Weakly Unambiguous Morphisms

Dominik D. Freydenberger[1], Hossein Nevisi[2], and Daniel Reidenbach[3]

1 Institut für Informatik, Goethe-Universität
Frankfurt am Main, Germany
freydenberger@em.uni-frankfurt.de

2 Department of Computer Science, Loughborough University
Loughborough, Leicestershire LE11 3TU, United Kingdom
H.Nevisi@lboro.ac.uk
Corresponding author

3 Department of Computer Science, Loughborough University
Loughborough, Leicestershire LE11 3TU, United Kingdom
D.Reidenbach@lboro.ac.uk

Abstract

A nonerasing morphism σ is said to be weakly unambiguous with respect to a word w if σ is the only nonerasing morphism that can map w to $\sigma(w)$, i.e., there does not exist any other nonerasing morphism τ satisfying $\tau(w) = \sigma(w)$. 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.

1998 ACM Subject Classification G.2.m [Discrete Mathematics]: Miscellaneous

Keywords and phrases Combinatorics on words, Morphisms, Ambiguity

Digital Object Identifier 10.4230/LIPIcs.STACS.2011.213

1 Introduction

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

$$\overbrace{a\ b\ b}^{\sigma(A)}\ \overbrace{a\ b\ b\ b}^{\sigma(B)}\ \overbrace{a\ b\ b\ b}^{\sigma(B)}\ \overbrace{a\ b\ b\ b\ b}^{\sigma(C)}\ \overbrace{a\ b\ b}^{\sigma(A)}\ \overbrace{a\ b\ b\ b\ b}^{\sigma(C)} = \underbrace{a\ b\ b\ a\ b}_{\tau(A)}\ \underbrace{b\ b\ a\ b}_{\tau(B)}\ \underbrace{b\ b\ a\ b}_{\tau(B)}\ \underbrace{b\ b\ b}_{\tau(C)}\ \underbrace{a\ b\ b\ a\ b}_{\tau(A)}\ \underbrace{b\ b\ b}_{\tau(C)}\,.$$

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

28th Symposium on Theoretical Aspects of Computer Science (STACS'11).
Editors: Thomas Schwentick, Christoph Dürr; pp. 213–224
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany
SYMPOSIUM ON THEORETICAL ASPECTS OF COMPUTER SCIENCE

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 [5]) and *pattern languages* (see Mateescu and Salomaa [7]). Regarding the latter topic, insights into the ambiguity of morphisms have been used to solve a number of prominent problems (see, e. g., Reidenbach [8, 9, 10]), 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 [3] 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, [3] introduces two types of unambiguity: The first type follows our intuitive definition given above; more precisely, a morphism σ is called *strongly unambiguous* with respect to a word s if it there exists no morphism τ 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 σ *weakly unambiguous* with respect to s if there is no *nonerasing* morphism τ (which means that τ must not map any symbol to the empty word) showing the above properties. Thus, e. g., our initial example morphism σ is weakly unambiguous with respect to $s' := A\,A\,B$, but it is not strongly unambiguous, since the morphism τ, given by $\tau(A) := \varepsilon$ and $\tau(B) := \sigma(s')$ (where ε 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. [3] 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 [4]) or being a shortest generator of a terminal-free E-pattern language. Freydenberger and Reidenbach [2], among other results, improve and deepen the techniques used in [3]. Schneider [12] studies the more general problem of the existence of arbitrary (i. e., possibly erasing) strongly unambiguous morphisms. While [12] 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 [11] 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 nonerasing 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. [3], 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 σ 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 [12] 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.

Note that, due to space constraints, almost all proofs and some related definitions and lemmas are omitted from this paper.

2 Definitions

Let $\mathbb{N} := \{1, 2, 3, ...\}$ and Σ be alphabets. We call any symbol in $\mathbb{N}$ a variable and any symbol in Σ a letter. 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 notion $|x|$ stands for the size of a set x or the length of a word x. We denote the empty word by ε, i. e., $|\varepsilon| = 0$. The symbol [...] is used to omit some canonically defined parts of a given word, e. g., $\alpha = 1 \cdot 2 \cdot [...] \cdot 5$ stands for $\alpha = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5$. In order to distinguish between a word over $\mathbb{N}$ and a word over Σ, we call the former a *pattern*. We name patterns with lower case letters from the beginning of the Greek alphabet such as α, β, γ. For every alphabet A, A^* is the set of all (empty and non-empty) words over A, and $A^+ := A^* \setminus \{\varepsilon\}$. We call a word $v \in A^*$ a *factor* of a word $w \in A^*$ if, for some $u_1, u_2 \in A^*$, $w = u_1 v u_2$; 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.

With regard to an arbitrary pattern α, $\mathrm{var}(\alpha)$ denotes the set of all variables occurring in α, and $|\alpha|_\beta$, $\beta \sqsubseteq \alpha$, shows the number of (possibly overlapping) occurrences of β in α.

A *morphism* is a mapping that is compatible with concatenation, i. e., for patterns $\alpha \in \mathbb{N}^+$ and $\beta \in \mathbb{N}^+$, a morphism $\sigma : \mathbb{N}^+ \to \Sigma^*$ satisfies $\sigma(\alpha \cdot \beta) = \sigma(\alpha) \cdot \sigma(\beta)$. A morphism σ is called *nonerasing* provided that, for every $i \in \mathbb{N}$, $\sigma(i) \neq \varepsilon$. The morphism σ is *length-increasing* if $|\sigma(\alpha)| > |\alpha|$, and it is called *1-uniform* if, for every $i \in \mathbb{N}$, $|\sigma(i)| = 1$.

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

We call any pattern $\alpha \in \mathbb{N}^+$ *prolix* if and only if, there exists a decomposition $\alpha = \beta_0\gamma_1\beta_1\gamma_2\beta_2[...]\beta_{n-1}\gamma_n\beta_n$ 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', $0 \leq k' \leq n$, $\mathrm{var}(\gamma_k) \cap \mathrm{var}(\beta_{k'}) = \emptyset$,
3. for every k, $1 \leq k \leq n$, there exists an $i_k \in \mathrm{var}(\gamma_k)$ such that $|\gamma_k|_{i_k} = 1$ and, for every k', $1 \leq k' \leq n$, if $i_k \in \mathrm{var}(\gamma_{k'})$ 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.

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 $j \in \mathrm{var}(\alpha)$, we define the following sets:

$$L_j := \{k \in \mathrm{var}(\alpha) \mid \alpha = ... \cdot k \cdot j \cdot ...\},$$
$$R_j := \{k \in \mathrm{var}(\alpha) \mid \alpha = ... \cdot j \cdot k \cdot ...\}.$$

Also, if $\alpha = j...$, then $\varepsilon \in L_j$, and if $\alpha = ...j$, 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 $i \in \mathrm{var}(\alpha)$ has *loyal neighbours (in α)* 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 $\alpha \in \mathbb{N}^+$, $|\alpha| > 1$, let S_α be the set of variables that have loyal neighbours and E_α be the set of variables that do not have loyal neighbours in α.

The following example clarifies the mentioned definitions.

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

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

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. [3] 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. [3])**.** *Let $\alpha \in \mathbb{N}^*$, let Σ be an alphabet, $|\Sigma| \geq 2$. There exists a strongly unambiguous nonerasing morphism $\sigma : \mathbb{N}^* \to \Sigma^*$ with respect to α if and only if α 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 α is prolix. In general, the converse of this statement does not hold true.*

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 lemmas 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 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.*

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 $\alpha \in \mathbb{N}^+$, $|\alpha| > 1$. Assume that $\sigma : \mathbb{N}^+ \to \Sigma^+$ is a morphism such that, for an $i \in \mathrm{var}(\alpha)$, $|\sigma(i)| > 1$ and, for every $x \in \mathrm{var}(\alpha) \setminus \{i\}$, $|\sigma(x)| = 1$. Moreover, assume that τ is a nonerasing morphism satisfying $\tau(\alpha) = \sigma(\alpha)$. If there exists a $j \in \mathrm{var}(\alpha)$ with $\tau(j) \neq \sigma(j)$, then $\tau(i) \sqsubset \sigma(i)$.*

The next lemma, which directly results from Definition 5, discusses those patterns having at least one square; more precisely, there exists an $i \in \mathbb{N}$ with $i^2 \sqsubset \alpha$.

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

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).

▶ **Theorem 13.** *Let $\alpha \in \mathbb{N}^+$ with $|\alpha| > 1$ and let $|\Sigma| \geq 3$. There is a weakly unambiguous length-increasing morphism $\sigma : \mathbb{N}^+ \to \Sigma^+$ with respect to α if and only if E_α is not empty.*

Proof. Let $\{a, b, c\} \subseteq \Sigma$.
We begin with the *if* direction. Assume that E_α 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 σ by $\sigma(x) := bc$ if $x = i$ and $\sigma(x) := a$ if $x \neq i$. So, $\sigma(i^2) = bcbc$. According to Lemma 11, any nonerasing morphism $\tau : \mathbb{N}^+ \to \Sigma^+$ with $\tau(\alpha) = \sigma(\alpha)$ and, for some $k \in \text{var}(\alpha)$, $\tau(k) \neq \sigma(k)$, must satisfy $\tau(i) \neq \sigma(i)$, and this means that $\tau(i)$ should 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 $\tau(\alpha) \neq \sigma(\alpha)$ and consequently, σ is weakly unambiguous with respect to α.

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 $j \in L_i$ such that $R_j \neq \{i\}$, and if $\varepsilon \notin R_i$, then there exits a $j' \in R_i$ such that $L_{j'} \neq \{i\}$.
Case 2.2: $\varepsilon \in L_i$ and $\varepsilon \in R_i$.

Let $\sigma : \mathbb{N}^+ \to \{a, b, c\}^+$ be the morphism defined in Case 1. Due to Lemma 11, any morphism $\tau : \mathbb{N}^+ \to \Sigma^+$ with $\tau(\alpha) = \sigma(\alpha)$ and, for some $k \in \text{var}(\alpha)$, $\tau(k) \neq \sigma(k)$, must satisfy $\tau(i) \neq \sigma(i)$, and this means that $\tau(i)$ should be a proper factor of $\sigma(i)$. Thus, $\tau(i) = b$ or $\tau(i) = c$.
With regard to Case 2.1, consider $\tau(i) = c$, and $\varepsilon \notin L_i$. Due to the number of c in $\sigma(\alpha)$, which equals the number of occurrences of i in α, and also due to $\sigma(i) = bc$, the positions of c of $\tau(i)$ should be at the same positions as c of $\sigma(i)$ in $\sigma(\alpha)$. So, to have $\tau(\alpha) = \sigma(\alpha)$, for every $l \in L_i$, b is a suffix of $\tau(l)$, and as a result 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, $\tau(\alpha) \neq \sigma(\alpha)$. Consider $\tau(i) = b$, and $\varepsilon \notin R_i$. Due to the number of b in $\sigma(\alpha)$, which equals the number of occurrences of i in α, and also due to $\sigma(i) = bc$, the positions of b of $\tau(i)$ should be at the same positions as b of $\sigma(i)$ in $\sigma(\alpha)$. Hence, to have $\tau(\alpha) = \sigma(\alpha)$, for every $r \in R_i$, c is a prefix of $\tau(r)$, and consequently, c is prefix of $\tau(j')$. However, since $L_{j'} \neq \{i\}$, 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 $\alpha = i\alpha' i$, $\alpha' \in \mathbb{N}^*$. So, $\sigma(\alpha) = bc\sigma(\alpha')bc$. As mentioned above, due to Lemma 11, $\tau(i) = b$ or $\tau(i) = c$. This implies that $\tau(\alpha)$ starts with b and finishes with b, or it starts with c and finishes 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 α.

We now prove the *only if* direction. In fact, we want to show that if there is a weakly unambiguous length-increasing morphism $\sigma : \mathbb{N}^+ \to \Sigma^+$ with respect to α, then E_α is not empty. Assume that σ maps one of the variables of α to a word of length more than 1, and let this variable be i. Also, let $\sigma(i) := a_1 a_2 [...] a_n$ with $n \geq 2$ and, for every q, $1 \leq q \leq n$, $a_q \in \Sigma$. Assume to the contrary that E_α 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

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

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, $i, l_k \notin \mathrm{var}(\alpha_{k'})$. Thus,

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

We now define the nonerasing morphism τ such that, for every k, $1 \leq k \leq m$, $\tau(l_k) := \sigma(l_k) a_1$, $\tau(i) := a_2 a_3 [...] a_n$ and, for all other variables in α, τ is identical to σ. 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 τ is nonerasing, σ is not weakly unambiguous.

Case 2: $\varepsilon \notin R_i$ and, for every $j \in R_i$, $L_j = \{i\}$.
We can directly conclude that $\alpha := \alpha_1 \cdot i \cdot r_1 \cdot \alpha_2 \cdot i \cdot r_2 \cdot [...] \cdot \alpha_m \cdot i \cdot r_m \cdot \alpha_{m+1}$, 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 $i, r_k \notin \mathrm{var}(\alpha_{k'})$. So,

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

We now define the nonerasing morphism τ such that, for every k, $1 \leq k \leq m$, $\tau(r_k) := a_n \sigma(r_k)$ and $\tau(i) := a_1 a_2 [...] a_{n-1}$ and, for all other variables in α, τ is identical to σ. As, for every k, $1 \leq k \leq m$, $L_{r_k} = \{i\}$, we can conclude that $\tau(\alpha) = \sigma(\alpha)$; since τ is nonerasing, σ is not weakly unambiguous. Hence, $E_\alpha = \emptyset$ implies that σ is not weakly unambiguous, which contradicts the assumption. Consequently, E_α is not empty. ◀

In order to illustrate Theorem 13, 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 σ by $\sigma(4) := bc$ and, for every other variable $j \in \mathrm{var}(\alpha)$, $\sigma(j) := a$. Due to Lemma 11, any morphism τ with $\tau(\alpha) = \sigma(\alpha)$ and, for a $k \in \mathrm{var}(\alpha)$, $\tau(k) \neq \sigma(k)$ 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, σ is weakly unambiguous with respect to α.

▶ **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, or in other words, $E_\alpha = \emptyset$. Hence, it follows from Theorem 13 that there is no weakly unambiguous length-increasing morphism $\sigma : \mathbb{N}^+ \to \Sigma^+$, $|\Sigma| \geq 3$, with respect to α.

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

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

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

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

▶ **Theorem 17.** *Let* $\alpha \in \mathbb{N}^+$ *with* $|\alpha| > 1$, *and let* $|\Sigma| \geq 3$. *The problem of whether there is a weakly unambiguous length-increasing morphism* $\sigma : \mathbb{N}^+ \to \Sigma^+$ *with respect to* α *is decidable in polynomial time.*

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. [3] and the complexity consideration by Holub [6]). In contrast to this, when we ask for the existence of strongly unambiguous erasing morphisms, the problem is NP-hard (according to Schneider [12]).

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

As we shall demonstrate below, our characterisation in Theorem 13 does not hold for binary target alphabets Σ (see Corollary 24). Hence, we have to study this case separately. The most significant result of our considerations is a necessary condition on the structure of those patterns α that satisfy $E_\alpha \neq \emptyset$, but nevertheless do not have a weakly unambiguous morphism $\sigma : \mathbb{N}^+ \to \Sigma^+$, $|\Sigma| = 2$.

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_α being empty for any given pattern α is a sufficient condition for α not having any weakly unambiguous length-increasing morphism:

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

Hence, if we wish to characterise those patterns with respect to which there is a weakly unambiguous morphism $\sigma : \mathbb{N}^+ \to \Sigma^+$, $|\Sigma| \leq 2$, then we can safely restrict our considerations to those patterns α where E_α 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 α must have a particular, and very simple, shape for all variables in S_α:

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

Thus, any weakly unambiguous length-increasing morphism with respect to a pattern α must not be length-increasing for the variables in S_α. 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_α is empty, then there is no weakly unambiguous length-increasing morphism $\sigma : \mathbb{N}^+ \to \Sigma^+$ with respect to α. In the next step, we give a strong *necessary* condition on the structure of those patterns α that satisfy $E_\alpha \neq \emptyset$, but nevertheless do *not* have a weakly unambiguous morphism $\sigma : \mathbb{N}^+ \to \Sigma^+$, $|\Sigma| = 2$.

▶ **Theorem 20.** *Let $\alpha \in \mathbb{N}^+$ such that E_α is nonempty. Let Σ be an alphabet, $|\Sigma| = 2$. If there is no weakly unambiguous length-increasing morphism $\sigma : \mathbb{N}^+ \to \Sigma^+$ with respect to α, 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 α.*

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 α, then there is a weakly unambiguous length-increasing morphism $\sigma : \mathbb{N}^+ \to \Sigma^+$, $\Sigma = \{a, b\}$, with respect to α. 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 26.

As can be concluded from Example 7 and Theorem 13, there is a weakly unambiguous length-increasing morphism $\sigma : \mathbb{N}^+ \to \Sigma^+$, $|\Sigma| \geq 3$, with respect to $\alpha = 1 \cdot 2 \cdot \mathbf{3} \cdot \mathbf{4} \cdot 5 \cdot 6 \cdot \mathbf{4} \cdot \mathbf{3} \cdot 7 \cdot 8$. We can define σ 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 α if $|\Sigma| = 2$. In order to substantiate this theorem, we need the following lemma.

▶ **Lemma 21.** *Let Σ be an alphabet with $|\Sigma| = 2$, and let $\sigma : \mathbb{N}^+ \to \Sigma^+$ be a morphism. For all $x_1, x_2 \in \mathbb{N}$, there exist $a_1 \sqsubseteq \sigma(x_1)$ and $a_2 \sqsubseteq \sigma(x_2)$ such that $a_1 a_2 \sqsubseteq \sigma(x_1 \cdot x_2)$ and $a_2 a_1 \sqsubseteq \sigma(x_2 \cdot x_1)$.*

The next result introduces a *sufficient* condition on the *non-existence* of weakly unambiguous length-increasing morphisms $\sigma : \mathbb{N}^+ \to \Sigma^+$, $|\Sigma| = 2$. 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 α. Hence, this requirement must be satisfied in the following theorem.

▶ **Theorem 22.** *Let $\alpha \in \mathbb{N}^+$ satisfying $E_\alpha \neq \emptyset$. Let Σ be an alphabet with $|\Sigma| = 2$. There is no weakly unambiguous length-increasing morphism $\sigma : \mathbb{N}^+ \to \Sigma^+$ with respect to α if*

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 α,*
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 α, 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 $L_{s'} = \{e\}$.*

In order to illustrate Theorem 22, we consider a few examples:

▶ **Example 23.** Let,

$$\begin{aligned} \alpha &:= 1 \cdot 2 \cdot \mathbf{3} \cdot \mathbf{4} \cdot 5 \cdot 6 \cdot \mathbf{4} \cdot \mathbf{3} \cdot 7 \cdot 8 \cdot \mathbf{3} \cdot 9 \cdot 10, \\ \beta &:= 1 \cdot 2 \cdot \mathbf{4} \cdot 5 \cdot 6 \cdot \mathbf{3} \cdot \mathbf{4} \cdot 7 \cdot 8 \cdot \mathbf{3} \cdot 9 \cdot 10 \cdot \mathbf{4} \cdot \mathbf{3} \cdot 11 \cdot 12, \\ \gamma &:= 1 \cdot 2 \cdot \mathbf{3} \cdot \mathbf{4} \cdot 5 \cdot 6 \cdot \mathbf{7} \cdot \mathbf{8} \cdot 9 \cdot 10 \cdot \mathbf{4} \cdot \mathbf{3} \cdot 11 \cdot 12 \cdot \mathbf{8} \cdot \mathbf{7} \cdot 13 \cdot 14. \end{aligned}$$

Then, according to Definition 6, E_α, E_β and E_γ are nonempty (the respective variables are typeset in bold face). However, since the patterns satisfy Theorem 22, there is no weakly unambiguous morphism $\sigma : \mathbb{N}^+ \to \Sigma^+$ with respect to them if $|\Sigma| = 2$.

Theorem 22 and Example 23 directly imply the insight mentioned above that Theorem 13 does not hold for binary alphabets Σ:

▶ **Corollary 24.** *Let Σ be an alphabet with $|\Sigma| = 2$. There is an $\alpha \in \mathbb{N}^+$ such that E_α is not empty and there is no weakly unambiguous length-increasing morphism $\sigma : \mathbb{N}^+ \to \Sigma^+$ with respect to α.*

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

▶ **Theorem 25.** *Let $\alpha \in \mathbb{N}^+$, and let Σ be an alphabet, $|\Sigma| = 2$. Also, assume 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 $i \in \mathrm{var}(\alpha)$. If a morphism $\sigma : \mathbb{N}^+ \to \Sigma^+$ satisfies
- *$|\sigma(e)| = 2$ and $|\sigma(e')| = 2$,*
- *for every $j \in \mathrm{var}(\alpha) \setminus \{e, e'\}$, $|\sigma(j)| = 1$, and*
- *there is no $x \in \Sigma$ with $x^2 \sqsubseteq \sigma(\alpha)$,*

then σ is weakly unambiguous with respect to α.

The main difference between Theorem 25 and Theorem 22 is that those patterns α being examined in Theorem 25 do not satisfy Condition 3 of Theorem 22. 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 25, we now give some examples. In contrast to Example 23, 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 26.** Let $\sigma : \mathbb{N}^+ \to \{a, b\}^+$ be a morphism. We define the morphism σ for the following patterns α (where the factors featured by Theorem 25 are typeset in bold face) as follows:
- $\alpha = 1 \cdot 2 \cdot \mathbf{5} \cdot \mathbf{3} \cdot \mathbf{4} \cdot 6 \cdot 7 \cdot 8 \cdot \mathbf{5} \cdot \mathbf{4} \cdot \mathbf{3} \cdot 9 \cdot 10$.
  σ 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 $\sigma(10) := a$.
- $\alpha = 1 \cdot 2 \cdot \mathbf{3} \cdot \mathbf{4} \cdot \mathbf{5} \cdot 6 \cdot 7 \cdot \mathbf{4} \cdot \mathbf{3} \cdot \mathbf{5} \cdot 8 \cdot 9$.
  σ 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 $\sigma(9) := a$.
- $\alpha = 1 \cdot 2 \cdot \mathbf{3} \cdot \mathbf{4} \cdot 5 \cdot 6 \cdot 7 \cdot \mathbf{8} \cdot \mathbf{3} \cdot \mathbf{4} \cdot 9 \cdot 10 \cdot 11 \cdot \mathbf{8} \cdot \mathbf{4} \cdot \mathbf{3} \cdot 12 \cdot 13$.
  σ 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 25, it can be easily verified that, in all above cases, σ is length-increasing and weakly unambiguous with respect to α.

The patterns in Example 26 further illustrate that the converse of Theorem 20 does not hold true. More precisely, although for every pattern α 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 α, there is a weakly unambiguous length-increasing morphism $\sigma : \mathbb{N}^+ \to \{a, b\}^+$ with respect to α.

Due to Theorems 22 and 25, 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 27.** *Let $\alpha \in \mathbb{N}^+$ with $|\alpha| > 1$, and let $|\Sigma| := 2$. The problem of whether there is a weakly unambiguous length-increasing morphism $\sigma : \mathbb{N}^+ \to \Sigma^+$ with respect to α 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 α must not be length-increasing for the variables in S_α. On the other hand, increasing the length of the morphic images of the variables in E_α under a morphism $\sigma : \mathbb{N}^+ \to \Sigma^+$, $|\Sigma| = 2$, seems to increase the chance of the existence of a morphism $\tau : \mathbb{N}^+ \to \Sigma^+$ satisfying $\tau(\alpha) = \sigma(\alpha)$ and, for some $i \in \mathrm{var}(\alpha)$, $\tau(i) \neq \sigma(i)$. Consequently, we believe that if all morphisms σ with, for every $e \in E_\alpha$ and an $x \in \mathbb{N}$, $|\sigma(e)| \leq x$ are not weakly unambiguous with respect to α, then there does not exist a weakly unambiguous morphism σ 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, Corollary 18 and Corollary 19 also hold for this special case, i. e., for any pattern α, every weakly unambiguous morphism must map the variables in S_α to words of length 1, which implies that such a morphism can only be length-increasing if E_α is not empty. Incorporating these observations, we now consider an example.

▶ **Example 28.** 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 $\sigma : \mathbb{N}^+ \to \{a\}^+$ by $\sigma(4) := aa$ and $\sigma(i) := a$, $i \in \mathbb{N} \setminus \{4\}$. It can be easily verified that σ is weakly unambiguous with respect to α_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 τ, 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, σ is not weakly unambiguous with respect to α_2.

Quite obviously, the fact that σ is unambiguous with respect to α_1 and ambiguous with respect to α_2 is due to 4 being the only variable in α_1 that has a single occurrence, whereas α_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 29.** *Let $\alpha \in \mathbb{N}^+$, $\mathrm{var}(\alpha) = \{1, 2, 3, ..., n\}$. There is no weakly unambiguous length-increasing morphism $\sigma : \mathbb{N}^+ \to \{a\}^+$ with respect to α if and only if, for every $i \in \mathrm{var}(\alpha)$, there exist $n_1, n_2, ..., n_n \in \mathbb{N} \cup \{0\}$, such that*

$$|\alpha|_i = n_1|\alpha|_1 + n_2|\alpha|_2 + [...] + n_{i-1}|\alpha|_{i-1} + n_{i+1}|\alpha|_{i+1} + [...] + n_n|\alpha|_n.$$

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 29 can exclusively refer to their number of occurrences.

7 Conclusion

In this paper, we have demonstrated that there is a weakly unambiguous length-increasing morphism $\sigma : \mathbb{N}^+ \to \Sigma^+$, $|\Sigma| \geq 3$, with respect to $\alpha \in \mathbb{N}^+$ if and only if E_α is not empty, where

$E_\alpha \subseteq \mathrm{var}(\alpha)$ 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 Σ. 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 $\sigma : \mathbb{N}^+ \to \Sigma^+$ 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 the anonymous referees, whose careful remarks and suggestions helped to correct Theorem 22 and the proof of Theorem 20.

References

1. J. Berstel and D. Perrin. *Theory of Codes*. Academic Press, Orlando, 1985.
2. D.D. Freydenberger and D. Reidenbach. The unambiguity of segmented morphisms. *Discrete Applied Mathematics*, 157:3055–3068, 2009.
3. 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.
4. 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`.
5. 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.
6. S. Holub. Polynomial-time algorithm for fixed points of nontrivial morphisms. *Discrete Mathematics*, 309:5069–5076, 2009.
7. 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.
8. D. Reidenbach. A non-learnable class of E-pattern languages. *Theoretical Computer Science*, 350:91–102, 2006.
9. 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.
10. D. Reidenbach. Discontinuities in pattern inference. *Theoretical Computer Science*, 397:166–193, 2008.
11. D. Reidenbach and J.C. Schneider. Restricted ambiguity of erasing morphisms. In *Proc. 14th International Conference on Developments in Language Theory, DLT 2010*, volume 6224 of *Lecture Notes in Computer Science*, pages 387–398, 2010.
12. J.C. Schneider. Unambiguous erasing morphisms in free monoids. *RAIRO Informatique théoretique et Applications*, 44:193–208, 2010.