Loughborough University

Thesis Access Form

Copy No…………..……………………**Location**……………………………………………….……………...…

Author…………...………………………………………………………………………………………..……

Title………..

Status of access OPEN / RESTRICTED / CONFIDENTIAL

Moratorium Period:…………………………………years, ending…………../…………200………………………

Conditions of access approved by (CAPITALS):…………………………………………………………………

Supervisor (Signature)………………………………………...…………………………………...

School of……………………………………………………………...…………………………………

Author's Declaration: *I agree the following conditions:*

Open access work shall be made available (in the University and externally) and reproduced as necessary at the discretion of the University Librarian or Dean of School. It may also be digitised by the British Library and made freely available on the Internet to registered users of the EThOS service subject to the EThOS supply agreements.

*The statement itself shall apply to **ALL** copies including electronic copies:*

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

Restricted/confidential work: All access and any photocopying shall be strictly subject to written permission from the University Dean of School and any external sponsor, if any.

Author's signature……………………………………….Date…………………………………...…………...……..

users declaration: for signature during any Moratorium period (Not Open work): ***I undertake to uphold the above conditions:***			
Date	Name (CAPITALS)	Signature	Address

4.1e

Conditions on the Existence of Unambiguous Morphisms

by

Hossein Nevisi

A Doctoral Thesis

Submitted in partial fulfilment
of the requirements for the award of

Doctor of Philosophy
of
Loughborough University

20th June 2012

Research Student Office, Academic Registry
Loughborough University, Leicestershire, LE11 3TU, UK
Switchboard: +44 (0)1509 263171 Fax: +44 (0)1509 223938

Loughborough University

Certificate of Originality

This is to certify that I am responsible for the work submitted in this thesis, that the original work is my own except as specified in acknowledgements or in footnotes, and that neither the thesis nor the original work contained therein has been submitted to this or any other institution for a higher degree.

. . . .

Hossein Nevisi

20th June 2012

Abstract

A morphism σ is *(strongly) unambiguous* with respect to a word α if there is no other morphism τ that maps α to the same image as σ. Moreover, σ is said to be *weakly unambiguous* with respect to a word α if σ is the only *nonerasing* morphism that can map α to $\sigma(\alpha)$, i. e., there does not exist any other nonerasing morphism τ satisfying $\tau(\alpha) = \sigma(\alpha)$. In the first main part of the present thesis, we wish to characterise those words with respect to which there exists a weakly unambiguous *length-increasing* morphism that maps 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.

The second main part of the present thesis studies the question of whether, for any given word, there exists a strongly unambiguous *1-uniform* morphism, i. e., a morphism that maps every letter in the word to an image of length 1. This problem shows some connections to previous research on *fixed points* of nontrivial morphisms, i. e., those words α for which there is a morphism ϕ satisfying $\phi(\alpha) = \alpha$ and, for a symbol x in α, $\phi(x) \neq x$. Therefore, we can expand our examination of the existence of unambiguous morphisms to a discussion of the question of whether we can reduce the number of different symbols in a word that is not a fixed point such that the resulting word is again not a fixed point. This problem is quite similar to the setting of Billaud's Conjecture, the correctness of which we prove for a special case.

Acknowledgements

First of all, I would like to express my sincere gratitude to my supervisor, Dr. Daniel Reidenbach, for the continuous support of my research, for his patience, motivation, enthusiasm, and immense knowledge. This thesis would not have been written without his experience, effort and significant advice. One simply could not wish for a better or friendlier supervisor.

Furthermore, I would like to thank Prof. Maxime Crochemore, Dr. Walter Hussak, and Dr. Paul Bell for reviewing this thesis and for their insightful comments and suggestions.

I would also like to acknowledge three of my previous tutors Prof. Hayedeh Ahrabian (R.I.P), Dr. Morteza Mohammad-Noori, and Prof. Mohammad Ardeshir for believing in me and encouraging me to do my PhD.

My family deserve a special mention for their constant support and encouragement. I will forever be indebted to them.

Finally, my special thanks are due to my wife Parisa for her understanding and love during the past few years. She shared my joy when gaining new results, and she was always curious about them. Undoubtedly, this thesis could not have been finished without her patience and support.

Contents

Chapter 1

Introduction

The main concept of this thesis is a word (or string), i. e., a finite or infinite sequence of symbols taken from a countable set. The starting point of formal studies on words goes back to the beginning of the last century. At that time, Axel Thue (1863-1922) did some mathematical research about repetitions in words and he wrote two papers on this topic, one in 1906 [38] and one in 1912 [39]. However, since his results were published in a rather fameless journal, some of his main results were reproved many years later (e. g., by Morse and Hedlund [26] in 1944).

The significance of doing research on words, as a topic in its own right, was truly appreciated by the beginning of the appearance of computers in the 1950s (although there were a few other papers on words before that time, e. g., [23], [24], and [25]). At that time, Schützenberger started a systematic research on theory of codes, see [37]. Moreover, Novikov and Adian developed the theory of words as a strong tool to find a fundamental solution to the Burnside Problem for groups, see [1].

The studies on words as a separate topic has grown rapidly due to the inevitable role of words in many aspects of computers and computing such as computer programs, logical formulas, and various kinds of application data; in fact, any sequence of bits in a computer is nothing but a word. As the first and still one of the most comprehensive books on words, Lothaire [18] needs to be mentioned, which was published in 1983 and covers many basic insights into combinatorial problems for words. After publishing this book, the title "Combinatorics on Words" of the book was chosen for the field of research dealing with combinatorial properties of words and operations on words. A second volume of that book "Algebraic Combination on Words" [19], was published in 2002 and a third volume "Applied Combinatorics on Words" [20], was published in 2005. In the latest Mathematic Subject Classification in the year 2010 (MSC2010), combinatorics on words is a topic of its own under the section discrete mathematics related to computer science. This classification results from the fact that the field of combinatorics on words does not only have many connections to several branches of mathematics such as semigroups, groups, number theory, probability, combinatorial topology, and dynamical systems, but also frequently occurs in problems of theoretical computer science, dealing with automata and formal languages.

A natural algebraic concept related to finite words is a free monoid or semigroup. In fact, the set $\mathcal{A}^*$ containing all finite words (including the empty word ε as the identity element) over some fixed set $\mathcal{A}$ of symbols – which is called an *alphabet* –, and the binary operation on words, which is the *concatenation*, establish a monoid. Also, the set $\mathcal{A}^+$ including all finite nonempty words over the set $\mathcal{A}$ and the operation concatenation establish a semigroup. Besides, words like "ab" and "ba" are not equal. Consequently, words can be seen as discrete combinatorial or algebraic objects in a noncommutative structure. Hence, noncommutativity and discreteness are fundamental features of words.

In addition to the operation of concatenation, a *morphism* is another fundamental operation on words. A morphism is a function that is compatible with the concatenation. More precisely, for any sets $\mathcal{A}, \mathcal{B}$ of symbols, a morphism from the monoid $\mathcal{A}^*$ into the monoid $\mathcal{B}^*$ is a mapping $\sigma : \mathcal{A}^* \to \mathcal{B}^*$ such that $\sigma(uv) = \sigma(u)\sigma(v)$ for all $u, v \in \mathcal{A}^*$, and $\sigma(\varepsilon) = \varepsilon$. This definition means that the function σ maps a word u over $\mathcal{A}$ to a word u' over $\mathcal{B}$ by mapping each symbol occurring in u to a word over $\mathcal{B}$, and concatenates these images in accordance with the order of the occurrences of the symbols in u. As an example of a well-known morphism, we can point to the morphism $\sigma' : \{a, b\}^* \to \{a, b\}^*$ that is defined by $\sigma'(a) := ab$ and $\sigma'(b) := a$. The morphism σ' is called the *Fibonacci morphism* due to the fact that the length of the words a, $\sigma'(a) = ab$, $\sigma'^2(a) = aba$, $\sigma'^3(a) = abaab$, etc. equals the Fibonacci number sequence 1, 2, 3, 5, etc. Also, as a consequence of the above definition of a morphism σ, we can say that for every word u over any set $\mathcal{A}$ of symbols and, for every monoid $\mathcal{B}^*$, u induces a partition of $\mathcal{B}^*$, depending on the question of whether, for any word $u' \in \mathcal{B}^*$, there exists a morphism σ with $\sigma(u) = u'$. This implies that some properties of u may be reflected by u' if u' is a possible morphic image of u. Consequently, many studies have been conducted about the properties of morphisms, and those morphisms are considered which lead to minimum loss of information about the preimage. These studies have a deep connection to *coding theory* [3]. Coding theory directly deals with the problem of the construction of a word over an alphabet $\mathcal{B}$ that contains as much information as possible about another word over an alphabet $\mathcal{A}$ (commonly satisfying $\mathcal{B} \subset \mathcal{A}$). For this purpose, coding theory uses a fixed *injective* morphism mapping the words in $\mathcal{A}^*$ onto selected words in $\mathcal{B}^*$.

However, in this thesis, we do not wish to consider single fixed morphisms that are applied to each word in some set. Instead, we are interested in a setting where several morphisms are applied to the same word. This concept is related to the field of *pattern languages*. A pattern language is the set of all morphic images of one fixed word – the common preimage of these morphic images is called a *pattern* (for more information on pattern languages see, e. g., Mateescu and Salomaa [22]). In the field of pattern languages, we might face the situation that two different morphisms applied to a given pattern generate the same morphic image. Referring to this observation, Freydenberger, Reidenbach and Schneider [11] define the concept of *ambiguity* of morphisms. For any alphabets $\mathcal{A}$ and $\mathcal{B}$, a morphism $\sigma : \mathcal{A}^* \to \mathcal{B}^*$ is said to be *unambiguous* with respect to a word α if there does not exist a second morphism $\tau : \mathcal{A}^* \to \mathcal{B}^*$ mapping α to the same image as σ. For example, if we consider $\mathcal{A} := \mathbb{N}$ (we always use the set of natural number $\mathbb{N}$ as an infinite domain alphabet), $\mathcal{B} := \{a, b\}$ and $\alpha_0 := 1 \cdot 2 \cdot 3 \cdot 1 \cdot 3 \cdot 2$ (where we separate the symbols in α_0 by a dot), then the morphism σ_0, defined by $\sigma_0(1) := a$, $\sigma_0(2) := a$, $\sigma_0(3) := b$, is not unambiguous with respect to α_0, since there exists a different morphism τ_0, given by $\tau_0(1) := \varepsilon$ (i. e., τ_0 maps 1 to the empty word), $\tau_0(2) := a$, $\tau_0(3) := ab$, satisfying $\tau_0(\alpha_0) = \sigma_0(\alpha_0)$:

$$\sigma_0(\alpha_0) = \underbrace{\overbrace{a}^{\sigma_0(1)}}_{\tau_0(2)}\underbrace{\overbrace{a}^{\sigma_0(2)}\overbrace{b}^{\sigma_0(3)}}_{\tau_0(3)}\underbrace{\overbrace{a}^{\sigma_0(1)}\overbrace{b}^{\sigma_0(3)}}_{\tau_0(3)}\underbrace{\overbrace{a}^{\sigma_0(2)}}_{\tau_0(2)} = \tau_0(\alpha_0)\,.$$

In contrast to this, e. g., the morphism $\sigma_1 : \{1, 2, 3\}^* \to \{a, b\}^*$ given by, $\sigma_1(1) := a$, $\sigma_1(2) := ab$, $\sigma_1(3) := b$, is unambiguous with respect to α_0, as can be verified with moderate effort.

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 [14]), *word equations* (see, e. g., Choffrut [5]) and, as mentioned, *pattern languages* (see Mateescu and Salomaa [22]). The equality set of two morphisms σ, τ is the set of all words α satisfying $\sigma(\alpha) = \tau(\alpha)$, and, thus, the famous undecidable Post Correspondence Problem (PCP) [30] is simply the emptiness problem for equality sets. In the terminology related to this problem, each word α in the equality set of σ and τ is said to be a solution to the PCP for σ and τ, and, hence, whenever we find a morphism σ such that σ is unambiguous with respect to α, then α is a non-solution to the PCP for σ and any other morphism τ.

In contrast to the broad and profound knowledge on coding theory, the ambiguity of morphisms has not been studied extensively, despite its connection to the PCP. Furthermore, insights into the ambiguity of morphisms have been used to solve a number of prominent problems with regard to the topic of pattern languages (see, e. g., Reidenbach [31, 32, 33]). This results from the fact that unambiguous morphisms have the ability to optimally encode information about the structure of the word (in a setting where various morphisms are applied to the same word). This shows an interesting contrast to the foundations of coding theory (see Berstel and Perrin [3]).

Since unambiguity can, thus, be seen as a desirable property of morphisms, the initial work on this topic by Freydenberger, Reidenbach and Schneider [11] and most of the subsequent papers have focused on the following question:

Question 1.1. *Let α be a word over an arbitrary alphabet. Does there* exist *a morphism with a finite target alphabet that is* unambiguous *with respect to α?*

In order to further qualify this question, [11] 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 α if there does not exist a morphism τ satisfying $\tau(\alpha) = \sigma(\alpha)$ and, for a symbol x occurring in α, $\tau(x) \neq \sigma(x)$. The second type slightly relaxes this requirement by calling σ *weakly unambiguous* with respect to α if there is no *nonerasing* morphism τ (i. e., τ must not map any symbol to the empty word) showing the above properties. Thus, e. g., our initial example morphism σ_0 is weakly unambiguous with respect to α_0, but it is not strongly unambiguous. 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 focused 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. [11] 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 (which is defined in the next paragraph; for additional explanations see Sections 2.3 and 3.2) or being a shortest generator of a terminal-free E-pattern language (see Section 2.4 for the definition and Section 3.3 for more explanations).

The present thesis studies Question 1.1 from two points of view. The first view deals with the existence of weakly unambiguous morphisms. However, since Question 1.1 is trivial if we allow a morphism to map every letter in a word to an image of length 1 (see Section 3.1 for additional explanations), we restrict ourselves to *length-increasing morphisms*, i. e., those morphisms that map a word to an image which is strictly longer than the word. The second view examines Question 1.1 while we restrict our considerations to *1-uniform morphisms*, i. e., morphisms that map every symbol in the word to an image of length 1. Furthermore, our studies regarding the ambiguity of 1-uniform morphisms lead to some interesting results on *fixed points* of nontrivial morphisms, i. e., a word α is a fixed point of ϕ if $\phi(\alpha) = \alpha$ and, for a symbol x in α, $\phi(x) \neq x$. These results show some connections to Billaud's Conjecture [4].

The present thesis is structured as follows: Chapter 2 introduces the basic definitions and notations that we shall use in this thesis. In Chapter 3, we describe the current state of knowledge regarding the ambiguity of morphisms, and we introduce our research questions which we shall study in this thesis. In Chapter 4, we investigate the existence of weakly unambiguous nonerasing morphisms. Subsequent to this, Chapter 5 studies the existence of strongly unambiguous 1-uniform morphisms for arbitrary words. Furthermore, the said chapter considers the concept of fixed points by answering a problem which is similar to Billaud's Conjecture. Additionally, in this chapter, we prove the correctness of Billaud's Conjecture for a special case not studied in the literature so far. Finally, Chapter 6 summarises the main statements of the present thesis and gives some problems that are left open.

Most major results of this thesis have been previously published in [8] (conference version: [9]), [28] and [27].

Chapter 2

Basic notations and definitions

In order to keep this thesis self-contained, we begin the formal part of it with some basic definitions and concepts of combinatorics on words and morphisms. A major part of our terminology is adopted from the research on pattern languages (cf. Mateescu and Salomaa [22]). Additionally, for notions not explained explicitly, we refer the reader to [6, 18, 19].

2.1 Words and patterns

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 ε. The notation $\mathcal{A}^*$ refers to the set of all (empty and nonempty) 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 $[\ldots]$ is used to omit some canonically defined parts of a given word, e. g., $\alpha = 1 \cdot 2 \cdot [\ldots] \cdot 5$ stands for $\alpha = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5$.

Let $\mathbb{N}$ be the set of natural numbers, and $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. In order to obtain unrestricted results, we often use $\mathbb{N}$ an as infinite alphabet. Also, to distinguish between a word over $\mathbb{N}$ and a word over a (possibly finite) alphabet Σ, we call the former a *pattern*. We call any symbol in $\mathbb{N}$ a variable and any symbol in Σ a letter – we often assume that $\Sigma := \{a, b, c, ...\}$. We name patterns with lower case letters from the beginning of the Greek alphabet such as α, β, γ. With regard to an arbitrary pattern α, $\mathrm{var}(\alpha)$ denotes the set of all variables occurring in α. We say that α is *in canonical form* if α is lexicographically minimal among all its renamings (which is formally defined in Section 2.2), where the lexicographic order is derived from the usual order on $\mathbb{N}$, i. e., $1 < 2 < 3 < \ldots$.

2.2 Morphisms and the concept of ambiguity

A *morphism* is a mapping that is compatible with concatenation, i. e., for any alphabets $\mathcal{A}, \mathcal{B}$, $\sigma : \mathcal{A}^* \to \mathcal{B}^*$ is a morphism if it satisfies $\sigma(\alpha \cdot \beta) = \sigma(\alpha) \cdot \sigma(\beta)$ for all $\alpha, \beta \in \mathcal{A}^*$. A morphism $\sigma : \mathcal{A}^* \to \mathcal{B}^*$ is called *nonerasing* provided that, for every $i \in \mathcal{A}$, $\sigma(i) \neq \varepsilon$; otherwise, σ is an *erasing* morphism. If σ is nonerasing, then we often indicate this by writing $\sigma : \mathcal{A}^+ \to \mathcal{B}^+$. Also, σ is said to be *injective* (on $\mathcal{A}^*$) providing that, for any words $\alpha, \beta \in \mathcal{A}^*$, the equality $\sigma(\alpha) = \sigma(\beta)$ implies $\alpha = \beta$. A morphism σ is *length-increasing (for α)* if $|\sigma(\alpha)| > |\alpha|$, and it is called *1-uniform* if, for every $i \in \mathcal{A}$, $|\sigma(i)| = 1$. Regarding 1-uniform morphisms, a 1-uniform morphism $\phi : \mathbb{N}^* \to \mathbb{N}^*$ is an *alphabet reduction (for α)* if ϕ maps α to an image containing a smaller number of different variables. A morphism is called a *renaming* if it is injective and 1-uniform. We additionally call any word v a renaming of a word w if there is a morphism ψ that is a renaming and satisfies $\psi(w) = v$. For any morphism σ, $\sigma : \mathcal{A}^* \to \mathcal{B}^*$, M_σ consists of those variables $i \in \mathcal{A}$ satisfying $\sigma^k(i) = \varepsilon$ for some $k \geq 1$. This set is called the set of *mortal* variables of σ. The *mortality exponent* of a morphism σ is defined to be the least integer $t \geq 0$ such that $\sigma^t(i) = \varepsilon$ for all $i \in M_\sigma$. We write the mortality exponent as $\exp(\sigma) = t$. Moreover, a variable i is said *expansive* variable if there exist $\beta, \gamma \in M_\sigma^*$ with $\sigma(i) = \beta a \gamma$, $a \in \mathcal{B}$, and $|\sigma(i)| \geq 2$. The set of all expansive variables of σ is denoted by E_σ.

For any alphabet Σ, for any morphism $\sigma : \mathbb{N}^* \to \Sigma^*$ and for any pattern $\alpha \in \mathbb{N}^+$, we call σ *(strongly) unambiguous* with respect to α if there is no morphism $\tau : \mathbb{N}^* \to \Sigma^*$ with $\tau(\alpha) = \sigma(\alpha)$ and, for some $q \in \mathrm{var}(\alpha)$, $\tau(q) \neq \sigma(q)$. Moreover, for any morphism $\sigma : \mathbb{N}^+ \to \Sigma^+$, σ is said to be *weakly unambiguous* with respect to α, if there is no morphism $\tau : \mathbb{N}^+ \to \Sigma^+$ with $\tau(\alpha) = \sigma(\alpha)$ and, for some $q \in \mathrm{var}(\alpha)$, $\tau(q) \neq \sigma(q)$. Additionally, we can call σ *ambiguous* with respect to α if it is not unambiguous (or, if applicable, weakly unambiguous), but we use this term in an informal context only.

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 $\alpha := x_1 \cdot x_2 \cdot [\ldots] \cdot x_n$, $x_k \in \mathbb{N}$, $1 \leq k \leq n$, and let $\sigma : \mathrm{var}(\alpha)^+ \to \Sigma^+$ and $\tau : \mathrm{var}(\alpha)^+ \to \Sigma^+$ 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

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

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

The following example illustrates this definition: Let $\alpha := 1 \cdot 2 \cdot 3$, and let the morphism $\sigma : \mathbb{N}^+ \to \{a, b\}^+$ be given by $\sigma(1) := a$, $\sigma(2) := b$ and $\sigma(3) := ab$. Furthermore, let the morphism $\tau : \mathbb{N}^+ \to \{a, b\}^+$ 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 $\sigma(2)$.

2.3 Fixed points, prolix patterns and succinct patterns

A pattern $\alpha \in \mathbb{N}^*$ is a *fixed point (of a nontrivial morphism)* if there is a nontrivial morphism ϕ satisfying $\phi(\alpha) = \alpha$ and, for a symbol x in α, $\phi(x) \neq x$. Note that the set of fixed points is equivalent to the set of prolix patterns, which is a vital concept for research on the unambiguity of morphisms, and it is defined as follows: We call any $\alpha \in \mathbb{N}^+$ *prolix* if and only if, there exists a factorisation $\alpha = \beta_0\gamma_1\beta_1\gamma_2\beta_2[\ldots]\gamma_n\beta_n$ 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.

Furthermore, the set of fixed points corresponds to the set of morphically imprimitive words: A pattern $\alpha \in \mathbb{N}^*$ is *morphically imprimitive* if there are a strictly shorter pattern β and morphisms $\phi, \phi' : \mathbb{N}^* \to \mathbb{N}^*$ satisfying $\phi(\beta) = \alpha$ and $\phi'(\alpha) = \beta$; otherwise, α is *morphically primitive.*

2.4 Pattern languages

The *pattern language* of a pattern is the set of all its morphic images in some fixed free monoid Σ^*, where Σ, as defined before, is an arbitrary alphabet (such as $\{a, b, c\}$). With regard to any $\alpha \in \mathbb{N}^+$, we distinguish between its *E-pattern language* $L_E(\alpha) := \{\sigma(\alpha) | \sigma : \mathbb{N}^* \to \Sigma^*\}$ and its *NE-pattern language* $L_{NE}(\alpha) := \{\sigma(\alpha) | \sigma : \mathbb{N}^+ \to \Sigma^+\}$. Note that this definition usually is referred to as *terminal-free* pattern languages, due to the fact that, a pattern commonly is seen as a word in $(\mathbb{N} \cup \Sigma)^+$, which means that the pattern also contains *terminal* symbols – arbitrary symbols in Σ. Therefore, in the general case, the pattern language of a pattern $\alpha \in (\mathbb{N} \cup \Sigma)^+$ is the set of all images of α under *terminal-preserving* morphisms $\sigma : (\mathbb{N} \cup \Sigma)^* \to \Sigma^*$ with $\sigma(a) = a$ for every $a \in \Sigma$. We write Pat_Σ for the set of all patterns and $\mathrm{Pat}_{\mathrm{tf}}$ denotes the set of all terminal-free patterns. Moreover, we can use ePAT_Σ (or ePAT for short) as an abbreviation for the full class of E-pattern languages and $\mathrm{ePAT}_{\mathrm{tf},\Sigma}$ (or $\mathrm{ePAT}_{\mathrm{tf}}$ for short) for the class of all terminal-free E-pattern languages (for more information on pattern languages see, e. g., Mateescu and Salomaa [22]).

A class $\mathcal{L}$ of languages is *indexable* if and only if there exists an *indexed family* $(L_i)_{i \in \mathbb{N}_0}$ such that $\mathcal{L} = \{L_i \mid i \in \mathbb{N}_0\}$. This means that there is a total and computable function which, given any pair of an index $i \in \mathbb{N}_0$ and a word $w \in \Sigma^*$, decides on whether or not $w \in L_i$.

Chapter 3

Related literature and research questions

In this chapter, we describe the current state of knowledge on the ambiguity of morphisms. Moreover, we formally introduce the main problems which we shall investigate in the present thesis.

3.1 Ambiguity of morphisms

As mentioned in Chapter 1, the ambiguity of morphisms is a new topic that has not been studied a lot. Nevertheless, there exist three important papers in this area that initiate a systematic research on the ambiguity of morphisms, namely by Freydenberger, Reidenbach and Schneider [11], Freydenberger and Reidenbach [10] and Schneider [36].

In [11], the authors introduce the question of determining for which patterns $\alpha \in \mathbb{N}^+$ there exists a nonempty word w in $\{a, b\}^*$ such that there is exactly one morphism σ with $\sigma(\alpha) = w$. In other words, for any pattern α over some alphabet this paper asks for the existence of a morphism σ such that σ is unambiguous with respect to α; to this end, it focuses on nonerasing morphisms σ. A first basic result on this question demonstrates that there is no single nonerasing morphism σ such that, for every $\alpha \in \mathbb{N}^+$, σ is strongly unambiguous with respect to α. Hence, strongly unambiguous nonerasing morphisms must be tailored to the structure of the respective preimages. The main result of [11] characterises those patterns with respect to which there is a strongly unambiguous nonerasing morphism:

Theorem 3.1 (Freydenberger et al. [11]). *Let $\alpha \in \mathbb{N}^*$, and 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.*

The injective morphism σ which is defined in the proof of the above theorem maps every variable $k \in \text{var}(\alpha)$, depending on some conditions, to one of the following images:

1. $ab^{3k}a\ \ ab^{3k+1}a\ \ ab^{3k+2}a$,
2. $ba^{3k}b\ \ ab^{3k+1}a\ \ ab^{3k+2}a$,
3. $ab^{3k}a\ \ ab^{3k+1}a\ \ ba^{3k+2}b$,
4. $ba^{3k}b\ \ ab^{3k+1}a\ \ ba^{3k+2}b$.

It is also worth noting that, in a sense, [11] complements the research on the nondeterminism of pattern languages that has been initiated by Mateescu and Salomma [21]. This is because [11] shows that for every pattern in some class, there exists at least one nonempty word in $\{a,b\}^*$ that has exactly one generating morphism – this generally holds true for research on the existence of unambiguous morphisms –, whereas, in a more general context, [21] examines the question whether, for an arbitrary upper bound $n \in \mathbb{N}$, there exists at least one pattern such that each of its morphic images has at most n distinct generating morphisms.

The paper [10] investigates the ambiguity of a fixed morphism with respect to the set of all patterns in $\mathbb{N}^+$, i. e., the authors ask for which patterns the morphism is strongly unambiguous. This paper presents the first approach to a characterisation of sets of patterns with respect to which certain fixed morphisms are unambiguous. To this end, the authors define a so-called segmented morphism $\sigma_n : \mathbb{N}^* \to \{a,b\}^*$, $n \in \mathbb{N}$, which maps each variable $x \in \mathbb{N}$ to a word that consists of n distinct factors in ab^+a:

$$\sigma_n(x) = ab^{nx-(n-1)}aab^{nx-(n-2)}a...ab^{nx-1}aab^{nx}a.$$

This paper then introduces the set $U(\sigma_n)$, which consists of all those $\alpha \in \mathbb{N}^+$ with respect to which σ_n is unambiguous, and it studies the relation of this set to any $U(\sigma_m)$, $m \neq n$. The studies of the paper are based on the following hypothesis:

Hypothesis 3.2. *For $0 \leq i < j$, $U(\sigma_i) \subseteq U(\sigma_j)$.*

The paper [10] shows that, in contrast to the above mentioned hypothesis, firstly, $U(\sigma_n) = U(\sigma_3)$ for all $n \geq 3$. Secondly, the sets $U(\sigma_0)$, $U(\sigma_1)$ and $U(\sigma_2)$ are strictly included in $U(\sigma_3)$ and, they are all incomparable. Also, the paper gives the following characterisation of $U(\sigma_n)$ for $n \geq 3$ by defining an *SCRN-partition* for $\alpha \in \mathbb{N}^*$, which is a partition of the variables of α into all disjoint sets S, C, R and N such that $\alpha \in (N^*SC^*R)^+N^*$:

Theorem 3.3 (Freydenberger et al. [10]). *For every $n \geq 3$,*

$$U(\sigma_n) = U(\sigma_3) = \{ \alpha \in \mathbb{N}^+ \mid \alpha \text{ is morphically primitive and } \alpha \text{ has no SCRN-partition} \}.$$

The systematic research on the ambiguity of erasing morphisms is initiated by Schneider [36], and it is continued in [35]. The paper [36] investigates the following question: For which patterns $\alpha \in \mathbb{N}^+$ does there exist an erasing unambiguous morphism $\sigma : \mathbb{N}^* \to \Sigma^*$? To study this question, the author defines the concept of an ambiguity partition:

Definition 3.4 (Schneider [36]). *Let $\alpha \in \mathbb{N}^+$. We inductively define an ambiguity partition (with respect to α):*

- $(\emptyset, \mathrm{var}(\alpha))$ *is an ambiguity partition with respect to α.*
- *If (E, N) is an ambiguity partition with respect to α and there exists a morphism $\phi : \mathbb{N}^* \to \mathbb{N}^*$ that is nontrivial for N and satisfies $\phi(\alpha) = \pi_N(\alpha)$ – $\pi_N : \mathbb{N}^* \to \mathbb{N}^*$ is a morphism with $\pi_N(x) := x$ if $x \in N$ and $\pi_N(x) := \varepsilon$ if $x \notin N$ – then (E', N') is an ambiguity partition with*

$$\begin{aligned} E' &:= E \cup \{x \in N \mid \phi(x) = \varepsilon\}, \\ N' &:= \{x \in N \mid \phi(x) \neq \varepsilon\}. \end{aligned}$$

The main results of [36] on ambiguity partitions show that the existence of an ambiguity partition with respect to a pattern strongly contributes to the ambiguity of morphisms applied to the pattern:

Theorem 3.5 (Schneider [36]). *Let Σ be an alphabet. Let $\alpha \in \mathbb{N}^+$ and let (E, N) be an ambiguity partition with respect to α. Then every morphism $\sigma : \mathbb{N}^* \to \Sigma^*$ satisfying $\sigma(x) \neq \varepsilon$ for an $x \in E$ is not unambiguous with respect to α.*

In the case of infinite target alphabet, the existence of an ambiguity partition characterises the ambiguity of erasing morphisms:

Theorem 3.6 (Schneider [36]). *Let Σ_∞ be an infinite alphabet and let $\alpha \in \mathbb{N}^+$. There is an unambiguous morphism $\sigma : \mathbb{N}^* \to \Sigma_\infty^*$ with respect to α if and only if $(\mathrm{var}(\alpha), \emptyset)$ is not an ambiguity partition with respect to α.*

The above theorem allows some conclusions to be drawn on the decidability of the question of whether or not there exists an unambiguous morphism $\sigma : \mathbb{N}^* \to \Sigma_\infty^*$ for an arbitrary pattern $\alpha \in \mathbb{N}^+$:

Corollary 3.7 (Schneider [36]). *Let Σ_∞ be an infinite alphabet. Then*

$$\{\alpha \in \mathbb{N}^+ \mid \text{there is no unambiguous morphism } \sigma : \mathbb{N}^* \to \Sigma_\infty^* \text{ with respect to } \alpha\}$$

is decidable and the decision problem is NP-complete.

Concerning finite target alphabets, the author shows that the problem of deciding the above mentioned question is NP-hard:

Corollary 3.8 (Schneider [36]). *Let Σ be an finite alphabet, $|\Sigma| \geq 2$. The problem of deciding*

$$\{\alpha \in \mathbb{N}^+ \mid \text{there is no unambiguous morphism } \sigma : \mathbb{N}^* \to \Sigma^* \text{ with respect to } \alpha\}$$

is NP-hard.

Also, with regard to finite target alphabets, the paper [36] shows that the existence of strongly unambiguous erasing morphisms for a given pattern can essentially depend on the size of the target alphabet Σ of the morphism (in contrast to Theorem 3.1):

Theorem 3.9 (Schneider [36]). *Let $k \in \mathbb{N}$ and Σ_k, Σ_{k+1} be finite alphabets with k and $k+1$ letters, respectively. There exists a pattern $\alpha \in \mathbb{N}^+$ such that:*

- $(\mathrm{var}(\alpha), \emptyset)$ *is not an ambiguity partition with respect to α,*
- *no morphism $\sigma : \mathbb{N}^* \to \Sigma_k^*$ is unambiguous with respect to α, and*
- *there exists an unambiguous morphism $\sigma' : \mathbb{N}^* \to \Sigma_{k+1}^*$ with respect to α.*

Besides, some sufficient conditions on the (non-)existence of unambiguous erasing morphisms are given in this paper.

Reidenbach and Schneider continue studying the ambiguity of erasing morphisms in [35]. To this end, they introduce *moderately ambiguous* morphisms, which are a special case of ambiguous morphisms:

Definition 3.10 (Reidenbach and Schneider [35]). *Let Σ be an alphabet, let $\alpha := i_1 \cdot i_2 \cdot [...] \cdot i_n$ with $n, i_1, i_2, ..., i_n \in \mathbb{N}$, and let $\sigma : \mathbb{N}^* \to \Sigma^*$ be a morphism satisfying $\sigma(\alpha) \neq \varepsilon$. Then σ is called moderately ambiguous with respect to α provided that there exist $l_2, l_3, ..., l_n, r_1, r_2, ..., r_{n-1} \in \mathbb{N}_0$ such that, for every morphism $\tau : \mathbb{N}^* \to \Sigma^*$ with $\tau(\alpha) = \sigma(\alpha)$,*

- *if* $\sigma(i_1) \neq \varepsilon$ *then* $r_1 \geq 1$*,*
- *if* $\sigma(i_n) \neq \varepsilon$ *then* $l_n \leq |\sigma(\alpha)|$*,*
- *for every* $k \in \{2, 3, ..., n-1\}$ *with* $\sigma(i_k) \neq \varepsilon$*,* $l_k \leq r_k$*,*
- *for every* k *with* $1 \leq k \leq n-1$*,* $|\tau(i_1 \cdot i_2 \cdot [...] \cdot i_k)| < l_{k+1}$*, and*
- *for every* k *with* $1 \leq k \leq n-1$*,* $|\tau(i_1 \cdot i_2 \cdot [...] \cdot i_k)| \geq r_k$*.*

σ *is called strongly ambiguous with respect to* α *if and only if it is not moderately ambiguous with respect to* α*.*

In order to state the main result of [35], we also need the following definition:

Definition 3.11 (Reidenbach and Schneider [35])**.** *Let* $\alpha \in \mathbb{N}^+$*. We call* α *morphically erasable if and only if* $(\mathrm{var}(\alpha), \emptyset)$ *is an ambiguity partition for* α*. Otherwise,* α *is called morphically unerasable.*

The paper [35] shows that concerning the ambiguity of erasing morphisms, the partition of patterns into morphically unerasable and erasable patterns has a similar importance as the partition into succinct and prolix patterns regarding the ambiguity of nonerasing morphisms; in other words, both partitions characterise the (non)existence of moderately ambiguous morphisms:

Theorem 3.12 (Reidenbach and Schneider [35])**.** *Let* Σ *be an alphabet,* $|\Sigma| \geq 2$*, let* $\alpha \in \mathbb{N}^+$*. There exists a morphism* $\sigma : \mathbb{N}^* \to \Sigma^*$ *that is moderately ambiguous with respect to* α *if and only if* α *is morphically unerasable.*

In addition to the above theorem, there is an interesting result in [35] with regard to the existence of patterns with only finitely many unambiguous morphisms. We now state this result by assuming that $\mathrm{UNAMB}_\Sigma(\alpha)$ denotes the set of all $\sigma(\alpha)$, where $\sigma : \mathbb{N}^* \to \Sigma^*$ is any morphism that is unambiguous with respect to α:

Theorem 3.13 (Reidenbach and Schneider [35])**.** *Let* $k \in \mathbb{N}$*. Let* Σ_k*,* Σ_{k+1}*,* Σ_{k+2} *be alphabets with* k*,* $k+1$*,* $k+2$ *letters, respectively. There exists an* $\alpha_k \in \mathbb{N}^+$ *such that*

- $|\mathrm{UNAMB}_{\Sigma_k}(\alpha_k)| = 0$*,*
- $|\mathrm{UNAMB}_{\Sigma_{k+1}}(\alpha_k)| = m$ *for an* $m \in \mathbb{N}$*, and*
- $\mathrm{UNAMB}_{\Sigma_{k+2}}(\alpha_k)$ *is an infinite set.*

This phenomenon differs from the research on unambiguous nonerasing morphisms, where every pattern has either infinitely many or not a single unambiguous morphism.

In the present thesis, we initially wish to investigate the existence of *weakly unambiguous* nonerasing morphisms in more detail. The paper [11] introduces the concept of weakly unambiguous morphisms, but it merely states the following trivial observation:

Proposition 3.14 (Freydenberger et al. [11])**.** *There is a nonerasing morphism* $\sigma : \mathbb{N}^+ \to \{a, b\}^+$ *such that, for every* $\alpha \in \mathbb{N}^+$*,* σ *is weakly unambiguous with respect to* α*.*

This proposition directly follows from the definitions, since 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 pattern α they are applied to, satisfy $|\sigma(\alpha)| > |\alpha|$. Hence, we wish to examine the following problem:

Problem 3.15. *Let* $\alpha \in \mathbb{N}^*$ *be a pattern, and let* Σ *be an alphabet. Does there exist a length-increasing nonerasing morphism* $\sigma : \mathbb{N}^+ \to \Sigma^+$ *that is weakly unambiguous with respect to* α*?*

Our results in the present thesis 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 Σ considered. However, unlike the above mentioned results by Schneider [36] on the existence of strongly unambiguous erasing morphisms (see, in the present thesis, Corollary 3.8 in conjunction with Theorem 3.9), we can give a compact and efficiently decidable characteristic condition on Problem 3.15, 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 3.15 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 3.15, we shall give a characteristic condition.

In addition to weakly unambiguous morphisms, in the present thesis, we study the existence of strongly unambiguous 1-uniform morphisms with respect to arbitrary patterns. More formally, we wish to investigate the following problem:

Problem 3.16. *Let $\alpha \in \mathbb{N}^*$ be a pattern, and let Σ be an alphabet. Does there exists a 1-uniform morphism $\sigma : \mathbb{N}^* \to \Sigma^*$ that is strongly unambiguous with respect to α?*

There are two main reasons why we study this question: Firstly, any insight into the existence of unambiguous 1-uniform morphisms improves the construction by Freydenberger et al. [11], which provides comprehensive results on the existence of unambiguous nonerasing morphisms (see, in the present thesis, Theorem 3.1), but is based on morphisms that are often much more involved than required. This can be illustrated using our initial example pattern α_0 on page 3 of Chapter 1. Here, the unambiguous morphism σ_1 – which is not 1-uniform, but still very simple – produces a morphic image of length 8, whereas the unambiguous morphism for α_0 defined in [11] (and shown on page 11) leads to a morphic image of length 162. This substantial complexity of known unambiguous morphisms has a severe effect on the runtime of inductive inference procedures for pattern languages (as to be described in Section 3.3). Thus, any insight into the existence of uncomplex unambiguous morphisms is not only of intrinsic interest, but is also important from a more applied point of view. Secondly, as shown by $\sigma_0(\alpha_0)$ (see Chapter 1), the images under 1-uniform morphisms have a structure that is very close to that of their preimages. This is because, whenever the pattern contains more different variables than there are letters in the target alphabet, a 1-uniform morphism reduces the complexity of the preimage by mapping certain variables to the same image. Thus, such a *morphic simplification* and its potential ambiguity are a very basic phenomenon in the combinatorial theory of morphisms. Our studies shall suggest that Problem 3.16 is nevertheless a challenging question, and we shall demonstrate that it is related to a number of other concepts and problems in combinatorics on words.

3.2 Fixed points

Fixed points are a vital concept for the research on the ambiguity of morphisms, as illustrated by, e. g., the following theorem:

Theorem 3.17 (Freydenberger et al. [11]). *Let $\alpha \in \mathbb{N}^*$ be a fixed point of a nontrivial morphism, and let Σ be any alphabet. Then every nonerasing morphism $\sigma : \mathbb{N}^* \to \Sigma^*$ is not strongly unambiguous with respect to α.*

In addition, we shall study fixed points separately in Section 5.3, and therefore we discuss them in a bit more detail in the present section.

Head [15] and Hamm and Shallit [13] characterise the language of fixed points of a given morphism in the following manner:

Theorem 3.18 (Head [15]). *Let $\phi : \mathbb{N}^* \to \mathbb{N}^*$ be a morphism. Then a finite pattern $\alpha \in \mathbb{N}^*$ is a fixed point of ϕ if and only if $\alpha \in F_\phi^*$.*

In the above theorem, F_ϕ is defined as follows:

$$F_\phi = \{\phi^t(i) : \ i \in A_\phi \text{ and } t = \exp(\phi)\}$$

where,

$$A_\phi = \{i \in \mathbb{N} : \ \exists \beta, \gamma \in \mathbb{N}^* \text{ such that } \phi(i) = \beta i \gamma \text{ and } \beta\gamma \in M_\phi^*\},$$

and M_ϕ is the set of mortal variables of ϕ.

According to the definition of prolix patterns, the above theorem can be stated as follows:

Theorem 3.19 (Freydenberger et al. [11]). *A pattern $\alpha \in \mathbb{N}^+$ is prolix if and only if it is a fixed point of a nontrivial morphism $\phi : \mathbb{N}^* \to \mathbb{N}^*$.*

Theorem 3.19 is extended in the paper [34] by Reidenbach and Schneider. In this paper, the authors demonstrate that the partition of $\mathbb{N}^*$ into the set of morphically primitive patterns and the set of morphically imprimitive patterns is characteristic for various aspects related to finite words and morphisms including fixed points:

Theorem 3.20 (Reidenbach and Schneider [34]). *Let $\alpha \in \mathbb{N}^*$. The following statements are equivalent:*

1. *α is morphically primitive.*
2. *α is not a fixed point of a nontrivial morphism $\phi : \mathbb{N}^* \to \mathbb{N}^*$.*
3. *α is a succinct pattern.*
4. *There is an unambiguous injective morphism $\sigma : \mathbb{N}^* \to \{a, b\}^*$ with respect to α.*

5. *α is a shortest generator of terminal-free E-pattern languages.*

In Section 3.3, we shall address point 5 of this theorem in more detail.

Moreover, regarding the complexity of finding out whether a given finite pattern is a fixed point of a nontrivial morphism, Holub [16] presents a polynomial-time algorithm.

Although the above mentioned studies have provided, e. g., a characterisation (see Theorem 3.18) and even a polynomial-time decision procedure (see Holub [16]), many fundamental properties and the actual fabric of those words that are not fixed points of a nontrivial morphism are not fully understood. This is epitomised by the fact Billaud's Conjecture (see [4]) is still largely unresolved. By assuming that $\delta_i : \mathbb{N}^* \to \mathbb{N}^*$ is a morphism defined by $\delta_i(i) := \varepsilon$ and $\delta_i(x) := x$ for $x \in \mathbb{N} \setminus \{i\}$, Billaud's Conjecture reads as follows:

Conjecture 3.21 (Billaud [4]). *Let $\alpha \in \mathbb{N}^*$ be a pattern with $|\operatorname{var}(\alpha)| \geq 3$. If α is not a fixed point of a nontrivial morphism, then there exists an $i \in \operatorname{var}(\alpha)$ such that $\delta_i(\alpha)$ is not a fixed point of a nontrivial morphism.*

Levé and Richomme [17] prove Conjecture 3.21 for a special case, where each morphism ϕ_i (defined in the following theorem) has only one expansive variable:

Theorem 3.22 (Levé and Richomme [17]). *Let $\alpha \in \mathbb{N}^*$ be a pattern with $|\operatorname{var}(\alpha)| \geq 3$. Assume that, for each $i \in \operatorname{var}(\alpha)$, the pattern $\delta_i(\alpha)$ is a fixed point of a nontrivial morphism $\phi_i : \mathbb{N}^* \to \mathbb{N}^*$ with $|E_{\phi_i}| = 1$. Then α is a fixed point of a nontrivial morphism $\phi : \mathbb{N}^* \to \mathbb{N}^*$ with $|E_\phi| = 1$.*

Also, regarding the validity of Conjecture 3.21, Zimmermann proves the conjecture for the case that $|\operatorname{var}(\alpha)| = 3$ (see [17]). Apart from that, little is known about this problem.

In the present thesis, we shall investigate whether alphabet reductions (i. e., 1-uniform morphisms that map a given pattern to an image containing a smaller number of different variables) can be given that map a pattern which is not a fixed point of a nontrivial morphism to a pattern which is not a fixed point, either:

Problem 3.23. *Let $\alpha \in \mathbb{N}^*$ be a pattern that is not a fixed point of a nontrivial morphism. Does there exist an alphabet reduction $\phi : \mathbb{N}^* \to \mathbb{N}^*$ such that $\phi(\alpha)$ is not a fixed point of a nontrivial morphism.*

For example, let $\alpha := 1 \cdot 2 \cdot 3 \cdot 4 \cdot 1 \cdot 3 \cdot 2 \cdot 4$; if $\phi : \{1, 2, 3, 4\}^* \to \{1, 2, 3, 4\}^*$ is a morphism with $\phi(1) := 1$, $\phi(2) := 2$, $\phi(3) := 2$ and $\phi(4) := 4$, then $\phi(\alpha)$ is a fixed point of a nontrivial morphism. On the other hand, $\psi(\alpha)$, where $\psi : \{1, 2, 3, 4\}^* \to \{1, 2, 3, 4\}^*$ is a morphism given by $\psi(1) := 1$, $\psi(2) := 1$, $\psi(3) := 3$

and $\psi(4) := 4$, is not a fixed point of a nontrivial morphism. Note that ϕ and ψ are two alphabet reductions for α.

Problem 3.23 appears to be very similar to Billaud's Conjecture (Conjecture 3.21), but the latter features a different type of morphism (which, intuitively, still can be seen as an alphabet reduction). Therefore, beside studying Problem 3.23, we examine the correctness of Billaud's Conjecture for a special case not studied in the literature so far. This special case consists of those patterns α in which every variable of α occurs exactly twice.

3.3 Pattern inference

Ambiguity of morphisms has some important applications in pattern languages. As the first applications, we can refer to Reidenbach [33, 31]. In these papers, the author investigates the inferrability of E-pattern languages from positive data in Gold's learning model [12]. In this model, a class of languages is said to be *inferrable from positive data* if and only if a computable device (the so-called learning strategy) which reads an arbitrary stream of words (fully enumerating the language) converges for every language after finitely many steps, and the output exactly represents the given language. In other words, the learning strategy is expected to extract a complete description of a language from finitely many examples for this language. Reidenbach [33] proves two theorems as vital tools for examining the learnability of the class of E-pattern languages. The first one characterises the structural properties of the shortest generators of terminal-free E-pattern languages by a factorisation. This factorisation is the same as the one used by Head [15] to characterise the set of fixed points of a nontrivial morphism. Also, the same factorisation is applied in [11] to characterise those patterns with respect to which there exists an unambiguous nonerasing morphism. Therefore, [33] proves the following theorem regarding the shortest generators of E-pattern languages and succinct patterns:

Theorem 3.24 (Reidenbach [33])**.** *A pattern $\alpha \in \mathbb{N}^+$ is succinct if and only if, for every $\beta \in \mathbb{N}^+$ with $L_E(\beta) = L_E(\alpha)$, $|\beta| \geq |\alpha|$.*

Before stating the second theorem of [33], we need to give some definitions. If there exists a set T_j satisfying the conditions of the following theorem, then it is called a *telltale* for L_j (with respect to $(L_i)_{i\in\mathbb{N}_0}$).

Theorem 3.25 (Angluin [2])**.** *Let $(L_i)_{i\in\mathbb{N}_0}$ be an indexed family of nonempty recursive languages. Then $(L_i)_{i\in\mathbb{N}_0}$ is inferrable from positive data if and only if there exists an effective procedure which, for every $j \in \mathbb{N}_0$, enumerates a set T_j such that*

- T_j *is finite,*
- $T_j \subseteq L_j$, *and*
- *there does not exist a* $j' \in \mathbb{N}_0$ *with* $T_j \subseteq L_{j'} \subset L_j$.

Using the concept of telltales, the second vital tool for the examination of the learnability of $\mathrm{ePAT}_{\mathrm{tf},\Sigma}$ in case of $\Sigma \geq 3$ is the following theorem:

Theorem 3.26 (Reidenbach [33]). *Let* Σ *be an alphabet,* $|\Sigma| \geq 2$, *and let* $\alpha \in \mathrm{Pat}_{\mathrm{tf}}$ *be a succinct pattern. Let* $T_\alpha = \{w_1, w_2, ..., w_n\} \subseteq L_\Sigma(\alpha)$, $n \geq 1$. *Then* T_α *is a telltale for* $L_\Sigma(\alpha)$ *with respect to* $\mathrm{ePAT}_{\mathrm{tf},\Sigma}$ *if and only if, for every* $x \in \mathrm{var}(\alpha)$ *there exists a* $w \in T_\alpha$ *such that, for every morphism* $\sigma : \mathrm{Pat}_{\mathrm{tf}} \to \Sigma^*$ *with* $\sigma(\alpha) = w$, *there is an* $u \in \Sigma$ *with* $|\sigma(x)|_u = 1$ *and* $|\sigma(\alpha)|_u = |\alpha|_x$.

The above theorem, for each word in a given set, examines all of its generating morphisms, and, hence, it deals with the ambiguity of words with respect to a fixed pattern. From an application of Theorem 3.26, Reidenbach [33] derives that the full class of terminal-free E-pattern languages is inferrable from positive data if and only if the corresponding terminal alphabet does not consist of exactly two distinct letters:

Theorem 3.27 (Reidenbach [33]). *Let* Σ *be an alphabet. Then* $\mathrm{ePAT}_{\mathrm{tf},\Sigma}$ *is inferrable from positive data if and only if* $|\Sigma| \neq 2$.

As the second important outcome of [33], the author proves that the positive result on terminal-free E-pattern languages over alphabets with three or four distinct letters cannot be extended to the class of general E-pattern languages:

Theorem 3.28 (Reidenbach [33]). *Let* Σ *be an alphabet with* $|\Sigma| \in \{3, 4\}$. *Then* ePAT_Σ *is not inferrable from positive data.*

As an another example of applications of ambiguity in pattern languages, we can refer to [32] by Reidenbach. In this paper, as the main result, the author disproves Ohlebusch and Ukkonen's Conjecture [29] on the equivalence problem for E-pattern languages.

Chapter 4

Weakly unambiguous morphisms

In the present chapter, we address Problem 3.15 (see page 15). Hence, we investigate the existence of weakly unambiguous nonerasing morphisms. In Proposition 3.14, Freydenberger et al. [11] discuss the existence of a nonerasing weakly unambiguous morphism $\sigma : \mathbb{N}^+ \to \{a, b\}^+$ with respect to arbitrary patterns. They state that by defining $|\sigma(i)| = 1$ for every $i \in \mathbb{N}$, σ is weakly unambiguous with respect to every $\alpha \in \mathbb{N}^+$. Obviously, this is a trivial observation. However, the problem is much more interesting if σ is more general, i. e., σ is a length-increasing morphism. Hence, we investigate the following question: For an arbitrary pattern $\alpha \in \mathbb{N}^+$, is there any weakly unambiguous length-increasing morphism $\sigma : \mathbb{N}^+ \to \Sigma^+$ with respect to α, for arbitrary target alphabets Σ? Our studies on this question shall lead to some significant results. The most remarkable point is probably that, in contrast to Proposition 3.14, which is satisfied for every target alphabet Σ with at least two letters, we have to distinguish between several sizes of Σ. Indeed, we shall demonstrate that our main result holds true for all morphisms with $|\Sigma| \geq 3$, but it does not hold for morphisms with binary or unary target alphabets.

We start this chapter by giving some important definitions, and after that we investigate the existence of weakly unambiguous length-increasing morphisms $\sigma : \mathbb{N}^+ \to \Sigma^+$ with respect to an arbitrary pattern, for $|\Sigma| \geq 3$, $|\Sigma| = 2$ and $|\Sigma| = 1$ in separate sections.

4.1 Loyal neighbours

We now 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 4.1. *Let $\alpha \in \mathbb{N}^+$. For every $j \in \mathrm{var}(\alpha)$, we define the following sets:*

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

Moreover, if $\alpha = j \ldots$, then $\varepsilon \in L_j$, and if $\alpha = \ldots j$, 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 4.2. *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 4.3. *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 4.4. *For any pattern $\alpha \in \mathbb{N}^+$, $|\alpha| \geq 2$, 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.*

Note that in Definition 4.4 the notations S_α and E_α are short for "*s*table" and "(possibly) *e*xpanding", 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 chapter (such as Theorem 4.10 and Corollary 4.16 below).

The following example clarifies these definitions.

Example 4.5. *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 4.1 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 4.3, the variables 3 and 4 do not have loyal neighbours. Thus, due to Definition 4.4, $S_\alpha = \{1, 2, 5, 6, 7, 8\}$ and $E_\alpha = \{3, 4\}$.

Our subsequent remark shows that having a variable with loyal neighbours is a sufficient, but not a necessary condition for a pattern being prolix (see Section 2.3 for the definition of prolix patterns).

Proposition 4.6. *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.*

Proof. Let $i \in S_\alpha$. According to Definition 4.3, 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 Σ be an alphabet. For every nonerasing morphism $\sigma : \mathbb{N}^* \to \Sigma^*$ over α, we define a morphism $\tau : \mathbb{N}^* \to \Sigma^*$ by, for every $x \in \mathrm{var}(\alpha)$,

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

It is easily verified that $\tau(\alpha) = \sigma(\alpha)$. Consequently, there is no strongly unambiguous nonerasing morphism σ with respect to α. So, according to Theorem 3.1, α is prolix.

For the second statement of Proposition 4.6, let $\alpha := 1 \cdot 2 \cdot 2$. Referring to the definition of prolix patterns (see Section 2.3), it can be verified with little effort that α is prolix, and $S_\alpha = \emptyset$. □

4.2 Weakly unambiguous morphisms with $|\Sigma| \geq 3$

We now make use of the concepts introduced in the previous section to comprehensively solve Problem 3.15 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 chapter. The first lemma is a general combinatorial insight that can be used in the proof of Lemma 4.8 – which, in turn, is a fundamental lemma in this chapter.

Lemma 4.7. *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 $v^n := v_1 \cdot v_2 \cdot [\ldots] \cdot v_n$ with, for every j, $1 \leq j \leq n$, $v_j = v$, and let $w^n := w_1 \cdot w_2 \cdot [\ldots] \cdot w_n$ 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

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

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, $w_{n-q+1}w_{n-q+2}w_{n-q+3}[\ldots]w_n$ 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_n, which is a contradiction. □

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

Lemma 4.8. *Let $\alpha \in \mathbb{N}^+$, $|\alpha| \geq 2$, and let Σ be an alphabet. Assume that $\sigma : \mathbb{N}^+ \to \Sigma^+$ is a morphism such that, for an $i \in \mathrm{var}(\alpha)$, $|\sigma(i)| \geq 2$ 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 $\tau(i) \sqsubset \sigma(i)$.*

Proof. Assume to the contrary that there exists a $j \in \mathrm{var}(\alpha)$ 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 4.8, there exists a $j \in \mathrm{var}(\alpha)$ with $\tau(j) \neq \sigma(j)$; hence, $j \neq i$. Since σ 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 α with $\tau(x) = \varepsilon$. This is a contradiction to the fact that morphism τ is nonerasing. If $|\tau(j)| = 1$, then this contradicts $\sigma(\alpha) = \tau(\alpha)$, since $\tau(j) \neq \sigma(j)$.

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

  Since σ maps all variables except i to a word of length 1 and due to the fact that τ is nonerasing, $|\tau(\alpha)| > |\sigma(\alpha)|$, and necessarily $\tau(\alpha) \neq \sigma(\alpha)$, which contradicts the assumption of Lemma 4.8.

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

  Assume that $\alpha = \alpha_1 \cdot i_1^{p_1} \cdot \alpha_2 \cdot i_2^{p_2} \cdot [\ldots] \cdot \alpha_n \cdot i_n^{p_n} \cdot \alpha_{n+1}$ where, $\alpha_2, \alpha_3, \ldots, \alpha_n \in \mathbb{N}^+$, $\alpha_1, \alpha_{n+1} \in \mathbb{N}^*$ 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 τ 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 4.7, and therefore $\tau(i_1) \sqsubset \sigma(i_1)$. However, this contradicts $\tau(i) \not\sqsubset \sigma(i)$. 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,

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

  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, $\tau(i) \sqsubset \sigma(i)$. □

The next lemma, which directly results from Definition 4.3 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 an $i \in \mathbb{N}$ with $i^2 \sqsubseteq \alpha$.

Lemma 4.9. *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 $x_1 \in \mathrm{var}(\alpha) \setminus \{i\}$ 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 $x_2 \in \mathrm{var}(\alpha) \setminus \{i\}$ 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 α. Hence, due to Definition 4.3, 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. □

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 chapter. 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 3.1 and Proposition 4.6).

Theorem 4.10. *Let $\alpha \in \mathbb{N}^+$ with $|\alpha| \geq 2$, and let Σ be an alphabet, $|\Sigma| \geq 3$. 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 4.3 and Lemma 4.9, 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, $\sigma(i^2) = bcbc$. Assume to the contrary that there is a morphism $\tau : \mathbb{N}^+ \to \Sigma^+$ with $\tau(\alpha) = \sigma(\alpha)$ and, for some $q \in \text{var}(\alpha)$, $\tau(q) \neq \sigma(q)$. According to Lemma 4.8, $\tau(i) \neq \sigma(i)$ 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, 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 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 $\sigma : \mathbb{N}^+ \to \{a, b, c\}^+$ be the morphism defined in Case 1. We assume to the contrary that there is a morphism $\tau : \mathbb{N}^+ \to \Sigma^+$ with $\tau(\alpha) = \sigma(\alpha)$ and, for some $q \in \text{var}(\alpha)$, $\tau(q) \neq \sigma(q)$. Lemma 4.8 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 α, 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 a 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 α, 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 a 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 4.8, $\tau(i) = b$ 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 α.

We now prove the *only if* direction. Hence, we shall demonstrate 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. Since σ is length-increasing, there exists a variable i that is mapped by σ to a word of length more than 1. Let $\sigma(i) := a_1 a_2[\ldots]a_n$ with $n \geq 2$ and, for every k, $1 \leq k \leq n$, $a_k \in \Sigma$. Assume to the contrary that E_α is empty. Thus, due to Lemma 4.9, $i^2 \not\sqsubseteq \alpha$. According to Definition 4.3, 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

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

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,

$$\begin{aligned} \sigma(\alpha) \quad = \quad & \sigma(\alpha_1)\sigma(l_1)\, a_1a_2[\ldots]a_n \cdot \sigma(\alpha_2)\sigma(l_2)a_1a_2[\ldots]a_n \\ & \cdot[\ldots] \cdot \sigma(\alpha_m)\sigma(l_m)\, a_1a_2[\ldots]a_n \cdot \sigma(\alpha_{m+1})\,. \end{aligned}$$

We now define a nonerasing morphism τ such that, for every k, $1 \leq k \leq m$, $\tau(l_k) := \sigma(l_k)a_1$, $\tau(i) := a_2a_3[\ldots]a_n$ 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, 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

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

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,

$$\begin{aligned} \sigma(\alpha) &= \sigma(\alpha_1)a_1a_2[\ldots]a_n\sigma(r_1) \cdot \sigma(\alpha_2)a_1a_2[\ldots]a_n\sigma(r_2) \\ &\quad \cdot[\ldots] \cdot \sigma(\alpha_m)a_1a_2[\ldots]a_n\sigma(r_m) \cdot \sigma(\alpha_{m+1}) \,. \end{aligned}$$

If we consider the nonerasing morphism τ that satisfies, for every k, $1 \leq k \leq m$, $\tau(r_k) := a_n\sigma(r_k)$ and $\tau(i) := a_1a_2[\ldots]a_{n-1}$ and that is identical to σ for all other variables in α, then 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 4.10 and its proof, we give two examples:

Example 4.11. *Let* $\alpha := 1 \cdot 2 \cdot 3 \cdot 4 \cdot 1 \cdot 2 \cdot 3$. *According to Definition 4.4,* $S_\alpha = \{1, 2, 3\}$ *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 4.8, 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 4.12. *Let* $\alpha := 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 4 \cdot 7 \cdot 8 \cdot 3$*. According to Definition 4.3, all variables have loyal neighbours; in other words,* $E_\alpha = \emptyset$*. Hence, it follows from Theorem 4.10 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 4.10 that is based on regular expressions.

Corollary 4.13. *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* $\alpha \in (N^*iR)^+N^*$*.*

Proof. According to the definition of loyal neighbours, it is easily verified that the first statement of Corollary 4.13 is equivalent to the first case of Definition 4.3, and the second one is equivalent to the second case of Definition 4.3. 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 $i \in \mathrm{var}(\alpha)$, one of the above statements being satisfied is equivalent to $E_\alpha = \emptyset$. Hence, Corollary 4.13 directly follows from Theorem 4.10. □

We conclude this section by determining the complexity of the decision problem resulting from Theorem 4.10.

Theorem 4.14. *Let $\alpha \in \mathbb{N}^+$ with $|\alpha| \geq 2$, and let Σ be an alphabet, $|\Sigma| \geq 3$. The problem of whether there is a length-increasing morphism $\sigma : \mathbb{N}^+ \to \Sigma^+$ that is weakly unambiguous with respect to α is decidable in polynomial time.*

Proof. According to Theorem 4.10, a procedure deciding on the problem in Theorem 4.14 needs to test whether E_α is empty. This can be accomplished by first producing the sets L_i and R_i for all $i \in \mathrm{var}(\alpha)$ 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 $O(|\mathrm{var}(\alpha)|^2)$ steps. Let $\alpha := a_1 \cdot a_2 \cdot [\ldots] \cdot a_n$ with, for every i, $1 \leq i \leq n$, $a_i \in \mathrm{var}(\alpha)$. For example, the following algorithm produces E_α as an output in polynomial time:

$a_0 \leftarrow \varepsilon$
$a_{n+1} \leftarrow \varepsilon$
for $i = 1$ to n **do**
  $L_{a_i} \leftarrow L_{a_i} \cup \{a_{i-1}\}$
  $R_{a_i} \leftarrow R_{a_i} \cup \{a_{i+1}\}$
end for
for $i = 1$ to n **do**
  if $\varepsilon \in L_{a_i}$ **then**
    $E' \leftarrow E' \cup \{a_i\}$
  else
    for all j such that $j \in L_{a_i}$ **do**
      if $R_j \neq \{a_i\}$ **then**
        $E' \leftarrow E' \cup \{a_i\}$
      end if
    end for
  end if
end for
for all k such that $k \in E'$ **do**
  if $\varepsilon \in R_k$ **then**
    $E \leftarrow E \cup \{k\}$
  else

    for all j such that $j \in R_k$ **do**
      if $L_j \neq \{k\}$ **then**
        $E \leftarrow E \cup \{k\}$
      end if
    end for
  end if
end for
return E

□

Hence, the complexity of Problem 3.15 is comparable to that of the equivalent problem for strongly unambiguous nonerasing morphisms (this is a consequence of Theorem 3.1 in conjunction with Theorem 3.19 and the complexity considerations by Holub [16]). In contrast to this, deciding on the existence of strongly unambiguous erasing morphisms is NP-hard (see Corollary 3.8).

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

As we shall demonstrate below, our characterisation in Theorem 4.10 does not hold for binary target alphabets Σ (see Corollary 4.26). Hence, we have to study this case separately. We do not give a characteristic condition on the existence of weakly unambiguous length-increasing morphisms with $|\Sigma| = 2$. Instead we shall present two criteria, namely Theorems 4.17 and 4.27, that can be interpreted as sufficient conditions on the existence of such morphisms, and one criterion, namely Theorem 4.24, that is a sufficient condition on 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 4.17, which requires an extensive reasoning that is based on Lemmata 4.18, 4.19, 4.20, and 4.21, and on Proposition 4.22. 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 4.2.

Despite being restricted to ternary or larger alphabets, Theorem 4.10 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 4.15. *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.*

Proof. The first statement of Corollary 4.15 directly follows from the proof of the *only if* direction of Theorem 4.10.

For the second statement of Corollary 4.15, 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 α. In Theorem 4.24, we demonstrate that, nevertheless, every length-increasing morphism $\sigma : \mathbb{N}^+ \to \{a, b\}^+$ is ambiguous with respect to α. $\square$

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 4.10 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 4.16. *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 $i \in S_\alpha$, $|\sigma(i)| = 1$.*

Proof. Corollary 4.16 directly follows from the proof of the *only if* direction of Theorem 4.10. $\square$

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 4.15, 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 4.17. *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.*

Before we can prove Theorem 4.17, we first need to introduce some technical lemmata. Referring to Section 4.2, if $i^2 \sqsubseteq \alpha$, $i \in \mathrm{var}(\alpha)$, then there is a weakly unambiguous length-increasing morphism $\sigma : \mathbb{N}^+ \to \Sigma^+$, $|\Sigma| \geq 3$, with respect to α; this is a direct consequence of Lemma 4.9 and Theorem 4.10. We now investigate this case for $|\Sigma| = 2$.

Lemma 4.18. *Let $\alpha \in \mathbb{N}^+$ such that, for an $i \in \mathbb{N}$, $i^2 \sqsubseteq \alpha$. Let Σ be an alphabet, $|\Sigma| = 2$. There is a weakly unambiguous length-increasing morphism $\sigma : \mathbb{N}^+ \to \Sigma^+$ with respect to α that maps i to an image of length more than 1 and every variable in $\mathrm{var}(\alpha) \setminus \{i\}$ to images of length 1 if*

(I) for every occurrence of i in α, the right or left neighbour of i is i, or

(II) for every $(i' \cdot i) \sqsubseteq \alpha$ with $i' \in \mathrm{var}(\alpha) \setminus \{i\}$, $(i \cdot i') \not\sqsubseteq \alpha$.

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

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

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

with $n \in \mathbb{N}$, $\alpha_2, \alpha_3, \ldots, \alpha_n \in (\mathbb{N} \setminus \{i\})^+$, $\alpha_1, \alpha_{n+1} \in (\mathbb{N} \setminus \{i\})^*$ 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 $\sigma : \mathbb{N}^+ \to \Sigma^+$ by, for every $x \in \mathbb{N}$,

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

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

We now assume to the contrary that σ is not weakly unambiguous with respect to α. Hence, there is a morphism $\tau : \mathbb{N}^+ \to \Sigma^+$ such that $\tau(\alpha) = \sigma(\alpha)$ and, for some $q \in \mathrm{var}(\alpha)$, $\tau(q) \neq \sigma(q)$. According to Lemma 4.8, 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, $\tau(\alpha) = \tau(\alpha_1) \cdot b^{p_1} \cdot \tau(\alpha_2) \cdot b^{p_2} \cdot [\ldots] \cdot \tau(\alpha_n) \cdot b^{p_n} \cdot \tau(\alpha_{n+1})$. 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 $z \in \mathrm{var}(\alpha_1)$ with $\tau(z) = \varepsilon$; however, this contradicts the fact that τ 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

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

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

It follows from our reasoning on the above cases that the morphism τ does not exist. Hence, if Condition (I) is satisfied, then σ is weakly unambiguous with respect to α.

We now prove that Condition (II) also implies the existence of a weakly unambiguous length-increasing morphism with respect to α. According to Condition (II), $(R_i \cap L_i) \setminus \{i\} = \emptyset$. So, by considering Condition (II), we can define a morphism $\sigma : \mathbb{N}^+ \to \Sigma^+$ 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 α 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 σ is not weakly unambiguous with respect to α. Hence, there is a morphism $\tau : \mathbb{N}^+ \to \Sigma^+$ satisfying $\tau(\alpha) = \sigma(\alpha)$ and for some $q \in \text{var}(\alpha)$, $\tau(q) \neq \sigma(q)$. According to Lemma 4.8, 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

$$|\tau(\alpha_1 \cdot l_1)| \geq |\sigma(\alpha_1 \cdot l_1)|. \tag{4.1}$$

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 τ 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 must be a $z \in (\mathrm{var}(\alpha_2) \cup \{r_1, l_2\})$ 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 τ 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 τ 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 (4.1), 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 τ 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 (4.1), 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 τ 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 τ is nonerasing and σ maps every variable of α' to words of length 1, $|\tau(\alpha')| \geq |\sigma(\alpha')|$. This result satisfies Relation (4.1). Hence, we can consider one of the above cases to investigate τ when applied to the first occurrence of i in α. 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,

$$\begin{aligned} |\tau(\alpha_1 \cdot l_1 \cdot i_1 \cdot r_1 \cdot \alpha_2 \cdot l_2)| &\geq |\sigma(\alpha_1 \cdot l_1 \cdot i_1 \cdot r_1 \cdot \alpha_2 \cdot l_2)| \text{ or} \\ |\tau(\alpha_1 \cdot l_1 \cdot i_1^{p_1} \cdot r_1 \cdot \alpha_2 \cdot l_2)| &\geq |\sigma(\alpha_1 \cdot l_1 \cdot i_1^{p_1} \cdot r_1 \cdot \alpha_2 \cdot l_2)|. \end{aligned} \tag{4.2}$$

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 (4.2), Relation (4.1) 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 α, by considering Cases 1, 2, 3, and 4, which can be extended over the other occurrences of i, and due to τ being nonerasing, $|\tau(\alpha)| > |\sigma(\alpha)|$. Thus, the morphism τ does not exist. This implies that σ is weakly unambiguous with respect to α. □

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

- *there exists an occurrence of* l *in* α *such that the right neighbour and the left neighbour of this occurrence are not* e, *and*
- *there exists an occurrence of* r *in* α *such that the right neighbour and the left neighbour of this occurrence are not* e.

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

Proof. Let $\alpha := \alpha_1 \cdot e_1 \cdot \alpha_2 \cdot e_2 \cdot [\ldots] \cdot \alpha_{n-1} \cdot e_{n-1} \cdot \alpha_n$ 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 σ is not weakly unambiguous with respect to α. So, there exists a morphism τ satisfying $\tau(\alpha) = \sigma(\alpha)$ and, for some $q \in \mathrm{var}(\alpha)$, $\tau(q) \neq \sigma(q)$. Lemma 4.8 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 α, 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 τ 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 α, $\tau(\alpha) \neq \sigma(\alpha)$.

- $\tau(e_j)$ 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 τ being nonerasing and $l \cdot e \cdot r$ being a factor of α,

$$\begin{aligned} \tau(l) &= v \cdot b_1,\ v \in \{a,b\}^* \text{ or} \\ \tau(r) &= b_2 \cdot v,\ v \in \{a,b\}^*. \end{aligned} \tag{4.3}$$

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

$$\tau(z_j) = \begin{cases} b_2 \text{ or } b_2 \cdot \sigma(z_j) \text{ or } b_2 \cdot \sigma(z_j) \cdot \sigma(z_{j+1}) \text{ or} & \\ \sigma(z_j) \text{ or } \sigma(z_j) \cdot \sigma(z_{j+1}), & \text{if } j = 1, \\ & \\ b_1 \text{ or } \sigma(z_j) \cdot b_1 \text{ or } \sigma(z_{j-1}) \cdot \sigma(z_j) \cdot b_1 \text{ or} & \\ \sigma(z_j) \text{ or } \sigma(z_{j-1}) \cdot \sigma(z_j), & \text{if } j = n, \\ & \\ \sigma(z_j) \text{ or } \sigma(z_{j+1}) \text{ or } \sigma(z_{j-1}) \text{ or } \sigma(z_j) \cdot \sigma(z_{j+1}) & \\ \text{ or } \sigma(z_{j-1}) \cdot \sigma(z_j) \text{ or } \sigma(z_{j-1}) \cdot \sigma(z_j) \cdot \sigma(z_{j+1}), & \text{if } 1 < j < n. \end{cases} \tag{4.4}$$

According to the assumption of Lemma 4.19, there exist an occurrence of l and an occurrence of r in α such that the right neighbour and the left neighbour of these occurrences are not e. So, by considering Condition (4.4), $\tau(l)$ and $\tau(r)$ cannot contain any factor b. This contradicts Condition (4.3). Hence, σ is weakly unambiguous with respect to α. □

Before we continue with the next two lemmata that are required to prove Theorem 4.17, we wish to briefly clarify their subject in an informal manner: Let $\alpha \in \mathbb{N}^+$, $|\alpha| \geq 2$, and let $\sigma : \mathbb{N}^+ \to \Sigma^+$ 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 τ is a nonerasing morphism satisfying $\tau(\alpha) = \sigma(\alpha)$. According to Lemma 4.8, if there exists a $j \in \text{var}(\alpha)$ 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 4.20. *Let $\alpha \in \mathbb{N}^+$ such that $E_\alpha \neq \emptyset$. Let $e \in E_\alpha$ with $L_e \cap R_e = \emptyset$. Let $\alpha = \alpha_1 \cdot e_1 \cdot \alpha_2 \cdot e_2 \cdot [\ldots] \cdot \alpha_{n-1} \cdot e_{n-1} \cdot \alpha_n$ 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 $\sigma : \mathbb{N}^+ \to \{a, b\}^+$ 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 τ 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,

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

According to Lemma 4.8, $\tau(e) = a$ 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 τ 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 τ 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 $z \in \mathrm{var}(\alpha_1)$, $\tau(z) = \varepsilon$, which again contradicts the fact that τ 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, $a_{k+1}, \ldots, a_{n-1}$ instead of a_k, $a_{k-1}, \ldots, a_1$ – 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). □

Lemma 4.20 and its proof enable us in the following lemma to investigate the morphism τ, which is defined in Lemma 4.20, for the variables occurring between two consecutive occurrences of e.

Lemma 4.21. *Let $\alpha \in \mathbb{N}^+$ such that $E_\alpha \neq \emptyset$. Let $e \in E_\alpha$ with $L_e \cap R_e = \emptyset$. Let $\alpha := \alpha_1 \cdot e_1 \cdot x_1 \cdot x_2 \cdot [\ldots] \cdot x_n \cdot e_2 \cdot \alpha_2$, $\alpha_1, \alpha_2 \in \mathbb{N}^*$, $e_1 = e_2 = e$, $n > 1$, and for every*

j, $1 \leq j \leq n$, $x_j \in \mathrm{var}(\alpha) \setminus \{e\}$. Let $\sigma : \mathbb{N}^+ \to \{a, b\}^+$ 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 $x \in \mathrm{var}(\alpha) \setminus (\{e\} \cup L_e \cup R_e)$. Then, for every morphism τ with $\tau(\alpha) = \sigma(\alpha)$ and, for some $j \in \mathrm{var}(\alpha)$, $\tau(j) \neq \sigma(j)$, 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 $\tau(x_n) = v \cdot \sigma(x_n)$, $v \in \{\sigma(x_{n-1}), \varepsilon\}$.

(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 $\tau(x_1) = \sigma(x_1) \cdot v$, $v \in \{\varepsilon, \sigma(x_2)\}$.

Proof. Assume that $\tau(\alpha) = \sigma(\alpha)$ and, for some $j \in \mathrm{var}(\alpha)$, $\tau(j) \neq \sigma(j)$. According to Lemmata 4.8 and 4.20, regardless of the number of occurrences of e in α_1 and α_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, $|\tau(x_1 \cdot x_2 \cdot [\ldots] \cdot x_n)| = n + 1$. So, as τ is nonerasing, $|\tau(x_i)| \leq 2$, $1 \leq i \leq n$.

  Hence, due to $\tau(\alpha) = \sigma(\alpha)$ and τ being nonerasing, it is required to define τ for the variables $x_1, x_2, \ldots, x_n$ 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, $|\tau(x_1 \cdot x_2 \cdot [\ldots] \cdot x_n)| = n + 1$, which, as τ is nonerasing, implies $|\tau(x_i)| \leq 2$, $1 \leq i \leq n$.

Therefore, since $\tau(\alpha) = \sigma(\alpha)$ and τ is nonerasing, τ needs to be defined for the variables $x_1, x_2, \ldots, x_n$ 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 4.17.

Proposition 4.22. *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 $\sigma : \mathbb{N}^+ \to \{a, b\}^+$ that is weakly unambiguous with respect to α.*

Proof. According to Definition 4.3, 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 the 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 α:

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

where $n := |\alpha|_e$, $\alpha_1, \alpha_2, \ldots, \alpha_{n+1} \in (\mathbb{N} \setminus \{e\})^*$, 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 $s' \in \mathrm{var}(\alpha) \cup \{\varepsilon\}$.

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 $\sigma : \mathbb{N}^+ \to \{a, b\}^+$, given by $\sigma(e) := aa$ and, for every $x \in \mathrm{var}(\alpha) \setminus \{e\}$, $\sigma(x) := b$. Assume to the contrary that there exists a morphism $\tau : \mathbb{N}^+ \to \{a, b\}^+$ satisfying $\tau(\alpha) = \sigma(\alpha)$ and, for a $q \in \mathrm{var}(\alpha)$, $\tau(q) \neq \sigma(q)$. According to Lemma 4.8, 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 α that $|\alpha|_s > |\alpha|_e$, and therefore $|\tau(\alpha)|_a > 2|\alpha|_e = |\sigma(\alpha)|_a$. This contradicts the assumption $\tau(\alpha) = \sigma(\alpha)$. □

Based on the preparatory work in Lemmata 4.18, 4.19, 4.20, 4.21 and Proposition 4.22, we can now verify Theorem 4.17:

Proof of Theorem 4.17. 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 α.

According to Proposition 4.22, since there is no weakly unambiguous length-increasing morphism σ with respect to α, 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 $x \in \mathrm{var}(\alpha) \setminus \{e\}$ 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 $x \in \mathrm{var}(\alpha) \setminus \{e\}$, $x \cdot e$ or $e \cdot x$ is not a factor of α, Case 1 satisfies the conditions of Lemma 4.18. Hence, there is a weakly unambiguous length-increasing morphism $\sigma : \mathbb{N}^+ \to \Sigma^+$ with respect to α. This contradicts the condition of Theorem 4.17, namely that there is no weakly unambiguous morphism σ with respect to α.

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 $x \in \mathrm{var}(\alpha) \setminus \{e\}$ 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 α; 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 α, and in Part (b) we assume that $l \cdot e \cdot r$ is not a factor of α.

Part (a) $l \cdot e \cdot r \sqsubseteq \alpha$.
We define a morphism $\sigma : \mathbb{N}^+ \to \{a,b\}^+$ by

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

According to Lemma 4.19, σ is weakly unambiguous with respect to α, which contradicts the condition of Theorem 4.17.

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 $\alpha = \ldots \cdot k \cdot l \cdot k' \cdot \ldots \cdot l \cdot e$ or $\alpha = e \cdot r \cdot \ldots \cdot k \cdot r \cdot k' \cdot \ldots$, $k, k' \in \mathrm{var}(\alpha) \setminus \{e\}$. We define a morphism $\sigma : \mathbb{N}^+ \to \{a,b\}^+$ by

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

Using Lemma 4.8, it can be easily verified that σ is weakly unambiguous with respect to α, 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 $\alpha = \ldots \cdot l \cdot e \cdot \ldots \cdot e \cdot r \cdot \ldots$. As mentioned above, there is no variable $x \in \mathrm{var}(\alpha)$ with $x \in L_e$ and $x \in R_e$. As a result, we can define a morphism σ by

$$\sigma(x) := \begin{cases} ab, & x = e, \\ b, & x \in L_e, \\ a, & x \in R_e. \end{cases} \tag{4.5}$$

For the other variables, we shall define the morphism σ later. Before we do this, we shall establish some insights into the structure of α. According to Definition (4.5), $\sigma(l) = b$ and $\sigma(r) = a$. Also, due to the condition of Theorem 4.17, there exists a nonerasing morphism τ with $\tau(\alpha) = \sigma(\alpha)$ and, for some $q \in \mathrm{var}(\alpha)$, $\tau(q) \neq \sigma(q)$. Moreover, since $\sigma(e)$ is the only image of length more than 1, Lemma 4.8 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 4.21 and $\tau(\alpha) = \sigma(\alpha)$ imply that $\tau(r) = b \cdot v$, $v \in \{\varepsilon, a\}$. However, according to Lemma 4.20, 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 4.21 and $\tau(\alpha) = \sigma(\alpha)$ imply that $\tau(l) = v \cdot a$, $v \in \{\varepsilon, b\}$. Due to Lemma 4.20, 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, without loss of generality regarding the different possibilities of the positions of l and r in α, let

$$\alpha := \alpha_1 \cdot e \cdot x_1 \cdot x_2 \cdot [\ldots] \cdot x_n \cdot r \cdot \alpha_2 \cdot l \cdot z_1 \cdot z_2 \cdot [\ldots] \cdot z_m \cdot e \cdot \alpha_3, \tag{4.6}$$

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 α, $y_1 \neq e$. Furthermore, if we assume that $y_1 = r$, then $rr \sqsubseteq \alpha$ and, in accordance with Lemma 4.9, $r \in E_\alpha$. Consequently, according to Case 1, the assumption of $y_1 = r$ leads to a contradiction. Hence, $y_1 \neq r$.

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

$i \leftarrow n$
while $\sigma(x_i) = b$ **do**
    $i \leftarrow i - 1$
end while
if $\sigma(x_i) = \mathit{null}$ **then**
    $\sigma(x_i) \leftarrow a$
end if
$i \leftarrow 1$
while $\sigma(z_i) = a$ **do**
    $i \leftarrow i + 1$
end while

if $\sigma(z_i) = null$ **then**
  $\sigma(z_i) \leftarrow b$
end if
if $\alpha_2 \neq \varepsilon$ and $\sigma(y_1) = null$ **then**
  $\sigma(y_1) \leftarrow b$
end if
for all $x \in \text{var}(\alpha)$ **do**
  if $\sigma(x) = null$ **then**
    $\sigma(x) \leftarrow a$
  end if
end for

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

Claim. The morphism σ is weakly unambiguous with respect to α.

Proof (Claim). We assume to the contrary that there exists a nonerasing morphism τ satisfying $\tau(\alpha) = \sigma(\alpha)$ and, for some $q \in \text{var}(\alpha)$, $\tau(q) \neq \sigma(q)$. It follows from Lemmata 4.8 and 4.20, 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 4.21,

$$\begin{gathered}\tau(e) = a \text{ implies that } \tau(r) = b \cdot v,\ v \in \{\sigma(r), \varepsilon\},\ \text{and} \\ b \text{ is a suffix of } \tau(l)\end{gathered} \tag{4.7}$$

and

$$\begin{gathered}\tau(e) = b \text{ implies that } \tau(l) = v \cdot a,\ v \in \{\sigma(l), \varepsilon\},\ \text{and} \\ a \text{ is a prefix of } \tau(r),\end{gathered} \tag{4.8}$$

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 α. Now, we consider the following cases:

- $\tau(e) = a$. As a result of Implication (4.7), b is a prefix of $\tau(r)$. We consider the factor $e \cdot x_1 \cdot x_2 \cdot [\ldots] \cdot x_n \cdot r$ of α. According to Lemma 4.21, $\tau(r) = \sigma(r)$ 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 4.21, $\tau(x_n)$ 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 4.20 and Lemma 4.21 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 σ 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 (4.8), 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 α. According to Lemma 4.21, $\tau(l) = \sigma(l)$ 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 σ and τ 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 4.20 and Lemma 4.21 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 4.21, $\tau(z_1) = \sigma(z_2)$, or, in other words, $\tau(z_1)$ is located at the position of $\sigma(z_2)$. 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 σ for all the variables $z_1, x_2, \dots, z_{m-1}$ 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 4.21, as τ 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 α. 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 4.21 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 (4.8). 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 α, 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 σ being weakly unambiguous with respect to α. Consequently, we cannot consider $\tau(e) = b$.

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

The above claim is a direct contradiction to the assumption of Theorem 4.17. In order to conclude our reasoning on Case 2.2, it is necessary to mention that, instead of Factorisation (4.6) of α, we can define α such that the variable l is located to the left of the position of r in α. More precisely, we can consider

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

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 (4.6) 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 $\sigma : \mathbb{N}^+ \to \{a, b\}^+$ 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 4.8, if σ is not weakly unambiguous, then there exists a nonerasing morphism $\tau : \mathbb{N}^+ \to \{a, b\}^+$ 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, σ is weakly unambiguous with respect to α. This contradicts the condition of Theorem 4.17, namely that there is no weakly unambiguous length-increasing morphism σ with respect to α. □

Theorem 4.17 (when compared to Theorem 4.10) 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 4.17 does not describe a sufficient condition for the non-existence of weakly unambiguous length-increasing morphisms in the case of $|\Sigma| = 2$; this is easily demonstrated by the pattern $1 \cdot 2 \cdot 1$ and further illustrated by Example 4.28.

As can be concluded from Example 4.5 and Theorem 4.10, there is a weakly unambiguous length-increasing morphism $\sigma : \mathbb{N}^+ \to \Sigma^+$, $|\Sigma| \geq 3$, with respect to $\alpha = 1 \cdot 2 \cdot \mathbf{3} \cdot \mathbf{4} \cdot 5 \cdot 6 \cdot \mathbf{4} \cdot \mathbf{3} \cdot 7 \cdot 8$, and 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 prove this theorem, we need the following lemma.

Lemma 4.23. *Let Σ be an alphabet, $|\Sigma| = 2$, and let $\sigma : \mathbb{N}^+ \to \Sigma^+$ be a morphism. For all $x_1, x_2 \in \mathbb{N}$, there exist $a_1, a_2 \in \Sigma$ 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 4.23 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 $a_1a_2 \sqsubseteq \sigma(x_1 \cdot x_2)$. Furthermore, since $\sigma(x_1) = a_2 \cdots a_1$, there must be a factor a_2a_1 in $\sigma(x_1)$, which directly implies that a_2a_1 is also a factor of $\sigma(x_2 \cdot x_1)$. Thus, Lemma 4.23 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_1a_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_2a_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_2a_1 and $\sigma(x_2)$ contains a factor a_1a_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 $\sigma : \mathbb{N}^+ \to \Sigma^+$, $|\Sigma| = 2$. According to Theorem 4.17, 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 4.24. *Let $\alpha \in \mathbb{N}^+$ satisfying $E_\alpha \neq \emptyset$. Let Σ be an alphabet, $|\Sigma| = 2$. There is no weakly unambiguous length-increasing morphism $\sigma : \mathbb{N}^+ \to \Sigma^+$ with respect to α if all of the following conditions hold:*

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 $L_{s'} = \{e\}$.*

Proof. We prove that there is no weakly unambiguous length-increasing morphism $\sigma : \mathbb{N}^+ \to \Sigma^+$, $|\Sigma| = 2$, with respect to α. This means that, for every morphism σ, there exists a morphism $\tau : \mathbb{N}^+ \to \Sigma^+$ satisfying $\tau(\alpha) = \sigma(\alpha)$ and, for some $q \in \text{var}(\alpha)$, $\tau(q) \neq \sigma(q)$. According to Corollary 4.16, if there exists a $j \in S_\alpha$ with $|\sigma(j)| > 1$, then σ is not weakly unambiguous with respect to α. Consequently, we can safely restrict our considerations to the set E_α, and we can assume that, for every $j \in S_\alpha$, $|\sigma(j)| = 1$. Hence, we choose an arbitrary variable e_1 from E_α, and we assume that $|\sigma(e_1)| > 1$. According to the conditions of Theorem 4.24, 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 α. 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 4.23, there exists a factor a_1a_2, $a_1, a_2 \in \Sigma$, such that $\sigma(e_1e_2) = u \cdot a_1a_2 \cdot v$, $u, v \in \Sigma^*$, $\sigma(e_2e_1) = u' \cdot a_2a_1 \cdot v'$, $u', v' \in \Sigma^*$, 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 τ 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 α 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 τ 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, σ is not weakly unambiguous with respect to the pattern α. Since the variable e_1 is an arbitrary variable of E_α, we can conclude that there is no weakly unambiguous length-increasing morphism σ with respect to α. □

In order to illustrate Theorem 4.24, we consider a few examples:

Example 4.25. *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 4.4, E_α, E_β and E_γ are nonempty (the respective variables are typeset in bold face). Since the patterns satisfy the conditions of Theorem 4.24, there is no length-increasing morphism $\sigma : \mathbb{N}^+ \to \Sigma^+$ that is weakly unambiguous with respect to them (provided that $|\Sigma| = 2$).

Theorem 4.24 and Example 4.25 directly imply the insight mentioned above that Theorem 4.10 does not hold for binary alphabets Σ:

Corollary 4.26. *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 length-increasing morphism $\sigma : \mathbb{N}^+ \to \Sigma^+$ that is weakly unambiguous with respect to α.*

In contrast to the previous theorem, 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 4.27. *Let $\alpha \in \mathbb{N}^+$, and let Σ 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 $\sigma : \mathbb{N}^+ \to \Sigma^+$ 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 σ *is weakly unambiguous with respect to* α.

Proof. Let $\Sigma := \{a, b\}$. We initially discuss the case where $i \cdot e \cdot e' \sqsubset \alpha$ and $i \cdot e' \cdot e \sqsubset \alpha$ are satisfied. We define a morphism $\sigma : \mathbb{N}^+ \to \Sigma^+$ such that the conditions of Theorem 4.27 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 σ is not weakly unambiguous with respect to α. Consequently, there is a nonerasing morphism $\tau : \mathbb{N}^+ \to \Sigma^+$ 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 $i \cdot e \cdot e' \sqsubset \alpha$, this implies that $\tau(i) = \alpha_1 b$, $\alpha_1 \in \Sigma^*$, and $\tau(e') = b\alpha_2$, $\alpha_2 \in \Sigma^*$. Due to $i \cdot e' \cdot e \sqsubset \alpha$, $\tau(i \cdot e' \cdot e) \sqsubset \tau(\alpha)$. However, $\tau(i \cdot e' \cdot e) = \alpha_1 b \cdot b\alpha_2 \cdot a$ and, this means that $b^2 \sqsubset \tau(\alpha)$, which contradicts $\tau(\alpha) = \sigma(\alpha)$.
  - $\tau(e) = b$. An analogous reasoning to that in the previous case leads to $a^2 \sqsubset \tau(\alpha)$, 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 τ is nonerasing, $|\tau(\alpha)| > |\sigma(\alpha)|$. This contradicts $\tau(\alpha) = \sigma(\alpha)$.
- $|\tau(e)| \geq 4$ or $|\tau(e')| \geq 4$. Since τ 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 $i \cdot e \cdot e' \sqsubset \alpha$ imply that $\tau(i) = \alpha_1 b$, $\alpha_1 \in \Sigma^*$, and $\tau(e') = b\alpha_2$, $\alpha_2 \in \Sigma^*$. Due to $i \cdot e' \cdot e \sqsubset \alpha$, $\tau(i \cdot e' \cdot e) \sqsubset \tau(\alpha)$. However, $\tau(i \cdot e' \cdot e) = \alpha_1 b \cdot b\alpha_2 \cdot aba$, and this means that $b^2 \sqsubset \tau(\alpha)$, which contradicts $\tau(\alpha) = \sigma(\alpha)$. If $\tau(e) = bab$, then the conditions $\tau(\alpha) = \sigma(\alpha)$ and $i \cdot e \cdot e' \sqsubset \alpha$ imply that $\tau(i) = \alpha_1 a$, $\alpha_1 \in \Sigma^*$, and $\tau(e') = a\alpha_2$, $\alpha_2 \in \Sigma^*$. Due to $i \cdot e' \cdot e \sqsubset \alpha$, $\tau(i \cdot e' \cdot e) \sqsubset \tau(\alpha)$. However, $\tau(i \cdot e' \cdot e) = \alpha_1 a \cdot a\alpha_2 \cdot bab$, and this means that $a^2 \sqsubset \tau(\alpha)$, 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 $j \in \mathrm{var}(\alpha) \setminus \{e, e'\}$, $|\tau(j)| = 1$. 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 $k \in \mathrm{var}(\alpha) \setminus \{e, e'\}$ with $\tau(k) = \varepsilon$, since otherwise $\tau(\alpha) \neq \sigma(\alpha)$. This contradicts the fact that τ 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 τ with $\tau(\alpha) = \sigma(\alpha)$ and, for some $q \in \mathrm{var}(\alpha)$, $\tau(q) \neq \sigma(q)$. Consequently, σ is weakly unambiguous with respect to α. Using the same reasoning as above, it can be demonstrated that Theorem 4.27 holds true for the case that $e \cdot e' \cdot i \sqsubset \alpha$ and $e' \cdot e \cdot i \sqsubset \alpha$. □

The main difference between Theorem 4.27 and Theorem 4.24 is that those patterns α being examined in Theorem 4.27 do not satisfy Condition 3 of Theorem 4.24. 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 4.27, we now consider some examples. In contrast to Example 4.25, 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 4.28. *We define a morphism $\sigma : \mathbb{N}^+ \to \{a, b\}^+$ for the given patterns α (where the factors featured by Theorem 4.27 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 4.27, it can be easily verified that, in all the above cases, σ is length-increasing and weakly unambiguous with respect to α.

The patterns in Example 4.28 further illustrate that the converse of Theorem 4.17 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 4.24 and 4.27, we expect that it is an extremely challenging task to find an equivalent to the characterisation in Theorem 4.10 for the binary case. From our understanding of the matter, we can therefore merely give the following conjecture on the decidability of Problem 3.15 for binary target alphabets.

Conjecture 4.29. *Let $\alpha \in \mathbb{N}^+$ with $|\alpha| \geq 2$, and let Σ be an alphabet, $|\Sigma| = 2$. 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 4.16, 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 $i \in \text{var}(\alpha)$, $\tau(i) \neq \sigma(i)$. 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.

4.4 Weakly unambiguous morphisms with $|\Sigma| = 1$

It is not surprising that most of our considerations in the previous sections of this chapter are not applicable to morphisms with a unary target alphabet. On the other hand, Corollaries 4.15 and 4.16 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 4.30. *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 weakly 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 3.15 for morphisms with unary target alphabets.

Theorem 4.31. *Let* $\alpha \in \mathbb{N}^+$, $\mathrm{var}(\alpha) = \{1, 2, 3, \ldots, n\}$. *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* $n_1, n_2, \ldots, n_{i-1}, n_{i+1}, \ldots, n_n \in \mathbb{N} \cup \{0\}$, *such that*

$$|\alpha|_i = \sum_{j \in \{1,2,\ldots,n\} \setminus \{i\}} n_j |\alpha|_j. \tag{4.9}$$

Proof. We begin with the *if* direction. Assume that, for every $i \in \mathrm{var}(\alpha)$, Equation (4.9) is satisfied. Also, assume that $\sigma : \mathbb{N}^+ \to \{a\}^+$ is an arbitrary length-increasing morphism with $|\sigma(i')| > 1$, $i' \in \mathrm{var}(\alpha)$. This means that $\sigma(i') = a^m$, $m \geq 2$ and, hence,

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

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

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

We need to demonstrate that

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

This is equivalent to:

$$\begin{aligned} |\alpha|_{i'} \quad = \quad & |\alpha|_1(|\tau(1)| - |\sigma(1)|) + |\alpha|_2(|\tau(2)| - |\sigma(2)|) + [\ldots] + \\ & |\alpha|_{i'-1}(|\tau(i'-1)| - |\sigma(i'-1)|) + |\alpha|_{i'+1}(|\tau(i'+1)| - |\sigma(i'+1)|) \\ & +[\ldots] + |\alpha|_n(|\tau(n)| - |\sigma(n)|). \end{aligned} \tag{4.10}$$

According to Equation (4.9), for Equation (4.10) to be satisfied, we define the morphism τ, for every $j \in \mathrm{var}(\alpha) \setminus \{i'\}$ such that $|\tau(j)| - |\sigma(j)| = n_j$, and this can be achieved by defining $\tau(j) := a^{(n_j + |\sigma(j)|)}$. Consequently, τ is given by

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

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

We now prove the *only if* direction. So, we assume that there is no weakly unambiguous length-increasing morphism $\sigma : \mathbb{N}^+ \to \{a\}^+$ with respect to α. Let i be an arbitrary variable of α. We define the morphism σ for the variables $x \in \mathrm{var}(\alpha)$ 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 τ satisfying $\tau(\alpha) = \sigma(\alpha)$ and, for some variables $q \in \mathrm{var}(\alpha)$, $\tau(q) \neq \sigma(q)$. According to Lemma 4.8, $\tau(i) \sqsubset \sigma(i)$ must be satisfied. Thus, $\tau(i) = a$. Consequently,

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

and

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

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

$$\begin{aligned} |\alpha|_1(|\tau(1)| - |\sigma(1)|) + |\alpha|_2(|\tau(2)| - |\sigma(2)|) + [\ldots] + & \\ (-|\alpha_i|) + [\ldots] + |\alpha|_n(|\tau(n)| - |\sigma(n)|) &= 0. \end{aligned}$$

This leads to

$$\begin{aligned} |\alpha|_i = {} & |\alpha|_1|(\tau(1)| - |\sigma(1)|) + |\alpha|_2(|\tau(2)| - |\sigma(2)|) \\ & +[\ldots] + |\alpha|_{i-1}(|\tau(i-1)| - |\sigma(i-1)|) + |\alpha|_{i+1}(|\tau(i+1)| - |\sigma(i+1)|) \\ & +[\ldots] + |\alpha|_n(|\tau(n)| - |\sigma(n)|). \end{aligned} \tag{4.11}$$

Consequently, for any variable $i \in \mathrm{var}(\alpha)$, there exists $n_1, n_2, \ldots, n_n \in \mathbb{N} \cup \{0\}$, such that Equation (4.9) is satisfied. □

Hence, we are able to provide a result on unary alphabets that is as strong as our result in Theorem 4.10 on ternary and larger alphabets. However, while

Theorem 4.10 needs to consider the order of variables in the patterns, it is evident that Theorem 4.31 can exclusively refer to their numbers of occurrences.

Chapter 5

Strongly unambiguous 1-uniform morphisms

In the present chapter, we investigate Problems 3.16 and 3.23 (see pages 16 and 18, respectively). To this end, we first study Problem 3.16, which deals with the existence of strongly unambiguous 1-uniform morphisms. Our analysis can make use of Theorem 3.17, which implies that an answer to Problem 3.16 is trivial for those patterns that are fixed points of nontrivial morphisms. Hence, in our investigation of Problem 3.16, we can safely restrict ourselves to those patterns that are not fixed points. Moreover, to investigate Problem 3.16, we consider two different settings: in Section 5.1 we assume that the size of Σ does not depend on the number of variables in the pattern, and in Section 5.2 we allow Σ to be arbitrarily chosen, subject to the number of different variables in the pattern α (provided that $|\Sigma| < |\operatorname{var}(\alpha)|$ remains satisfied).

Subsequent to this, in Section 5.3, we examine Problem 3.23 and its relation to Conjecture 3.21 (Billaud's Conjecture). Finally, we prove the correctness of Conjecture 3.21 for a special case.

Note that, in this chapter, instead of using the term "strongly unambiguous", we use "unambiguous" for short. Furthermore, as already used above and in contrast to Chapter 4, we distinguish between these patterns that are a fixed point and those that are not a fixed point , instead of using the equivalent partition into prolix and succinct patterns (see Theorem 3.19 for the equivalence of the concepts). This is because, we shall expand our studies on the ambiguity of morphisms in Sections 5.1 and 5.2 to a discussion of questions that make use of the definition of fixed points (see Section 5.3).

5.1 Fixed target alphabets

In the present section, we describe a number of conditions on the existence of unambiguous 1-uniform morphisms $\sigma : \mathbb{N}^* \to \Sigma^*$ with a *fixed* target alphabet Σ, i. e., the size of Σ does not depend on the number of variables occurring in a given pattern. While Theorem 3.1 demonstrates that the set of patterns with an unambiguous *nonerasing* morphisms is independent of the size of Σ (provided that $|\Sigma| \geq 2$), all patterns $\alpha_m := 1 \cdot 1 \cdot 2 \cdot 2 \cdot [\ldots] \cdot m \cdot m$ with $m \geq 4$ do not have an unambiguous *1-uniform* morphism $\sigma : \mathbb{N}^* \to \Sigma^*$ for binary alphabets Σ. However, such morphisms can be given for ternary (and, thus, larger) alphabets:

Theorem 5.1. *Let $m \in \mathbb{N}$, $m \geq 4$, let Σ be an alphabet, and let $\alpha_m := 1 \cdot 1 \cdot 2 \cdot 2 \cdot [\ldots] \cdot m \cdot m$. There exists a 1-uniform morphism $\sigma : \mathbb{N}^* \to \Sigma^*$ that is unambiguous with respect to α_m if and only if $|\Sigma| \geq 3$.*

Proof. Since squares cannot be avoided over unary and binary alphabets, it can be shown with very limited effort that there is no unambiguous 1-uniform morphism $\sigma : \mathbb{N}^* \to \Sigma^*$ with respect to any α_m if Σ does not contain at least three letters.

Thue [38] gives an infinite square-free word over a ternary alphabet. Let this word be w. Thus,

$$w = abcacbabcbacabcacbaca \cdots .$$

We define a word w' by repeating every letter of w twice. Consequently,

$$w' = aabbccaaccbbaabbccbbaaccaabbccaaccbbaaccaa \cdots .$$

We now define a 1-uniform morphism $\sigma : \mathbb{N}^* \to \{a, b, c\}^*$ such that $\sigma(\alpha_m)$ is a prefix of w'. Since w is square-free, the only square factors of w' are aa, bb and cc. Hence, it can be easily verified that σ is unambiguous with respect to α_m. □

Thus – and just as for the equivalent problem on unambiguous *erasing* morphisms (see Theorem 3.9) – any characteristic condition on the existence of unambiguous 1-uniform morphisms needs to incorporate the size of Σ, which suggests that such criteria might be involved. This is further strengthened by the following result, which establishes an analogous phenomenon for the transition from $|\Sigma| = 3$ to $|\Sigma| \geq 4$:

Theorem 5.2. *There exists an $\alpha \in \mathbb{N}^+$ such that*

- *every 1-uniform morphism $\sigma : \mathbb{N}^* \to \{a, b, c\}^*$ is not unambiguous with respect to α and*
- *there is a 1-uniform morphism $\sigma : \mathbb{N}^* \to \{a, b, c, d\}^*$ that is unambiguous with respect to α.*

Proof. Let $\alpha := 1 \cdot 2^2 \cdot 3^2 \cdot 4^2 \cdot 5^2 \cdot 6^2 \cdot 1 \cdot 2^2 \cdot 3^2 \cdot 4^2 \cdot 5^2 \cdot 6^2 \cdot 2^2$. We begin with the first statement of Theorem 5.2: Assume to the contrary that there is a 1-uniform morphism $\sigma : \mathbb{N}^* \to \{a, b, c\}^*$ that is unambiguous with respect to α. If $\sigma(3 \cdot 4 \cdot 5 \cdot 6)$ contains at most two different symbols, then there exists a morphism τ with $\tau(\alpha) = \sigma(\alpha)$ and $\tau(x) = \varepsilon$ for an $x \in \{3, 4, 5, 6\}$ (since $\sigma(3 \cdot 4 \cdot 5 \cdot 6)$ then necessarily contains a square), which is a contradiction.

Hence, $\sigma(3 \cdot 4 \cdot 5 \cdot 6)$ must be a word over $\{a, b, c\}$. This implies that there is an $x \in \{3, 4, 5, 6\}$ satisfying $\sigma(x) = \sigma(2)$. We now consider the morphism $\tau : \mathbb{N}^* \to \{a, b, c\}^*$ given by

$$\tau(i) := \begin{cases} \sigma(1 \cdot 2^2 \cdot [\ldots] \cdot (x-1)^2), & i = 1, \\ \sigma(i), & i = 2 \text{ or } x + 1 \leq i \leq 6, \\ \varepsilon, & 3 \leq i \leq x. \end{cases}$$

Hence, $\tau(\alpha) = \sigma(\alpha)$ and $\tau(x) = \varepsilon \neq \sigma(x)$. Thus, σ is not unambiguous with respect to α, which is a contradiction.

Regarding the second statement of Theorem 5.2, we define a morphism $\sigma_{a,b,c,d} : \mathbb{N}^* \to \{a, b, c, d\}^*$ by

$$\sigma_{a,b,c,d}(i) := \begin{cases} a, & i \in \{1, 4, 6\}, \\ b, & i = 2, \\ c, & i = 3, \\ d, & i = 5. \end{cases}$$

An exhaustive search demonstrates that there is no morphism $\tau : \mathbb{N}^* \to \{a, b, c, d\}^*$ with $\tau(\alpha) = \sigma_{a,b,c,d}(\alpha)$ and $\tau(x) \neq \sigma_{a,b,c,d}(x)$ for an $x \in \text{var}(\alpha)$. Thus, $\sigma_{a,b,c,d}$ is unambiguous with respect to α. □

Using a different construction, Reidenbach [27] demonstrates that any characteristic condition on those patterns that have unambiguous 1-uniform morphisms needs to incorporate the target alphabet size:

Theorem 5.3 (Reidenbach [27]). *For every $k \in \mathbb{N}$ and for every alphabet Σ with $|\Sigma| \leq k$, there exist an $\alpha_k \in \mathbb{N}^+$ and an alphabet Σ' with $k < |\Sigma'| < |\text{var}(\alpha_k)|$ such that*

- *there is no 1-uniform morphism $\sigma : \mathbb{N}^* \to \Sigma^*$ that is unambiguous with respect to α_k and*
- *there is a 1-uniform morphism $\sigma : \mathbb{N}^* \to \Sigma'^*$ that is unambiguous with respect to α_k.*

Therefore, due to Theorems 5.1, 5.2 and 5.3, our further results regarding the existence of unambiguous 1-uniform morphisms are restricted to *sufficient* conditions.

Our first criterion is based on (un)avoidable patterns and is, thus, related to the above-mentioned property of the patterns α_m:

Theorem 5.4. *Let $n \in \mathbb{N}$, $\beta := r_1 \cdot r_2 \cdot [\ldots] \cdot r_{\lceil n/2 \rceil}$ and $\alpha := 1^{r_1} \cdot 2^{r_1} \cdot 3^{r_2} \cdot 4^{r_2} \cdot [\ldots] \cdot n^{(r_{\lceil n/2 \rceil})}$ with $r_i \geq 2$ for every i, $1 \leq i \leq \lceil n/2 \rceil$. If β is square-free, then there exists a 1-uniform morphism $\sigma : \mathbb{N}^* \to \{a, b\}^*$ that is unambiguous with respect to α.*

Proof. For any $n \in \mathbb{N}$, let $A := \{1, 2, 3, \ldots, n\}$. For every $q \in A$, we define a 1-uniform morphism σ by $\sigma(q) := a$ if q is odd and $\sigma(q) := b$ if q is even. Thus, $\sigma(\alpha) = a^{r_1} b^{r_1} \cdot a^{r_2} b^{r_2} \cdot [\ldots] \cdot x^{(r_{\lceil n/2 \rceil})}$ with $x \in \{a, b\}$. We claim that σ is unambiguous with respect to α if β is square-free. Assume to the contrary that σ is ambiguous. Consequently, there is a morphism $\tau : A^* \to \{a, b\}^*$ with $\tau(\alpha) = \sigma(\alpha)$ and, for an $i \in A$, $\tau(i) \neq \sigma(i)$. Without loss of generality, we assume that for any $i' < i$, $\tau(i') = \sigma(i')$. Thus, we can define $u \in \{a, b\}^*$ such that $\sigma(\alpha) = u \cdot \tau(i) \cdot \cdots$. Let $B := \{r_1, r_2, \ldots, r_{\lceil n/2 \rceil}\}$ and assume that y is the maximum number in B.

Claim. $\sigma(\alpha)$ does not contain any factor v^2 such that $v \in \{a^p b^p | p \in B\}^+$.

Proof (Claim). Since β is square-free, every subpattern of it is square-free. So, by considering the structure of $\sigma(\alpha)$, this implies that $\sigma(\alpha)$ does not contain any factor v^2 such that $v \in \{a^p b^p | p \in B\}^+$. □(Claim)

Let, $\tau(i) = a^j \cdot b^k \cdot v \cdot a^l \cdot b^m$, $v \in \{a^p b^p | p \in B\}^*$, $0 \leq j \leq y$, $0 \leq k \leq y$, $0 \leq l \leq y$ and $0 \leq m \leq y$. Furthermore, since $r_i \geq 2$, $\tau(i)^2$ is a factor of $\tau(\alpha)$. Hence,

$$\tau(\alpha) = u \cdot (a^j b^k v a^l b^m)^2 \cdots ,$$

$u \neq \cdots a$ if $j \neq 0$ and, $u \neq \cdots b$ if $j = 0$. We now consider the following cases:

1. $j \neq k$, $j \neq 0$ and $k \neq 0$.

   (a) $v \neq \varepsilon$. So, $\tau(\alpha) = u \cdot a^j b^k v \cdots$. However, the factor $u \cdot a^j b^k v$ does not occur in $\sigma(\alpha)$, because $j \neq k$.

   (b) $v = \varepsilon$.

      i. $l = m = 0$. Hence, $\tau(\alpha) = u \cdot a^j b^k a^j b^k \cdots$. However, due to $j \neq k$, $\sigma(\alpha)$ does not have the factor $u \cdot a^j b^k a^j$, and this contradicts the assumption $\tau(\alpha) = \sigma(\alpha)$.

ii. $l = 0$ and $m \neq 0$. We have $\tau(\alpha) = u \cdot a^j b^k b^m a^j b^k b^m \cdots$; in other words, $\tau(\alpha)$ contains the factor $u \cdot a^j b^{k+m} a^j b^{k+m}$. Let $v' = a^j b^{k+m}$. Since $\tau(\alpha) = \sigma(\alpha)$, $j = (k + m)$ and $v' \in \{a^p b^p | p \in B\}$. So, $\tau(\alpha)$ contains the factor $v'v'$ which contradicts the Claim.

iii. $l \neq 0$. So, $\tau(\alpha) = u \cdot a^j b^k a^l b^m a^j b^k a^l b^m \cdots$. However, the factor $u \cdot a^j b^k a^l$ does not occur in $\sigma(\alpha)$, because $j \neq k$. Hence, $\tau(\alpha) \neq \sigma(\alpha)$ and this again contradicts the assumption.

2. $j = k \neq 0$.

(a) $l \neq m$, $l \neq 0$ and $m \neq 0$. Thus, $\tau(\alpha) = u \cdot a^j b^j v a^l b^m \cdot a^j b^j v a^l b^m \cdots$. This means that $\tau(\alpha)$ contains the factor $b^j v \cdot a^l b^m \cdot a^j$. Due to $l \neq m$, this factor does not occur in $\sigma(\alpha)$, and this contradicts the assumption $\sigma(\alpha) = \tau(\alpha)$.

(b) $l = m = 0$. Hence, $\tau(\alpha) = u \cdot a^j b^j v \cdot a^j b^j v \cdots$. Let $v' = a^j b^j v$. Thus, $v'v'$ is a factor of $\tau(\alpha)$ which implies that $v' \in \{a^p b^p | p \in B\}^+$. However, this contradicts the above mentioned claim.

(c) $l = m \neq 0$ and $l \neq 1$. We can conclude that $\tau(\alpha) = u \cdot a^j b^j v a^l b^l \cdot a^j b^j v a^l b^l \cdots$. We can infer from the factor $b^j v \cdot a^l b^l \cdot a^j$ that $a^l b^l \in \{a^p b^p | p \in B\}$. Let $v' = a^j b^j v a^l b^l$. So, $v'v'$ is a factor of $\tau(\alpha)$ while $v' \in \{a^p b^p | p \in B\}^+$, which again contradicts the mentioned claim.

(d) $l = m = 1$. So, $\tau(\alpha)$ contains the factor $b^j v a^1 b^1 \cdot a^j$ which does not occur in $\sigma(\alpha)$.

(e) $l \neq 0$ and $m = 0$. Hence, $\tau(\alpha) = u \cdot a^j b^j v \cdot a^{l+j} b^j \cdot v a^l \cdots$. However, this contradicts the assumption $\sigma(\alpha) = \tau(\alpha)$, because of $(l + j) \neq j$.

(f) $l = 0$ and $m \neq 0$. This means that $\tau(\alpha)$ has the factor $u \cdot a^j b^j v b^m \cdot a^j$.

i. $v = \varepsilon$. So, $u \cdot a^j b^{j+m} \cdot a^j$ is a factor of $\tau(\alpha)$, and this contradicts $\sigma(\alpha) = \tau(\alpha)$ due to $j \neq (j + m)$.

ii. $v \neq \varepsilon$. Thus, we have the factor $b^j \cdot v b^m \cdot a^j$ in $\tau(\alpha)$. However, the number of repetitions of the last b in v plus m is larger than the repetitions of its previous a, and such a factor does not occur in $\sigma(\alpha)$.

3. $j \neq 0$ and $k = 0$.

(a) $v \neq \varepsilon$. So, $a^j v a^l b^m \cdot a^j$ is a factor of $\tau(\alpha)$. However, the number of repetitions of the first a in v plus j is larger than the number of the subsequent b, and this contradicts the structure of $\sigma(\alpha)$.

(b) $v = \varepsilon$. This implies $\tau(\alpha) = u \cdot a^{j+l} b^m \cdot a^{j+l} b^m \cdots$.

i. $m \neq 0$. Since $\tau(\alpha) = \sigma(\alpha)$, we can infer from the factor $u \cdot a^{j+l}b^m \cdot a^{j+l}$ that $j+l = m$ and as a result $a^{j+l}b^m \in \{a^p b^p | p \in B\}$. Let $v' = a^{j+l}b^m$. So, $v'v'$ is a factor of $\tau(\alpha)$; however, this again contradicts the mentioned claim.

ii. $m = 0$. We can conclude that $(a^{j+l})^{r_i}$ is a factor of $\tau(\alpha)$. However, $\sigma(\alpha)$ does not contain this factor, because we know that after r_i occurrences of a in $\sigma(\alpha)$, we have b or ε.

4. $j = 0$ and $k \neq 0$.

(a) $l \neq 0$. Hence, $\tau(\alpha) = u \cdot b^k v a^l b^m \cdot b^k v a^l b^m \cdots$ and, consequently, $\tau(\alpha)$ contains the factor $b^k v a^l b^{m+k} v a^l$. Because of $\tau(\alpha) = \sigma(\alpha)$, we can conclude that $l = (m+k)$ and $a^l b^{m+k} \in \{a^p b^p | p \in B\}$ and also $\tau(\alpha) = u \cdot (b^k v a^l b^{m+k} v a^l b^m) \cdot (b^k) \cdot \cdots$. Let $v' = v a^l b^{m+k}$. So, $v'v'$ is a factor of $\tau(\alpha)$ while $v' \in \{a^p b^p | p \in B\}^+$. This again contradicts the mentioned claim.

(b) $l = 0$. So, $\tau(\alpha) = u \cdot b^k v b^m \cdot b^k v b^m \cdots$.

i. $v \neq \varepsilon$. As a result, $b^k \cdot v b^{k+m} \cdot v b^m$ is a factor of $\tau(\alpha)$. However, the number of repetitions of the last b in v plus $k + m$ is larger than the repetitions of its previous a, and such a factor does not occur in $\sigma(\alpha)$.

ii. $v = \varepsilon$. We can conclude that $(b^{k+m})^{r_i}$ is a factor of $\tau(\alpha)$. However, $\sigma(\alpha)$ does not contain this factor, because we know that after r_i occurrences of b in $\sigma(\alpha)$, we have a or ε.

5. $\tau(i) = \varepsilon$. Due to $\tau(\alpha) = \sigma(\alpha)$, there exists an $i' > i$ with $|\tau(i')| > 1$. So, we can consider $\tau(i') = a^j \cdot b^k \cdot v \cdot a^l \cdot b^m$, which leads to the above cases.

Consequently, in all cases, assuming the existence of a morphism τ with $\tau(\alpha) = \sigma(\alpha)$ and, for an $i \in \text{var}(\alpha)$, $\tau(i) \neq \sigma(i)$ leads to a contradiction. Thus, σ is unambiguous with respect to α. □

Our second criterion again holds for binary (and, thus, all larger) alphabets Σ. It features a rather restricted class of patterns, which, however, are minimal with regard to their length.

Theorem 5.5. *Let $n \in \mathbb{N}$, $n \geq 2$. If n is even, then let*

$$\alpha := 1 \cdot 2 \cdot [\ldots] \cdot n \cdot (n/2+1) \cdot 1 \cdot (n/2+2) \cdot 2 \cdot [\ldots] \cdot n \cdot n/2,$$

and if n is odd, then let

$$\alpha := 1 \cdot 1 \cdot 2 \cdot 3 \cdot [\ldots] \cdot n \cdot (\lceil n/2 \rceil + 1) \cdot 2 \cdot (\lceil n/2 \rceil + 2) \cdot 3 \cdot [\ldots] \cdot n \cdot \lceil n/2 \rceil.$$

Then α is a shortest pattern with $|\operatorname{var}(\alpha)| = n$ that is not a fixed point of a nontrivial morphism, and there exists a 1-uniform morphism $\sigma : \mathbb{N}^* \to \{a, b\}^*$ that is unambiguous with respect to α.

Proof. We first briefly explain why any pattern α' with $|\operatorname{var}(\alpha')| = n$ and $|\alpha'| < |\alpha|$ must be a fixed point of a nontrivial morphism: If $|\operatorname{var}(\alpha')| = n$ and $|\alpha'| < |\alpha|$, then α' must contain at least one variable z with just a single occurrence, because all variables in α have exactly two occurrences. We can then define a morphism $\phi : \mathbb{N}^* \to \mathbb{N}^*$ by $\phi(z) := \alpha'$ and $\phi(z') := \varepsilon$ for all $z' \in \operatorname{var}(\alpha') \setminus \{z'\}$. Since $n \geq 2$, ϕ is nontrivial, and obviously $\phi(\alpha') = \alpha'$. Hence, α' is a fixed point of ϕ. At the end of this proof, we shall show that α is not a fixed point of a nontrivial morphism, which will then complete the proof of the first statement of the theorem.

We now consider the second statement of the theorem. We define the morphism σ by

$$\sigma(x) := \begin{cases} a, & \text{if } 1 \leq x \leq \lceil n/2 \rceil, \\ b, & \text{else }. \end{cases}$$

Assume to the contrary that σ is ambiguous with respect to α. Consequently, there exists a morphism $\tau : \mathbb{N}^* \to \{a, b\}^*$ satisfying $\tau(\alpha) = \sigma(\alpha)$ and, for some $q \in \operatorname{var}(\alpha)$, $\tau(q) \neq \sigma(q)$.
Let n be even. So,

$$\alpha := 1 \cdot 2 \cdot [\ldots] \cdot n \cdot (n/2 + 1) \cdot 1 \cdot (n/2 + 2) \cdot 2 \cdot [\ldots] \cdot n \cdot n/2.$$

As a result,

$$\sigma(\alpha) = a^{n/2} \cdot b^{n/2} \cdot (ba)^{n/2}.$$

Assume that $\alpha = \beta_1 \beta_2$ with

$$\beta_1 = 1 \cdot 2 \cdot [\ldots] \cdot n$$

and,

$$\beta_2 = (n/2 + 1) \cdot 1 \cdot (n/2 + 2) \cdot 2 \cdot [\ldots] \cdot n \cdot n/2.$$

Due to the structure of α and $|\tau(\alpha)| = |\sigma(\alpha)|$, it is easily verified that $|\tau(\beta_1)| = |\sigma(\beta_1)|$ and, hence, $\tau(\beta_1) = \sigma(\beta_1)$. Besides, $\tau(\beta_1) = \sigma(\beta_1)$ implies that $\tau(\beta_2) = \sigma(\beta_2)$. Since σ is a 1-uniform morphism, there exists a $q \in \operatorname{var}(\alpha)$ such that $|\tau(q)| \geq 2$. Due to $\tau(\beta_1) = \sigma(\beta_1)$, we have one of the following cases:

1. $a^k \sqsubseteq \tau(q)$ with $k \geq 2$. Since q has an occurrence in β_2 and $a^k \not\sqsubseteq \sigma(\beta)$, $\tau(\beta) \neq \sigma(\beta)$, and as a result, $\tau(\alpha) \neq \sigma(\alpha)$, which is a contradiction.

2. $b^k \sqsubseteq \tau(q)$ with $k \geq 2$. Using the same reasoning as above, this leads to a

contradiction.

3. $\tau(q) = ab$. We consider the following cases:

   - $q < n/2$. Then, due to $\tau(\beta_1) = \sigma(\beta_1)$, there exists a $q' < q$ satisfying $\tau(q') = a^k$ with $k \geq 2$, which according to Case 1 leads to a contradiction.
   - $q = n/2$. Due to the facts that $n/2$ is the last variable occurring in α and ba must be a suffix of $\tau(\alpha)$, this leads to a contradiction.
   - $q = n/2 + 1$. Since $\tau(\beta_2) = \sigma(\beta_2)$, ba must be a prefix of $\tau(\beta_2)$. However, the variable $n/2 + 1$ is the first variable of β_2. Consequently, this contradicts $\tau(\alpha) = \sigma(\alpha)$.
   - $q > n/2 + 1$. Then, due to $\tau(\beta_1) = \sigma(\beta_1)$, there exists a $q' > q$ satisfying $\tau(q') = b^k$ with $k \geq 2$, which according to Case 2 leads to a contradiction.

Hence, all above cases contradict the assumption of $\tau(\alpha) = \sigma(\alpha)$.
However, if n is odd,

$$\alpha := 1 \cdot 1 \cdot 2 \cdot 3 \cdot [\ldots] \cdot n \cdot (\lceil n/2 \rceil + 1) \cdot 2 \cdot (\lceil n/2 \rceil + 2) \cdot 3 \cdot [\ldots] \cdot n \cdot \lceil n/2 \rceil.$$

Thus,

$$\sigma(\alpha) = aa \cdot a^{\lfloor n/2 \rfloor} \cdot b^{\lfloor n/2 \rfloor} \cdot (ba)^{\lfloor n/2 \rfloor}.$$

Due to the structure of α and $\tau(\alpha) = \sigma(\alpha)$, it is easily verified that $\tau(1) = \sigma(1) = a$. This implies that an analogous reasoning to the case when n is even can also be used for the case that n is odd. Consequently, we can conclude that σ is unambiguous with respect to α.

What remains to explain is why α is not a fixed point of a nontrivial morphism. Since σ is nonerasing and, as shown above, unambiguous with respect to α, this directly follows from the contraposition of Theorem 3.17. □

The following examples illustrates Theorem 5.5 and its proof: For $n := 6$, $\alpha := 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 4 \cdot 1 \cdot 5 \cdot 2 \cdot 6 \cdot 3$, and the 1-uniform morphism $\sigma : \mathbb{N}^* \to \{a, b\}^*$ with $\sigma(1) := \sigma(2) := \sigma(3) := a$ and $\sigma(4) := \sigma(5) := \sigma(6) := b$ is unambiguous with respect to α. For $n := 5$, $\alpha := 1 \cdot 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 4 \cdot 2 \cdot 5 \cdot 3$, and the respective unambiguous morphism is given by $\sigma(1) := \sigma(2) := \sigma(3) := a$ and $\sigma(4) := \sigma(5) := b$.

From Theorem 5.5 we can conclude that patterns α with unambiguous 1-uniform morphisms using a binary target alphabet exist for every cardinality of $\mathrm{var}(\alpha)$ and that corresponding examples can be given where every variable occurs just twice.

5.2 Variable target alphabets

In order to continue our examination of Problem 3.16, we now relax one of the requirements of Section 5.1: We no longer investigate criteria on the existence of unambiguous 1-uniform morphisms for a fixed target alphabet Σ, but we permit Σ to depend on the number of variables in the pattern α in question. Regarding this question, we conjecture the following statement to be true:

Conjecture 5.6. *Let α be a pattern with $|\operatorname{var}(\alpha)| \geq 4$. There exists an alphabet Σ satisfying $|\Sigma| < |\operatorname{var}(\alpha)|$ and a 1-uniform morphism $\sigma : \mathbb{N}^* \to \Sigma^*$ that is unambiguous with respect to α if and only if α is not a fixed point of a nontrivial morphism.*

Due to Theorem 3.1 and also the fact that the set of succinct patterns is equivalent to the set of patterns which are not a fixed point of a nontrivial morphism (see Theorem 3.19), the *only if* direction of Conjecture 5.6 is trivial. Moreover, this conjecture would be trivially true if we allowed Σ to satisfy $|\Sigma| \geq |\operatorname{var}(\alpha)|$. That explains why we exclusively study the case where the number of letters in the target alphabet is smaller than the number of variables in the pattern. From Theorem 5.1, it directly follows that an analogous conjecture would not be true if we considered fixed binary target alphabets (as is done in Section 5.1), since none of the patterns α_m is a fixed point of a nontrivial morphism – this can be easily verified using Theorem 3.20 and the definition of prolix patterns. Hence, characteristic criteria must necessarily look different in such a context. It can also be effortlessly understood that Conjecture 5.6 would be incorrect if we dropped the condition that α needs to contain at least 4 distinct variables, since not only σ_0, but all 1-uniform morphisms $\sigma : \mathbb{N}^* \to \Sigma^*$ with $|\Sigma| \leq 2$ are ambiguous with respect to our example pattern $\alpha_0 = 1 \cdot 2 \cdot 3 \cdot 1 \cdot 3 \cdot 2$ discussed in Chapter 1.

Technically, many of our subsequent considerations are based on the following generic morphisms:

Definition 5.7. *Let Σ be an infinite alphabet, and let $\sigma : \mathbb{N}^* \to \Sigma^*$ be a renaming. For any $i, j \in \mathbb{N}$ with $i \neq j$ and for every $x \in \mathbb{N}$, let the morphism $\sigma_{i,j}$ be given by*

$$\sigma_{i,j}(x) := \begin{cases} \sigma(i), & \text{if } x = j, \\ \sigma(x), & \text{if } x \neq j. \end{cases}$$

Thus, $\sigma_{i,j}$ maps exactly two variables to the same image, and therefore, for any pattern α with at least two different variables, $\sigma_{i,j}(\alpha)$ is a word over $|\operatorname{var}(\alpha)| - 1$ distinct letters. Using this definition, we can now state a more specific version of Conjecture 5.6:

Conjecture 5.8. *Let α be a pattern with $|\operatorname{var}(\alpha)| \geq 4$. There exist $i, j \in \operatorname{var}(\alpha)$, $i \neq j$, such that $\sigma_{i,j}$ is unambiguous with respect to α if and only if α is not a fixed point of a nontrivial morphism.*

Before we study Conjectures 5.6 and 5.8 in more detail, we establish that they are equivalent. To this end, and also for many of our subsequent technical considerations, we need the following concept:

Definition 5.9. *Let $\alpha \in \mathbb{N}^*$. For any $i, j \in \mathbb{N}$ with $i \neq j$ and, for every $x \in \mathbb{N}$, let the morphism $\phi_{i,j} : \operatorname{var}(\alpha)^* \to \operatorname{var}(\alpha)^*$ be given by*

$$\phi_{i,j}(x) := \begin{cases} i, & \text{if } x = j, \\ x, & \text{if } x \neq j, \end{cases}$$

and let $\alpha_{i,j} := \phi_{i,j}(\alpha)$. Note that $\phi_{i,j}$ is an alphabet reduction for α.

For example, let $\alpha := 1 \cdot 2 \cdot 3 \cdot 3 \cdot 1 \cdot 4 \cdot 2 \cdot 4$. If we consider $i := 2$ and $j := 4$, then $\alpha_{2,4} = \phi_{2,4}(\alpha) = 1 \cdot 2 \cdot 3 \cdot 3 \cdot 1 \cdot 2 \cdot 2 \cdot 2$.

Using Definition 5.9, we can now address the relation between Conjectures 5.6 and 5.8:

Proposition 5.10. *Let α be a pattern with $|\operatorname{var}(\alpha)| \geq 4$. There exists an alphabet Σ satisfying $|\Sigma| < |\operatorname{var}(\alpha)|$ and a 1-uniform morphism $\sigma : \mathbb{N}^* \to \Sigma^*$ that is unambiguous with respect to α if and only if there exist $i, j \in \operatorname{var}(\alpha)$, $i \neq j$, such that $\sigma_{i,j}$ is unambiguous with respect to α.*

Proof. Since the size of the target alphabet of $\sigma_{i,j}$ equals $|\operatorname{var}(\alpha)| - 1$, the *if* direction is trivially true.

We now prove the *only if* direction. So, we assume that there exists a 1-uniform morphism $\sigma : \mathbb{N}^* \to \Sigma^*$, $|\Sigma| < |\operatorname{var}(\alpha)|$, that is unambiguous with respect to α. This means that there does not exist any morphism $\tau : \mathbb{N}^* \to \Sigma^*$ such that $\tau(\alpha) = \sigma(\alpha)$ and, for a variable q occurring in α, $\tau(q) \neq \sigma(q)$. Let $V := \{v \in \operatorname{var}(\alpha) \mid |\sigma(\alpha)|_{\sigma(v)} \neq |\alpha|_v\}$. If $|V| = 2$, then the *only if* direction holds immediately. Otherwise, we choose two arbitrary variables i, j from V satisfying $\sigma(i) = \sigma(j)$. We define a morphism $\phi : \Sigma^* \to \mathbb{N}^*$ by

$$\phi(x) := \begin{cases} i, & \text{if } x = \sigma_{i,j}(i), \\ \sigma_{i,j}^{-1}(x), & \text{else.} \end{cases}$$

The morphism ϕ exists due to the definition of $\sigma_{i,j}$, and we can directly conclude the correctness of the following statement:

Claim 1. $\phi \circ \sigma_{i,j}(\alpha) = \alpha_{i,j}$.

According to Definition 5.9, $\phi_{i,j} : \mathbb{N}^* \to \mathbb{N}^*$ is given by

$$\phi_{i,j}(x) := \begin{cases} i, & \text{if } x = j, \\ x, & \text{else.} \end{cases}$$

Since, by Definition 5.9, $\alpha_{i,j}$ equals $\phi_{i,j}(\alpha)$, we can prove the following vital fact:

Claim 2. $\sigma(\alpha) = \sigma(\alpha_{i,j})$.

Proof (Claim 2). Due to our choice of i and j, we know that $\sigma(i) = \sigma(j)$ is satisfied. Furthermore, $\phi_{i,j}(i) = \phi_{i,j}(j) = i$, and therefore $\sigma(\phi_{i,j}(i)) = \sigma(\phi_{i,j}(j)) = \sigma(i) = \sigma(j)$. Hence, and since the definition of $\phi_{i,j}$ directly implies $\sigma(x) = \sigma(\phi_{i,j}(x))$ for every $x \in \text{var}(\alpha) \setminus \{i, j\}$, we can conclude $\sigma(\alpha) = \sigma(\phi_{i,j}(\alpha))$. Since $\phi_{i,j}(\alpha) = \alpha_{i,j}$, this proves $\sigma(\alpha) = \sigma(\alpha_{i,j})$. □(Claim 2)

We now assume to the contrary that $\sigma_{i,j}(\alpha)$ is ambiguous. Hence, there is a morphism $\tau_{i,j} : \mathbb{N}^* \to \Sigma^*$ satisfying $\tau_{i,j}(\alpha) = \sigma_{i,j}(\alpha)$ and, for a variable q occurring in α, $\tau_{i,j}(q) \neq \sigma_{i,j}(q)$. Since $\sigma_{i,j}$ is 1-uniform, this implies that there exists a variable $q' \in \text{var}(\alpha)$ with $\tau_{i,j}(q') = \varepsilon$.

The following diagram illustrates all morphisms, patterns and words introduced so far:

$$\begin{tikzcd}
\alpha \arrow[rr, "\sigma"] \arrow[d, shift right, "\tau_{i,j}"'] \arrow[d, shift left, "\sigma_{i,j}"] \arrow[drr, "\phi_{i,j}"] & & \sigma(\alpha) \\
\sigma_{i,j}(\alpha) \arrow[rr, "\phi"] & & \alpha_{i,j} \arrow[u, "\sigma"']
\end{tikzcd}$$

We now define the morphism $\tau : \text{var}(\alpha)^* \to \Sigma^*$ by

$$\tau := \sigma \circ \phi \circ \tau_{i,j}.$$

Since we assume that $\tau_{i,j}(\alpha)$ equals $\sigma_{i,j}(\alpha)$, Claims 1 and 2 facilitate the following reasoning:

$$\begin{aligned} \tau(\alpha) &= \sigma \circ \phi \circ \tau_{i,j}(\alpha) \\ &= \sigma \circ \phi \circ \sigma_{i,j}(\alpha) \\ &= \sigma(\alpha_{i,j}) \\ &= \sigma(\alpha). \end{aligned}$$

Consequently, $\tau(\alpha) = \sigma(\alpha)$. As stated above, there exists a variable $q' \in \mathrm{var}(\alpha)$ that satisfies $\tau_{i,j}(q') = \varepsilon$, and therefore $\tau(q') = \sigma \circ \phi \circ \tau_{i,j}(q') = \varepsilon$. On the other hand, σ is 1-uniform, and therefore $\sigma(q') \neq \varepsilon$. Hence, the existence of τ implies that σ is ambiguous with respect to α, and this is a contradiction to the initial assumption of our proof for the *only if* direction. Thus, $\sigma_{i,j}$ is unambiguous with respect to α. □

Thus, our two conjectures are equivalent:

Corollary 5.11. *Conjecture 5.6 is true if and only if Conjecture 5.8 is true.*

Proof. Follows directly from Proposition 5.10. □

Due to Theorem 3.17, the *only if* directions of Conjectures 5.6 and 5.8 hold true immediately. In the remainder of this section, we shall therefore exclusively study those patterns that are not fixed points. Our corresponding results yield large classes of such patterns that have an unambiguous 1-uniform morphism, but we have to leave the overall correctness of our conjectures open.

Conjecture 5.8 suggests that the examination of the existence of unambiguous 1-uniform morphisms for a pattern α may be reduced to finding suitable variables i and j such that $\sigma_{i,j}$ is unambiguous with respect to α. In this regard, one particular choice can be ruled out immediately:

Proposition 5.12. *Let α be a pattern, and let $i, j \in \mathrm{var}(\alpha)$, $i \neq j$. If $\sigma_{i,j}(\alpha)$ is a fixed point of a nontrivial morphism, then $\sigma_{i,j}$ is not unambiguous with respect to α.*

Proof. If $\sigma_{i,j}(\alpha)$ is a fixed point of a nontrivial morphism, then, by definition, there is a morphism ϕ satisfying $\phi(\sigma_{i,j}(\alpha)) = \sigma_{i,j}(\alpha)$ and, for a letter a in $\sigma_{i,j}(\alpha)$, $\phi(a) \neq a$. This implies that there must be a letter in α that is mapped by ϕ to the empty word; without loss of generality, we simply assume $\phi(a) := \varepsilon$. If we now define $\tau := \phi \circ \sigma_{i,j}$, then $\tau(\alpha) = \sigma_{i,j}(\alpha)$ and $\tau(x) = \varepsilon \neq \sigma_{i,j}(x)$, where x is a variable in α satisfying $\sigma_{i,j}(x) = a$. Thus, $\sigma_{i,j}$ is ambiguous with respect to α. □

For example, if we consider the pattern $\alpha_1 := 1 \cdot 2 \cdot 3 \cdot 4 \cdot 1 \cdot 4 \cdot 3 \cdot 2$ (which is not a fixed point) and define $\Sigma := \{a, b, c\}$, then $\sigma_{2,4}(\alpha_1)$ equals *abcbabcb* (or any renaming thereof), which is a fixed point of the morphism ϕ given by $\phi(a) := abcb$ and $\phi(b) := \phi(c) := \varepsilon$. Thus, $\sigma_{2,4}$ is ambiguous with respect to α_1. However, Proposition 5.12 does not provide a characteristic condition on the ambiguity of $\sigma_{i,j}$, since $\sigma_{2,3}(\alpha_1) = abbcacbb$ is not a fixed point, but still $\sigma_{2,3}$ is ambiguous with respect to α_1. Furthermore, while the ambiguity of $\sigma_{2,3}$ results from the fact that α_1 contains the factors $2 \cdot 3$ and $3 \cdot 2$, and is therefore easy to comprehend, there are more difficult examples of morphisms $\sigma_{i,j}$ that are ambiguous although they do not lead to a morphic image that is a fixed point. This is illustrated by the example $\alpha_2 := 1 \cdot 2 \cdot 3 \cdot 3 \cdot 4 \cdot 4 \cdot 1 \cdot 2 \cdot 3 \cdot 3 \cdot 4 \cdot 4 \cdot 2$. Here, $\sigma_{2,4}(\alpha_1) = abccbbabccbbb$ again is not a fixed point, but $\sigma_{2,4}$ is nevertheless ambiguous with respect to α_2, since the morphism τ given by $\tau(1) := abccb$, $\tau(2) := b$ and $\tau(3) := \tau(4) := \varepsilon$ satisfies $\tau(\alpha_2) = \sigma_{2,4}(\alpha_2)$. We therefore conclude that it seems not to be a straightforward task to find amendments that could turn Proposition 5.12 into a characteristic condition.

We now show that Conjecture 5.8 is correct for several types of patterns. To this end, we need the following simple sufficient condition (which uses Definition 4.3) on a pattern being a fixed point:

Lemma 5.13. *Let $\alpha \in \mathbb{N}^+$. If there exists a variable $i \in \mathrm{var}(\alpha)$ such that i has loyal neighbours, then α is a fixed point of a nontrivial morphism.*

Proof. Assume that Condition 1 of Definition 4.3 is satisfied. So, without loss of generality, let

$$\alpha := \alpha_1 \cdot l_1 \cdot i_1 \cdot \alpha_2 \cdot l_2 \cdot i_2 \cdot \alpha_3 \cdot [\ldots] \cdot \alpha_n \cdot l_n \cdot i_n \cdot \alpha_{n+1}$$

where $i_1, i_2, \ldots, i_n$ are all occurrences of the variable i in α and, for every j, $1 \leq j \leq n$, $\alpha_j \in \mathbb{N}^*$, $\alpha_{n+1} \in \mathbb{N}^*$ and $l_j \in \mathbb{N}$. Also, Condition 1 of Definition 4.3 implies that, for every j, $1 \leq j \leq n$ and for every j', $1 \leq j' \leq n+1$, $l_j \neq i$, $l_j \not\sqsubseteq \alpha_{j'}$. We define a morphism $\phi : \mathbb{N}^+ \to \mathbb{N}^*$ by:

$$\phi(x) := \begin{cases} l_j i, & \text{if } x = l_j,\ 1 \leq j \leq n, \\ \varepsilon, & \text{if } x = i, \\ x, & \text{else.} \end{cases}$$

Hence, $\phi(\alpha) = \alpha$ which means that α is a fixed point of a nontrivial morphism ϕ. Using an analogous reasoning as above, we can show that the lemma also holds true when Condition 2 of Definition 4.3 is satisfied. □

Using this lemma, we can now establish a class of patterns for which Conjecture 5.8 holds true. All variables in these patterns have the same number of occurrences, and for one pair of variables they do not contain any factors as discussed above with respect our example $\sigma_{2,3}$:

Theorem 5.14. *Let $m \in \mathbb{N}$, $m \geq 2$. Let $\alpha \in \mathbb{N}^+$ be a pattern that is not a fixed point of a nontrivial morphism and satisfies, for every $x \in \mathrm{var}(\alpha)$, $|\alpha|_x = m$. If there are $i, j \in \mathrm{var}(\alpha)$, $i \neq j$, such that*

- *there is no* $k \in \mathrm{var}(\alpha)$ *with* $\{i,j\} \subseteq L_k$ *or* $\{i,j\} \subseteq R_k$*, and*
- $\alpha \neq \alpha_1 \cdot \mathbf{i} \cdot \mathbf{j} \cdot \alpha_2 \cdot \mathbf{j} \cdot \mathbf{i} \cdot \alpha_3,\ \alpha_1, \alpha_2, \alpha_3 \in \mathbb{N}^*$*,*

then $\sigma_{i,j}$ *is unambiguous with respect to* α*.*

Proof. Assume to the contrary that $\sigma_{i,j}$ is ambiguous. So, there exists a morphism $\tau : \mathbb{N}^+ \to \Sigma^*$ such that $\tau(\alpha) = \sigma_{i,j}(\alpha)$ and, for some $x \in \mathrm{var}(\alpha)$, $\tau(x) \neq \sigma_{i,j}(x)$. Since $\sigma_{i,j}$ is a 1-uniform morphism, there exists a $k \in \mathrm{var}(\alpha)$ with $|\tau(k)| \geq 2$. Let $uv \sqsubseteq \tau(k)$, $u, v \in \Sigma$. Due to the fact that k occurs m times in α, $\sigma_{i,j}(\alpha) = \tau(\alpha) = w_1 \cdot uv \cdot w_2 \cdot uv \cdot [\ldots] \cdot w_m \cdot uv \cdot w_{m+1}$ with, for every q, $1 \leq q \leq m+1$, $w_q \in \Sigma^*$. We now consider the following cases:

- $\sigma_{i,j}(i) \neq u$ and $\sigma_{i,j}(i) \neq v$. This implies that there exist the variables $x_1, x_2 \in \mathrm{var}(\alpha)$, $x_1, x_2 \neq i$ and $x_1, x_2 \neq j$, such that $\alpha = \alpha_1 \cdot x_1x_2 \cdot \alpha_2 \cdot x_1x_2 \cdot [\ldots] \cdot \alpha_m \cdot x_1x_2 \cdot \alpha_{m+1}$, for every q, $1 \leq q \leq m+1$, $\alpha_q \in \mathbb{N}^*$, and $\sigma_{i,j}(x_1) = u$ and $\sigma_{i,j}(x_2) = v$. Due to $|\alpha|_{x_1} = |\alpha|_{x_2} = m$, $x_1, x_2 \not\sqsubseteq \alpha_q$, for every q, $1 \leq q \leq m+1$. This implies that $R_{x_1} = \{x_2\}$ and $L_{x_2} = \{x_1\}$. Then, according to Lemma 5.13, α is a fixed point of a nontrivial morphism which is a contradiction to the assumption of the theorem.

- $\sigma_{i,j}(i) = \sigma_{i,j}(j) = u$, and $u \neq v$. So, we assume that $\alpha = \alpha_1 \cdot x_1x' \cdot \alpha_2 \cdot x_2x' \cdot [\ldots] \cdot \alpha_m \cdot x_mx' \cdot \alpha_{m+1}$ with, $x' \in \mathrm{var}(\alpha)$ and, for every q, $1 \leq q \leq m+1$, $x_q \in \mathrm{var}(\alpha)$, $\alpha_q \in \mathbb{N}^*$, and $\sigma_{i,j}(x_q) = u$ and $\sigma_{i,j}(x') = v$. Additionally, since $\sigma_{i,j}(x') = v$ and $u \neq v$, we can conclude that $x' \neq i$ and $x' \neq j$. We now consider the following cases:

  1. For every q, $1 \leq q \leq m$, $x_q = i$. This implies, using the same reasoning as above, that α is a fixed point of a nontrivial morphism which is a contradiction.
  2. There exists q, q', $1 \leq q, q' \leq m$ and $q \neq q'$, such that $x_q = i$ and $x_{q'} = j$. This means that $\{i,j\} \subseteq L_{x_2}$, which contradicts the first condition of the theorem.

- $\sigma_{i,j}(i) = v$, and $u \neq v$. The reasoning is analogous to that in the previous case.

- $\sigma_{i,j}(i) = \sigma_{i,j}(j) = u$ and $v = u$. Hence, we may assume that $\alpha = \alpha_1 \cdot x_1x'_1 \cdot \alpha_2 \cdot x_2x'_2 \cdot [\ldots] \cdot \alpha_m \cdot x_mx'_m \cdot \alpha_{m+1}$ with, for every q, $1 \leq q \leq m+1$, $\alpha_q \in \mathbb{N}^*$, $x_q, x'_q \in \mathrm{var}(\alpha)$ and $\sigma_{i,j}(x_q) = \sigma_{i,j}(x'_q) = u$. Due to the conditions of the theorem, the factors $i \cdot i \cdot j$, $i \cdot j \cdot j$, $j \cdot i \cdot i$ and $j \cdot j \cdot i$ could not be the factors of α. Moreover, it can be observed that $u \cdot u \cdot u \not\sqsubseteq \tau(k)$; otherwise, since

$\tau(\alpha) = \sigma_{i,j}(\alpha)$, then $|\alpha|_i > m$ or $\alpha_j > m$. This implies that $i \cdot j \cdot i$ and $j \cdot i \cdot j$ are not the factors of α. We now consider the following cases:

1. For every q, $1 \leq q \leq m$, $x_q = i$ and $x'_q = j$. As a result, $R_i = \{j\}$ and $L_j = \{i\}$. According to Lemma 5.13, α is a fixed point of a nontrivial morphism.
2. For every q, $1 \leq q \leq m$, $x_q = j$ and $x'_q = i$. As a result, $R_j = \{i\}$ and $L_i = \{j\}$. Referring to Lemma 5.13, α is a fixed point of a nontrivial morphism.
3. There exists a q, q', $1 \leq q, q' \leq m$ and $q \neq q'$, such that $x_q \cdot x'_q = i \cdot j$ and $x_{q'} \cdot x'_{q'} = j \cdot i$. This case contradicts the second condition of the theorem.
4. There exists a q, q', $1 \leq q, q' \leq m$ and $q \neq q'$, such that $x_q \cdot x'_q = i \cdot j$ and, $x_{q'} \cdot x'_{q'} = i \cdot i$ or $x_{q'} \cdot x'_{q'} = j \cdot j$. This means that $\{i, j\} \subseteq R_i$ or $\{i, j\} \subseteq L_j$ which is a contradiction to the first condition of the theorem.
5. There exists a q, q', $1 \leq q, q' \leq m$ and $q \neq q'$, such that $x_q \cdot x'_q = j \cdot i$ and, $x_{q'} \cdot x'_{q'} = i \cdot i$ or $x_{q'} \cdot x'_{q'} = j \cdot j$. This implies that $\{i, j\} \subseteq L_i$ or $\{i, j\} \subseteq R_j$ which contradicts the first condition of the theorem.
6. There exist q, q', $1 \leq q, q' \leq m$, $q' \neq q$, such that $x_q \cdot x'_q = i \cdot i$ and $x_{q'} \cdot x'_{q'} = j \cdot j$. Since $uu \sqsubseteq \tau(k)$ and due to the conditions of the theorem, it results from $\tau(\alpha) = \sigma_{i,j}(\alpha)$ that $k \neq i$ and $k \neq j$, in other words, $\tau(i) \neq uu$ and $\tau(j) \neq uu$; otherwise, $|\tau(\alpha)|_u > |\sigma_{i,j}(\alpha)|_u$. Moreover, we may observe that if $\sigma_{i,j}(k) \sqsubseteq \tau(k)$, then this implies that there exists an $x \in \mathrm{var}(\alpha) \setminus \{i, j\}$, with $\{i, j\} \subseteq L_x$ or $\{i, j\} \subseteq R_x$, which is a contradiction. Thus, $\sigma_{i,j}(k) \not\sqsubseteq \tau(k)$. Since $\tau(\alpha) = \sigma_{i,j}(\alpha)$, there must be a $k' \in \mathrm{var}(\alpha)$, $k' \neq k, i, j$, such that $\sigma_{i,j}(k) \sqsubseteq \tau(k')$, which means that $|\tau(k')| \geq 2$ or we can extend the reasoning over the other variables. Consequently, since $\tau(\alpha) = \sigma(\alpha)$, this argumentation implies the existence of a $k'' \in \mathrm{var}(\alpha)$, $k'' \neq k, i, j$, such that $|\tau(k'')| \geq 2$, which, according to the above cases, leads to a contradiction.

Hence, in all cases, our assumption leads to a contradiction, and this proves the theorem. □

We wish to point out that Theorem 5.14 does not only demonstrate the correctness of Conjecture 5.8 for the given class of patterns, but additionally provides an efficient way of finding an unambiguous morphism $\sigma_{i,j}$. For example, we can immediately conclude from it that $\sigma_{1,4}$ is unambiguous with respect to our above example pattern α_1 (see page 68). Furthermore, the theorem also holds for patterns with less than four different variables.

We now consider those patterns that are not a fixed point and, moreover, contain all of their variables exactly twice (note that some of these "shortest" patterns that are not fixed points are also studied in Theorem 5.5). We wish to demonstrate that Theorem 5.14 implies the existence of an unambiguous $\sigma_{i,j}$ for *every* such pattern. This insight is based on the following lemma:

Lemma 5.15. *Let $\alpha \in \mathbb{N}^+$ be a pattern with $|\operatorname{var}(\alpha)| > 6$ and, for every $x \in \operatorname{var}(\alpha)$, $|\alpha|_x = 2$. Then there exist $i, j \in \operatorname{var}(\alpha)$, $i \neq j$, such that*

- *there is no $k \in \operatorname{var}(\alpha)$ with $\{i, j\} \subseteq L_k$ or $\{i, j\} \subseteq R_k$, and*
- *$\alpha \neq \alpha_1 \cdot \mathbf{i} \cdot \mathbf{j} \cdot \alpha_2 \cdot \mathbf{j} \cdot \mathbf{i} \cdot \alpha_3$, $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{N}^*$.*

Proof. Let $n := |\operatorname{var}(\alpha)|$. Since every variable occurs exactly twice in α, it directly follows that, for every $x \in \operatorname{var}(\alpha)$, $|R_x| \leq 2$ and $|L_x| \leq 2$. By omitting the neighbourhood sets containing ε, we have at most $2n-2$ sets of size 2. Besides, it can be verified with little effort that α contains at most $n-1$ different factors $i \cdot j$, $i, j \in \operatorname{var}(\alpha)$, $i \neq j$, such that $j \cdot i \sqsubseteq \alpha$ (e. g., for $n := 4$, $\alpha := 1 \cdot 2 \cdot 3 \cdot 4 \cdot 4 \cdot 3 \cdot 2 \cdot 1$ has 3 different factors $i \cdot j$, $i, j \in \operatorname{var}(\alpha)$, $i \neq j$, satisfying $j \cdot i \sqsubseteq \alpha$). Assume to the contrary that, for every $i, j \in \operatorname{var}(\alpha)$, one of the following cases is satisfied:

- there exists $k \in \operatorname{var}(\alpha)$ with $\{i, j\} \subseteq L_k$ or $\{i, j\} \subseteq R_k$, or
- $\alpha = \alpha_1 \cdot \mathbf{i} \cdot \mathbf{j} \cdot \alpha_2 \cdot \mathbf{j} \cdot \mathbf{i} \cdot \alpha_3$, $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{N}^*$.

As mentioned above, the maximum number of pairs that are covered by the first case is $2n-2$, and for the second case it is $n-1$. On the other hand, since $|\operatorname{var}(\alpha)| = n$, there exist $\binom{n}{2}$ different pairs of variables. However, for $n > 6$, we have

$$\binom{n}{2} > (2n-2) + (n-1),$$

which contradicts the assumption. $\square$

Hence, whenever a pattern α is not a fixed point, the conditions of Theorem 5.14 are automatically satisfied if α contains at least seven distinct variables and all of its variables occur exactly twice. Using a less elegant reasoning than the one on Lemma 5.15, we can extend this insight to all such patterns over at least four distinct variables. This yields the following result:

Theorem 5.16. *Let $\alpha \in \mathbb{N}^+$ be a pattern with $|\operatorname{var}(\alpha)| > 3$ and, for every $x \in \operatorname{var}(\alpha)$, $|\alpha|_x = 2$. If α is not a fixed point of a nontrivial morphism, then there exist $i, j \in \operatorname{var}(\alpha)$, $i \neq j$, such that $\sigma_{i,j}$ is unambiguous with respect to α.*

Proof. Let $n := |\operatorname{var}(\alpha)|$. For $n > 6$, it directly follows from Theorem 5.14 and Lemma 5.15 that Theorem 5.16 is satisfied. Hence, we consider the following cases:

- $|\operatorname{var}(\alpha)| = 4$. The only patterns that do not satisfy the conditions of Theorem 5.14 are:

$$\alpha_1 := 1 \cdot 2 \cdot 3 \cdot 4 \cdot 4 \cdot 1 \cdot 3 \cdot 2,\ \alpha_2 := 1 \cdot 2 \cdot 3 \cdot 4 \cdot 4 \cdot 2 \cdot 1 \cdot 3,$$
$$\alpha_3 := 1 \cdot 2 \cdot 3 \cdot 4 \cdot 2 \cdot 1 \cdot 4 \cdot 3,\ \alpha_4 := 1 \cdot 2 \cdot 3 \cdot 3 \cdot 4 \cdot 4 \cdot 2 \cdot 1,$$
$$\alpha_5 := 1 \cdot 2 \cdot 3 \cdot 3 \cdot 4 \cdot 1 \cdot 4 \cdot 2,\ \alpha_6 := 1 \cdot 2 \cdot 3 \cdot 3 \cdot 1 \cdot 4 \cdot 2 \cdot 4,$$
$$\alpha_7 := 1 \cdot 2 \cdot 3 \cdot 3 \cdot 4 \cdot 2 \cdot 1 \cdot 4,\ \alpha_8 := 1 \cdot 2 \cdot 3 \cdot 2 \cdot 4 \cdot 4 \cdot 1 \cdot 3,$$
$$\alpha_9 := 1 \cdot 2 \cdot 1 \cdot 3 \cdot 4 \cdot 4 \cdot 2 \cdot 3,\ \alpha_{10} := 1 \cdot 2 \cdot 3 \cdot 1 \cdot 4 \cdot 4 \cdot 3 \cdot 2.$$

  It can be verified with little effort that

  - $\sigma_{3,4}$ is unambiguous with respect to α_1, α_2, α_5, α_9 and α_{10},
  - $\sigma_{2,3}$ is unambiguous with respect to α_3, α_6 and α_7,
  - $\sigma_{1,4}$ is unambiguous with respect to α_4, α_8.

- $|\operatorname{var}(\alpha)| \in \{5, 6\}$. Assume to the contrary that for every $i, j \in \operatorname{var}(\alpha)$, $i \neq j$, $\sigma_{i,j}$ is ambiguous with respect to α. This implies that the conditions of Theorem 5.14 are not satisfied. Consequently, for every $i, j \in \operatorname{var}(\alpha)$, one of the following cases is satisfied:

  - there is a $k \in \operatorname{var}(\alpha)$ with $\{i, j\} \subseteq L_k$ or $\{i, j\} \subseteq R_k$, or
  - $\alpha = \cdots i \cdot j \cdots j \cdot i \cdots$.

  It directly follows from the proof of Lemma 5.15 that, if $\operatorname{var}(\alpha) = n$, then the maximum number of pairs of variables satisfying the first case is $2n - 2$. On the other hand, the number of different pairs i, j which must satisfy the above cases is $\binom{n}{2}$, consequently, for any n, $n \geq 5$, there exist

$$\binom{n}{2} - (2n - 2)$$

  pairs which must satisfy the second case. So, for $n = 5$, since

$$\binom{5}{2} - (2 * 5 - 2) = 2,$$

  there exist at least two different pairs of i, j satisfying $\alpha = \cdots i \cdot j \cdots j \cdot i \cdots$. For $n = 6$, that amount increases to 5, due to:

$$\binom{6}{2} - (2 * 6 - 2) = 5.$$

By investigating the all patterns α with $\text{var}(\alpha) = 5$ which are containing 2 different pairs of i, j that satisfy the second case, we can conclude that there exist $i, j \in \text{var}(\alpha)$, $i \neq j$, such that $\sigma_{i,j}$ is unambiguous with respect to α. Moreover, the only pattern α with $|\text{var}(\alpha)| = 6$ that is not a fixed point of a nontrivial morphism and contains 5 different pairs of i, j satisfying the second case is $\alpha = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$, with respect to which there exists an unambiguous 1-uniform morphisms $\sigma_{1,6}$.

Hence, in both cases, the results contradict the assumption. □

Theorem 5.16 does not only directly prove the correctness of Conjecture 5.8 for all patterns that contain all their variables exactly twice, but it also allows a large set of patterns to be constructed for which the conjecture holds true as well. This construction is specified as follows:

Theorem 5.17. *Let $\alpha := \alpha_1 \cdot \beta \cdot \alpha_2$ and $\gamma := \alpha_1 \cdot \alpha_2$ be patterns with $\alpha_1, \alpha_2, \beta \in \mathbb{N}^*$, such that*

- *γ and β are not a fixed point of a nontrivial morphism,*
- *$|\text{var}(\gamma)| > 3$ and, for every $x \in \text{var}(\gamma)$, $|\gamma|_x = 2$, or $|\text{var}(\beta)| > 3$ and, for every $x \in \text{var}(\beta)$, $|\beta|_x = 2$, and*
- *$\text{var}(\gamma) \cap \text{var}(\beta) = \emptyset$.*

Then there exist $i, j \in \text{var}(\alpha)$, $i \neq j$, such that $\sigma_{i,j}$ is unambiguous with respect to α.

Proof. Assume that $|\text{var}(\gamma)| > 3$ and, for every $x \in \text{var}(\gamma)$, $|\gamma|_x = 2$. So, since γ satisfies the conditions of Theorem 5.16, there exist $i, j \in \text{var}(\gamma)$, $i \neq j$, such that $\sigma_{i,j}$ with target alphabet Σ_1 is unambiguous with respect to γ. Also, due to β not being a fixed point of a nontrivial morphism, there is an unambiguous 1-uniform morphism $\sigma : \mathbb{N}^* \to \Sigma_2^*$, $|\Sigma_2| = |\text{var}(\beta)|$, with respect to β. Let $\Sigma_1 \cap \Sigma_2 := \emptyset$.

We now assume to the contrary that $\sigma_{i,j}$ with target alphabet $\Sigma_1 \cup \Sigma_2$ is ambiguous with respect to α. This implies that there is a morphism $\tau : \mathbb{N}^* \to \Sigma^*$ satisfying $\tau(\alpha) = \sigma_{i,j}(\alpha)$ and, for some $q \in \text{var}(\alpha)$, $\tau(q) \neq \sigma(q)$.

Claim 1. There does not exist an $x \in \text{var}(\alpha)$ satisfying $|\tau(x)| \geq 2$ and $v_1 v_2 \sqsubseteq \tau(x)$, $v_1 \in \Sigma_1$ and $v_2 \in \Sigma_2$, or $v_1 \in \Sigma_2$ and $v_2 \in \Sigma_1$.

Proof (Claim 1). Assume to the contrary that there is an $x \in \text{var}(\alpha)$ such that $|\tau(x)| \geq 2$ and $v_1 v_2 \sqsubseteq \tau(x)$, $v_1 \in \Sigma_1$ and $v_2 \in \Sigma_2$, or $v_1 \in \Sigma_2$ and $v_2 \in \Sigma_1$. Since x occurs at least twice in α, $\tau(\alpha) = \cdots \cdot v_1 v_2 \cdot \cdots \cdot v_1 v_2 \cdot \cdots$. However, because of $\alpha := \alpha_1 \cdot \beta \cdot \alpha_2$ and $\text{var}(\gamma) \cap \text{var}(\beta) = \emptyset$, this contradicts $\sigma_{i,j}(\alpha) = \tau(\alpha)$. □(Claim 1)

Claim 2. There exists an $x \in \text{var}(\beta)$ such that $\tau(x) \in \Sigma_1^+$.

Proof (Claim 2). Assume to the contrary that, for every $x \in \text{var}(\beta)$, $\tau(x) \notin \Sigma_1^+$. Due to Claim 1, it results from $\tau(\alpha) = \sigma_{i,j}(\alpha)$ and $\sigma_{i,j}$ being unambiguous with respect to γ and β that there exist some $x' \in \text{var}(\gamma)$ such that $\tau(x') \in \Sigma_2^+$. Let $A \subseteq \text{var}(\gamma)$ be the set of all variables x' with $\tau(x') \in \Sigma_2^+$. We can now define a morphism $\sigma' : \mathbb{N}^* \to \Sigma_1^*$ such that, for every $k \in \text{var}(\gamma) \setminus A$, $\sigma'(k) = \tau(k)$ and, for every $x' \in A$, $\sigma'(x') = \varepsilon$. Consequently, due to the fact that there is no $k \in \text{var}(\gamma)$ with $\tau(k) \in (\Sigma_1 \cup \Sigma_2)^* \setminus (\Sigma_1^* \cup \Sigma_2^*)$, $\sigma'(\gamma) = \sigma_{i,j}(\alpha)$, which means that $\sigma_{i,j}$ is ambiguous with respect to γ. This is a contradiction. □(Claim 2)

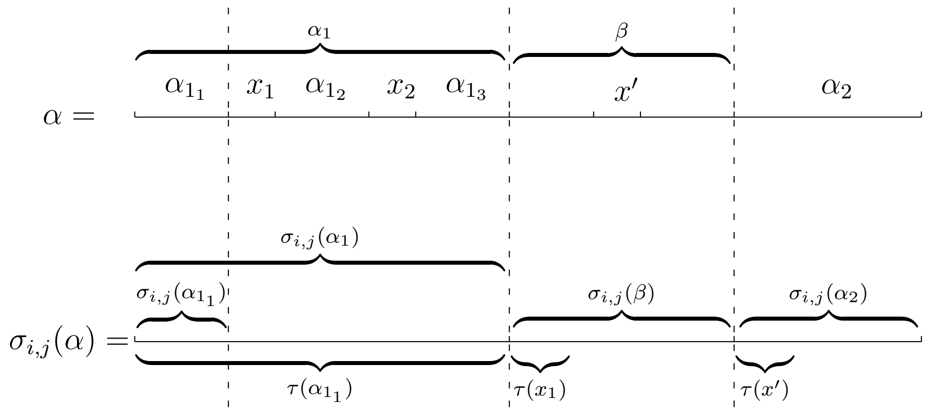
Claim 3. There exists an $x \in \text{var}(\gamma)$ satisfying $\tau(x) \in \Sigma_2^+$.

Proof (Claim 3). Assume to the contrary that, for every $x \in \text{var}(\gamma)$, $\tau(x) \notin \Sigma_2^+$. Because of Claim 1, $\tau(\alpha) = \sigma_{i,j}(\alpha)$ and $\sigma_{i,j}$ being unambiguous with respect to γ and β imply that there exists a nonempty set $A \subseteq \text{var}(\beta)$ such that, for every $x' \in A$, $\tau(x') \in \Sigma_1^+$. We can now define a morphism $\sigma' : \mathbb{N}^* \to \Sigma_2^*$ such that, for every $k \in \text{var}(\beta) \setminus \{x'\}$, $\sigma'(k) = \tau(k)$ and, for every $x' \in A$, $\sigma'(x') = \varepsilon$. Consequently, due to the fact that there is no $k \in \text{var}(\beta)$ with $\tau(k) \in (\Sigma_1 \cup \Sigma_2)^* \setminus (\Sigma_1^* \cup \Sigma_2^*)$, $\sigma'(\beta) = \sigma(\beta)$, which contradicts σ being unambiguous with respect to β. □(Claim 3)

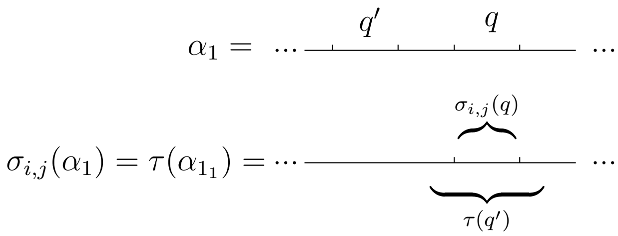
Claim 4. If $|\tau(q)| \geq 2$, $q \in \text{var}(\gamma)$, and $\tau(q) \in \Sigma_1^+$, then $\sigma_{i,j}(i) \sqsubseteq \tau(q)$.

Proof (Claim 4). Assume to the contrary that $\sigma_{i,j}(i) \not\sqsubseteq \tau(q)$. Let $v_1v_2 \sqsubseteq \tau(q)$, $v_1, v_2 \in \Sigma_1 \setminus \{\sigma_{i,j}(i)\}$. Due to $|\gamma|_q = 2$, $\tau(\alpha) = \cdots \cdot v_1v_2 \cdot \cdots \cdot v_1v_2 \cdot \cdots$. Since $\Sigma_1 \cap \Sigma_2 := \emptyset$ and $\tau(\alpha) = \sigma_{i,j}(\alpha)$, we can conclude that $\gamma = \cdots \cdot x_1x_2 \cdot \cdots \cdot x_1x_2 \cdot \cdots$, $x_1, x_2 \in \text{var}(\gamma) \setminus \{i, j\}$. Because of $|\gamma|_{x_1} = 2$ and $|\gamma|_{x_2} = 2$, Lemma 5.13 implies that γ is a fixed point of a nontrivial morphism, which contradicts the assumption. □(Claim 4)

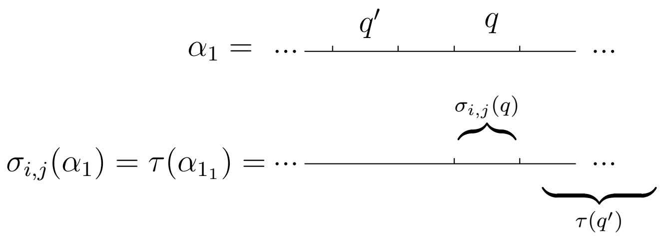
According to Claims 1, 2 and 3, there exists an $x \in \text{var}(\gamma)$ such that $\tau(x) \in \Sigma_2^+$, and there exists an $x' \in \text{var}(\beta)$ with $\tau(x') \in \Sigma_1^+$. The two occurrences of x are both in α_1 or both in α_2; otherwise, there does not exist an $x' \in \text{var}(\beta)$ such that $\tau(x') \in \Sigma_1^+$. Without loss of generality, we assume that both occurrences of x are in α_1, and we also assume that x is the leftmost variable in α_1 satisfying $\tau(x) \in \Sigma_2^+$ and x' is the leftmost variable in β with $\tau(x') \in \Sigma_1^+$. Let x_1 be the first occurrence of x, and let x_2 be the second occurrence of x. So, $\alpha_1 = \alpha_{1_1} \cdot x_1 \cdot \alpha_{1_2} \cdot x_2 \cdot \alpha_{1_3}$, $\alpha_{1_1}, \alpha_{1_2}, \alpha_{1_3} \in \mathbb{N}^*$. Consequently, $\tau(\alpha_{1_1}) = \sigma_{i,j}(\alpha_1)$.



Before we proceed with our proof, we define two notations. If, for variables q, q' in α_1 (q and q' have a same position or q' occurs to the left of q in α_1) $\sigma_{i,j}(q) \sqsubseteq \tau(q')$ and $\tau(q')$ in $\tau(\alpha_{1_1})$ is located at the position of $\sigma_{i,j}(q)$ in $\sigma_{i,j}(\alpha_1)$, then we write $\sigma_{i,j}(q) \downarrow \tau(q')$. This is illustrated by the following diagram (where we assume that the occurrence of q' is to the left of the occurrence of q):



If the position of $\tau(q')$ in $\tau(\alpha_{1_1})$ is located to the right of the position of $\sigma_{i,j}(q)$ in $\sigma_{i,j}(\alpha_1)$, then we write $\sigma_{i,j}(q) \mapsto \tau(q')$. We again give a diagram (assuming that the occurrence of q' is to the left of the occurrence of q) that illustrates the setting where we use this notation:

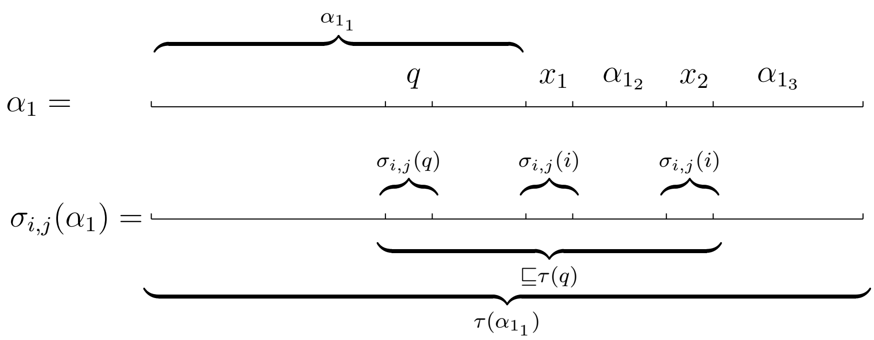


We return to our proof and recollect that $\alpha_1 = \alpha_{1_1} \cdot x_1 \cdot \alpha_{1_2} \cdot x_2 \cdot \alpha_{1_3}$, $x_1 = x_2 = x$, and $\tau(\alpha_{1_1}) = \sigma_{i,j}(\alpha_1)$. This implies that we have to consider the following cases:

Case 1. $x = i$ or $x = j$

Due to $\tau(\alpha_{1_1}) = \sigma_{i,j}(\alpha_1)$, one of the following cases holds true:

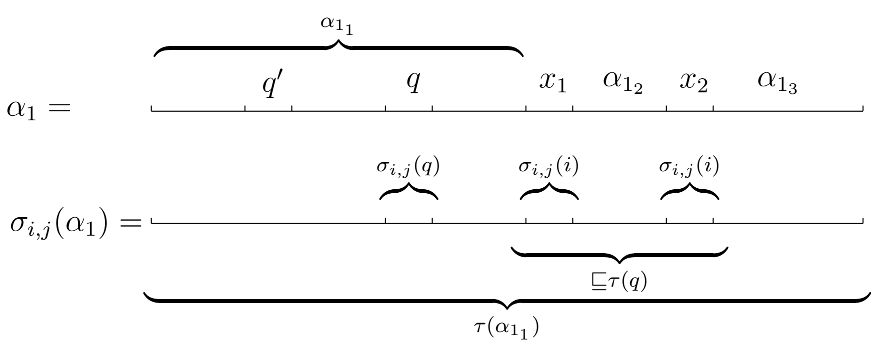
Case 1.1. There exists a variable $q \in \mathrm{var}(\alpha_{1_1})$ to the left of x_1 satisfying $|\tau(q)|_{\sigma_{i,j}(i)} \geq 2$ and $\sigma_{i,j}(i) \downarrow \tau(q)$.
Assume that $\sigma_{i,j}(q) \downarrow \tau(q)$.



Let A be a set of those variables $k \in \mathrm{var}(\gamma) \setminus \{q\}$ satisfying $\sigma_{i,j}(k) \sqsubseteq \tau(q)$. We define a morphism $\sigma' : \mathbb{N}^* \to \Sigma_1^*$ such that, for every $k' \in \mathrm{var}(\gamma)$,

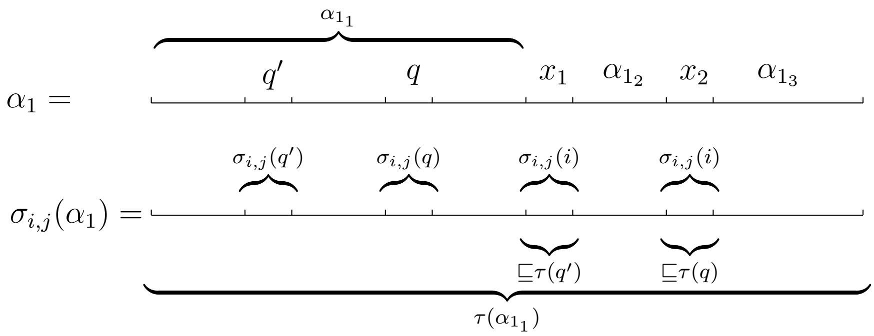
$$\sigma'(k') := \begin{cases} \varepsilon, & \text{if } k' \in A, \\ \tau(q), & \text{if } k' = q, \\ \sigma_{i,j}(k'), & \text{else.} \end{cases}$$

Due to the facts that, for all $k, k' \in \mathrm{var}(\gamma)$, $k \neq k'$, $|\gamma|_k = 2$, and if $k \neq i$ and $k' \neq j$, then $\sigma_{i,j}(k) \neq \sigma_{i,j}(k')$, it can be verified that $\sigma'(\gamma) = \sigma_{i,j}(\gamma)$, which is a contradiction to $\sigma_{i,j}$ being unambiguous with respect to γ.
If $\sigma_{i,j}(q) \mapsto \tau(q)$, then, due to $\tau(\alpha_{1_1}) = \sigma_{i,j}(\alpha_1)$, there exists a variable $q' \in \mathrm{var}(\alpha_{1_1})$ to the left of q satisfying $|\tau(q')| \geq 2$.



According to Claim 4, $\sigma_{i,j}(i) \sqsubseteq \tau(q')$. Besides, $|\gamma|_{q'} = 2$. On the other hand, $|\gamma|_q = 2$ and we assume $|\tau(q)|_{\sigma_{i,j}(i)} \geq 2$ in the present case. Consequently, $|\tau(\alpha)|_{\sigma_{i,j}(i)} > 4$, which contradicts $\tau(\alpha) = \sigma_{i,j}(\alpha)$.

Case 1.2. There exist variables $q, q' \in \mathrm{var}(\alpha_{1_1})$ to the left of x_1 satisfying $\sigma_{i,j}(i) \downarrow \tau(q)$ and $\sigma_{i,j}(i) \downarrow \tau(q')$.



Therefore, due to $\tau(\alpha_{1_1}) = \sigma_{i,j}(\alpha_1)$ and $\tau(x) \in \Sigma_2^+$, we can conclude that $\sigma_{i,j}(q) \mapsto \tau(q)$. If $\sigma_{i,j}(q') \downarrow \tau(q')$, then $\sigma_{i,j}(q) \downarrow \tau(q')$. This implies that $\sigma_{i,j}(q') \cdot w \cdot \sigma_{i,j}(q) \cdot w' \cdot \sigma_{i,j}(i) \sqsubseteq \tau(q')$, $w, w' \in \Sigma_1^*$. Due to $|\gamma|_{q'} = 2$, it can be verified that $\gamma = \gamma_1 \cdot q' \cdot \gamma_2 \cdot q \cdot \gamma_3 \cdot q' \cdot \gamma_2 \cdot q \cdot \gamma_4$ with $\gamma_1, \gamma_2, \gamma_3, \gamma_4 \in \mathbb{N}^*$ and $\sigma_{i,j}(\gamma_2) = w$. Without loss of generality, we assume that $x = i$. This implies that $q \neq i$, $q' \neq i$ and $i \notin \mathrm{var}(\gamma_2)$. Also, for every $k \in \mathrm{var}(\gamma)$, $|\gamma|_k = 2$. Consequently, $(\{q, q'\} \cup \mathrm{var}(\gamma_2)) \cap (\mathrm{var}(\gamma_1) \cup \mathrm{var}(\gamma_3) \cup \mathrm{var}(\gamma_4)) = \emptyset$. So, the structure of γ satisfies Lemma 5.13, which implies that γ is a fixed point of a nontrivial morphism. This is a contradiction. As a result, $\sigma_{i,j}(q') \mapsto \tau(q')$; in addition, as mentioned, $\sigma_{i,j}(q) \mapsto \tau(q)$. Therefore, and again because of $\tau(\alpha_{1_1}) = \sigma_{i,j}(\alpha_1)$, there exists a variable $q'' \in \mathrm{var}(\alpha_{1_1})$ to the left of q' satisfying $|\tau(q'')| \geq 2$. According to Claim 4, $\sigma_{i,j}(i) = \sigma_{i,j}(j) \sqsubseteq \tau(q'')$. Without loss of generality, assume that $x = i$. Hence, it results from $\sigma_{i,j}(j) \sqsubseteq \tau(q'')$, $|\tau(q'')| \geq 2$ and $|\gamma|_{q''} = 2$ that there is a factor $k \cdot j \sqsubseteq \gamma$ or $j \cdot k \sqsubseteq \gamma$, $k \in \mathrm{var}(\alpha_1)$, $k \neq i$ and $k \neq j$, which occurring twice in γ. Consequently, we can assume $\gamma = \gamma_1 \cdot k \cdot j \cdot \gamma_2 \cdot k \cdot j \cdot \gamma_3$ or $\gamma = \gamma_1 \cdot j \cdot k \cdot \gamma_2 \cdot j \cdot k \cdot \gamma_3$ where $\gamma_1, \gamma_2, \gamma_3 \in \mathbb{N}^*$ and $k, j \notin \mathrm{var}(\gamma_1\gamma_2\gamma_3)$. According to Lemma 5.13, this implies that γ is a fixed point of a nontrivial morphism, which is a contradiction.

Case 2. $x \neq i$ and $x \neq j$.

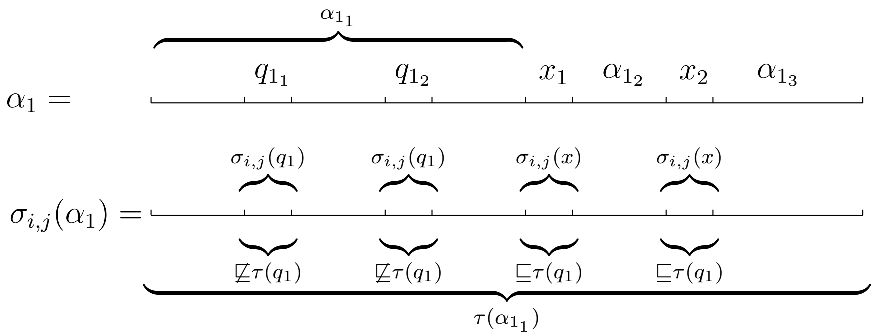
Since $\tau(\alpha_{1_1}) = \sigma_{i,j}(\alpha_1)$, one of the following cases holds true:

Case 2.1. There exists a variable $q \in \mathrm{var}(\alpha_{1_1})$ to the left of x_1 satisfying $|\tau(q)|_{\sigma_{i,j}(x)} = 2$. Since $|\gamma|_q = 2$, $|\tau(\alpha)|_{\sigma_{i,j}(x)} > 2$, which contradicts $\tau(\alpha) = \sigma_{i,j}(\alpha)$.

Case 2.2. There exist variables $q, q' \in \mathrm{var}(\alpha_{1_1})$, $q \neq q'$, to the left of x_1 satisfying $\sigma_{i,j}(x) \downarrow \tau(q)$ and $\sigma_{i,j}(x) \downarrow \tau(q')$. It results from $|\gamma|_q = 2$ and $|\gamma|_{q'} = 2$ that $|\tau(\alpha)|_{\sigma_{i,j}(x)} > 2$, which is a contradiction to $\tau(\alpha) = \sigma_{i,j}(\alpha)$.

Case 2.3. There exists a variable $q_1 \in \mathrm{var}(\alpha_{1_1})$, with two occurrences named q_{1_1} and q_{1_2}, to the left of x_1 satisfying $|\tau(q_1)|_{\sigma_{i,j}(x)} = 1$, $\sigma_{i,j}(x_1) \downarrow \tau(q_{1_1})$ and

$\sigma_{i,j}(x_2) \downarrow \tau(q_{1_2})$. Due to $\tau(\alpha_{1_1}) = \sigma_{i,j}(\alpha_1)$, $\tau(x) \in \Sigma_2^+$ and the two occurrences of q_1 being to the left of x_1, we can conclude that $\sigma_{i,j}(q_1) \mapsto \tau(q_1)$.



We first demonstrate that the overall condition of Case 2 does not only hold for x, but also for q_1:

Claim 5. $q_1 \neq i$ and $q_1 \neq j$.

Proof (Claim 5). Assume to the contrary that $q_1 = i$ or $q_1 = j$. Without loss of generality let $q_1 := i$. Thus, $q_{1_1} = q_{1_2} = i$. On the other hand, as mentioned, $\sigma_{i,j}(q_1) \mapsto \tau(q_1)$. Thus, again because of $\tau(\alpha_{1_1}) = \sigma_{i,j}(\alpha_1)$, there exists a variable $k \in \text{var}(\alpha_{1_1})$ to the left of q_{1_1} satisfying $|\tau(k)| \geq 2$. According to Claim 4, $\sigma_{i,j}(j) \sqsubseteq \tau(k)$. This implies that due to $|\gamma|_k = 2$ there is a factor $k' \cdot j \sqsubseteq \gamma$ or $j \cdot k' \sqsubseteq \gamma$, $k' \in \text{var}(\alpha_1)$, $k' \neq i$ and $k' \neq j$, which occurs twice in γ. Consequently, we can assume $\gamma = \gamma_1 \cdot k' \cdot j \cdot \gamma_2 \cdot k' \cdot j \cdot \gamma_3$ or $\gamma = \gamma_1 \cdot j \cdot k' \cdot \gamma_2 \cdot j \cdot k' \cdot \gamma_3$, where $\gamma_1, \gamma_2, \gamma_3 \in \mathbb{N}^*$ and $k', j \notin \text{var}(\gamma_1\gamma_2\gamma_3)$. According to Lemma 5.13, this implies that γ is a fixed point of a nontrivial morphism, which is a contradiction.
If we assume to the contrary that $q_1 = j$, then the same reasoning as above leads to a contradiction. □(Claim 5)

The following statement shall be the core argument of our reasoning on Case 2.3.

Claim 6. There exists a variable to the left of q_{1_1} in α_{1_1} satisfying the condition of Case 2.3.

Proof (Claim 6). According to Claim 5, $q_1 \neq i$ and $q_1 \neq j$. Besides, as mentioned in Case 2.3, $\sigma_{i,j}(q_1) \mapsto \tau(q_1)$. Consequently, applying Case 2 leads to the existence of a variable q_2 to the left of q_{1_1} satisfying $\sigma_{i,j}(q_1) \downarrow \tau(q_2)$. However, a same reasoning as in Cases 2.1, 2.2 (considering q_1 instead of x) leads to a contradiction. As a result, q_2 must satisfy the condition of Case 2.3. □(Claim 6)

Therefore, according to Claim 6 and Case 2.3, there exists a $q_2 \in \alpha_{1_1}$ with two occurrences named q_{2_1} and q_{2_2}, to the left of q_{1_1} with $|\tau(q_2)|_{\sigma(i,j)(q_1)} = 1$ and $\sigma_{i,j}(q_2) \mapsto \tau(q_2)$. Furthermore, due to a same reasoning as in Claim 5, $q_2 \neq i$ and

$q_2 \neq j$. Hence, we can again apply Claim 6. Consequently, this reasoning finally leads to a contradiction based on Case 2.1 or 2.2 since the length of α_1 is finite, which means that, by a continued application of Claim 6, there is a $q_n \in \mathrm{var}(\alpha_{1_1})$ not satisfying Case 2.3.

Now, assume the case that $|\,\mathrm{var}(\beta)| > 3$ and, for every $x \in \mathrm{var}(\beta)$, $|\beta|_x = 2$. It can be verified that this case satisfies Claims 1, 2 and 3. Consequently, using an analogous reasoning as previous case leads to a contradiction again.

Hence, there is no morphism τ satisfying $\tau(\alpha) = \sigma_{i,j}(\alpha)$ and $\tau(x) \neq \sigma_{i,j}(x)$, for an $x \in \mathrm{var}(\alpha)$, and this implies that $\sigma_{i,j}$ is unambiguous with respect to α. □

In order to illustrate the above statement, we consider the following example. Let

$$\alpha := 1 \cdot 2 \cdot 1 \cdot 3 \cdot 2 \cdot 3 \cdot 2 \cdot 4 \cdot 5 \cdot 6 \cdot 7 \cdot 5 \cdot 7 \cdot 8 \cdot 6 \cdot 8 \cdot 4 \cdot 2 \cdot 9 \cdot 3 \cdot 9 \cdot 2.$$

We now define

$$\begin{aligned} \alpha_1 &:= 1 \cdot 2 \cdot 1 \cdot 3 \cdot 2 \cdot 3 \cdot 2 \cdot 4, \\ \alpha_2 &:= 4 \cdot 2 \cdot 9 \cdot 3 \cdot 9 \cdot 2, \\ \beta &:= 5 \cdot 6 \cdot 7 \cdot 5 \cdot 7 \cdot 8 \cdot 6 \cdot 8, \end{aligned}$$

which implies $\alpha = \alpha_1 \cdot \beta \cdot \alpha_2$. Using Theorem 3.20 and the definition of prolix patterns, it can be effortlessly verified that both β and $\gamma = \alpha_1 \cdot \alpha_2$ are not a fixed point of a nontrivial morphism. Furthermore, β contains four different variables, and every $x \in \mathrm{var}(\beta)$ satisfies $|\beta|_x = 2$. Therefore, we can apply Theorem 5.17, which says that there are $i, j \in \mathrm{var}(\alpha)$ such that $\sigma_{i,j}$ is unambiguous with respect to α; from the proofs of Theorems 5.14, 5.16 and 5.17, we can conclude that, for example, $i := 5$ and $j := 7$ are a suitable choice for the definition of $\sigma_{i,j}$.

In the remainder of this section, we shall not directly address the morphism $\sigma_{i,j}$ any longer. Hence, we focus on Conjecture 5.6, and we use an approach that differs quite significantly from those above: we consider words that cannot be morphic images of a pattern under any ambiguous 1-uniform morphism, and we construct suitable morphic preimages from these words. This method yields another major set of patterns for which Conjectures 5.6 and 5.8 are satisfied.

Our corresponding technique is based on the well-known concept of *de Bruijn sequences*. Since de Bruijn sequences are cyclic, which does not fit with our subject, we introduce a non-cyclic variant:

Definition 5.18. *A* non-cyclic De Bruijn sequence *(of order n) is a word over a given alphabet Σ (of size k) for which all possible words of length n in Σ^* appear* *exactly once as factors of this sequence. We denote the set of all non-cyclic De Bruijn sequences of order n by $B'(k, n)$. A $w \in B'(k, n)$ is said to be* in canonical form *if it is lexicographically minimal (with regard to any fixed order on Σ) among all its renamings in $B'(k, n)$.*

For example, the word $w_0 := aabacbbcca$ is a non-cyclic de Bruijn sequence in $B'(3, 2)$ if we assume $\Sigma := \{a, b, c\}$. Furthermore, w_0 is in canonical form if we assume Σ to be ordered alphabetically. The introduction of a canonical form is needed at the end of this section, where we shall provide a lower bound on the number of patterns with unambiguous 1-uniform morphisms that can be derived from de Bruijn sequences.

It can now be easily understood that a non-cyclic de Bruijn sequence cannot be a morphic image of any pattern under ambiguous 1-uniform morphisms:

Theorem 5.19. *Let Σ be an alphabet, and let $\alpha \in \mathbb{N}^+$ be a pattern satisfying, for every $x \in \text{var}(\alpha)$, $|\alpha|_x \geq 2$. Let $\sigma : \mathbb{N}^* \to \Sigma^*$ be a 1-uniform morphism such that, for every $u_1u_2 \sqsubseteq \sigma(\alpha)$, $u_1, u_2 \in \Sigma$, the factor u_1u_2 occurs in $\sigma(\alpha)$ exactly once. Then σ is unambiguous with respect to α.*

Proof. Assume to the contrary that σ is ambiguous with respect to α. Consequently, there exists a morphism $\tau : \mathbb{N}^* \to \Sigma^*$ satisfying $\tau(\alpha) = \sigma(\alpha)$ and, for some $q \in \text{var}(\alpha)$, $\tau(q) \neq \sigma(q)$. Since σ is a 1-uniform morphism, there exists a $q \in \text{var}(\alpha)$ satisfying $|\tau(q)| \geq 2$. Hence, let $v_1v_2 \sqsubseteq \tau(q)$, $v_1, v_2 \in \Sigma$. Due to $|\alpha|_q \geq 2$, this implies that $\tau(\alpha) = \cdots \cdot v_1v_2 \cdot \cdots \cdot v_1v_2 \cdot \cdots$. However, this contradicts the condition of the theorem stating that, for every $u_1u_2 \sqsubseteq \sigma(\alpha)$, $u_1, u_2 \in \Sigma$, the factor u_1u_2 occurs in $\sigma(\alpha)$ exactly once. So, σ is unambiguous with respect to α. □

This insight implies that if a pattern can be mapped by a 1-uniform morphism to a de Bruijn sequence and has at least two occurrences of each of its variables, then this pattern necessarily is not a fixed point. Thus, for such patterns, Conjecture 5.6 holds true:

Corollary 5.20. *Let Σ be an alphabet, and let $\alpha \in \mathbb{N}^+$ be a pattern satisfying, for every $x \in \text{var}(\alpha)$, $|\alpha|_x \geq 2$. Let $\sigma : \mathbb{N}^* \to \Sigma^*$ be a 1-uniform morphism such that, for every $u_1u_2 \sqsubseteq \sigma(\alpha)$, $u_1, u_2 \in \Sigma$, the factor u_1u_2 occurs in $\sigma(\alpha)$ exactly once. Then α is not a fixed point of a nontrivial morphism.*

Proof. According to Theorem 5.19, σ is unambiguous with respect to α. Since σ, by definition, is nonerasing, the corollary directly follows from Theorem 3.17. □

We now show how we can construct patterns that satisfy the conditions of Theorem 5.19 and Corollary 5.20:

Definition 5.21. *Let $\Sigma := \{a_1, a_2, \dots, a_k\}$. Let $B'(k,2)$ be the set of non-cyclic de Bruijn sequences of order 2 over Σ. Then $\Pi_{DB}(k) \subseteq \mathbb{N}^*$ is the set of all patterns that can be constructed as follows: For every $w \in B'(k,2)$ and every letter a_j in w, all n_j occurrences of a_j are replaced by $\lfloor n_j/2 \rfloor$ different variables from a set $N_j := \{x_{j_1}, x_{j_2}, \dots, x_{j_{\lfloor n_j/2 \rfloor}}\} \subseteq \mathbb{N}$, such that the following conditions are satisfied:*

- *for every $x \in N_j$, $|\alpha|_x > 1$,*
- *for all i, i', $1 \leq i, i' \leq k$, with $i \neq i'$, $N_i \cap N_{i'} = \emptyset$, and*
- *for all i, $1 \leq i \leq k$, the variables in N_i are assigned to occurrences of a_i in a way such that the resulting pattern is in canonical form.*

For instance, with regard to our above example word $w_0 = aabacbbcca \in B'(3,2)$, Definition 5.21 says that, e. g., the pattern $1 \cdot 1 \cdot 2 \cdot 3 \cdot 4 \cdot 2 \cdot 2 \cdot 4 \cdot 4 \cdot 3$ is contained in $\Pi_{DB}(3)$.

From this construction, it directly follows that Conjecture 5.6 holds true for every pattern in $\Pi_{DB}(k)$:

Theorem 5.22. *Let $\Sigma := \{a_1, a_2, \dots, a_k\}$, $k \geq 3$. Then, for every $\alpha \in \Pi_{DB}(k)$,*

- $\mathrm{var}(\alpha)$ *contains at least $k+1$ elements, and*
- *there exists a 1-uniform morphism $\sigma : \mathbb{N}^* \to \Sigma^*$ that is unambiguous with respect to α.*

Proof. We begin this proof with the first statement of the theorem: It is obvious that there are k^2 different words of length 2 over Σ. The shortest word that contains k^2 factors of length 2 has length k^2+1, which means that this is the length of any word $w \in B'(k,2)$. Thus, there must be at least one letter in w that has at least $\lceil (k^2+1)/k \rceil$ occurrences. Since we assume $k \geq 3$, this means that this letter has at least 4 occurrences. From Definition 5.21 it then follows that this letter is replaced by at least two different variables when a pattern $\alpha \in \Pi_{DB}(k)$ is generated from w. Since all other letters in w must be replaced by at least one variable, this shows that $|\mathrm{var}(\alpha)| \geq k+1$. Note that from the proof of Theorem 5.23 it can be derived that, more precisely, $|\mathrm{var}(\alpha)| = (k-1)\lfloor k/2 \rfloor + \lfloor (k+1)/2 \rfloor$.

Concerning the second statement, we define σ by, for every j, $1 \leq j \leq k$, and for every $x \in N_j$, $\sigma(x) := a_j$. Thus, σ is 1-uniform, and $\sigma(\alpha) \in B'(k,2)$. This implies that, for every $u_1u_2 \sqsubseteq \sigma(\alpha)$, $u_1, u_2 \in \Sigma$, the factor u_1u_2 occurs in $\sigma(\alpha)$ exactly once. Consequently, according to Theorem 5.19, σ is unambiguous with respect to α. $\square$

We conclude this section with a statement on the cardinality of $\Pi_{DB}(k)$, demonstrating that the use of de Bruijn sequences indeed leads to a rich class of patterns α with unambiguous 1-uniform morphisms, and that these morphisms, in general, can even have a target alphabet of size much less than $\mathrm{var}(\alpha) - 1$ (as featured by Theorem 5.22):

Theorem 5.23. *Let $k \in \mathbb{N}$. Then $|\Pi_{DB}(k)| \geq k!^{(k-1)}$, and, for every $\alpha \in \Pi_{DB}(k)$,*

$$|\mathrm{var}(\alpha)| = (k-1)\lfloor k/2 \rfloor + \lfloor (k+1)/2 \rfloor .$$

Proof. Let $B(k,n)$ be the set of all distinct De Bruijn sequences of order n over alphabet Σ, and let $B'(k,n)$ be the set of all distinct non-cyclic De Bruijn sequences over Σ.

Claim 1. Every element of $B'(k,n)$ has length $k^n + n - 1$, and $|B'(k,n)| = k!^{k^{n-1}}$.

Proof (Claim 1). According to [7],

- every element of $B(k,n)$ has length k^n, and
- $|B(k,n)| = k!^{k^{n-1}}/k^n$.

Let $w \in B(k,n)$. Therefore, $|w| = k^n$. Assume that $w = a_1a_2[\ldots]a_m$, $m = k^n$. Since all words of length n over alphabet Σ appear exactly once in the cyclic sequence w, this implies that, for every v,

$$\begin{aligned} v \in \quad & \{a_{m-(n-2)}a_{m-(n-3)}[\ldots]a_m a_1, a_{m-(n-3)}a_{m-(n-4)}[\ldots]a_m a_1 a_2, [\ldots], \\ & a_m a_1 a_2[\ldots]a_{n-1}\}, \end{aligned}$$

$v \not\sqsubseteq w$. Consequently, by defining $w' := a_1a_2[\ldots]a_m a_1 a_2 \cdots a_{n-1}$, w' satisfies Definition 5.18, and as a result, $w' \in B'(k,n)$. Thus, $|w'| = |w| + (n-1)$, and this implies that, for every $w' \in B'(k,n)$,

$$|w'| = k^n + (n-1)$$

Besides, since w is a cyclic sequence, all words in

$$W := \{a_1a_2[\ldots]a_{k^n}, a_2a_3[\ldots]a_{k^n}a_1, \ldots, a_{k^n}a_1a_2, \ldots, a_{k^n-1}\}$$

are equivalent, and they are counted as one sequence of $B(k,n)$. Consequently, to find the number of distinct non-cyclic De Bruijn sequences $B'(k,n)$, it is sufficient to multiply $|W| = k^n$ to the number of distinct De Bruijn sequences $B(k,n)$. Thus,

$$|B'(k,n)| = k^n \, \frac{k!^{k^{n-1}}}{k^n} = k!^{k^{n-1}}.$$

□(Claim 1)

Now, let $B''(k, n)$ be the set of non-cyclic De Bruijn sequences in canonical form of order n.

Claim 2. $|B''(k, n)| = k!^{(k^{n-1}-1)}$

Proof (Claim 2). Let $\Sigma := \{a_1, a_2, \ldots, a_k\}$ and let $w \in B'(k, n)$. According to Definition 5.18, w is in canonical form if it is lexicographically minimal with regard to Σ, $a_1 < a_2 < \ldots < a_k$. However, by renaming w, it can be verified that there exist $k! - 1$ other sequences in $B'(k, n)$; in other words, we can consider w as a representative of $k!$ element of $B'(k, n)$. So, it directly follows from Claim 1 that the number of non-cyclic De Bruijn sequences in canonical form of order n over Σ is

$$\frac{k!^{k^{n-1}}}{k!} = k!^{(k^{n-1}-1)}.$$

□(Claim 2)

Consequently, according to Definition 5.21,

$$|\Pi_{DB}(k)| \geq k!^{(k-1)}.$$

We continue to prove the second part of Theorem 5.23 by the following claim:

Claim 3. Let $\Sigma := \{a_1, a_2, \ldots, a_k\}$. *Let* $B''(k, 2)$ *be the set of non-cyclic De Bruijn sequences in canonical form of order* 2 *over* Σ. *Then, for every* $w \in B''(k, 2)$, $|w|_{a_1} = k + 1$ *and, for every* j, $2 \leq j \leq k$, $|w|_{a_j} = k$.

Proof (Claim 3). Let a_i, $i \neq 1$, be an arbitrary element of Σ. According to Definition 5.18, for every $w \in B''(k, 2)$, $a_i a_1, a_i a_2, \ldots, a_i a_i, a_i a_{i+1}, \ldots, a_i a_k \sqsubseteq w$. Hence, without loss of generality regarding the order of letters in Σ , we can assume one of the following cases to be satisfied:

- $w = w_1 a_i a_1 \cdot w_2 a_i a_2 \cdot [\ldots] \cdot w_i a_i a_i \cdot w_{i+1} a_i a_{i+1} \cdot [\ldots] \cdot w_k a_i a_k \cdot w_{k+1}$, or
- $w = w_1 a_i a_1 \cdot w_2 a_i a_2 \cdot [\ldots] \cdot w_i a_i a_i a_{i+1} \cdot w_{i+1} a_i a_{i+2} \cdot [\ldots] \cdot w_{k-1} a_i a_k \cdot w_k$,

where, for every j, $1 \leq j \leq k + 1$, $w_j \in \Sigma^*$ and $a_i \not\sqsubseteq w_j$. Since $i \neq 1$ and w is in canonical form, then $w_1 \neq \varepsilon$.

In the first case, a_i occurs $k + 1$ times. Since $w_1 \neq \varepsilon$ and every word of length 2 over Σ appears exactly once in w, $|L_{a_i}| = k + 1$, $\varepsilon \notin L_{a_i}$. Consequently, we can conclude that there exist a sequence ua_i, $u \in \Sigma$, occurring more than once in w. This contradict the fact that $w \in B''(k, 2)$. Thus, in accordance with the second case, $|w|_{a_i} = k$. As a result, for every j, $2 \leq j \leq k$, $|w|_{a_j} = k$. Hence, for every $w \in B''(k, 2)$, $|w| - |w|_{a_1} = (k - 1)k$. On the other hand, Claim 1 implies that, for every $w \in B''(k, 2)$, $|w| = k^2 + 1$. This means that

$$|w|_{a_1} = (k^2 + 1) - ((k - 1)k) = k + 1.$$

□(Claim 3)

Consequently, according to Definition 5.21, for every $\alpha \in \Pi_{DB}(k)$, $k+1$ occurrences of a_1 are replaced by $\lfloor (k+1)/2 \rfloor$ different variables from N_1 and, for every j, $2 \leq j \leq k$, k occurrences of a_j are replaced by $\lfloor k/2 \rfloor$ different variables from N_j. Therefore,

$$|\operatorname{var}(\alpha)| = (k - 1)\lfloor k/2 \rfloor + \lfloor (k + 1)/2 \rfloor ,$$

and this proofs the theorem. □

Although we have established major sets of patterns for which Conjectures 5.6 and 5.8 hold true, we are unable to prove or refute these conjectures. However, we can point out that they show some connections to Problem 3.23. These shall be discussed in the next section.

5.3 Alphabet reductions and fixed points

We now turn our attention to Problem 3.23, i. e., we study whether there exists an alphabet reduction (i. e., a 1-uniform morphism that maps a given pattern to an image containing a smaller number of different variables) that maps a pattern that is not a fixed point to a pattern that is not a fixed point. Therefore, in contrast to the previous sections, we consider the set of natural number $\mathbb{N}$ both as domain and target alphabets of our morphisms. As an example of an alphabet reduction, we can mention $\phi_{i,j}$, that is defined by Definition 5.9, and we shall use this alphabet reduction in our next considerations many times.

We start with a general observation (which is a general case of Proposition 5.12), that links the research on ambiguity of morphisms to the question of whether a morphic image is not a fixed point of a nontrivial morphism:

Proposition 5.24. *Let $\alpha \in \mathbb{N}^+$. If $\phi : \mathbb{N}^* \to \mathbb{N}^*$ is unambiguous with respect to α then $\phi(\alpha)$ is not a fixed point of a nontrivial morphism.*

Proof. Assume to the contrary that $\phi(\alpha)$ is a fixed point of a nontrivial morphism. So, there must be a morphism $\psi : \mathbb{N}^* \to \mathbb{N}^*$ satisfying $\psi(\phi(\alpha)) = \phi(\alpha)$ and, for a variable u in $\phi(\alpha)$, $\psi(u) \neq u$. Consequently, we have the following relation:

$$\begin{tikzcd} \alpha \arrow[r, "\phi"] & \phi(\alpha) \arrow[loop above, "\psi"] \end{tikzcd}$$

We now define a morphism $\varphi : \mathbb{N}^* \to \mathbb{N}^*$ by $\varphi := \psi \circ \phi$. Hence,

$$\begin{aligned} \varphi(\alpha) &= \psi \circ \phi(\alpha) \\ &= \phi(\alpha). \end{aligned}$$

Since ψ is not an identity morphism, φ is different from ϕ. This contradicts the assumption of ϕ being unambiguous. Hence, $\phi(\alpha)$ is not a fixed point of a nontrivial morphism. $\square$

In general, the converse of the above proposition does not hold true. For example, let $\alpha := 1 \cdot 2 \cdot 3 \cdot 4 \cdot 4 \cdot 3 \cdot 2 \cdot 1$. Thus, $\phi_{1,2}(\alpha) = 1 \cdot 1 \cdot 3 \cdot 4 \cdot 4 \cdot 3 \cdot 1 \cdot 1$ that is not a fixed point of a nontrivial morphism. However, $\phi_{1,2}$ is ambiguous with respect to α, because we can define a morphism φ satisfying $\varphi(\alpha) = \phi_{1,2}(\alpha)$ by $\varphi(1) := \phi_{1,2}(1) \cdot \phi_{1,2}(1)$, $\varphi(2) := \varepsilon$, $\varphi(3) := \phi_{i,j}(3)$ and $\varphi(4) := \phi_{i,j}(4)$.

If Conjecture 5.8 is correct, then Problem 3.23 can be answered in the affirmative. This is a direct consequence of the following application of Proposition 5.24:

Corollary 5.25. *Let $\alpha \in \mathbb{N}^+$ and assume that there exist $i, j \in \mathrm{var}(\alpha)$, $i \neq j$, such that $\phi_{i,j}$ is unambiguous with respect to α. Then, $\alpha_{i,j}$ is not a fixed point of a nontrivial morphism.*

Proof. Directly from Proposition 5.24. $\square$

Hence, if Conjecture 5.8 is correct then it is stronger than Proposition 5.24.

The above approach does not only facilitate a direct application of our results in Section 5.2 on the existence of unambiguous 1-uniform morphisms to Problem 3.16, but it also has the advantage of providing a chance of a constructive method that might reveal which variables to map to the same image in an alphabet reduction in order to have both preimage and image not being a fixed point of a nontrivial morphism. However, since we are unable to prove Conjecture 5.8, we now present in Theorem 5.28 below a non-constructive answer to Problem 3.23. This is based on two lemmata, the first of which is a basic insight into fixed points of nontrivial morphisms:

Lemma 5.26. *Let α be a fixed point of a nontrivial morphism. Then there exists a nontrivial morphism $\phi : \mathrm{var}(\alpha)^* \to \mathrm{var}(\alpha)^*$ such that $\phi(\alpha) = \alpha$ and, for every $x \in \mathrm{var}(\alpha)$, if $\phi(x) \neq \varepsilon$, then $x \sqsubseteq \phi(x)$.*

Proof. According to Section 2.3, since α is a fixed point of a nontrivial morphism, there exists a factorisation $\alpha = \beta_0\gamma_1\beta_1\gamma_2\beta_2[...]\beta_{n-1}\gamma_n\beta_n$ 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 $\gamma_k = \gamma_{k'}$.

For every k, $1 \leq k \leq n$, let $i_k \in \mathrm{var}(\gamma_k)$ be the variable satisfying Condition 3 for γ_k. We now define a morphism $\phi : \mathrm{var}(\alpha)^* \to \mathrm{var}(\alpha)^*$ by, for every $x \in \mathrm{var}(\alpha)$,

$$\phi(x) := \begin{cases} \gamma_k, & x = i_k,\ 1 \leq k \leq n, \\ \varepsilon, & x = \mathrm{var}(\gamma_k) \setminus \{i_k\},\ 1 \leq k \leq n, \\ x, & x \in \mathrm{var}(\beta_k),\ 1 \leq k \leq n. \end{cases}$$

Referring to Condition 1 of the above decomposition, the morphism ϕ is not trivial. Also, due to Condition 2 and Condition 3, ϕ is indeed a morphism. Therefore, the definition of ϕ implies that $\phi(\alpha) = \alpha$ and, for every $x \in \mathrm{var}(\alpha)$, if $\phi(x) \neq \varepsilon$, then $x \sqsubseteq \phi(x)$. □

Using Lemma 5.26, we can now prove the following technical observation on the pattern $\alpha_{i,j}$ as introduced in Definition 5.9, which is required in the proof of Theorem 5.28:

Lemma 5.27. *Let α not be a fixed point of a nontrivial morphism. For any $i, j \in \mathrm{var}(\alpha)$, $i \neq j$, if $\alpha_{i,j}$ is a fixed point of a nontrivial morphism $\phi : \mathrm{var}(\alpha)^* \to \mathrm{var}(\alpha)^*$, then $\phi(i) = \varepsilon$.*

Proof. Since $\alpha_{i,j}$ is a fixed point of a nontrivial morphism ϕ, $\phi(\alpha_{i,j}) = \alpha_{i,j}$. According to Lemma 5.26, we can assume that ϕ is a nontrivial morphism satisfying, for every $x \in \mathrm{var}(\alpha_{i,j})$, if $\phi(x) \neq \varepsilon$, then $x \sqsubseteq \phi(x)$. Assume to the contrary that $\phi(i) \neq \varepsilon$. As a result, $i \sqsubseteq \phi(i)$. Also, due to $\phi(\alpha_{i,j}) = \alpha_{i,j}$, $|\phi(i)|_i = 1$. Let $n := |\alpha_{i,j}|_i$ and $\alpha_{i,j} := \alpha_1 i_1 \alpha_2 i_2 [...] \alpha_{n-1} i_n \alpha_n$, where, for every k, $1 \leq k \leq n$, $\alpha_k \in (\mathrm{var}(\alpha_{i,j}) \setminus \{i\})^*$ and $i_k = i$. We now define a nontrivial morphism $\phi' : \mathrm{var}(\alpha)^* \to \mathrm{var}(\alpha)^*$ by, for every $x \in \mathrm{var}(\alpha)$,

$$\phi'(x) := \begin{cases} \phi(x), & x \neq j, \\ \psi(\phi(i)), & x = j, \end{cases}$$

where $\psi : \mathbb{N}^* \to \mathbb{N}^*$ is a morphism given by $\psi(i) := j$ and $\psi(x) := x$, $x \in \mathbb{N} \setminus \{i\}$. According to the definition of $\alpha_{i,j}$, for every occurrence of j in α, there exists a k, $1 \leq k \leq n$, such that j occurs in α at the same position as i_k in $\alpha_{i,j}$. Moreover, for every k, $1 \leq k \leq n$, $i \not\sqsubseteq \alpha_k$, and $|\phi(i)|_i = 1$. Therefore, $\phi'(\alpha) = \alpha$, which is a contradiction to the fact that α is not a fixed point of a nontrivial morphism. As a result, $\phi(i) = \varepsilon$. □

We now provide a comprehensive and affirmative answer to Problem 3.23 for all alphabets that have at least six distinct variables. As mentioned above, our corresponding proof is non-constructive, which means that it does not provide any direct insights into the character of alphabet reductions that preserve not being a fixed point. On the other hand, the applicability of our technique to Billaud's Conjecture (see below) can therefore easily be examined, and the fact that it is not applicable allows some conclusions to be drawn on the complexity of that conjecture.

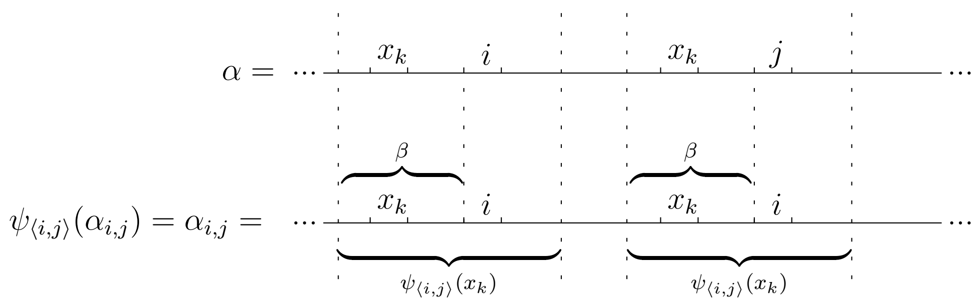
Theorem 5.28. *Let α be a pattern with $|\operatorname{var}(\alpha)| > 5$. If α is not a fixed point of a nontrivial morphism, then there exist $i, j \in \operatorname{var}(\alpha)$, $i \neq j$, such that $\alpha_{i,j}$ is not a fixed point of a nontrivial morphism.*

Proof. Assume to the contrary that, for every $i, j \in \operatorname{var}(\alpha)$, $\alpha_{i,j}$ is a fixed point of a nontrivial morphism. Therefore, due to Lemma 5.26, for every i, j, there exists a nontrivial morphism $\psi_{\langle i,j\rangle} : \operatorname{var}(\alpha)^* \to \operatorname{var}(\alpha)^*$ satisfying $\psi_{\langle i,j\rangle}(\alpha_{i,j}) = \alpha_{i,j}$ and, for every $x \in \operatorname{var}(\alpha_{i,j})$, if $\psi_{\langle i,j\rangle}(x) \neq \varepsilon$, then $x \sqsubseteq \psi_{\langle i,j\rangle}(x)$. On the other hand, it results from Lemma 5.27 that $\psi_{\langle i,j\rangle}(i) = \varepsilon$. Consequently, for every occurrence of i in $\alpha_{i,j}$, there exists a variable $x \in \operatorname{var}(\alpha_{i,j}) \setminus \{i\}$ with $i \sqsubseteq \psi_{\langle i,j\rangle}(x)$ and $x \sqsubseteq \psi_{\langle i,j\rangle}(x)$. We assume that there exist m different variables x in $\alpha_{i,j}$ and we denote them by $x_1, x_2, [...], x_m$. Since α is not a fixed point of a nontrivial morphism, for every k, $1 \leq k \leq m$, $|\alpha_{i,j}|_{x_k} \geq 2$. As a result, for every k, $1 \leq k \leq m$, $|\psi_{\langle i,j\rangle}(\alpha_{i,j})|_{\psi_{\langle i,j\rangle}(x_k)} \geq 2$.

Claim. There exists an x_k, $1 \leq k \leq m$, with at least two occurrences of $\psi_{\langle i,j\rangle}(x_k)$ in $\psi_{\langle i,j\rangle}(\alpha_{i,j})$ such that

- *one of them contains an occurrence of i as nth variable, $1 \leq n \leq |\psi_{\langle i,j\rangle}(x_k)|$, which is at the same position in $\alpha_{i,j}$ as an occurrence of i in α, and*
- *the other one contains an occurrence of i as nth variable, which is at the same position in $\alpha_{i,j}$ as an occurrence of j in α.*

We illustrate the Claim in the following diagram, where β is a prefix of $\psi_{\langle i,j\rangle}(x_k)$ with length $(n-1)$.



Proof (Claim). We denote those occurrences of i in $\alpha_{i,j}$ that are at the same positions as j in α with i_j. We assume to the contrary that there does not exist any x_k, $1 \leq k \leq m$, with at least two occurrences of $\psi_{\langle i,j\rangle}(x_k)$ in $\psi_{\langle i,j\rangle}(\alpha_{i,j})$ satisfying the following conditions:

- one of them contains an occurrence of i as nth variable, $1 \leq n \leq |\psi_{\langle i,j\rangle}(x_k)|$, and
- the other one contains an occurrence of i_j as nth variable.

Let X_j be a set of those variables $q \in \mathrm{var}(\alpha_{i,j}) \setminus \{i\}$ satisfying $|\psi_{\langle i,j\rangle}(q)| \geq 2$ and $i_j \sqsubset \psi_{\langle i,j\rangle}(q)$. Due to the above conditions, there does not exist any $q' \in X_j$ with at least two occurrences of $\psi_{\langle i,j\rangle}(q')$ in $\psi_{\langle i,j\rangle}(\alpha_{i,j})$ such that one of them contains an occurrence of i at the same position as an occurrence of i_j in the other one. Therefore, we can define a nontrivial morphism $\phi : \mathrm{var}(\alpha)^* \to \mathrm{var}(\alpha)^*$ over α by, for every $y \in \mathrm{var}(\alpha)$,

$$\phi(y) := \begin{cases} \varepsilon, & y = j, \\ \varphi_{\langle i,j\rangle}(\psi_{\langle i,j\rangle}(y)), & y \in X_j, \\ \psi_{\langle i,j\rangle}(y), & \text{else}, \end{cases}$$

where $\varphi_{\langle i,j\rangle} : \mathbb{N}^* \to \mathbb{N}^*$ is a morphism with, for every $y' \in \mathrm{var}(\alpha_{i,j})$,

$$\varphi_{\langle i,j\rangle}(y') = \begin{cases} j, & y' = i_j, \\ y', & \text{else}. \end{cases}$$

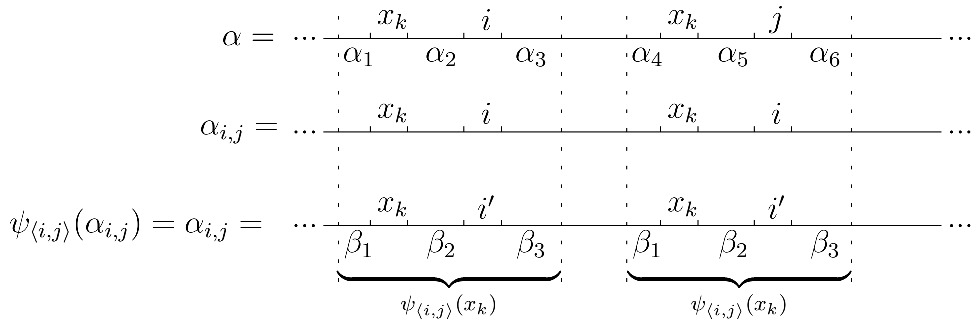
Due to $\psi_{\langle i,j\rangle}(i) = \varepsilon$, because of the definition of $\varphi_{\langle i,j\rangle}$, and since there does not exist any x_k, $1 \leq k \leq m$, satisfying the above mentioned conditions, it can be verified that $\phi(\alpha) = \alpha$, which contradicts the fact that α is not a fixed point of a nontrivial morphism. Therefore, the Claim holds true. □(Claim)

Henceforth, we denote those occurrences of i in $\psi_{\langle i,j\rangle}(x_k)$ satisfying the conditions of the Claim by i'. Consequently, according to the Claim, there exists an x_k, $1 \leq k \leq m$, with $i' \sqsubseteq \psi_{\langle i,j\rangle}(x_k)$. Furthermore, if we wish to refer to the relation between x_k on the one hand and the variables i, j on the other hand as described by the Claim, we say that x_k is *responsible for* the pair $\langle i,j\rangle$.

We now study the following question: Is x_k responsible for any pair of variables of α except $\langle i,j\rangle$ (we do not distinguish between the pairs $\langle i,j\rangle$ and $\langle j,i\rangle$, in other words, $\langle i,j\rangle$ and $\langle j,i\rangle$ are the same pairs)? If the answer is yes, for how many pairs can this happen?

In order to answer this question, we consider the following cases:

1. The variable i' occurs to the right of x_k in $\psi_{\langle i,j\rangle}(x_k)$. So, we can assume that $\alpha = \ldots \cdot \alpha_1 \cdot x_k \cdot \alpha_2 \cdot i \cdot \alpha_3 \cdot \ldots \cdot \alpha_4 \cdot x_k \cdot \alpha_5 \cdot j \cdot \alpha_6 \cdot \ldots$, where, for every k', $1 \leq k' \leq 6$, $\alpha_{k'} \in \mathrm{var}(\alpha)^*$, and $\psi_{\langle i,j\rangle}(x_k) := \beta_1 \cdot x_k \cdot \beta_2 \cdot i' \cdot \beta_3$, $\beta_1, \beta_2, \beta_3 \in \mathrm{var}(\alpha_{i,j})^*$.



We now examine the mentioned question for the pair $\langle l, r\rangle$, $l, r \in \mathrm{var}(\alpha)$ and $\langle l, r\rangle \neq \langle i, j\rangle$, by assuming that $\alpha_{l,r}$ is a fixed point of a nontrivial morphism $\psi_{\langle l,r\rangle}$. According to our discussion for $\langle i, j\rangle$, if x_k is responsible for $\langle l, r\rangle$, we need to have l' (defined analogously to i') in $\psi_{\langle l,r\rangle}(x_k)$.

We assume that l' occurs to the right of x_k in $\psi_{\langle l,r\rangle}(x_k)$. Therefore, one of the following cases needs to be satisfied:

- l' occurs to the right of i'. As a result, due to $\langle l, r\rangle \neq \langle i, j\rangle$, in one occurrence of $\psi_{\langle l,r\rangle}(x_k)$ in $\psi_{\langle l,r\rangle}(\alpha_{l,r})$, we have an occurrence of i, and in the other occurrence of $\psi_{\langle l,r\rangle}(x_k)$ at the same position as i, we have j, which is a contradiction.
- l' occurs in β_2. Then, because of $\langle l, r\rangle \neq \langle i, j\rangle$, there exists an occurrence of $\psi_{\langle i,j\rangle}(x_k)$ in $\psi_{\langle i,j\rangle}(\alpha_{i,j})$ such that its β_2 factor is different from the factor β_2 of the other occurrences of $\psi_{\langle i,j\rangle}(x_k)$ in $\psi_{\langle i,j\rangle}(\alpha_{i,j})$, which is again a contradiction.
- l' occurs at the same position as i'. However, this contradicts the fact that $\langle l, r\rangle \neq \langle i, j\rangle$.

Consequently, x_k can be responsible for $\langle l, r\rangle$ iff l' occurs to the left of x_k in $\psi_{\langle l,r\rangle}(x_k)$. By investigating the responsibility of x_k for any other pair of variables $\langle q, z\rangle$, $q, z \in \mathrm{var}(\alpha)$, $\langle q, z\rangle \neq \langle i, j\rangle$ and $\langle q, z\rangle \neq \langle l, r\rangle$, we can conclude with the same reasoning as above that q' cannot occur to the right of x_k in $\psi_{\langle q,z\rangle}(x_k)$. Also, by assuming that l' occurs to the left of x_k in $\psi_{\langle l,r\rangle}(x_k)$, an analogous reasoning as above leads to the fact that q' cannot occur to the left of x_k in $\psi_{\langle q,z\rangle}(x_k)$. Consequently, x_k cannot be responsible for any other pairs $\langle q, z\rangle$, $q, z \in \mathrm{var}(\alpha)$, $\langle q, z\rangle \neq \langle i, j\rangle$ and $\langle q, z\rangle \neq \langle l, r\rangle$.

2. The variable i' occurs to the left of x_k in $\psi_{\langle i,j\rangle}(x_k)$. An analogous reasoning to that in the previous case implies that, firstly, x_k can be responsible for another pair of variables $\langle l, r\rangle$, $\langle l, r\rangle \neq \langle i, j\rangle$, iff l' occurs to the right of x_k in $\psi_{\langle l,r\rangle}(x_k)$. Secondly, x_k is not responsible for any other pairs $\langle q, z\rangle$, $q, z \in \mathrm{var}(\alpha)$, $\langle q, z\rangle \neq \langle i, j\rangle$ and $\langle q, z\rangle \neq \langle l, r\rangle$.

Consequently, due to the above cases, we can conclude that every variable $x \in \alpha$ can at most be responsible for two pairs of variables. On the other hand, if $|\mathrm{var}(\alpha)| = n$, the number of pairs of variables of α is $\binom{n}{2}$. Referring to the assumption of the theorem, $n > 5$. Therefore,

$$\binom{n}{2} > 2n.$$

This implies that there is a pattern $\alpha_{i,j}$, $i, j \in \mathrm{var}(\alpha)$ such that there does not exist any variable $x \in \mathrm{var}(\alpha_{i,j}) \setminus \{i\}$ that is responsible for the pair $\langle i, j\rangle$, which is a contradiction to the Claim. Thus, there exist variables $i, j \in \mathrm{var}(\alpha)$ such that $\alpha_{i,j}$ is not a fixed point of a nontrivial morphism. □

Theorem 5.28 shows that the structural property of a pattern α that eliminates the existence of a nontrivial morphism ψ satisfying $\psi(\alpha) = \alpha$ is strong enough to also eliminate the existence of a nontrivial morphism ψ' satisfying $\psi'(\phi_{i,j}(\alpha)) = \phi_{i,j}(\alpha)$ for an appropriate choice of the alphabet reduction $\phi_{i,j}$ (see Definition 5.9). However, if we consider a different notion of an alphabet reduction, namely a morphism $\delta_i : \mathbb{N}^* \to \mathbb{N}^*$ defined by $\delta_i(i) := \varepsilon$ and $\delta_i(x) := x$ for $x \in \mathbb{N} \setminus \{i\}$, then Theorem 5.28 and its proof are not sufficient to establish a result that is equivalent to Theorem 5.28. Hence, we have to study Billaud's Conjecture (given as Conjecture 3.21 in the present thesis) separately. As mentioned in Section 3.2, Theorem 3.22 provides a confirmation of the contraposition of Conjecture 3.21 for a special case, but, apart from that, little is known about this problem. The final result of our thesis shall demonstrate that Conjecture 3.21 is correct if patterns are considered that contain each of their variables exactly twice:

Theorem 5.29. *Let α be a pattern with $|\mathrm{var}(\alpha)| \geq 3$ that is not a fixed point of a nontrivial morphism. If, for every $x \in \mathrm{var}(\alpha)$, $|\alpha|_x = 2$, then there exists an $i \in \mathrm{var}(\alpha)$ such that $\delta_i(\alpha)$ is not a fixed point of a nontrivial morphism.*

Proof. Assume to the contrary that, for every $i \in \mathrm{var}(\alpha)$, $\delta_i(\alpha)$ is a fixed point of a nontrivial morphism. This implies that, for every $\delta_i(\alpha)$, there exists a morphism $\phi_i : \mathbb{N}^* \to \mathbb{N}^*$ satisfying $\phi_i(\delta_i(\alpha)) = \delta_i(\alpha)$ and, for a variable q in $\delta_i(\alpha)$, $\phi_i(q) \neq q$. As a result, there exists a variable $y \in \mathrm{var}(\alpha) \setminus \{i\}$ with $|\phi_i(y)| \geq 2$. We assume that $xx' \sqsubseteq \phi_i(y)$, $x, x' \in \text{var}(\alpha) \setminus \{i\}$ and $x \neq x'$. Thus,

$$\delta_i(\alpha) \quad = \quad ... \cdot xx' \cdot ... \cdot xx' \cdot ...,$$

where $|\delta_i(\alpha)|_{xx'} = 2$. Since α is not a fixed point of a nontrivial morphism, and every variable occurs twice in α, $\alpha \neq ... \cdot xx' \cdot ... \cdot xx' \cdot ...$. Therefore, if we, without loss of generality, focus on just one possible order of factors in the equations below, we can assume α to satisfy one of the following equations (note that in the present proof, to emphasise some variables, we show them in a bold face; so, e. g., $\mathbf{x}' = x'$):

$$\alpha = \quad \alpha_1 \cdot x\mathbf{x}' \cdot \alpha_2 \cdot xi\mathbf{x}' \cdot \alpha_3, \text{ where } \alpha_1, \alpha_2, \alpha_3 \in (\text{var}(\alpha) \setminus \{x, x'\})^* \qquad (5.1)$$

$$\alpha = \quad \alpha_1 \cdot x\mathbf{x}' \cdot \alpha_2 \cdot xii\mathbf{x}' \cdot \alpha_3, \text{ where } \alpha_1, \alpha_2, \alpha_3 \in (\text{var}(\alpha) \setminus \{x, x', i\})^* \qquad (5.2)$$

We now investigate $\delta_x(\alpha)$. Referring to the assumption, $\phi_x(\delta_x(\alpha)) = \delta_x(\alpha)$ and there must exist a variable $y' \in \text{var}(\alpha)$ with $|\phi_x(y')| \geq 2$. Let $u_1u_2 \sqsubseteq \phi_x(y')$, $u_1, u_2 \in \text{var}(\alpha) \setminus \{x\}$. Since α is not a fixed point of a nontrivial morphism, and y' has two occurrences in $\delta_x(\alpha)$, we can conclude that one of the following cases must be satisfied:

1. $u_1xu_2 \sqsubseteq \alpha$ and $u_1u_2 \sqsubseteq \alpha$, or
2. $u_1xxu_2 \sqsubseteq \alpha$ and $u_1u_2 \sqsubseteq \alpha$.

However, Case 2 does not hold true, since, according to Equations (5.1) and (5.2), $xx \not\sqsubseteq \alpha$. Moreover, since there does not exist any variable in Equation (5.2) satisfying Case 1, α cannot be factorised as described by Equation (5.2). Consequently, when applying Equation (5.1) to Case 1, and if we again, without loss of generality, focus on just one possible order of factors in the equations below, one of the following equations needs to be satisfied:

$$\alpha \quad = \quad ... \cdot ixx' \cdot ... \cdot xix' \cdot ..., \text{ where } u_1 = i \text{ and } u_2 = x', \text{ or} \qquad (5.3)$$

$$\alpha \quad = \quad ... \cdot \mathbf{x}''i \cdot ... \cdot xx' \cdot ... \cdot \mathbf{x}''xix'..., \text{where } u_1 = x'', u_2 = i, \text{ and } x'' \in \text{var}(\alpha) \setminus \{i, x, x'\}. \qquad (5.4)$$

In the next step, we consider $\delta_{x'}(\alpha)$. Using an analogous reasoning to the one above, we can conclude that there must exist a variable $y'' \in \text{var}(\alpha)$ with $|\phi_{x'}(y'')| \geq 2$. Let $u_3u_4 \sqsubseteq \phi_{x'}(y'')$, $u_3, u_4 \in \text{var}(\alpha) \setminus \{x'\}$. Similarly to our explanations above, and since in the above equations $x'x' \not\sqsubseteq \alpha$, we can conclude that $u_3x'u_4 \sqsubseteq \alpha$ and $u_3u_4 \sqsubseteq \alpha$. However, this condition does not hold true in Equation (5.3), due to the fact that every variable must occur exactly twice in α. Therefore, considering α as factorised in Equation (5.3) leads to a contradiction. In Equation (5.4), in order to satisfy the said condition, we need another variable $x''' \in \mathrm{var}(\alpha) \setminus \{i, x, x''\}$ such that, without loss of generality regarding the order of factors,

$$\alpha \quad = \quad \ldots \cdot x''i\mathbf{x}''' \cdot \ldots \cdot xx' \cdot \ldots \cdot x''xix'\mathbf{x}''' \ldots \,, \tag{5.5}$$

where $u_3 = i$ and $u_4 = x'''$. In other words, $ix''' \sqsubseteq \phi_{x'}(y'')$.

We now consider $\delta_{x''}(\alpha)$. Using the same reasoning as above, but applied to Equation (5.5), we need another variable $x'''' \in \mathrm{var}(\alpha) \setminus \{i, x, x'', x'''\}$ such that, without loss of generality regarding the order of factors,

$$\alpha \quad = \quad \ldots \cdot x''ix''' \cdot \ldots \cdot \mathbf{x}''''xx' \cdot \ldots \cdot \mathbf{x}''''x''xix'x''' \ldots \,, \tag{5.6}$$

where there must exist a variable $y''' \in \mathrm{var}(\alpha)$ with $x''''x \sqsubseteq \phi_{x''}(y''')$.

Consequently, by continuing this reasoning, we can conclude that if we wish to construct a pattern α that satisfies our assumptions, in each step, we have to add a new variable to α, shown in a bold face in each step. This implies that the length of α is infinite, which is a contradiction. Therefore, there exists an $i \in \mathrm{var}(\alpha)$ such that $\delta_i(\alpha)$ is not a fixed point of a nontrivial morphism. $\square$

We expect that even a moderate extension of Theorem 5.29 would require a substantially more involved reasoning. We therefore conclude that the actual nature of patterns that are not a fixed point of a nontrivial morphism, despite our almost comprehensive result in Theorem 5.28 and the strong insights that are mentioned in Section 3.2, is not really understood. This view is further substantiated by the fact that another property of those patterns that are not a fixed point, namely their frequency, is largely unresolved as well (see Reidenbach and Schneider [34]).

Chapter 6

Conclusions

In Chapter 4 of the present thesis, 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 $E_\alpha \subseteq \text{var}(\alpha)$ consists of those variables that have neighbour variables which are not loyal. We have shown 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. Consequently, the following problem has not been completely solved in Chapter 4:

Open Problem 6.1. *Let Σ be a binary alphabet. For which patterns is there a weakly unambiguous length-increasing morphism $\sigma : \mathbb{N}^* \to \Sigma^*$? For which patterns is there no such morphism?*

Regarding the decidability of the above problem, we have given a conjecture in Section 4.3, which we now state as an open problem:

Open Problem 6.2. *Let $\alpha \in \mathbb{N}^+$ with $|\alpha| \geq 2$, and let Σ be a binary alphabet. Is the problem of whether there is a weakly unambiguous length-increasing morphism $\sigma : \mathbb{N}^+ \to \Sigma^+$ with respect to α decidable by testing a finite number of morphisms?*

In Chapter 5, we have investigated the question of whether, for a given pattern in $\mathbb{N}^*$, there exists a strongly unambiguous 1-uniform morphism $\sigma : \mathbb{N}^* \to \Sigma^*$. To this end, we have considered two different settings: in Section 5.1 we have assumed

Σ to be *fixed*, i. e., $|\Sigma|$ does not depend on the number of variables in the pattern, and in Section 5.2 we have allowed Σ to be *arbitrarily chosen*, subject to the number of different variables in the pattern α in question (provided that $|\Sigma| < |\operatorname{var}(\alpha)|$). Our results in Section 5.1 have revealed that it is impossible to give a characteristic condition on those patterns that have a strongly unambiguous 1-uniform morphism if this condition does not incorporate the size of target alphabet Σ. Therefore, for fixed target alphabets, we have given some sufficient conditions on the existence of such morphisms. With regard to variable alphabets Σ, we have given two equivalent conjectures in Section 5.2, which say that such morphisms exist if and only if the pattern is not a fixed point of a nontrivial morphism. Our corresponding results have established major sets of patterns for which these conjectures hold true, but we have left the overall correctness of our conjectures open. We now state one of these conjectures as an open problem:

Open Problem 6.3. *Let α be a pattern with $|\operatorname{var}(\alpha)| \geq 4$. Do there exist an alphabet Σ satisfying $|\Sigma| < |\operatorname{var}(\alpha)|$ and a 1-uniform morphism $\sigma : \mathbb{N}^* \to \Sigma^*$ that is strongly unambiguous with respect to α if and only if α is not a fixed point of a nontrivial morphism?*

Moreover, in Section 5.3, we have studied whether there exists an alphabet reduction that maps a pattern that is not a fixed point of a nontrivial morphism to a pattern that is not a fixed point of a nontrivial morphism, either. Theorem 5.28 has provided a comprehensive and affirmative answer to this problem for all alphabets that have at least six distinct letters. Additionally, since there exist some connections between our studies and Billaud's Conjecture, as a final result of this thesis, we have proved the correctness of this conjecture for those patterns in which every variable occurs exactly twice.

References

[1] S.I. Adian. *The Burnside Problem and Identities in Groups*. Springer-Verlag, Berlin, 1979.

[2] D. Angluin. Inductive inference of formal languages from positive data. *Information and Control*, 45:117–135, 1980.

[3] J. Berstel and D. Perrin. *Theory of Codes*. Academic Press, Orlando, 1985.

[4] M. Billaud. A problem with words. Letter in Newsgroup Comp.Theory, 1993.
https://groups.google.com/d/topic/comp.theory/V_xDDtoR9a4/discussion.

[5] C. Choffrut. Equations in words. In M. Lothaire, editor, *Combinatorics on Words*, chapter 9, pages 164–183. Cambbridge University Press, 1997.

[6] C. Choffrut and J. Karhumäki. Combinatorics of words. In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Languages*, volume 1, chapter 6, pages 329–438. Springer, 1997.

[7] N.G. de Bruijn. A combinatorial problem. *Proc. Koninklijke Nederlandse Akademie v. Wetenschappen*, 49:758–764, 1946.

[8] D.D. Freydenberger, H. Nevisi, and D. Reidenbach. Weakly unambiguous morphisms. *Theoretical Computer Science*. To appear.

[9] 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.

[10] D.D. Freydenberger and D. Reidenbach. The unambiguity of segmented morphisms. *Discrete Applied Mathematics*, 157:3055–3068, 2009.

[11] 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.

[12] E.M. Gold. Language identification in the limit. *Information and Control*, 10:447–474, 1967.

[13] 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.

[14] 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.

[15] T. Head. Fixed languages and the adult languages of 0L schemes. *International Journal of Computer Mathematics*, 10:103–107, 1981.

[16] S. Holub. Polynomial-time algorithm for fixed points of nontrivial morphisms. *Discrete Mathematics*, 309:5069–5076, 2009.

[17] F. Levé and G. Richomme. On a conjecture about finite fixed points of morphisms. *Theoretical Computer Science*, 339:103–128, 2005.

[18] M. Lothaire. *Combinatorics on Words*. Addison-Wesley, Reading, MA, 1983.

[19] M. Lothaire. *Algebraic Combinatorics on Words*. Cambridge University Press, Cambridge, New York, 2002.

[20] M. Lothaire. *Applied Combinatorics on Words*. Cambridge University Press, Cambridge, New York, 2005.

[21] A. Mateescu and A. Salomaa. Finite degrees of ambiguity in pattern languages. *RAIRO – Theoretical Informatics and Applications*, 28:233–253, 1994.

[22] 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.

[23] M. Morse. Recurrent geodesics on a surface of negative curvature. *Transactions of the American Mathematical Society*, 22:84–100, 1921.

[24] M. Morse. A solution of the problem of infinite play in chess. *Bulletin of the American Mathematical Society*, 44:632, 1938.

[25] M. Morse and G. Hedlund. Symbolic dynamics. *American Journal of Mathematics*, 60:815–866, 1938.

[26] M. Morse and G. Hedlund. Unending chess, symbolic dynamics and a problem in semigroups. *Duke Mathematical Journal*, 11:1–7, 1944.

[27] H. Nevisi and D. Reidenbach. Morphic primitivity and alphabet reductions. In *Proc. 16th International Conference on Developments in Language Theory, DLT 2012*. To appear.

[28] H. Nevisi and D. Reidenbach. Unambiguous 1-uniform morphisms. In *Proc. 8th International Conference on Words, WORDS 2011*, volume 63 of *EPTCS*, pages 158–167, 2011.

[29] E. Ohlebusch and E. Ukkonen. On the equivalence problem for E-pattern languages. *Theoretical Computer Science*, 186:231–248, 1997.

[30] E.L. Post. A variant of a recursively unsolvable problem. *Bulletin of the American Mathematical Society*, 52:264–268, 1946.

[31] D. Reidenbach. A non-learnable class of E-pattern languages. *Theoretical Computer Science*, 350:91–102, 2006.

[32] 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.

[33] D. Reidenbach. Discontinuities in pattern inference. *Theoretical Computer Science*, 397:166–193, 2008.

[34] D. Reidenbach and J.C. Schneider. Morphically primitive words. *Theoretical Computer Science*, 410:2148–2161, 2009.

[35] D. Reidenbach and J.C. Schneider. Restricted ambiguity of erasing morphisms. *Theoretical Computer Science*, 412:3510–3523, 2011.

[36] J.C. Schneider. Unambiguous erasing morphisms in free monoids. *RAIRO – Theoretical Informatics and Applications*, 44:193–208, 2010.

[37] M.P. Schützenberger. On an application of semi groups methods to some problems in coding. *IRE Transactions on Information Theory*, 2:47–60, 1956.

[38] A. Thue. Über unendliche Zeichenreihen. *Kra. Vidensk. Selsk. Skrifter. I Mat. Nat. Kl.*, 7, 1906.

[39] A. Thue. Über die gegenseitige Lage gleicher Teile gewisser Zeichenreihen. *Kra. Vidensk. Selsk. Skrifter. I Mat. Nat. Kl.*, 1, 1912.