb{N}^*$ ,  $e_1 = e_2 = e$ ,  $n > 1$ , and for every  $j$ ,  $1 \leq j \leq n$ ,  $x_j \in \text{var}(\alpha) \setminus \{e\}$ . Let  $\sigma : \mathbb{N}^+ \rightarrow \{a, b\}^+$  be a morphism satisfying*

$$\sigma(x) = \begin{cases} ab, & x = e, \\ b, & x \in L_e, \\ a, & x \in R_e, \end{cases}$$

and  $|\sigma(x)| = 1$  for every  $x \in \text{var}(\alpha) \setminus (\{e\} \cup L_e \cup R_e)$ . Then, for every morphism  $\tau$  with  $\tau(\alpha) = \sigma(\alpha)$  and, for some  $j \in \text{var}(\alpha)$ ,  $\tau(j) \neq \sigma(j)$ , one of the following cases is satisfied:

- (I) For every  $i$ ,  $1 < i < n$ ,  $\tau(x_i) = \sigma(x_i)$ , or  $\tau(x_i) = \sigma(x_{i-1}) \cdot v$ ,  $v \in \{\sigma(x_i), \varepsilon\}$ .  
If  $i = 1$ , then  $\tau(x_1) = b \cdot v$ ,  $v \in \{\sigma(x_1), \varepsilon\}$ , and if  $i = n$ , then  $\tau(x_n) = \sigma(x_{n-1}) \cdot \sigma(x_n)$ .
- (II) For every  $i$ ,  $1 < i < n$ ,  $\tau(x_i) = \sigma(x_i)$ , or  $\tau(x_i) = v \cdot \sigma(x_{i+1})$ ,  $v \in \{\sigma(x_i), \varepsilon\}$ .  
If  $i = n$ , then  $\tau(x_n) = v \cdot a$ ,  $v \in \{\sigma(x_n), \varepsilon\}$ , and if  $i = 1$ , then  $\tau(x_1) = \sigma(x_1) \cdot \sigma(x_2)$ .

*Proof.* Assume that  $\tau(\alpha) = \sigma(\alpha)$  and, for some  $j \in \text{var}(\alpha)$ ,  $\tau(j) \neq \sigma(j)$ . According to Lemmata 11 and 23, regardless of the number of occurrences of  $e$  in  $\alpha_1$  and  $\alpha_2$ , one of the following cases is satisfied:

- $\tau(e_1) = a$ , and this letter is located at the same position as the letter  $a$  of  $\sigma(e_1)$  in  $\sigma(\alpha)$ ; in addition to this,  $\tau(e_2) = a$ , and this letter is located at the same position as the letter  $a$  of  $\sigma(e_2)$  in  $\sigma(\alpha)$ . Thus,  $|\tau(x_1 \cdot x_2 \cdot [\dots] \cdot x_n)| = n + 1$ . So, as  $\tau$  is nonerasing,  $|\tau(x_i)| \leq 2$ ,  $1 \leq i \leq n$ .

Hence, due to  $\tau(\alpha) = \sigma(\alpha)$  and  $\tau$  being nonerasing, it is required to define  $\tau$  for the variables  $x_1, x_2, \dots, x_n$  such that

- $\tau(x_1) = b \cdot v$ ,  $v \in \{\varepsilon, \sigma(x_1)\}$ , and
- for  $2 \leq j \leq n - 1$ , if  $\tau(x_{j-1})$  is not located at the position of  $\sigma(x_{j-1})$  in  $\sigma(\alpha)$ , then  $\tau(x_j) = \sigma(x_{j-1}) \cdot v$ ,  $v \in \{\varepsilon, \sigma(x_j)\}$ ; otherwise,  $\tau(x_j) = \sigma(x_j)$ , and
- if  $\tau(x_{n-1})$  is not located at the position of  $\sigma(x_{n-1})$  in  $\sigma(\alpha)$ , then  $\tau(x_n) = \sigma(x_{n-1}) \cdot \sigma(x_n)$ ; otherwise,  $\tau(x_n) = \sigma(x_n)$ .

This implies that, for every  $i$ ,  $1 \leq i \leq n$ ,  $\tau(i)$  satisfies Condition (I) of the lemma.

- $\tau(e_1) = b$ , and this letter is located at the same position as the letter  $b$  of  $\sigma(e_1)$  in  $\sigma(\alpha)$ ; furthermore,  $\tau(e_2) = b$ , and this letter is located at the same position as the letter  $b$  of  $\sigma(e_2)$  in  $\sigma(\alpha)$ . Thus,  $|\tau(x_1 \cdot x_2 \cdot [\dots] \cdot x_n)| = n + 1$ , which, as  $\tau$  is nonerasing, implies  $|\tau(x_i)| \leq 2$ ,  $1 \leq i \leq n$ .

Therefore, since  $\tau(\alpha) = \sigma(\alpha)$  and  $\tau$  is nonerasing,  $\tau$  needs to be defined for the variables  $x_1, x_2, \dots, x_n$  such that

- $\tau(x_1) = \sigma(x_1) \cdot v$ ,  $v \in \{\varepsilon, \sigma(x_2)\}$ , and
- for  $2 \leq j \leq n - 1$ , if  $\tau(x_{j-1})$  is not located at the position of  $\sigma(x_j)$  in  $\sigma(\alpha)$ , then  $\tau(x_j) = \sigma(x_j) \cdot v$ ,  $v \in \{\varepsilon, \sigma(x_{j+1})\}$ ; otherwise,  $\tau(x_j) = \sigma(x_{j+1})$ , and
- if  $\tau(x_{n-1})$  is not located at the position of  $\sigma(x_n)$  in  $\sigma(\alpha)$ , then  $\tau(x_n) = \sigma(x_n) \cdot a$ ; otherwise,  $\tau(x_n) = a$ .

Consequently, for every  $i$ ,  $1 \leq i \leq n$ ,  $\tau(i)$  satisfies Condition (II) of the lemma. □

In the following proposition, we establish a sufficient condition on the existence of weakly unambiguous length-increasing morphisms that we shall use in the proof of Theorem 20.

**Proposition 25.** *Let  $\alpha \in \mathbb{N}^+$ . If there exists an  $s \in S_\alpha$  satisfying, for an  $e \in E_\alpha$ ,  $s \cdot e \sqsubseteq \alpha$  and  $e \cdot s \sqsubseteq \alpha$ , then there is a length-increasing morphism  $\sigma : \mathbb{N}^+ \rightarrow \{a, b\}^+$  that is weakly unambiguous with respect to  $\alpha$ .*

*Proof.* According to Definition 5, since  $s \in S_\alpha$ , one of the following cases is satisfied:

1.  $\varepsilon \notin L_s$  and, for every  $i \in L_s$ ,  $R_i = \{s\}$ , or
2.  $\varepsilon \notin R_s$  and, for every  $i \in R_s$ ,  $L_i = \{s\}$ .

Without loss of generality, we only consider the first case (since a same reasoning can be applied for the second case). The conditions of the proposition and of Case 1 imply that there exists the following unique factorisation of  $\alpha$ :

$$\alpha = \alpha_1 \cdot \beta_1 \cdot \alpha_2 \cdot \beta_2 \cdot \alpha_3 \cdot \dots \cdot \alpha_n \cdot \beta_n \cdot \alpha_{n+1},$$

where  $n := |\alpha|_e$ ,  $\alpha_1, \alpha_2, \dots, \alpha_{n+1} \in (\mathbb{N} \setminus \{e\})^*$ , and, for every  $k$  with  $1 \leq k \leq n$ ,

- $\beta_k = s \cdot e \cdot s$  or
- $\beta_k = s' \cdot e \cdot s$  for an  $s' \in \text{var}(\alpha) \cup \{\varepsilon\}$ .

Note that, due to the conditions  $s \cdot e \sqsubseteq \alpha$  and  $e \cdot s \sqsubseteq \alpha$ , there must exist at least one  $k'$ ,  $1 \leq k' \leq n$ , with  $\beta_{k'} = s \cdot e \cdot s$ .

We now consider the length-increasing morphism  $\sigma : \mathbb{N}^+ \rightarrow \{a, b\}^+$ , given by  $\sigma(e) := aa$  and, for every  $x \in \text{var}(\alpha) \setminus \{e\}$ ,  $\sigma(x) := b$ . Assume to the contrary that there exists a morphism  $\tau : \mathbb{N}^+ \rightarrow \{a, b\}^+$  satisfying  $\tau(\alpha) = \sigma(\alpha)$  and, for a variable  $q \in \text{var}(\alpha)$ ,  $\tau(q) \neq \sigma(q)$ . According to Lemma 11, we can conclude that this implies  $\tau(e) = a$ . Furthermore, due to  $s \cdot e \cdot s \sqsubseteq \alpha$ ,  $\tau(s)$  needs to contain the letter  $a$  as a factor. However, it follows from the above factorisation of  $\alpha$  that  $|\alpha|_s > |\alpha|_e$ , and therefore  $|\tau(\alpha)|_a > 2|\alpha|_e = |\sigma(\alpha)|_a$ . This contradicts the assumption  $\tau(\alpha) = \sigma(\alpha)$ .  $\square$

Based on the preparatory work in Lemmata 21, 22, 23, 24 and Proposition 25, we can now verify Theorem 20:

*Proof of Theorem 20.* We assume to the contrary that there exists an  $e \in E_\alpha$  such that, for every  $e' \in E_\alpha$  with  $e' \neq e$ ,  $e \cdot e'$  or  $e' \cdot e$  is not a factor of  $\alpha$ .

According to Proposition 25, since there is no weakly unambiguous length-increasing morphism  $\sigma$  with respect to  $\alpha$ , there exists no variable  $s \in S_\alpha$  with  $s \cdot e \sqsubseteq \alpha$  and  $e \cdot s \sqsubseteq \alpha$ . Thus, and due to our assumption, there is *no* variable  $x \in \text{var}(\alpha) \setminus \{e\}$  satisfying both  $x \in L_e$  and  $x \in R_e$ . Since  $e \in E_\alpha$ , we can therefore conclude that at least one of the following cases is satisfied:

1.  $ee \sqsubseteq \alpha$ ,
2. if  $\varepsilon \notin L_e$ , then there exists an  $l \in L_e$  with  $R_l \neq \{e\}$  and  $e \notin L_l$ , and if  $\varepsilon \notin R_e$ , then there exists an  $r \in R_e$  with  $L_r \neq \{e\}$  and  $e \notin R_r$ , or

3.  $\varepsilon \in L_e$  and  $\varepsilon \in R_e$ .

Due to the fact that, for every  $x \in \text{var}(\alpha) \setminus \{e\}$ ,  $x \cdot e$  or  $e \cdot x$  is not a factor of  $\alpha$ , Case 1 satisfies the conditions of Lemma 21. Hence, there is a weakly unambiguous length-increasing morphism  $\sigma : \mathbb{N}^+ \rightarrow \Sigma^+$  with respect to  $\alpha$ . This contradicts the condition of Theorem 20, namely that there is no weakly unambiguous morphism  $\sigma$  with respect to  $\alpha$ .

Our investigation of Case 2 is based on the assumption that Case 1 is not satisfied. This implies that  $l \neq e$  and  $r \neq e$ . As mentioned, there is no variable  $x \in \text{var}(\alpha) \setminus \{e\}$  satisfying  $x \in L_e$  and  $x \in R_e$ . Consequently, it follows from Case 2 that  $e \cdot l$  and  $r \cdot e$  are not factors of  $\alpha$ ; in other words,  $e \notin L_l$  and  $e \notin R_r$ . Also, we can conclude that  $l \neq r$ . We divide Case 2 into two parts, Part (a) and Part (b). In Part (a) we assume that  $l \cdot e \cdot r$  is a factor of  $\alpha$ , and in Part (b) we assume that  $l \cdot e \cdot r$  is not a factor of  $\alpha$ .

Part (a)  $l \cdot e \cdot r \sqsubseteq \alpha$ .

We define a morphism  $\sigma : \mathbb{N}^+ \rightarrow \{a, b\}^+$  by

$$\sigma(x) := \begin{cases} bb, & x = e, \\ a, & x \neq e. \end{cases}$$

According to Lemma 22,  $\sigma$  is weakly unambiguous with respect to  $\alpha$ , which contradicts the condition of Theorem 20.

Part (b)  $l \cdot e \cdot r \not\sqsubseteq \alpha$ .

We now consider the following cases:

*Case 2.1.*  $|\alpha|_e = 1$ .

Hence, according to Case 2 and  $l \cdot e \cdot r \not\sqsubseteq \alpha$ , we can assume that  $\alpha = \dots \cdot k \cdot l \cdot k' \cdot \dots \cdot l \cdot e$  or  $\alpha = e \cdot r \cdot \dots \cdot k \cdot r \cdot k' \cdot \dots$ ,  $k, k' \in \text{var}(\alpha) \setminus \{e\}$ . We define a morphism  $\sigma : \mathbb{N}^+ \rightarrow \{a, b\}^+$  by

$$\sigma(x) := \begin{cases} bb, & x = e, \\ a, & x \neq e. \end{cases}$$

Using Lemma 11, it can be easily verified that  $\sigma$  is weakly unambiguous with respect to  $\alpha$ , which is a contradiction.

*Case 2.2.*  $|\alpha|_e > 1$ .

Consequently, according to Case 2 and  $l \cdot e \cdot r \not\sqsubseteq \alpha$ , there exists an  $l \in L_e$  with  $R_l \neq \{e\}$  and  $e \notin L_l$ , and there exists an  $r \in R_e$  with  $L_r \neq \{e\}$  and  $e \notin R_r$ . Therefore, we can assume that  $\alpha = \dots \cdot l \cdot e \cdot \dots \cdot e \cdot r \cdot \dots$ . As mentioned above, there is no variable  $x \in \text{var}(\alpha)$  with  $x \in L_e$  and  $x \in R_e$ . As a result, we can define a morphism  $\sigma$  by

$$\sigma(x) := \begin{cases} ab, & x = e, \\ b, & x \in L_e, \\ a, & x \in R_e. \end{cases} \quad (5)$$

For the other variables, we shall define the morphism  $\sigma$  later. Before we do this, we shall establish some insights into the structure of  $\alpha$ . According to Definition (5),  $\sigma(l) = b$  and  $\sigma(r) = a$ . Also, due to the condition of Theorem 20, there exists a nonerasing morphism  $\tau$  with  $\tau(\alpha) = \sigma(\alpha)$  and, for some  $q \in \text{var}(\alpha)$ ,  $\tau(q) \neq \sigma(q)$ . Moreover, since  $\sigma(e)$  is the only image of length more than 1, Lemma 11 implies that  $\tau(e) = a$  or  $\tau(e) = b$ . We first consider two special cases as follows:

- Let there be an occurrence of  $r$  (denoted by  $r'$ ) such that  $\alpha = \alpha_1 \cdot r' \cdot \alpha_2$ ,  $\alpha_1 \in \mathbb{N}^*$ ,  $\alpha_2 \in \mathbb{N}^+$  and  $e \not\sqsubseteq \alpha_1$ . By considering the factor  $e \cdot r$ , if  $\tau(e) = a$ , then Lemma 24 and  $\tau(\alpha) = \sigma(\alpha)$  imply that  $\tau(r) = b \cdot v$ ,  $v \in \{\varepsilon, a\}$ . However, according to Lemma 23, the letters  $a$  which are produced by  $\tau(e)$  are located at the same positions as those letters  $a$  produced by  $\sigma(e)$  in  $\sigma(\alpha)$ , and since the length of images of all variables except  $e$  is 1,  $\tau(r') = \sigma(r') = a$  must be satisfied in order to obtain  $\tau(\alpha) = \sigma(\alpha)$ . This means that  $\tau(r) \neq \tau(r')$ , which is a contradiction.
- Let there be an occurrence of  $l$  (denoted by  $l'$ ) such that  $\alpha = \alpha_1 \cdot l' \cdot \alpha_2$ ,  $\alpha_1 \in \mathbb{N}^+$ ,  $\alpha_2 \in \mathbb{N}^*$  and  $e \not\sqsubseteq \alpha_2$ . If we consider the factor  $l \cdot e$ , and if we assume  $\tau(e) = b$ , then Lemma 24 and  $\tau(\alpha) = \sigma(\alpha)$  imply that  $\tau(l) = v \cdot a$ ,  $v \in \{\varepsilon, b\}$ . Due to Lemma 23, the letters  $b$  which are produced by  $\tau(e)$  are located at the same positions as those letters  $b$  produced by  $\sigma(e)$  in  $\sigma(\alpha)$ , and since the length of images of all variables except  $e$  is 1,  $\tau(l') = \sigma(l') = b$  must hold true. Thus,  $\tau(l) \neq \tau(l')$ , and this is a contradiction.

By considering the above special cases, and without loss of generality regarding the different possibilities of the positions of  $l$  and  $r$  in  $\alpha$ , let

$$\alpha := \alpha_1 \cdot e \cdot x_1 \cdot x_2 \cdot [\dots] \cdot x_n \cdot r \cdot \alpha_2 \cdot l \cdot z_1 \cdot z_2 \cdot [\dots] \cdot z_m \cdot e \cdot \alpha_3, \quad (6)$$

with  $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{N}^*$ , for every  $i$ ,  $1 \leq i \leq n$  and, for every  $j$ ,  $1 \leq j \leq m$ ,  $x_i \in \text{var}(\alpha)$ ,  $x_i \neq e$ ,  $x_i \neq r$ ,  $z_j \in \text{var}(\alpha)$ ,  $z_j \neq e$  and  $z_j \neq l$ . Also, let  $\alpha_2 := y_1 \alpha'_2$  with  $y_1 \in \text{var}(\alpha) \cup \{\varepsilon\}$  and  $\alpha'_2 \in \mathbb{N}^*$ . Since  $r \cdot e$  is not a factor of  $\alpha$ ,  $y_1 \neq e$ . Furthermore, if we assume that  $y_1 = r$ , then  $rr \sqsubseteq \alpha$  and, in accordance with Lemma 12,  $r \in E_\alpha$ . Consequently, according to Case 1, the assumption of  $y_1 = r$  leads to a contradiction. Hence,  $y_1 \neq r$ .

Now, we define  $\sigma$  for the other variables using the following algorithm, where, for any variable  $x$ , the notation  $\sigma(x) = \text{null}$  shall refer to the fact that  $\sigma(x)$  has not been defined yet.

```

1:  $i \leftarrow n$ 
2: while  $\sigma(x_i) = b$  do
3:    $i \leftarrow i - 1$ 
4: end while
5: if  $\sigma(x_i) = \text{null}$  then
6:    $\sigma(x_i) \leftarrow a$ 
7: end if
8:  $i \leftarrow 1$ 
9: while  $\sigma(z_i) = a$  do

```



```

10:    $i \leftarrow i + 1$ 
11: end while
12: if  $\sigma(z_i) = \text{null}$  then
13:    $\sigma(z_i) \leftarrow b$ 
14: end if
15: if  $\alpha_2 \neq \varepsilon$  and  $\sigma(y_1) = \text{null}$  then
16:    $\sigma(y_1) \leftarrow b$ 
17: end if
18: for all  $x \in \text{var}(\alpha)$  do
19:   if  $\sigma(x) = \text{null}$  then
20:      $\sigma(x) \leftarrow a$ 
21:   end if
22: end for

```

We now show that this definition of  $\sigma$  and the conditions of Case 2.2 lead to the following contradictory statement:

*Claim.* *The morphism  $\sigma$  is weakly unambiguous with respect to  $\alpha$ .*

*Proof (Claim).* We assume to the contrary that there exists a nonerasing morphism  $\tau$  satisfying  $\tau(\alpha) = \sigma(\alpha)$  and, for some  $q \in \text{var}(\alpha)$ ,  $\tau(q) \neq \sigma(q)$ . It follows from Lemmata 11 and 23, that  $\tau(e) = a$  or  $\tau(e) = b$  which is located at the same position as that letter  $a$  or  $b$  produced by  $\sigma(e)$  in  $\sigma(\alpha)$ . Due to the factors  $e \cdot r$  and  $l \cdot e$  and due to Lemma 24,

$$\begin{aligned} \tau(e) = a \text{ implies that } \tau(r) = b \cdot v, v \in \{\sigma(r), \varepsilon\}, \text{ and} \\ b \text{ is a suffix of } \tau(l) \end{aligned} \tag{7}$$

and

$$\begin{aligned} \tau(e) = b \text{ implies that } \tau(l) = v \cdot a, v \in \{\sigma(l), \varepsilon\}, \text{ and} \\ a \text{ is a prefix of } \tau(r), \end{aligned} \tag{8}$$

since otherwise  $\tau(\alpha) \neq \sigma(\alpha)$ . On the other hand, we know that there exist factors  $x_n \cdot r$  and  $l \cdot z_1$  in  $\alpha$ . Now, we consider the following cases:

- $\tau(e) = a$ . As a result of Implication (7),  $b$  is a prefix of  $\tau(r)$ . We consider the factor  $e \cdot x_1 \cdot x_2 \cdot [\dots] \cdot x_n \cdot r$  of  $\alpha$ . According to Lemma 24,  $\tau(r) = \sigma(r)$  or  $\tau(r) = \sigma(x_n) \cdot v$ ,  $v \in \{\sigma(r), \varepsilon\}$ . Since  $\sigma(r) = a$  and  $b$  is a prefix of  $\tau(r)$ ,  $\tau(r) = \sigma(r)$  cannot be satisfied. Hence,  $\tau(r) = \sigma(x_n) \cdot v$ ,  $v \in \{a, \varepsilon\}$ . Since  $b$  is a prefix of  $\tau(r)$ ,  $\sigma(x_n) = b$ . However, this implies that  $\sigma(x_n)$  has been assigned before running the algorithm, and this leads to the fact that  $x_n \in L_e$ . According to the proof of Lemma 24,  $\tau(x_n)$  must be located at the position of  $\sigma(x_{n-1})$ , or in other words,  $\tau(x_n) = \sigma(x_{n-1})$ . Thus, if  $\sigma(x_{n-1}) = a$ , then  $\tau(x_n) = a$ , while Lemma 23 and Lemma 24 imply that, due to  $x_n \in L_e$  and  $\tau(e) = a$ ,  $b$  is a suffix of  $\tau(x_n)$ . So,  $\sigma(x_{n-1})$  must equal  $b$ , which means that  $x_{n-1} \in L_e$ . This argument can then be extended to  $\tau(x_{n-1}) = \sigma(x_{n-2})$ . If the value of  $\sigma$  for

all variables  $x_n, x_{n-1}, \dots, x_2$  equals  $b$ , since  $\sigma(x_1) = a$ , we finally get a contradiction, because  $\tau(x_2) = \sigma(x_1) = a$ , while  $x_2 \in L_e$ , which means that  $b$  is a suffix of  $\tau(x_2)$ . Hence,  $\tau(e)$  cannot equal  $a$ .

- $\tau(e) = b$ . Because of Implication (8),  $a$  is a suffix of  $\tau(l)$ . We consider the factor  $l \cdot z_1 \cdot z_2 \cdot [\dots] \cdot z_m \cdot e$  of  $\alpha$ . According to Lemma 24,  $\tau(l) = \sigma(l)$  or  $\tau(l) = v \cdot \sigma(z_1)$ ,  $v \in \{\sigma(l), \varepsilon\}$ . Due to  $\sigma(l) = b$ ,  $\tau(l)$  cannot equal  $\sigma(l)$ , because we know that the factor  $a$  is a suffix of  $\tau(l)$ . Hence,  $\tau(l) = v \cdot \sigma(z_1)$ ,  $v \in \{b, \varepsilon\}$ . Since the factor  $a$  is a suffix of  $\tau(l)$ ,  $\sigma(z_1) = a$  follows; in other words,  $\tau(l) = v \cdot a$ ,  $v \in \{\sigma(l), \varepsilon\}$ . For the other variables  $z_j$ ,  $1 \leq j \leq m$ , we investigate the morphisms  $\sigma$  and  $\tau$  as follows:

*Assumption 1.* Assume that, for every  $j$ ,  $1 \leq j \leq m$ ,  $\sigma(z_j)$  is not defined by line 6 of the algorithm.

By considering this assumption, it follows from  $\sigma(z_1) = a$  that  $\sigma(z_1)$  has been defined before running the algorithm, and this means that  $z_1 \in R_e$ . So, Lemma 23 and Lemma 24 imply that, due to  $z_1 \in R_e$  and  $\tau(e) = b$ ,  $a$  is a prefix of  $\tau(z_1)$ . Moreover, as mentioned above,  $\tau(l) = v \cdot \sigma(z_1)$ ,  $v \in \{\sigma(l), \varepsilon\}$ . According to Lemma 24,  $\tau(z_1) = \sigma(z_2)$ , or, in other words,  $\tau(z_1)$  is located at the position of  $\sigma(z_1)$ . If  $\sigma(z_2) = b$ , then  $\tau(z_1) = b$ , which contradicts the fact that  $a$  is a prefix of  $\tau(z_1)$ . Consequently,  $\sigma(z_2)$  must equal  $a$ , which means that  $z_2 \in R_e$ . This discussion can be continued for  $\tau(z_2) = \sigma(z_3)$ . If the value of  $\sigma$  for all variables  $z_1, z_2, \dots, z_{m-1}$  equals  $a$ , since  $\sigma(z_m) = b$ , we finally get a contradiction, because  $\tau(z_{m-1}) = \sigma(z_m) = b$ , while  $z_{m-1} \in R_e$ , which means that  $a$  is a prefix of  $\tau(z_{m-1})$ . Hence,  $\tau(e)$  cannot equal  $b$ .

*Assumption 2.* Assume that there exists a  $j$ ,  $1 \leq j \leq m-1$ , such that  $\sigma(z_j)$  is defined by line 6 of the algorithm.

This means that  $\sigma(z_j) = a$ . Since line 6 of our algorithm just runs once, if  $\sigma(z_{j+1}) = a$ , then  $z_{j+1} \in R_e$  and we can use the above argument, which again leads to a contradiction. So, this implies that  $\sigma(z_{j+1}) = b$ . According to Lemma 24, as  $\tau$  is nonerasing and  $\tau(\alpha) = \sigma(\alpha)$ ,  $\tau(z_j) = \sigma(z_{j+1}) = b$ , or, in other words,  $\tau(z_j)$  is located at the position of  $\sigma(z_{j+1})$ . On the other hand, Assumption 2 means that  $z_j$  has another occurrence to the left of  $r$  in  $\alpha$ . In fact, there exists a  $k$ ,  $1 \leq k \leq n$ , with  $x_k = z_j$ . Hence,  $\tau(x_k) = \tau(z_j) = b$  and  $\sigma(x_k) = \sigma(z_j) = a$ . According to Lemma 24 and its proof, since  $\sigma(x_k) = a$  and  $\tau(x_k) = b$ , for every  $q$ ,  $k \leq q \leq (n-1)$ ,  $\tau(x_q) = \sigma(x_{q+1})$ , and  $\tau(x_n) = \sigma(r)$  and  $\tau(r) = \sigma(y_1)$  if  $\alpha_2 \neq \varepsilon$ ; otherwise,  $\tau(r) = \sigma(l)$ . If  $k = n$ , then  $\tau(x_k) = \sigma(r) = a$ , and this contradicts  $\tau(x_k) = \tau(z_j) = b$ . As a result,  $k < n$ . If  $\tau(r) = \sigma(l) = b$  or  $\tau(r) = \sigma(y_1) = b - \sigma(y_1) = b$  follows from line 16 of our algorithm; then this contradicts the fact that  $a$  is a prefix of  $\tau(r)$ , which follows from Implication (8). However, if  $\sigma(y_1) = a$ , then this implies that  $y_1 \in R_e$  or  $y_1 = x_k$ . Also, since  $\sigma(x_k)$  is assigned by line 6 of our algorithm, and due to  $k < n$ , for every  $q$ ,  $k \leq q \leq (n-1)$ ,  $x_q \in L_e$ . As a result,  $x_n \in L_e$ .

We now consider the factor  $x_n \cdot r \cdot y_1$ . It follows from

$$\begin{aligned} y_1 &\in R_e \text{ or } y_1 = x_k, \ k < n, \text{ and} \\ x_n &\in L_e \end{aligned}$$

that  $r \in E_\alpha$ , and  $\sigma(y_1) = a$  and  $\sigma(x_n) = b$  imply that  $y_1 \neq x_n$ . We now denote  $r$ ,  $x_n$  and  $y_1$  by  $e'$ ,  $l'$  and  $r'$ , respectively; thus,  $l' \neq r'$ . Since  $e' \in E_\alpha$ , if  $r' = e'$ , then  $e'e' \sqsubseteq \alpha$  and we can consider Case 1 of our proof, which leads to a contradiction. So,  $r' \neq e'$ . Moreover, according to the definition of  $\alpha$ , for every  $i$ ,  $1 \leq i \leq n$ ,  $x_i \neq r$ . Consequently,  $x_n \neq r$  and, hence,  $e' \neq l'$ . Then, since  $l' \cdot e' \cdot r' \sqsubseteq \alpha$ , we can consider Part (a) of Case 2 of our proof with

$$\sigma(x) := \begin{cases} bb, & x = e', \\ a, & x \neq e', \end{cases}$$

which leads to a contradiction, due to  $\sigma$  being weakly unambiguous with respect to  $\alpha$ . Consequently, we cannot consider  $\tau(e) = b$ .

It follows from the above cases that we cannot define a morphism  $\tau$  satisfying  $\tau(\alpha) = \sigma(\alpha)$ . Consequently,  $\sigma$  is weakly unambiguous and this concludes the proof of the Claim.  $\square$  (*Claim*)

This Claim is a direct contradiction to the assumption of Theorem 20. In order to conclude our reasoning on Case 2.2, it is necessary to mention that, instead of Factorisation (6) of  $\alpha$ , we can define  $\alpha$  such that the variable  $l$  is located to the left of the position of  $r$  in  $\alpha$ . More precisely, we can consider

$$\alpha := \alpha_1 \cdot e \cdot x_1 \cdot x_2 \cdot [\dots] \cdot x_k \cdot l \cdot x_{k+1} \cdot x_{k+2} \cdot [\dots] \cdot x_n \cdot r \cdot z_1 \cdot z_2 \cdot [\dots] \cdot z_m \cdot e \cdot \alpha_2,$$

with  $\alpha_1, \alpha_2 \in \mathbb{N}^*$ , for every  $i$ ,  $1 \leq i \leq n$ ,  $x_i \neq e \neq l$  and, for every  $j$ ,  $1 \leq j \leq m$ ,  $z_j \neq e \neq r$ . However, for this factorisation a simplified version of our above reasoning on Factorisation (6) can be used in order to obtain a contradiction.

In order to investigate Case 3, we assume that Cases 1 and 2 are not satisfied. Since  $\varepsilon \in L_e$  and  $\varepsilon \in R_e$ , we can write  $\alpha := e \cdot \alpha_1 \cdot e$ . We define a length-increasing morphism  $\sigma : \mathbb{N}^+ \rightarrow \{a, b\}^+$  by

$$\sigma(x) := \begin{cases} ab, & x = e, \\ a, & \text{else.} \end{cases}$$

Thus,  $\sigma(\alpha) = ab \cdot \sigma(\alpha_1) \cdot ab$ . According to Lemma 11, if  $\sigma$  is not weakly unambiguous, then there exists a nonerasing morphism  $\tau : \mathbb{N}^+ \rightarrow \{a, b\}^+$  satisfying  $\tau(\alpha) = \sigma(\alpha)$ , while  $\tau(e) \sqsubset \sigma(e)$ . This implies that  $\tau(e) = a$  or  $\tau(e) = b$ . Consequently,  $\tau(\alpha) = a \cdot \tau(\alpha_1) \cdot a$  or  $\tau(\alpha) = b \cdot \tau(\alpha_1) \cdot b$  which contradicts  $\tau(\alpha) = \sigma(\alpha)$ . Hence,  $\sigma$  is weakly unambiguous with respect to  $\alpha$ . This contradicts the condition of Theorem 20, namely that there is no weakly unambiguous length-increasing morphism  $\sigma$  with respect to  $\alpha$ .  $\square$

Theorem 20 (when compared to Theorem 13) provides deep insights into the difference between binary and ternary target alphabets if the weak unambiguity of morphisms is studied. In addition to this, it implies that whenever, for a given pattern  $\alpha \in \mathbb{N}^+$  with  $E_\alpha \neq \emptyset$ , there exists an  $e \in E_\alpha$  such that, for every  $e' \in E_\alpha$  with  $e' \neq e$ , the factors  $e \cdot e'$  or  $e' \cdot e$  do not occur in  $\alpha$ , then there is a weakly unambiguous length-increasing morphism  $\sigma : \mathbb{N}^+ \rightarrow \Sigma^+$ ,  $\Sigma = \{a, b\}$ , with respect to  $\alpha$ . It must be noted, though, that Theorem 20 does not describe a sufficient condition for the non-existence of weakly unambiguous length-increasing morphisms in case of  $|\Sigma| = 2$ ; this is easily demonstrated by the pattern  $1 \cdot 2 \cdot 1$  and further illustrated by Example 31.

As can be concluded from Example 7 and Theorem 13, there is a weakly unambiguous length-increasing morphism  $\sigma : \mathbb{N}^+ \rightarrow \Sigma^+$ ,  $|\Sigma| \geq 3$ , with respect to  $\alpha = 1 \cdot 2 \cdot \mathbf{3} \cdot 4 \cdot 5 \cdot 6 \cdot \mathbf{4} \cdot \mathbf{3} \cdot 7 \cdot 8$ , and we can define  $\sigma$  by  $\sigma(3) := bc$  and, for every  $j \neq 3$ ,  $\sigma(j) := a$ . In contrast to this, the next theorem implies that there is no weakly unambiguous morphism with respect to  $\alpha$  if  $|\Sigma| = 2$ . In order to prove this theorem, we need the following lemma.

**Lemma 26.** *Let  $\Sigma$  be an alphabet,  $|\Sigma| = 2$ , and let  $\sigma : \mathbb{N}^+ \rightarrow \Sigma^+$  be a morphism. For all  $x_1, x_2 \in \mathbb{N}$ , there exist  $a_1, a_2 \in \Sigma$  with  $a_1 \sqsubseteq \sigma(x_1)$  and  $a_2 \sqsubseteq \sigma(x_2)$  such that  $a_1 a_2 \sqsubseteq \sigma(x_1 \cdot x_2)$  and  $a_2 a_1 \sqsubseteq \sigma(x_2 \cdot x_1)$ .*

*Proof.* If  $a_1$  is a prefix and a suffix of  $\sigma(x_1)$  and  $a_2$  is a prefix and a suffix of  $\sigma(x_2)$ , then Lemma 26 holds trivially true. We can therefore restrict this proof to a situation where the first and the last letters of  $\sigma(x_1)$  differ or the first and the last letters of  $\sigma(x_2)$  differ. Let  $\Sigma := \{a, b\}$ . Without loss of generality, we can exclusively consider  $\sigma(x_1) = a \cdots b$ , since all other cases can be dealt with in an analogous manner.

Regarding  $\sigma(x_2)$ , we now consider the following cases:

- $\sigma(x_2)$  starts with  $a$ .

We define  $a_1 := b$  and  $a_2 := a$ . Then  $a_1 \sqsubseteq \sigma(x_1)$  and  $a_2 \sqsubseteq \sigma(x_2)$ , and  $a_1 a_2 \sqsubseteq \sigma(x_1 \cdot x_2)$ . Furthermore, since  $\sigma(x_1) = a_2 \cdots a_1$ , there must be a factor  $a_2 a_1$  in  $\sigma(x_1)$ , which directly implies that  $a_2 a_1$  is also a factor of  $\sigma(x_2 \cdot x_1)$ . Thus, Lemma 26 holds true for this choice of  $a_1$  and  $a_2$ .

- $\sigma(x_2)$  starts with  $b$  and ends with  $b$ .

We define  $a_1 := a$  and  $a_2 := b$ . This again implies that  $a_1 \sqsubseteq \sigma(x_1)$  and  $a_2 \sqsubseteq \sigma(x_2)$ . Since  $\sigma(x_1) = a_1 \cdots a_2$ , there must be a factor  $a_1 a_2$  in  $\sigma(x_1)$ , and, hence, in  $\sigma(x_1 \cdot x_2)$ . Finally, when considering the last letter of  $\sigma(x_2)$  and the first letter of  $\sigma(x_1)$ , we can immediately observe that  $a_2 a_1$  is a factor of  $\sigma(x_2 \cdot x_1)$ .

- $\sigma(x_2)$  starts with  $b$  and ends with  $a$ .

We define  $a_1 := b$  and  $a_2 := a$ , which means that  $a_1 \sqsubseteq \sigma(x_1)$  and  $a_2 \sqsubseteq \sigma(x_2)$ . Since  $\sigma(x_1) = a_2 \cdots a_1$  and  $\sigma(x_2) = a_1 \cdots a_2$ ,  $\sigma(x_1)$  contains a factor  $a_2 a_1$  and  $\sigma(x_2)$  contains a factor  $a_1 a_2$ . Consequently, both  $\sigma(x_1 \cdot x_2)$  and  $\sigma(x_2 \cdot x_1)$  contain these factors as well.

□

The next result introduces a *sufficient* condition on the *non-existence* of weakly unambiguous length-increasing morphisms  $\sigma : \mathbb{N}^+ \rightarrow \Sigma^+$ ,  $|\Sigma| = 2$ . According to Theorem 20, it is necessary for the non-existence of such morphisms, with respect to a given pattern  $\alpha \in \mathbb{N}^+$  that, for every  $e \in E_\alpha$ , there exists an  $e' \in E_\alpha$ ,  $e' \neq e$ , such that  $e \cdot e'$  and  $e' \cdot e$  are factors of  $\alpha$ . Hence, this requirement must be satisfied in the following theorem.

**Theorem 27.** *Let  $\alpha \in \mathbb{N}^+$  satisfying  $E_\alpha \neq \emptyset$ . Let  $\Sigma$  be an alphabet,  $|\Sigma| = 2$ . There is no weakly unambiguous length-increasing morphism  $\sigma : \mathbb{N}^+ \rightarrow \Sigma^+$  with respect to  $\alpha$  if all of the following four conditions are satisfied:*

1. *for every  $e \in E_\alpha$ ,  $e^2 \not\sqsubseteq \alpha$ , and there is exactly one  $e' \in E_\alpha \setminus \{e\}$  such that  $e' \in L_e$  or  $e' \in R_e$ ,  $e' \cdot e \cdot e' \not\sqsubseteq \alpha$ , and there are  $s_1, s_2, s_3, s_4 \in S_\alpha$  such that  $s_1 \cdot e \cdot e' \cdot s_2$  and  $s_3 \cdot e' \cdot e \cdot s_4$  are factors of  $\alpha$ ,*
2. *for every  $e \in E_\alpha$ ,  $\varepsilon \notin R_e$  and  $\varepsilon \notin L_e$ ,*
3. *for any  $s, s' \in S_\alpha$  and  $e, e' \in E_\alpha$ , if  $(s \cdot e \cdot e' \cdot s') \sqsubset \alpha$ , then, for all occurrences of  $s$  and  $s'$  in  $\alpha$ , the right neighbour of  $s$  is the factor  $e \cdot e'$  and the left neighbour of  $s'$  is the factor  $e \cdot e'$ , and*
4. *for any  $s, s' \in S_\alpha$  and  $e \in E_\alpha$ , if  $(s \cdot e \cdot s') \sqsubset \alpha$ , then  $R_s = \{e\}$  and  $L_{s'} = \{e\}$ .*

*Proof.* We prove that there is no weakly unambiguous length-increasing morphism  $\sigma : \mathbb{N}^+ \rightarrow \Sigma^+$ ,  $|\Sigma| = 2$ , with respect to  $\alpha$ . This means that, for every morphism  $\sigma$ , there exists a morphism  $\tau : \mathbb{N}^+ \rightarrow \Sigma^+$  satisfying  $\tau(\alpha) = \sigma(\alpha)$  and, for some  $q \in \text{var}(\alpha)$ ,  $\tau(q) \neq \sigma(q)$ . According to Corollary 19, if there exists a variable  $j \in S_\alpha$  with  $|\sigma(j)| > 1$ , then  $\sigma$  is not weakly unambiguous with respect to  $\alpha$ . Consequently, we can safely restrict our considerations to the set  $E_\alpha$ , and we can assume that, for every  $j \in S_\alpha$ ,  $|\sigma(j)| = 1$ . Hence, we choose an arbitrary variable  $e_1$  from  $E_\alpha$ , and we assume that  $|\sigma(e_1)| > 1$ . According to the conditions of Theorem 27, there is exactly one  $e_2 \in E_\alpha$  such that  $e_2 \in L_{e_1}$  or  $e_2 \in R_{e_1}$ . Moreover, it follows from the conditions that  $s_1 \cdot e_1 \cdot e_2 \cdot s_2$  and  $s_3 \cdot e_2 \cdot e_1 \cdot s_4$ , with  $s_1, s_2, s_3, s_4 \in S_\alpha$ , are factors of  $\alpha$ . Let,

$$\alpha := \alpha_1 \cdot s_1 \cdot e_1 \cdot e_2 \cdot s_2 \cdot \alpha_2 \cdot s_3 \cdot e_2 \cdot e_1 \cdot s_4 \cdot \alpha_3,$$

$\alpha_1, \alpha_2, \alpha_3 \in \mathbb{N}^*$ . So,

$$\sigma(\alpha) = \sigma(\alpha_1) \cdot \sigma(s_1) \sigma(e_1 \cdot e_2) \sigma(s_2) \cdot \sigma(\alpha_2) \cdot \sigma(s_3) \sigma(e_2 \cdot e_1) \sigma(s_4) \cdot \sigma(\alpha_3).$$

In accordance with Lemma 26, there exists a factor  $a_1 a_2$ ,  $a_1, a_2 \in \Sigma$ , such that  $\sigma(e_1 e_2) = u \cdot a_1 a_2 \cdot v$ ,  $u, v \in \Sigma^*$ ,  $\sigma(e_2 e_1) = u' \cdot a_2 a_1 \cdot v'$ ,  $u', v' \in \Sigma^*$ , and  $a_1 \sqsubseteq \sigma(e_1)$  and  $a_2 \sqsubseteq \sigma(e_2)$ . Also, since  $|\sigma(e_1)| > 1$ ,  $uv \neq \varepsilon$  and  $u'v' \neq \varepsilon$ . We define a nonerasing morphism  $\tau$  by  $\tau(e_1) := a_1$ ,  $\tau(e_2) := a_2$ ,  $\tau(s_1) := \sigma(s_1)u$ ,  $\tau(s_2) := v\sigma(s_2)$ ,  $\tau(s_3) := \sigma(s_3)u$  and  $\tau(s_4) := v\sigma(s_4)$ . Consequently,  $\tau(s_1 \cdot e_1 \cdot e_2 \cdot s_2) = \sigma(s_1 \cdot e_1 \cdot e_2 \cdot s_2)$  and  $\tau(s_3 \cdot e_2 \cdot e_1 \cdot s_4) = \sigma(s_3 \cdot e_2 \cdot e_1 \cdot s_4)$ . Due to the assumption,  $e_1$  and  $e_2$  can occur in  $\alpha$  in accordance with the following cases:

- $s \cdot e_1 \cdot e_2 \cdot s'$ .

If we consider  $\tau(s) := \sigma(s)u$  and  $\tau(s') := v\sigma(s_2)$ , then  $\tau(s \cdot e_1 \cdot e_2 \cdot s') = \sigma(s \cdot e_1 \cdot e_2 \cdot s')$ .

- $s \cdot e_2 \cdot e_1 \cdot s'$ .

If we consider  $\tau(s) := \sigma(s)u'$  and  $\tau(s') := v'\sigma(s_2)$ , then  $\tau(s \cdot e_2 \cdot e_1 \cdot s') = \sigma(s \cdot e_2 \cdot e_1 \cdot s')$ .

- $s \cdot e_1 \cdot s'$ .

The definition  $\tau(s) := \sigma(s)u$  implies that  $\tau(s \cdot e_1 \cdot s') = \sigma(s \cdot e_1 \cdot s')$ .

- $s \cdot e_2 \cdot s'$ .

Defining  $\tau(s) := \sigma(s)u'$ , we have  $\tau(s \cdot e_2 \cdot s') = \sigma(s \cdot e_2 \cdot s')$ .

Also, we define  $\tau$  for every  $j \in \text{var}(\alpha) \setminus \{e_1, e_2\}$  with  $j \notin L_{e_1}, L_{e_2}, R_{e_1}, R_{e_2}$  by  $\tau(j) := \sigma(j)$ . Hence, conditions 1, 2, 3 and 4 imply  $\tau(\alpha) = \sigma(\alpha)$ , while  $\tau(e_1) \neq \sigma(e_1)$ . Consequently,  $\sigma$  is not weakly unambiguous with respect to the pattern  $\alpha$ . Since the variable  $e_1$  is an arbitrary variable of  $E_\alpha$ , we can conclude that there is no weakly unambiguous length-increasing morphism  $\sigma$  with respect to  $\alpha$ .  $\square$

In order to illustrate Theorem 27, we consider a few examples:

**Example 28.** *Let,*

$$\alpha := 1 \cdot 2 \cdot \mathbf{3} \cdot \mathbf{4} \cdot 5 \cdot 6 \cdot \mathbf{4} \cdot \mathbf{3} \cdot 7 \cdot 8 \cdot \mathbf{3} \cdot 9 \cdot 10,$$

$$\beta := 1 \cdot 2 \cdot \mathbf{4} \cdot 5 \cdot 6 \cdot \mathbf{3} \cdot \mathbf{4} \cdot 7 \cdot 8 \cdot \mathbf{3} \cdot 9 \cdot 10 \cdot \mathbf{4} \cdot \mathbf{3} \cdot 11 \cdot 12,$$

$$\gamma := 1 \cdot 2 \cdot \mathbf{3} \cdot \mathbf{4} \cdot 5 \cdot 6 \cdot \mathbf{7} \cdot \mathbf{8} \cdot 9 \cdot 10 \cdot \mathbf{4} \cdot \mathbf{3} \cdot 11 \cdot 12 \cdot \mathbf{8} \cdot \mathbf{7} \cdot 13 \cdot 14.$$

*Then, according to Definition 6,  $E_\alpha$ ,  $E_\beta$  and  $E_\gamma$  are nonempty (the respective variables are typeset in bold face). Since the patterns satisfy the conditions of Theorem 27, there is no length-increasing morphism  $\sigma : \mathbb{N}^+ \rightarrow \Sigma^+$  that is weakly unambiguous with respect to them (provided that  $|\Sigma| = 2$ ).*

Theorem 27 and Example 28 directly imply the insight mentioned above that Theorem 13 does not hold for binary alphabets  $\Sigma$ :

**Corollary 29.** *Let  $\Sigma$  be an alphabet with  $|\Sigma| = 2$ . There is an  $\alpha \in \mathbb{N}^+$  such that  $E_\alpha$  is not empty and there is no length-increasing morphism  $\sigma : \mathbb{N}^+ \rightarrow \Sigma^+$  that is weakly unambiguous with respect to  $\alpha$ .*

In contrast to the previous theorems, the following result features a *sufficient* condition on the *existence* of weakly unambiguous length-increasing morphisms  $\sigma : \mathbb{N}^+ \rightarrow \Sigma^+$ ,  $|\Sigma| = 2$ , with respect to a given pattern. This phenomenon partly depends on the question of whether we can avoid short squares in the morpic image.

**Theorem 30.** *Let  $\alpha \in \mathbb{N}^+$ , and let  $\Sigma$  be an alphabet,  $|\Sigma| = 2$ . Suppose that*

- $i \cdot e \cdot e' \sqsubset \alpha$  and  $i \cdot e' \cdot e \sqsubset \alpha$ , or
- $e \cdot e' \cdot i \sqsubset \alpha$  and  $e' \cdot e \cdot i \sqsubset \alpha$ ,

with  $e, e' \in E_\alpha$  and  $i \in \text{var}(\alpha) \setminus \{e, e'\}$ . If a morphism  $\sigma : \mathbb{N}^+ \rightarrow \Sigma^+$  satisfies

- $|\sigma(e)| = 2$  and  $|\sigma(e')| = 2$ ,
- for every  $j \in \text{var}(\alpha) \setminus \{e, e'\}$ ,  $|\sigma(j)| = 1$ , and
- there is no  $x \in \Sigma$  with  $x^2 \sqsubseteq \sigma(\alpha)$ ,

then  $\sigma$  is weakly unambiguous with respect to  $\alpha$ .

*Proof.* Let  $\Sigma := \{a, b\}$ . We initially discuss the case where  $i \cdot e \cdot e' \sqsubseteq \alpha$  and  $i \cdot e' \cdot e \sqsubseteq \alpha$  are satisfied. We define a morphism  $\sigma : \mathbb{N}^+ \rightarrow \Sigma^+$  such that the conditions of Theorem 30 are satisfied. This implies that  $\sigma(\alpha) = (ab)^n \cdot v$ ,  $v \in \{a, \varepsilon\}$ , or  $\sigma(\alpha) = (ba)^n \cdot v$ ,  $v \in \{b, \varepsilon\}$ ; moreover,  $\sigma(e) = ab$  and  $\sigma(e') = ab$  or, alternatively,  $\sigma(e) = ba$  and  $\sigma(e') = ba$ . Consequently,  $\sigma(i \cdot e \cdot e') = b \cdot ab \cdot ab$ , or  $\sigma(i \cdot e \cdot e') = a \cdot ba \cdot ba$ .

Assume to the contrary that  $\sigma$  is not weakly unambiguous with respect to  $\alpha$ . Consequently, there is a nonerasing morphism  $\tau : \mathbb{N}^+ \rightarrow \Sigma^+$  with  $\tau(\alpha) = \sigma(\alpha)$  and, for some  $q \in \text{var}(\alpha)$ ,  $\tau(q) \neq \sigma(q)$ . Hence, if  $\sigma(e) = ab$  and  $\sigma(e') = ab$ , then one of the following cases is satisfied:

- $|\tau(e)| < |\sigma(e)|$ , which leads to the following sub-cases:
  - $\tau(e) = a$ . Since  $\tau(\alpha) = \sigma(\alpha)$  and  $i \cdot e \cdot e' \sqsubseteq \alpha$ , this implies that  $\tau(i) = \alpha_1 b$ ,  $\alpha_1 \in \Sigma^*$ , and  $\tau(e') = b\alpha_2$ ,  $\alpha_2 \in \Sigma^*$ . Due to  $i \cdot e' \cdot e \sqsubseteq \alpha$ ,  $\tau(i \cdot e' \cdot e) \sqsubseteq \tau(\alpha)$ . However,  $\tau(i \cdot e' \cdot e) = \alpha_1 b \cdot b\alpha_2 \cdot a$  and, this means that  $b^2 \sqsubseteq \tau(\alpha)$ , which contradicts  $\tau(\alpha) = \sigma(\alpha)$ .
  - $\tau(e) = b$ . An analogous reasoning to that in the previous case leads to  $a^2 \sqsubseteq \tau(\alpha)$ , which is a contradiction.
- $|\tau(e')| < |\sigma(e')|$ . The reasoning is analogous to that in the previous case.
- $|\tau(e)| \geq 3$  and  $|\tau(e')| \geq 3$ . Since  $\tau$  is nonerasing,  $|\tau(\alpha)| > |\sigma(\alpha)|$ . This contradicts  $\tau(\alpha) = \sigma(\alpha)$ .
- $|\tau(e)| \geq 4$  or  $|\tau(e')| \geq 4$ . Since  $\tau$  is nonerasing,  $|\tau(\alpha)| > |\sigma(\alpha)|$ . This again contradicts  $\tau(\alpha) = \sigma(\alpha)$ .
- $|\tau(e)| = 3$ . If  $\tau(e) = aba$ , then the conditions  $\tau(\alpha) = \sigma(\alpha)$  and  $i \cdot e \cdot e' \sqsubseteq \alpha$  imply that  $\tau(i) = \alpha_1 b$ ,  $\alpha_1 \in \Sigma^*$ , and  $\tau(e') = b\alpha_2$ ,  $\alpha_2 \in \Sigma^*$ . Due to  $i \cdot e' \cdot e \sqsubseteq \alpha$ ,  $\tau(i \cdot e' \cdot e) \sqsubseteq \tau(\alpha)$ . However,  $\tau(i \cdot e' \cdot e) = \alpha_1 b \cdot b\alpha_2 \cdot aba$ , and this means that  $b^2 \sqsubseteq \tau(\alpha)$ , which contradicts  $\tau(\alpha) = \sigma(\alpha)$ . If  $\tau(e) = bab$ , then the conditions  $\tau(\alpha) = \sigma(\alpha)$  and  $i \cdot e \cdot e' \sqsubseteq \alpha$  imply that  $\tau(i) = \alpha_1 a$ ,  $\alpha_1 \in \Sigma^*$ , and  $\tau(e') = a\alpha_2$ ,  $\alpha_2 \in \Sigma^*$ . Due to  $i \cdot e' \cdot e \sqsubseteq \alpha$ ,  $\tau(i \cdot e' \cdot e) \sqsubseteq \tau(\alpha)$ . However,  $\tau(i \cdot e' \cdot e) = \alpha_1 a \cdot a\alpha_2 \cdot bab$ , and this means that  $a^2 \sqsubseteq \tau(\alpha)$ , which again contradicts  $\tau(\alpha) = \sigma(\alpha)$ .
- $|\tau(e')| = 3$ . The reasoning is analogous to that in the previous case.

- $\tau(e) = \tau(e') = ba$ . Consequently, since  $\tau(\alpha) = \sigma(\alpha)$ , for every  $j \in \text{var}(\alpha) \setminus \{e, e'\}$ ,  $|\tau(j)| = 1$ . As a result  $|\tau(i)| = 1$  and due to  $x^2 \not\sqsubseteq \sigma(\alpha)$ ,  $x \in \Sigma$ ,  $\tau(i) = a$ . So,  $\tau(i \cdot e \cdot e') = \tau(i \cdot e' \cdot e) = ababa$ , while  $\sigma(i \cdot e \cdot e') = \sigma(i \cdot e' \cdot e) = babab$ . This implies that there exists at least one variable  $k \in \text{var}(\alpha) \setminus \{e, e'\}$  with  $\tau(k) = \varepsilon$ , since otherwise  $\tau(\alpha) \neq \sigma(\alpha)$ . This contradicts the fact that  $\tau$  is nonerasing.

The extension of this reasoning to the case where  $\sigma(e) = ba$  and  $\sigma(e') = ba$  are satisfied is straightforward. Hence, there is no morphism  $\tau$  with  $\tau(\alpha) = \sigma(\alpha)$  and, for some  $q \in \text{var}(\alpha)$ ,  $\tau(q) \neq \sigma(q)$ . Consequently,  $\sigma$  is weakly unambiguous with respect to  $\alpha$ . Using the same reasoning as above, it can be demonstrated that Theorem 30 holds true for the case that  $e \cdot e' \cdot i \sqsubset \alpha$  and  $e' \cdot e \cdot i \sqsubset \alpha$ .  $\square$

The main difference between Theorem 30 and Theorem 27 is that those patterns  $\alpha$  being examined in Theorem 30 do not satisfy Condition 3 of Theorem 27. Thus, the two theorems demonstrate what subtleties in the structure of a pattern can determine whether or not it has a weakly unambiguous morphism with a binary target alphabet.

In order to illustrate Theorem 30, we now consider some examples. In contrast to Example 28, the factors  $3 \cdot 4$  and  $4 \cdot 3$  of the patterns in the following example have an identical right neighbour or an identical left neighbour.

**Example 31.** We define a morphism  $\sigma : \mathbb{N}^+ \rightarrow \{a, b\}^+$  for the given patterns  $\alpha$  (where the factors featured by Theorem 30 are typeset in bold face) as follows:

- $\alpha = 1 \cdot 2 \cdot \mathbf{5} \cdot \mathbf{3} \cdot 4 \cdot 6 \cdot 7 \cdot 8 \cdot \mathbf{5} \cdot 4 \cdot \mathbf{3} \cdot 9 \cdot 10$ .  
 $\sigma$  is defined by  $\sigma(1) := a$ ,  $\sigma(2) := b$ ,  $\sigma(5) := a$ ,  $\sigma(3) := ba$ ,  $\sigma(4) := ba$ ,  $\sigma(6) := b$ ,  $\sigma(7) := a$ ,  $\sigma(8) := b$ ,  $\sigma(9) := b$  and  $\sigma(10) := a$ .
- $\alpha = 1 \cdot 2 \cdot \mathbf{3} \cdot 4 \cdot \mathbf{5} \cdot 6 \cdot 7 \cdot 4 \cdot \mathbf{3} \cdot \mathbf{5} \cdot 8 \cdot 9$ .  
 $\sigma$  is defined by  $\sigma(1) := a$ ,  $\sigma(2) := b$ ,  $\sigma(3) := ab$ ,  $\sigma(4) := ab$ ,  $\sigma(5) := b$ ,  $\sigma(6) := a$ ,  $\sigma(7) := b$ ,  $\sigma(8) := b$  and  $\sigma(9) := a$ .
- $\alpha = 1 \cdot 2 \cdot \mathbf{3} \cdot 4 \cdot 5 \cdot 6 \cdot 7 \cdot \mathbf{8} \cdot \mathbf{3} \cdot 4 \cdot 9 \cdot 10 \cdot 11 \cdot \mathbf{8} \cdot 4 \cdot \mathbf{3} \cdot 12 \cdot 13$ .  
 $\sigma$  is defined by  $\sigma(1) := b$ ,  $\sigma(2) := a$ ,  $\sigma(3) := ba$ ,  $\sigma(4) := ba$ ,  $\sigma(5) := b$ ,  $\sigma(6) := a$ ,  $\sigma(7) := b$ ,  $\sigma(8) := a$ ,  $\sigma(9) := b$ ,  $\sigma(10) := a$ ,  $\sigma(11) := b$ ,  $\sigma(12) := b$  and  $\sigma(13) := a$ .

With reference to Theorem 30, it can be easily verified that, in all above cases,  $\sigma$  is length-increasing and weakly unambiguous with respect to  $\alpha$ .

The patterns in Example 31 further illustrate that the converse of Theorem 20 does not hold true. More precisely, although for every pattern  $\alpha$  in this example, for every  $e \in E_\alpha$  there exists an  $e' \in E_\alpha$ ,  $e' \neq e$ , such that  $e \cdot e'$  and  $e' \cdot e$  are factors of  $\alpha$ , there is a weakly unambiguous length-increasing morphism  $\sigma : \mathbb{N}^+ \rightarrow \{a, b\}^+$  with respect to  $\alpha$ .

Due to Theorems 27 and 30, we expect that it is an extremely challenging task to find an equivalent to the characterisation in Theorem 13 for the binary case. From our understanding of the matter, we can therefore merely give the following conjecture on the decidability of Problem 2 for binary target alphabets.



**Conjecture 32.** *Let  $\alpha \in \mathbb{N}^+$  with  $|\alpha| \geq 2$ , and let  $\Sigma$  be an alphabet,  $|\Sigma| = 2$ . The problem of whether there is a weakly unambiguous length-increasing morphism  $\sigma : \mathbb{N}^+ \rightarrow \Sigma^+$  with respect to  $\alpha$  is decidable by testing a finite number of morphisms.*

The above conjecture is based on the fact that according to the Corollary 19, any weakly unambiguous length-increasing morphism with respect to a pattern  $\alpha$  must not be length-increasing for the variables in  $S_\alpha$ . On the other hand, increasing the length of the morphic images of the variables in  $E_\alpha$  under a morphism  $\sigma : \mathbb{N}^+ \rightarrow \Sigma^+$ ,  $|\Sigma| = 2$ , seems to increase the chance of the existence of a morphism  $\tau : \mathbb{N}^+ \rightarrow \Sigma^+$  satisfying  $\tau(\alpha) = \sigma(\alpha)$  and, for some  $i \in \text{var}(\alpha)$ ,  $\tau(i) \neq \sigma(i)$ . Consequently, we believe that if all morphisms  $\sigma$  with, for every  $e \in E_\alpha$  and an  $x \in \mathbb{N}$ ,  $|\sigma(e)| \leq x$  are not weakly unambiguous with respect to  $\alpha$ , then there does not exist a weakly unambiguous morphism  $\sigma$  with  $|\sigma(e)| > x$  for some  $e \in E_\alpha$ , either. For all patterns, we expect a value of  $x = 2$  to be a sufficiently large bound for the morphisms to be tested.

## 6. Weakly unambiguous morphisms with $|\Sigma| = 1$

It is not surprising that most of our considerations in the previous sections are not applicable to morphisms with a unary target alphabet. On the other hand, Corollaries 18 and 19 also hold for this special case, i. e., for any pattern  $\alpha$ , every weakly unambiguous morphism must map the variables in  $S_\alpha$  to words of length 1, which implies that such a morphism can only be length-increasing if  $E_\alpha$  is not empty. Incorporating these observations, we now consider an example.

**Example 33.** *Let  $\alpha_1 := 1 \cdot 2 \cdot 3 \cdot 4 \cdot 1 \cdot 2 \cdot 3$ . Consequently,  $E_{\alpha_1} = \{4\}$ . We define a morphism  $\sigma : \mathbb{N}^+ \rightarrow \{a\}^+$  by  $\sigma(4) := aa$  and  $\sigma(i) := a$ ,  $i \in \mathbb{N} \setminus \{4\}$ . It can be easily verified that  $\sigma$  is weakly unambiguous with respect to  $\alpha_1$ . Now let  $\alpha_2 := 1 \cdot 2 \cdot 3 \cdot 4 \cdot 1 \cdot 2 \cdot 3 \cdot 5 \cdot 6$ . As a result,  $E_{\alpha_2} = \{4\}$ . If we now consider the morphism  $\tau$ , given by  $\tau(4) := a$ ,  $\tau(5) := aa$  and  $\tau(i) := \sigma(i)$ ,  $i \in \mathbb{N} \setminus \{4, 5\}$ , then we may conclude  $\tau(\alpha_2) = \sigma(\alpha_2)$ . Thus,  $\sigma$  is not weakly unambiguous with respect to  $\alpha_2$ .*

Quite obviously, the fact that  $\sigma$  is unambiguous with respect to  $\alpha_1$  and ambiguous with respect to  $\alpha_2$  is due to 4 being the only variable in  $\alpha_1$  that has a single occurrence, whereas  $\alpha_2$  also has single occurrences of the variables 5 and 6. This aspect is reflected by the following characterisation that completely solves Problem 2 for morphisms with unary target alphabets.

**Theorem 34.** *Let  $\alpha \in \mathbb{N}^+$ ,  $\text{var}(\alpha) = \{1, 2, 3, \dots, n\}$ . There is no weakly unambiguous length-increasing morphism  $\sigma : \mathbb{N}^+ \rightarrow \{a\}^+$  with respect to  $\alpha$  if and only if, for every  $i \in \text{var}(\alpha)$ , there exist  $n_1, n_2, \dots, n_{i-1}, n_{i+1}, \dots, n_n \in \mathbb{N} \cup \{0\}$ , such that*

$$|\alpha|_i = \sum_{j \in \{1, 2, \dots, n\} \setminus \{i\}} n_j |\alpha|_j. \quad (9)$$

*Proof.* We begin with the *if* direction. Assume that, for every  $i \in \text{var}(\alpha)$ , Equation (9) is satisfied. Also, assume that  $\sigma : \mathbb{N}^+ \rightarrow \{a\}^+$  is an arbitrary length-increasing morphism with  $|\sigma(i')| > 1$ ,  $i' \in \text{var}(\alpha)$ . This means that  $\sigma(i') = a^m$ ,  $m \geq 2$  and, hence,

$$|\sigma(\alpha)| = |\sigma(1)||\alpha|_1 + |\sigma(2)||\alpha|_2 + [\dots] + m|\alpha|_{i'} + [\dots] + |\sigma(n)||\alpha|_n.$$

Due to  $|\Sigma| = 1$ , we can prove that  $\sigma$  is not weakly unambiguous with respect to  $\alpha$  by defining a morphism  $\tau : \mathbb{N}^+ \rightarrow \{a\}^+$  with  $|\tau(\alpha)| = |\sigma(\alpha)|$  and, for some  $q \in \text{var}(\alpha)$ ,  $|\tau(q)| \neq |\sigma(q)|$ . We define the morphism  $\tau$  such that  $\tau(i') := a^{(m-1)}$ , and as a result,

$$|\tau(\alpha)| = |\tau(1)||\alpha|_1 + |\tau(2)||\alpha|_2 + [\dots] + (m-1)|\alpha|_{i'} + [\dots] + |\tau(n)||\alpha|_n.$$

We need to demonstrate that

$$|\tau(\alpha)| - |\sigma(\alpha)| = 0.$$

This is equivalent to:

$$\begin{aligned} |\alpha|_{i'} &= |\alpha|_1(|\tau(1)| - |\sigma(1)|) + |\alpha|_2(|\tau(2)| - |\sigma(2)|) + [\dots] + \\ &|\alpha|_{i'-1}(|\tau(i'-1)| - |\sigma(i'-1)|) + |\alpha|_{i'+1}(|\tau(i'+1)| - |\sigma(i'+1)|) \\ &+ [\dots] + |\alpha|_n(|\tau(n)| - |\sigma(n)|). \end{aligned} \quad (10)$$

According to Equation (9), for Equation (10) to be satisfied, we define the morphism  $\tau$ , for every  $j \in \text{var}(\alpha) \setminus \{i'\}$  such that  $|\tau(j)| - |\sigma(j)| = n_j$ , and this can be achieved by defining  $\tau(j) := a^{(n_j + |\sigma(j)|)}$ . Consequently,  $\tau$  is given by

$$\tau(i) := \begin{cases} a^{|\sigma(i)|-1}, & i = i', \\ a^{(n_i + |\sigma(i)|)}, & i \in \text{var}(\alpha) \setminus \{i'\}, \end{cases}$$

which implies that  $\tau$  is nonerasing,  $\tau(i') \neq \sigma(i')$ , and  $|\tau(\alpha)| = |\sigma(\alpha)|$ . This means that  $\sigma$  is not weakly unambiguous with respect to  $\alpha$ .

We now prove the *only if* direction. So, we assume that there is no weakly unambiguous length-increasing morphism  $\sigma : \mathbb{N}^+ \rightarrow \{a\}^+$  with respect to  $\alpha$ . Let  $i$  be an arbitrary variable of  $\alpha$ . We define the morphism  $\sigma$  for the variables  $x \in \text{var}(\alpha)$  by

$$\sigma(x) := \begin{cases} aa, & x = i, \\ a, & x \neq i. \end{cases}$$

The assumption of the *only if* direction implies that there exists a morphism  $\tau$  satisfying  $\tau(\alpha) = \sigma(\alpha)$  and, for some variables  $q \in \text{var}(\alpha)$ ,  $\tau(q) \neq \sigma(q)$ . According to Lemma 11,  $\tau(i) \sqsubset \sigma(i)$  must be satisfied. Thus,  $\tau(i) = a$ . Consequently,

$$|\sigma(\alpha)| = |\sigma(1)||\alpha|_1 + |\sigma(2)||\alpha|_2 + [\dots] + 2|\alpha|_i + [\dots] + |\sigma(n)||\alpha|_n$$

and

$$|\tau(\alpha)| = |\tau(1)||\alpha|_1 + |\tau(2)||\alpha|_2 + [\dots] + |\alpha|_i + [\dots] + |\tau(n)||\alpha|_n.$$

It follows from  $|\tau(\alpha)| = |\sigma(\alpha)|$ , that  $|\tau(\alpha)| - |\sigma(\alpha)| = 0$ . Thus,

$$|\alpha|_1(|\tau(1)| - |\sigma(1)|) + |\alpha|_2(|\tau(2)| - |\sigma(2)|) + [\dots] + (-|\alpha_i|) + [\dots] + |\alpha|_n(|\tau(n)| - |\sigma(n)|) = 0.$$

This leads to

$$\begin{aligned} |\alpha|_i &= |\alpha|_1(|\tau(1)| - |\sigma(1)|) + |\alpha|_2(|\tau(2)| - |\sigma(2)|) \\ &+ [\dots] + |\alpha|_{i-1}(|\tau(i-1)| - |\sigma(i-1)|) + |\alpha|_{i+1}(|\tau(i+1)| - |\sigma(i+1)|) \\ &+ [\dots] + |\alpha|_n(|\tau(n)| - |\sigma(n)|). \end{aligned} \tag{11}$$

Consequently, for any variable  $i \in \text{var}(\alpha)$ , there exists  $n_1, n_2, \dots, n_n \in \mathbb{N} \cup \{0\}$ , such that Equation (9) is satisfied.  $\square$

Hence, we are able to provide a result on unary alphabets that is as strong as our result in Theorem 13 on ternary and larger alphabets. However, while Theorem 13 needs to consider the order of variables in the patterns, it is evident that Theorem 34 can exclusively refer to their numbers of occurrences.

## 7. Conclusion

In this paper, we have demonstrated that there is a weakly unambiguous length-increasing morphism  $\sigma : \mathbb{N}^+ \rightarrow \Sigma^+$ ,  $|\Sigma| \geq 3$ , with respect to  $\alpha \in \mathbb{N}^+$  if and only if  $E_\alpha$  is not empty, where  $E_\alpha \subseteq \text{var}(\alpha)$  consists of those variables that have special, namely *illoyal* neighbour variables. We have demonstrated that this condition is not characteristic, but only necessary for the case  $|\Sigma| = 2$ , which leads to an interesting difference between binary and all other target alphabets  $\Sigma$ . We have not been able to characterise the existence of weakly unambiguous length-increasing morphisms with binary target alphabets, but we have found strong conditions that are either sufficient or necessary. Finally, for  $|\Sigma| = 1$ , we have been able to demonstrate that the existence of weakly unambiguous length-increasing morphisms  $\sigma : \mathbb{N}^+ \rightarrow \Sigma^+$  solely depends on particular equations that the numbers of occurrences of the variables in the corresponding pattern need to satisfy.

**Acknowledgements** The authors are indebted to Johannes C. Schneider and the anonymous referees of both the conference and the full version of this paper, whose careful remarks and suggestions helped to improve this paper substantially.

## References

- [1] J. Berstel and D. Perrin. *Theory of Codes*. Academic Press, Orlando, 1985.
- [2] D.D. Freydenberger, H. Nevisi, and D. Reidenbach. Weakly unambiguous morphisms. In *Proc. 28th Annual Symposium on Theoretical Aspects of Computer Science, STACS 2011*, volume 9 of *LIPICs*, pages 213–224, 2011.
- [3] D.D. Freydenberger and D. Reidenbach. The unambiguity of segmented morphisms. *Discrete Applied Mathematics*, 157:3055–3068, 2009.

- [4] D.D. Freydenberger, D. Reidenbach, and J.C. Schneider. Unambiguous morphic images of strings. *International Journal of Foundations of Computer Science*, 17:601–628, 2006.
- [5] D. Hamm and J. Shallit. Characterization of finite and one-sided infinite fixed points of morphisms on free monoids. Technical Report CS-99-17, Dep. of Computer Science, University of Waterloo, 1999. <http://www.cs.uwaterloo.ca/~shallit/papers.html>.
- [6] T. Harju and J. Karhumäki. Morphisms. In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Languages*, volume 1, chapter 7, pages 439–510. Springer, 1997.
- [7] S. Holub. Polynomial-time algorithm for fixed points of nontrivial morphisms. *Discrete Mathematics*, 309:5069–5076, 2009.
- [8] A. Mateescu and A. Salomaa. Patterns. In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Languages*, volume 1, chapter 4.6, pages 230–242. Springer, 1997.
- [9] D. Reidenbach. A non-learnable class of E-pattern languages. *Theoretical Computer Science*, 350:91–102, 2006.
- [10] D. Reidenbach. An examination of Ohlebusch and Ukkonen’s conjecture on the equivalence problem for E-pattern languages. *Journal of Automata, Languages and Combinatorics*, 12:407–426, 2007.
- [11] D. Reidenbach. Discontinuities in pattern inference. *Theoretical Computer Science*, 397:166–193, 2008.
- [12] D. Reidenbach and J.C. Schneider. Morphically primitive words. *Theoretical Computer Science*, 410:2148–2161, 2009.
- [13] D. Reidenbach and J.C. Schneider. Restricted ambiguity of erasing morphisms. *Theoretical Computer Science*, 412:3510–3523, 2011.
- [14] J.C. Schneider. Unambiguous erasing morphisms in free monoids. *RAIRO – Theoretical Informatics and Applications*, 44:193–208, 2010.