

Loughborough University Institutional Repository

Unambiguous morphic images of strings

This item was submitted to Loughborough University's Institutional Repository by the/an author.

Citation: FREYDENBERGER, D.D., REIDENBACH, D. and SCHNEIDER, J.C., 2006. Unambiguous morphic images of strings. *International Journal of Foundations of Computer Science*, 17 (3), pp. 601-628

Additional Information:

- This is a journal article. It was published in the journal, *International Journal of Foundations of Computer Science* [© World Scientific Publishing Company]. The definitive version is available at: <http://www.cs.ucsb.edu/~ijfcs/>

Metadata Record: <https://dspace.lboro.ac.uk/2134/3464>

Publisher: © World Scientific Publishing Company

Please cite the published version.

This item was submitted to Loughborough's Institutional Repository by the author and is made available under the following Creative Commons Licence conditions.



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

International Journal of Foundations of Computer Science
© World Scientific Publishing Company

UNAMBIGUOUS MORPHIC IMAGES OF STRINGS

DOMINIK D. FREYDENBERGER*

DANIEL REIDENBACH†

JOHANNES C. SCHNEIDER

*Fachbereich Informatik, Technische Universität Kaiserslautern,
Postfach 3049, 67653 Kaiserslautern, Germany
{freydenb, reidenba, j_schne}@informatik.uni-kl.de*

Received (Day Month Year)
Accepted (Day Month Year)
Communicated by (xxxxxxxxxx)

We study a fundamental combinatorial problem on morphisms in free semigroups: With regard to any string α over some alphabet we ask for the existence of a morphism σ such that $\sigma(\alpha)$ is unambiguous, i. e. there is no morphism τ with $\tau(i) \neq \sigma(i)$ for some symbol i in α and, nevertheless, $\tau(\alpha) = \sigma(\alpha)$. As a consequence of its elementary nature, this question shows a variety of connections to those topics in discrete mathematics which are based on finite strings and morphisms such as pattern languages, equality sets and, thus, the Post Correspondence Problem.

Our studies demonstrate that the existence of unambiguous morphic images essentially depends on the structure of α : We introduce a partition of the set of all finite strings into those that are decomposable (referred to as *prolix*) in a particular manner and those that are indecomposable (called *succinct*). This partition, that is also known to be of major importance for the research on pattern languages and on finite fixed points of morphisms, allows to formulate our main result according to which a string α can be mapped by an injective morphism onto an unambiguous image if and only if α is succinct.

Keywords: Combinatorics on words; morphisms; ambiguity; pattern languages.

1. Introduction

In the past decades a lot of effort has been spent on investigating the properties of those *morphisms* which map a string over some alphabet Δ onto a string over a second alphabet Σ (cf., e. g., Lothaire [11], Choffrut and Karhumäki [3], Harju and Karhumäki [6]). In this context, many problems only arise if Δ contains more symbols than Σ , and therefore—in order to address these difficulties—we assume, for the remainder of our paper, that $\Delta = \mathbb{N}$ and $\Sigma = \{\mathbf{a}, \mathbf{b}\}$. Consequently, we regard

*New address: Facultat de Lletres, Universitat Rovira i Virgili, Pl. Imperial Tàrraco 1, 43005 Tarragona, Spain.

†Corresponding author.

the set of morphisms mapping the strings in an infinitely generated free semigroup to the strings in a free monoid with two generators. According to the closely related research on *pattern languages* (cf. Mateescu and Salomaa [13]) we call an element of \mathbb{N}^+ a *pattern* and an element of $\{\mathbf{a}, \mathbf{b}\}^*$ a *word*. We separate all symbols in a pattern by a dot (see, e.g., the example pattern α' below) so as to avoid any confusion.

Quite a number of the basic questions on such mappings deal with morphisms which, in spite of the resulting alphabet reduction, preserve the structure of their input string as far as possible. In other words, those morphisms are sought which lead to a minimum loss of information about the preimage; this is manifested, e.g., in the theory of codes (cf. Berstel and Perrin [2]). Obviously, any consideration on this problem strongly depends on the formal definition of what is considered to be a “structure preserving” morphism. Nevertheless, from a very intuitive point of view one surely would agree that, for instance, the shape of the pattern $\alpha' := 1 \cdot 2 \cdot 3 \cdot 4 \cdot 1 \cdot 4 \cdot 3 \cdot 2$ is not adequately reflected by its morphic image $w_1 := \mathbf{a}^{20}$. Typically, for such a task, a morphism like $\sigma'(i) := \mathbf{a} \mathbf{b}^i$, $i \in \mathbb{N}$, is the more appropriate choice since σ' is a code—which, in our sense, simply means that it is an *injective* morphism—such that an inverse morphism can be given mapping $w_2 := \sigma'(\alpha') = \mathbf{a} \mathbf{b} \mathbf{a} \mathbf{b}^2 \mathbf{a} \mathbf{b}^3 \mathbf{a} \mathbf{b}^4 \mathbf{a} \mathbf{b} \mathbf{a} \mathbf{b}^4 \mathbf{a} \mathbf{b}^3 \mathbf{a} \mathbf{b}^2$ onto α' again. Consequently, injectivity can be considered a necessary condition for morphisms satisfying our vague yet well-founded requirement.

So far, our considerations can be fully subsumed under the principles of coding, but from now on we break with that theory. Hence, we do not assume to be confronted with a single injective morphism σ such that, in an algebraic interpretation of coding theory, we can analyse the structure preserving capacity of σ by examining the properties of the free subsemigroup which σ , when applied to the patterns in Δ^+ , induces in Σ^* . Instead of this, we focus on those problems where we have to deal with a variety of morphic images of one and the same pattern, and, thus, we tackle the features of morphisms in a rather combinatorial manner. Among the prominent research fields colligated to this view we are mainly concerned with the area of *pattern languages* as introduced by Angluin [1]—basically, the pattern language of a pattern α is simply the set of *all* morphic images of α in some fixed free monoid. Originally, pattern languages have been derived from considerations on *inductive inference* (cf., e.g., Gold [4]), i.e. algorithmic language learning from examples. Thus, inductive inference of pattern languages asks for the computability of an (unknown) pattern from the set of its morphic images, and therefore it is manifest that this task strongly depends on the existence of a morphism which reproduces the shape of the pattern as precisely as possible. However, since a pattern language in general is by no means a free monoid—in fact, it is not even closed under concatenation—there is no fixed factorisation of the image of a pattern under some code. Consequently, inductive inference of pattern languages cannot rely so much on the existence of an inverse morphism mapping a word onto a pattern, and therefore, in this context, we cannot regard injectivity as a sufficient condition for a morphism to satisfy our requirements. Of course, with respect to the demands

under consideration, there is still some combinatorial value associated to the injectivity of a morphism (after all, the word $w_2 = \mathbf{a b a b^2 a b^3 a b^4 a b a b^4 a b^3 a b^2}$ is surely a more instructive instance of α' than $w_1 = \mathbf{a^{20}}$), but, in addition, a second potential property of a morphic image significantly gains in importance, namely its *unambiguity*. We call a word w *unambiguous* with respect to a pattern α provided that there is exactly one morphism σ with $\sigma(\alpha) = w$; correspondingly, w is said to be *ambiguous* with respect to α if there are at least two morphisms σ, τ such that, for some symbol i in α , $\sigma(i) \neq \tau(i)$, but nevertheless $\sigma(\alpha) = w = \tau(\alpha)$. The great impact of ambiguity of words on inductive inference of pattern languages (recently shown by Reidenbach [16]) can be illustrated easily by our example word w_2 which, in fact, is ambiguous with respect to α' : In addition to σ' , it can, e. g., also be generated by the morphism τ' with $\tau'(1) := \mathbf{a b a b^2}$, $\tau'(2) := \varepsilon$, $\tau'(3) := \mathbf{a b^3 a b^2}$ and $\tau'(4) := \mathbf{b^2}$, where ε is the empty word:

$$\begin{array}{ccccccc}
 \begin{array}{c} \sigma'(1) \\ \underbrace{\mathbf{a b}} \\ \tau'(1) \end{array} & \begin{array}{c} \sigma'(2) \\ \underbrace{\mathbf{a b b}} \\ \tau'(3) \end{array} & \begin{array}{c} \sigma'(3) \\ \underbrace{\mathbf{a b b b}} \\ \tau'(3) \end{array} & \begin{array}{c} \sigma'(4) \\ \underbrace{\mathbf{a b b b b}} \\ \tau'(4) \end{array} & \begin{array}{c} \sigma'(1) \\ \underbrace{\mathbf{a b}} \\ \tau'(1) \end{array} & \begin{array}{c} \sigma'(4) \\ \underbrace{\mathbf{a b b b b}} \\ \tau'(4) \end{array} & \begin{array}{c} \sigma'(3) \\ \underbrace{\mathbf{a b b b}} \\ \tau'(3) \end{array} & \begin{array}{c} \sigma'(2) \\ \underbrace{\mathbf{a b b}} \\ \tau'(3) \end{array}
 \end{array}$$

Consequently, w_2 does not adequately substantiate the existence of the symbol 2 in α' since this symbol is not needed for generating w_2 and, hence, in our context w_2 does not sufficiently reflect the structure of α' . But even if we restrict our examination to nonerasing morphisms, i. e. if we use the free semigroup $\{\mathbf{a, b}\}^+$ instead of $\{\mathbf{a, b}\}^*$ as the range of the morphisms, the multitude of potential generating morphisms blurs the evidence of α in w_2 . In contrast to this, an unambiguous word allows for more reliable conclusions about its preimage, and therefore several recent results on inductive inference of pattern languages are based on insights related to the existence of unambiguous (or at least: “not overly” ambiguous) words (cf., e. g., Reidenbach [15, 16]). With regard to our example pattern α' , a shortest word showing such a property is $w_3 := \mathbf{a a b b b a a b a b a b}$, which can only be generated by the non-injective morphism σ'' with $\sigma''(1) := \mathbf{a}$, $\sigma''(2) := \mathbf{a b}$, $\sigma''(3) := \mathbf{b}$, $\sigma''(4) := \mathbf{b a}$.

Unfortunately, contrary to the broad and profound knowledge on coding theory and, thus, the injectivity of morphisms, very little is known about the ambiguity of morphic images. It seems, however, that in general many morphic images of patterns are ambiguous, and, effortlessly, examples can be given—such as the pattern $1 \cdot 2$ —for which there is no morphism at all leading to an unambiguous word. On the other hand, it is by no means obvious for which patterns there exist unambiguous morphic images. In the present paper, we initiate the systematic analysis of this combinatorial problem of intrinsic interest. To this end, we concentrate on the ambiguity of those words that are images of injective morphisms, and we explicitly distinguish between the general case where the set of all morphisms $\tau_E : \mathbb{N}^+ \longrightarrow \{\mathbf{a, b}\}^*$ is considered and the restricted case that focuses on nonerasing morphisms $\tau_{NE} : \mathbb{N}^+ \longrightarrow \{\mathbf{a, b}\}^+$.

Our paper is organised as follows: After the formal definitions we first collect a number of rather evident yet vital preliminary results on the (non-)existence of

unambiguous images of patterns under nonerasing morphisms τ_{NE} . We proceed with a more detailed examination of the analogous problem for general morphisms τ_{E} which yields the insight that a rich and natural class of patterns is characterised by its ability of morphically generating unambiguous words. An extensive proof for this main result, a summary and a list of some potential future research directions conclude the paper.

Due to the simple definition of the question under consideration, there exists a number of fields in combinatorics on morphisms which are related to our research. According to our above remarks, this is evident for the concept of pattern languages. In this regard, we show that in every pattern language (note that we use this term in a restricted way, see Section 2) there exists an unambiguous word with respect to some particular generating pattern. This interpretation of our main result answers a question posed in [16]. Additionally, it seems worth mentioning that, in a sense, our work complements a research that has been initiated by Mateescu and Salomaa [12] on the *nondeterminism of pattern languages*. As explained above we show that, for *every* pattern in some class, there exists *at least one* non-empty word in $\{\mathbf{a}, \mathbf{b}\}^*$ that has exactly one generating morphism, whereas, in a more general context, [12] 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. Thus, roughly speaking, we ask for the “minimum ambiguity” in a pattern language, whereas [12] enquires its “maximum ambiguity” and seeks after patterns for which this ambiguity is bounded. In our restricted setting, for all patterns with occurrences of at least two different symbols, this question has a trivial answer in the negative. In addition to this close relation to pattern languages, our work shows some evident connections to equality sets (cf., e. g., Harju and Karhumäki [6]); the *equality set* of two morphisms σ, τ is the set of all patterns α with $\sigma(\alpha) = \tau(\alpha)$, and, thus, the famous undecidable Post Correspondence Problem (PCP)—named after E. L. Post [14]—is simply the emptiness problem for equality sets. In the terminology related to this problem, each pattern 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 $\sigma(\alpha)$ is unambiguous then α is a *non-solution* to the PCP for σ and any other morphism τ . Consequently, in this context, our paper complements the research on primitive solutions to the PCP as, e. g., given by Lipponen and Păun [10]. Finally, our studies also lead to some insights on finite *fixed points* of nontrivial morphisms, i. e. those patterns for which there is some morphism $\phi : \mathbb{N}^* \rightarrow \mathbb{N}^*$ such that $\phi(\alpha) = \alpha$ and, for a symbol i in α , $\phi(i) \neq i$ (cf. Hamm and Shallit [5]). The foundations of this rather unexpected analogy are introduced in the subsequent section.

2. Definitions and Basic Notes

We begin the formal part of this paper with a number of basic definitions. A major part of our terminology is adopted from the research on pattern languages (cf. Ma-

teescu and Salomaa [13]). Additionally, for notions not explained explicitly, we refer the reader to Choffrut and Karhumäki [3].

Let $\mathbb{N} := \{1, 2, 3, \dots\}$ and $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. Let \mathcal{A} be an *alphabet*, i. e. an enumerable set of symbols. We regard two different alphabets: \mathbb{N} and $\{\mathbf{a}, \mathbf{b}\}$ with $\mathbf{a} \neq \mathbf{b}$. Henceforth we call any symbol in \mathbb{N} a *variable* and any symbol in $\{\mathbf{a}, \mathbf{b}\}$ a *letter*. A *string (over \mathcal{A})* is a finite sequence of symbols from \mathcal{A} . For the *concatenation* of two strings w_1, w_2 we write $w_1 \cdot w_2$ or simply $w_1 w_2$. The string that results from the n -fold concatenation of a string w is denoted by w^n . The notation $|x|$ stands for the size of a set x or the length of a string x . We denote the *empty string* by ε , i. e. $|\varepsilon| = 0$. In order to distinguish between a string over \mathbb{N} and a string over $\{\mathbf{a}, \mathbf{b}\}$, we call the former a *pattern* and the latter a *word*. 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 , $\text{var}(\alpha)$ denotes the set of all variables occurring in α . For every alphabet \mathcal{A} , \mathcal{A}^* is the set of all (empty and non-empty) strings over \mathcal{A} , and $\mathcal{A}^+ := \mathcal{A}^* \setminus \{\varepsilon\}$. We call a string $v \in \mathcal{A}^*$ a *substring* of a string $w \in \mathcal{A}^*$ if, for some $u_1, u_2 \in \mathcal{A}^*$, $w = u_1 v u_2$; moreover, if v is a substring of w then we say that w *contains* v and, conversely, that v *occurs* in w . A string $v \in \mathcal{A}^*$ is a *scattered substring* of a string $w \in \mathcal{A}^*$ if and only if there exist $u_0, u_1, \dots, u_n \in \mathcal{A}^*$ and $v_1, v_2, \dots, v_n \in \mathcal{A}^*$, $n \in \mathbb{N}$, such that $w = u_0 v_1 u_1 v_2 u_2 \dots v_n u_n$ and $v = v_1 v_2 \dots v_n$. Subject to the concrete alphabet considered, we call a substring a *subword* or a *sub-pattern*. If, for some $w, v_1, v_2 \in \mathcal{A}^*$, $w = v_1 v_2$ then v_1 is a *prefix* of w and v_2 is a *suffix* of w . Additionally, we use the notations $w = \dots v \dots$ if v is a substring of w , $w = v \dots$ if v is a prefix of w , and $w = \dots v$ if v is a suffix of w . In contrast to this, if we wish to omit some parts of a canonically given string then we henceforth use the symbol $[\dots]$, i. e., e. g., $\alpha = 1 \cdot 2 \cdot [\dots] \cdot 5$ stands for $\alpha = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5$. Finally, $|w|_v$ denotes the number of occurrences of a substring v in a string w . We do not use this notion for substrings with overlapping occurrences.

Since we deal with word semigroups, a *morphism* σ is a mapping that is compatible with the concatenation, i. e. for patterns $\alpha, \beta \in \mathbb{N}^+$, a morphism $\sigma : \mathbb{N}^+ \longrightarrow \{\mathbf{a}, \mathbf{b}\}^*$ satisfies $\sigma(\alpha \cdot \beta) = \sigma(\alpha) \cdot \sigma(\beta)$. Hence, a morphism is fully explained as soon as it is declared for all variables in \mathbb{N} . Note that we restrict ourselves to total morphisms, even though we normally declare a morphism only for those variables explicitly that, in the respective context, are relevant.

Let $\sigma : \mathbb{N}^+ \longrightarrow \{\mathbf{a}, \mathbf{b}\}^*$ be a morphism. Then σ is called *nonerasing* provided that, for every $i \in \mathbb{N}$, $\sigma(i) \neq \varepsilon$. Moreover, σ is *injective* if, for any patterns $\alpha, \beta \in \mathbb{N}^+$, $\sigma(\alpha) = \sigma(\beta)$ implies $\alpha = \beta$. Note that σ necessarily is nonerasing if it is injective.

For any pattern $\alpha \in \mathbb{N}^+$ and for any word $w \in \{\mathbf{a}, \mathbf{b}\}^+$, we call w *weakly unambiguous with respect to α* if there exists a nonerasing morphism $\sigma : \mathbb{N}^+ \longrightarrow \{\mathbf{a}, \mathbf{b}\}^+$ such that $\sigma(\alpha) = w$ and there is no nonerasing morphism $\tau : \mathbb{N}^+ \longrightarrow \{\mathbf{a}, \mathbf{b}\}^+$ with $\tau(\alpha) = \sigma(\alpha)$ and, for some $i \in \text{var}(\alpha)$, $\tau(i) \neq \sigma(i)$. Additionally, w is said to be (*strongly*) *unambiguous with respect to α* if there exists an (arbitrary) morphism $\sigma : \mathbb{N}^+ \longrightarrow \{\mathbf{a}, \mathbf{b}\}^*$ such that $\sigma(\alpha) = w$ and there is no (arbitrary) morphism $\tau : \mathbb{N}^+ \longrightarrow$

6 *D. D. Freydenberger, D. Reidenbach, J. C. Schneider*

$\{\mathbf{a}, \mathbf{b}\}^*$ with $\tau(\alpha) = \sigma(\alpha)$ and, for some $i \in \text{var}(\alpha)$, $\tau(i) \neq \sigma(i)$. Obviously, if w is strongly unambiguous with respect to α then it is weakly unambiguous with respect to α as well. Finally, w is *ambiguous with respect to α* if and only if there exists a morphism $\sigma : \mathbb{N}^+ \longrightarrow \{\mathbf{a}, \mathbf{b}\}^+$ with $\sigma(\alpha) = w$ and w is not weakly unambiguous with respect to α . We omit the supplement “with respect to α ” provided that the pattern under consideration is understood, i. e., for any pattern α and for any morphism σ , if we simply say that $\sigma(\alpha)$ is *weakly unambiguous*, *(strongly) unambiguous* or *ambiguous* then this means that $\sigma(\alpha)$ is weakly unambiguous w. r. t. α , strongly unambiguous w. r. t. α or ambiguous w. r. t. α .

As mentioned above, our subject is closely related to pattern languages (cf., e. g., Mateescu and Salomaa [13]), and therefore we consider it useful to provide an adequate background for some explanatory remarks. The *pattern language* of a pattern is the set of all of its morphic images in some fixed free monoid Σ^* (in our case $\Sigma := \{\mathbf{a}, \mathbf{b}\}$). More precisely and with regard to any $\alpha \in \mathbb{N}^+$, we distinguish between its *E-pattern language* $L_E(\alpha) := \{\sigma(\alpha) \mid \sigma : \mathbb{N}^+ \longrightarrow \Sigma^*\}$ and its *NE-pattern language* $L_{NE}(\alpha) := \{\sigma(\alpha) \mid \sigma : \mathbb{N}^+ \longrightarrow \Sigma^+\}$.

Note that this definition implies that both the *class of E-pattern languages* and the *class of NE-pattern languages*, i. e. the sets $\{L_E(\alpha) \mid \alpha \in \mathbb{N}^+\}$ and $\{L_{NE}(\alpha) \mid \alpha \in \mathbb{N}^+\}$, considered in this paper merely cover a special case which, in literature, usually is referred to as *terminal-free* (or: *pure*) pattern languages. This is due to the fact that, contrary to our view, a pattern commonly is seen as a string in $(\mathbb{N} \cup \Sigma)^+$ —thus potentially containing so-called *terminal symbols*, i. e. arbitrary letters in Σ —and not just in \mathbb{N}^+ . 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 $\sigma : (\mathbb{N} \cup \Sigma)^* \longrightarrow \Sigma^*$ with $\sigma(a) = a$ for every $a \in \Sigma$. It is an open problem to which extent our subsequent considerations can be adopted to general pattern languages.

We conclude the definitions in this section with a partition of the set of all patterns subject to the following criterion:

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

We call $\alpha \in \mathbb{N}^+$ succinct if and only if it is not prolix.

Example 2. *Obviously, any pattern α , $|\alpha| \geq 2$, necessarily is prolix if there is a variable $i \in \mathbb{N}$ such that $|\alpha|_i = 1$. Our initial example $\alpha' = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 1 \cdot 4 \cdot 3 \cdot 2$ and the pattern $\alpha_1 := 1 \cdot 1$ are succinct, whereas $\alpha_2 := 1 \cdot 2 \cdot 3 \cdot 3 \cdot 1 \cdot 2 \cdot 3$ is prolix with $\beta_0 = \varepsilon$, $\gamma_1 = 1 \cdot 2$, $\beta_1 = 3 \cdot 3$, $\gamma_2 = 1 \cdot 2$ and $\beta_2 = 3$. Note that this obligatory*

decomposition of a prolix pattern does not have to be unique. Additional and more complex examples can be found in the subsequent sections and in [15].

Definition 1 is crucial for our studies in Section 4. Additionally, the above partition of \mathbb{N}^+ is known to be of major importance in other domains dealing with morphisms and finite strings as well: First, the succinct patterns are the shortest generators for their respective E-pattern language—this explains the terms “succinct” and “prolix”:

Theorem 3 (Reidenbach [15]) *A pattern $\alpha \in \mathbb{N}^+$ is succinct if and only if, for every $\beta \in \mathbb{N}^+$ with $L_E(\beta) = L_E(\alpha)$, $|\beta| \geq |\alpha|$.*

Consequently, $\{L_E(\alpha) \mid \alpha \in \mathbb{N}^+, \alpha \text{ is succinct}\} = \{L_E(\alpha) \mid \alpha \in \mathbb{N}^+\}$, although the set of succinct patterns is a proper subset of \mathbb{N}^+ .

Second, the set of prolix patterns exactly corresponds to the class of finite *fixed points* of nontrivial morphisms:

Theorem 4 (Head [7]) *A pattern α is prolix if and only if there exists a morphism $\phi : \mathbb{N}^* \rightarrow \mathbb{N}^*$ such that, for an $i \in \text{var}(\alpha)$, $\phi(i) \neq i$ and yet $\phi(\alpha) = \alpha$.*

For more insights into the latter subject, see, e. g., Hamm and Shallit [5] and Levé and Richomme [9].

Finally note that our restriction to morphic images in $\{\mathbf{a}, \mathbf{b}\}^*$ does not lead to likewise restricted results:

Remark 5. *All of the results in this paper hold for morphisms to arbitrary finitely generated free monoids with three or more generators instead of $\{\mathbf{a}, \mathbf{b}\}^*$ as well. With regard to the positive results, this follows by definition, and the proofs of the negative results can be adapted with little effort.*

Nevertheless, we expect that a detailed analysis of the ambiguity of images in free monoids with more than two generators might result in challenging combinatorial questions and numerous technical considerations strongly differing from those established in the present paper.

3. Weakly Unambiguous Words

We begin our examination with some significant statements on weakly unambiguous morphic images. The first is an evident yet strong positive result:

Proposition 6. *There is a nonerasing morphism $\sigma : \mathbb{N}^+ \rightarrow \{\mathbf{a}, \mathbf{b}\}^+$ such that, for every $\alpha \in \mathbb{N}^+$, $\sigma(\alpha)$ is weakly unambiguous.*

Proof. For every $i \in \mathbb{N}$, let $|\sigma(i)| = 1$. Then, for every $\alpha \in \mathbb{N}^+$, $|\sigma(\alpha)| = |\alpha|$ and, consequently, $\sigma(\alpha)$ is weakly unambiguous. \square

8 *D. D. Freydenberger, D. Reidenbach, J. C. Schneider*

In Proposition 6 we restrict ourselves to nonerasing morphisms, and therefore it corresponds to the concept of NE-pattern languages. Indeed, the weak unambiguity of the words referred to in the proof is of major importance for inductive inference of NE-pattern languages: Due to the fact given in Proposition 6, for every NE-pattern language L a pattern α with $L = L_{\text{NE}}(\alpha)$ can be inferred from the set of all the shortest words in this language (shown by Lange and Wiehagen [8]). With regard to E-pattern languages, however, this is provably impossible (cf. Reidenbach [16]) since, in general, these words are not strongly unambiguous. Consequently, in respect of pattern inference, the unambiguity of certain words—which are not generated by an injective morphism—is surprisingly powerful.

For the main goal of our approach (see Section 1), however, injectivity of morphisms is vital. Unfortunately, for these morphisms the outcome significantly differs from Proposition 6:

Theorem 7. *There is no injective morphism $\sigma : \mathbb{N}^+ \longrightarrow \{\mathbf{a}, \mathbf{b}\}^+$ such that, for every $\alpha \in \mathbb{N}^+$, $\sigma(\alpha)$ is weakly unambiguous.*

Proof. Assume to the contrary that there is such a morphism σ . Since σ is injective, $\sigma(\alpha) \neq \sigma(\beta)$ for every $\alpha \neq \beta$. In particular, this implies $\sigma(i) \neq \sigma(i')$ for every $i, i' \in \mathbb{N}$ with $i \neq i'$. Hence, there must be a $j \in \mathbb{N}$ with $\sigma(j) = w_1 w_2$ for some $w_1, w_2 \in \{\mathbf{a}, \mathbf{b}\}^+$. Now, for an arbitrary $j' \neq j$, let $\alpha := j \cdot j'$. Then, for the morphism $\tau : \mathbb{N}^+ \longrightarrow \{\mathbf{a}, \mathbf{b}\}^+$ given by $\tau(j) := w_1$ and $\tau(j') := w_2 \sigma(j')$, $\tau(\alpha) = \sigma(\alpha)$, and, thus, $\sigma(\alpha)$ is ambiguous. This contradicts the assumption. \square

Obviously, Theorem 7 includes the analogous result for strong unambiguity. Consequently, there is no single injective morphism which, when applied to arbitrary patterns, always leads to unambiguous words. Thus, two natural questions arise from Theorem 7: Is there a significant subclass of all patterns for which the opposite of Theorem 7 holds true? Is it possible to find for every pattern a tailor-made injective morphism that leads to an unambiguous word? In the following section we examine these questions with regard to strong unambiguity.

4. Strongly Unambiguous Words

Bearing the consequences of Theorem 7 in mind the present section deals with strongly unambiguous words. We begin with the observation that the example pattern α in the proof of Theorem 7 is prolix. However, if we focus on succinct patterns then the analogue turns out to be true; as we now ask for strong unambiguity we can even prove the negation of the equivalent of Proposition 6:

Proposition 8. *There is no nonerasing morphism $\sigma : \mathbb{N}^+ \longrightarrow \{\mathbf{a}, \mathbf{b}\}^+$ such that, for every succinct $\alpha \in \mathbb{N}^+$, $\sigma(\alpha)$ is strongly unambiguous.*

Proof. Assume to the contrary that there is such a morphism. Then there exist $j, j' \in \mathbb{N}$, $j \neq j'$, and a $\mathbf{c} \in \{\mathbf{a}, \mathbf{b}\}$ with $\sigma(j) = v \mathbf{c}$ and $\sigma(j') = v' \mathbf{c}$, $v, v' \in \Sigma^*$. For

some $k, k' \in \mathbb{N}$, $k \neq k'$, $j \neq k \neq j'$ and $j \neq k' \neq j'$, we then regard the pattern $\alpha := j \cdot k \cdot j \cdot k' \cdot j' \cdot k \cdot j' \cdot k'$. According to Definition 1, α is succinct. Now consider the morphism τ , given by $\tau(j) := v$, $\tau(j') := v'$, $\tau(k) := \mathbf{c}\sigma(k)$, $\tau(k') := \mathbf{c}\sigma(k')$. Then evidently $\sigma(\alpha) = \tau(\alpha)$, but, e. g., $\sigma(j) \neq \tau(j)$. This is a contradiction. \square

Consequently, for every succinct pattern, it is necessary to give an individual injective morphism that leads to a strongly unambiguous word—provided that such a morphism exists. In this regard, the search for suitable candidates cannot *a priori* ignore certain morphisms since *every* (and, thus, even a completely inappropriate looking) nonerasing morphism maps some nontrivial succinct pattern onto an unambiguous image:

Proposition 9. *For every nonerasing morphism $\sigma : \mathbb{N}^+ \longrightarrow \{\mathbf{a}, \mathbf{b}\}^+$ there exists a succinct $\alpha \in \mathbb{N}^+$, $|\text{var}(\alpha)| \geq 2$, such that $\sigma(\alpha)$ is strongly unambiguous.*

Proof. Since our argumentation solely deals with the length of the morphic images of the variables, we can utilise the following fact on linear combinations:

Claim 1. *For all $p, q \in \mathbb{N}$ there exist $r, s \in \mathbb{N}$, $r, s \geq 2$ such that there are no $p', q' \in \mathbb{N}_0 \setminus \{p, q\}$ satisfying $rp + sq = rp' + sq'$.*

Proof (Claim 1). Choose r, s such that $r > q$, $s > p$, $\text{gcd}(r, s) = 1$. Assume to the contrary that there are $p', q' \in \mathbb{N} \setminus \{p, q\}$ such that $rp + sq = rp' + sq'$. Then $r(p - p') = s(q' - q)$. As r and s are relatively prime we have $r|q' - q$ and $s|p - p'$. Hence, there is a $t \neq 0$ with $rt = q' - q$; furthermore, $st = p - p'$ because of $srt = s(q' - q) = r(p - p')$. This finally leads to $p' = p - st$ and $q' = q + rt$. Now, if $t > 0$ then $p' < 0$, if $t < 0$ then $q' < 0$. This contradicts the assumption $p', q' \in \mathbb{N}$.

\square (Claim 1)

Now, for some $i, j \in \mathbb{N}$, $i \neq j$, let $p := |\sigma(i)|$, $q := |\sigma(j)|$. Furthermore, let $\alpha := i^r \cdot j^s$ with r, s derived from Claim 1. Obviously, α is succinct. Assume to the contrary that there is a morphism $\tau : \mathbb{N}^+ \longrightarrow \{\mathbf{a}, \mathbf{b}\}^*$ with $\tau(\alpha) = \sigma(\alpha)$ and, for some $k \in \{i, j\}$, $\tau(k) \neq \sigma(k)$. Then τ must satisfy $|\tau(i)| \neq |\sigma(i)|$, $|\tau(j)| \neq |\sigma(j)|$ and $|\tau(\alpha)| = |\sigma(\alpha)|$. Consequently, with $p' := |\tau(i)|$, $q' := |\tau(j)|$, $rp + sq = |\sigma(\alpha)| = |\tau(\alpha)| = rp' + sq'$. This contradicts Claim 1. \square

Before we go further into this matter of strongly unambiguous morphic images for succinct patterns (see Theorem 16), we turn our attention to prolix patterns. Here we can easily give a definite answer, which follows from the fact that every prolix pattern is a fixed point of some nontrivial morphism (cf. Theorem 4):

Theorem 10. *For any prolix $\alpha \in \mathbb{N}^+$ and for any nonerasing morphism $\sigma : \mathbb{N}^+ \longrightarrow \{\mathbf{a}, \mathbf{b}\}^+$, $\sigma(\alpha)$ is not strongly unambiguous.*

Proof. If α is prolix then, by Theorem 4, there exists a morphism ϕ with $\phi(\alpha) = \alpha$ and, for some $i \in \text{var}(\alpha)$, $\phi(i) \neq i$; w.l.o.g. we may assume that $\phi(i) = \varepsilon$.

Then, for every nonerasing morphism σ , $\sigma(\alpha) = \sigma(\phi(\alpha))$, but $\sigma(i) \neq \varepsilon = \sigma(\phi(i))$. Consequently, $\sigma(\alpha)$ is not strongly unambiguous. \square

Thus, for every prolix pattern there is no strongly unambiguous word at all—at least as long as we restrict ourselves to the images of nonerasing morphisms. If this requirement is omitted then we face a fairly intricate situation:

Example 11. *Let $\alpha_1 := 1 \cdot 1 \cdot 2 \cdot 2 \cdot 3 \cdot 4 \cdot 3 \cdot 4$, $\beta_1 := 1 \cdot 2 \cdot 2 \cdot 1 \cdot 3 \cdot 4 \cdot 3 \cdot 4$, and $\alpha_2 := 1 \cdot 2 \cdot 3 \cdot 3 \cdot 1 \cdot 2$, $\beta_2 := 1 \cdot 2 \cdot 2 \cdot 3 \cdot 3 \cdot 1 \cdot 2 \cdot 2$. The patterns are prolix. For α_1 and α_2 there is no morphism σ such that $\sigma(\alpha_1)$ or $\sigma(\alpha_2)$ are unambiguous. Contrary to this, for β_1 and β_2 there exist suitable words such as \mathbf{abba} and \mathbf{baab} .*

As shown in Example 11, there are prolix patterns for which we can unambiguously map certain scattered subpatterns onto strings in $\{\mathbf{a}, \mathbf{b}\}^*$, whereas for other, quite similar appearing patterns this is impossible. Furthermore, these scattered subpatterns can consist of parts of some β_k as well as parts of some γ_k in the required decomposition of the patterns (cf. Definition 1). We now briefly discuss this phenomenon, and we begin with the immediate conclusion that both types of patterns introduced in Example 11 can be extended easily:

Proposition 12. *For every $n \geq 2$ there exists a prolix pattern α , $|\text{var}(\alpha)| = n$, for which there is no morphism $\sigma : \mathbb{N}^+ \rightarrow \{\mathbf{a}, \mathbf{b}\}^*$ such that $\sigma(\alpha)$ is strongly unambiguous. For every $n \geq 2$, $|\text{var}(\alpha)| = n$, there exists a prolix pattern β and a morphism $\sigma : \mathbb{N}^+ \rightarrow \{\mathbf{a}, \mathbf{b}\}^*$ such that $\sigma(\beta)$ is strongly unambiguous.*

Proof. Proposition 12 holds, e. g. for every $\alpha := 1 \cdot 1 \cdot 2 \cdot 2 \cdot 3 \cdot 4 \cdot [\dots] \cdot n \cdot 3 \cdot 4 \cdot [\dots] \cdot n$ and $\beta := 1 \cdot 2 \cdot 2 \cdot 1 \cdot 3 \cdot 4 \cdot [\dots] \cdot n \cdot 3 \cdot 4 \cdot [\dots] \cdot n$, $n \geq 4$. With regard to $n = 3$, see the patterns α_2 and β_2 as given in Example 11, and, concerning $n = 2$, Proposition 12 is proven by $\alpha := 1 \cdot 2$ and $\beta := 1 \cdot 2 \cdot 2$. \square

We proceed with a criterion which covers prolix as well as succinct patterns. Each of these patterns must contain a distinct variable, the number of occurrences of which does not equal any weighted sum of the number of occurrences of other variables in the pattern:

Condition 13. *A pattern $\alpha \in \mathbb{N}^+$ satisfies Condition 13 if and only if there exists an $i \in \text{var}(\alpha)$ such that, for $n = |\text{var}(\alpha)| - 1$, for all $j_1, j_2, \dots, j_n \in \text{var}(\alpha) \setminus \{i\}$ and for all $k_1, k_2, \dots, k_n \in \mathbb{N}_0$, $|\alpha|_i \neq k_1|\alpha|_{j_1} + k_2|\alpha|_{j_2} + \dots + k_n|\alpha|_{j_n}$.*

For those patterns satisfying Condition 13 we can give a positive result:

Proposition 14. *For every $\alpha \in \mathbb{N}^+$ satisfying Condition 13 there exists a morphism $\sigma : \mathbb{N}^+ \rightarrow \{\mathbf{a}, \mathbf{b}\}^*$ such that $\sigma(\alpha)$ is strongly unambiguous.*

Proof. Let $\sigma(i) := \mathbf{a}$ (i as defined in Condition 13) and, for all $j \in \mathbb{N}$ with $j \neq i$, $\sigma(j) := \varepsilon$. Then Proposition 14 follows immediately. \square

For prolix patterns with exactly two different variables, Condition 13 even characterises the subclass for which there are strongly unambiguous words:

Proposition 15. *Let $\alpha \in \mathbb{N}^+$ be prolix, $\text{var}(\alpha) := \{i, j\}$. Then there exists a morphism $\sigma : \mathbb{N}^+ \rightarrow \{\mathbf{a}, \mathbf{b}\}^*$ such that $\sigma(\alpha)$ is strongly unambiguous if and only if $|\alpha|_i \neq |\alpha|_j$.*

Proof. For the *if* part, w. l. o. g. assume $|\alpha|_i < |\alpha|_j$. Then the existence of a morphism σ such that $\sigma(\alpha)$ is strongly unambiguous is guaranteed by Proposition 14. We proceed with the *only if* part: Let $|\alpha|_i = |\alpha|_j$. Then, since α is prolix, α can only be of the form $(i \cdot j)^n$ (or $(j \cdot i)^n$), $n \in \mathbb{N}$. Thus, there is no strongly unambiguous morphic image for α . \square

Note that Propositions 14 and 15 utilise a morphism that is non-empty for a single variable only. In general, of course, one might wish to find a morphism that assigns non-empty words to a preferably large number of variables in a prolix pattern and, nevertheless, leads to a strongly unambiguous word (cf. Example 11 and Proposition 12). However, as soon as the number of variables to be mapped onto non-empty words exceeds the number of generators of the target monoid (i. e., in our case, is larger than 2), we consider it an extraordinarily challenging problem to find reasonably strong criteria.

We now return to the remaining crucial question of this paper left open after Proposition 8, Proposition 9 and Theorem 10, namely the existence of *injective* morphisms generating strongly unambiguous words for *succinct* patterns. Particularly the proof of Proposition 8 suggests that a finitely generated free monoid might not be rich enough to include strongly unambiguous morphic images for all succinct patterns. On the other hand, the proof of the comprehensive negative result for prolix patterns (cf. Theorem 10) strongly utilises the fact that these patterns exactly correspond to the fixed points of nontrivial morphisms, and, indeed, our main result shows the negation of Theorem 10 to be true for succinct patterns:

Theorem 16. *For every succinct $\alpha \in \mathbb{N}^+$, there is an injective morphism $\sigma : \mathbb{N}^+ \rightarrow \{\mathbf{a}, \mathbf{b}\}^+$ such that $\sigma(\alpha)$ is strongly unambiguous.*

The proof of Theorem 16 is given in Section 4.1.

Consequently, for every succinct string in an infinitely generated free semigroup there is a morphic image in a free monoid with two generators that—in accordance with our requirements explained in Section 1—preserves its structure. As every strongly unambiguous word by definition is weakly unambiguous as well, this statement holds for both types of definitions.

With regard to pattern languages, Theorem 16 proves that in every E-pattern language there is an unambiguous word with respect to any shortest generating pattern. In terms of equality sets (cf. Section 1), it shows that for every succinct pattern α there exists a morphism σ such that α is not a solution to the Post

Correspondence Problem for σ and any other morphism τ . Moreover, Theorem 4 explains that Theorem 16 can be rephrased in respect of fixed points of morphisms simply by substituting “non-fixed-point” for “succinct”.

Finally, we can immediately use Theorem 10 and Theorem 16 for a characterisation of succinctness:

Corollary 17. *Let $\alpha \in \mathbb{N}^+$. Then α is succinct if and only if there exists an injective morphism $\sigma : \mathbb{N}^+ \rightarrow \{\mathbf{a}, \mathbf{b}\}^+$ such that $\sigma(\alpha)$ is strongly unambiguous.*

Thus, in addition to the statements given in Theorem 3 and Theorem 4, we have a third natural problem area where the set of all finite strings is exactly partitioned into the classes “prolix” and “succinct”.

4.1. Proof of Theorem 16

The present section proves our main result on the existence of an unambiguous word for every succinct pattern. The decisive arguments verifying Theorem 16 are given in Lemma 28 and Lemma 29, and the additional lemmata, definitions, claims, examples and informal remarks are meant to facilitate or, alternatively, illustrate our formal reasoning.

We begin our argumentation with a closer look at the proof of Proposition 8 that utilises a particular phenomenon which obviously leads to the undesirable ambiguity of images under injective morphisms. In this proof, we can observe that, for the abstract succinct example pattern $\alpha = j \cdot k \cdot j \cdot k' \cdot j' \cdot k \cdot j' \cdot k'$, the ambiguity of the regarded word is caused by the fact that all occurrences of certain variables (namely k, k') have *left* neighbours (the variables j, j'), the morphic images of which have the same *suffix* (namely a word $w = \mathbf{c} \in \{\mathbf{a}, \mathbf{b}\}^+$ of length 1). Before we discuss the question of how a tailor-made morphism leading to an unambiguous image in such a case might look like, we first formalise our understanding of “neighbourship” of variables. Additionally, we anticipate that an analogous proof can be given focussing on sets of *right* neighbours of some variables, the images of which have the same *prefix*, and therefore we regard both left and right neighbours of any variable:

Definition 18. *Let $\alpha \in \mathbb{N}^+$. For every $j \in \text{var}(\alpha)$, we define the following sets: $L_j := \{k \mid \alpha = \dots \cdot k \cdot j \cdot \dots\}$ and $R_j := \{k \mid \alpha = \dots \cdot j \cdot k \cdot \dots\}$.*

Thus, L_j consists of all “left neighbours” of a variable j in a pattern α and R_j of all “right neighbours”.

As mentioned above, the identical suffixes of the images of the variables in some L_i entail the ambiguity of the corresponding word. We call a set $V \subseteq \mathbb{N}$ of variables (*morphically*) *s-homogeneous* (with respect to a morphism σ) if and only if there exists a word $w \in \{\mathbf{a}, \mathbf{b}\}^+$ such that, for every $i \in V$, $\sigma(i) = \dots w$; correspondingly, we say that V is (*morphically*) *p-homogeneous* (with respect to a morphism σ) if and only if there exists a word $w \in \{\mathbf{a}, \mathbf{b}\}^+$ such that, for every $i \in V$, $\sigma(i) = w \dots$. Additionally, V is *s-heterogeneous* if and only if it is not s-homogeneous, and it is

p -heterogeneous if and only if it is not p -homogeneous. Provided that the context is understood we simply speak of the *heterogeneity* of V .

Our initial conclusions drawn from the proof of Proposition 8 now suggest that it might be sufficient for the strong unambiguity of the image of a pattern α under a morphism σ if, for every $i \in \text{var}(\alpha)$, L_i is s -heterogeneous w. r. t. σ and R_i is p -heterogeneous w. r. t. σ . Unfortunately, however, there are patterns for which we cannot find any morphism σ showing such a property:

Example 19. Let $\alpha := 1 \cdot 2 \cdot 3 \cdot 2 \cdot 1 \cdot 3 \cdot 1$. Note that α is succinct. Thus, $L_1 = \{2, 3\}$, $L_2 = \{1, 3\}$ and $L_3 = \{1, 2\}$. Then, for a binary alphabet, there is no morphism σ such that, at a time, L_1 , L_2 and L_3 are morphically s -heterogeneous with respect to σ .

Nevertheless, morphic heterogeneity is the key to unambiguity as a thorough combinatorial examination of the subject (to be given in Lemma 29) reveals that, for a special type of morphism to be introduced in Definition 21 and discussed in Example 23 and Lemmata 24 – 28, it is sufficient to guarantee this property for certain selected L_i and R_i . More precisely, it can be shown that whenever, for sets $L_{j_1}, L_{j_2}, \dots, L_{j_n}$, $n \in \mathbb{N}$, it is $L_{j_1} \cap L_{j_2} \neq \emptyset$, $L_{j_2} \cap L_{j_3} \neq \emptyset$, \dots , $L_{j_{n-1}} \cap L_{j_n} \neq \emptyset$ then it suffices if one of these sets is s -heterogeneous (and the same holds for any $R_{j_1}, R_{j_2}, \dots, R_{j_n}$ with respect to p -heterogeneity). Thus, in an equivalent view and for every succinct pattern α , we now restrict our considerations to particular equivalence classes on $\text{var}(\alpha)$:

Definition 20. Let $\alpha \in \mathbb{N}^+$. Let, for every $j \in \text{var}(\alpha)$, the sets L_j and R_j be given according to Definition 18. With these sets, construct two equivalence relations \sim_l and \sim_r on $\text{var}(\alpha)$: For all $k, k' \in \text{var}(\alpha)$

- $k \sim_l k'$ if and only if there are $j_1, j_2, \dots, j_s \in \text{var}(\alpha)$, $s \geq 1$, such that
 - (1) $L_{j_1} \cap L_{j_2} \neq \emptyset$, $L_{j_2} \cap L_{j_3} \neq \emptyset$, \dots , $L_{j_{s-1}} \cap L_{j_s} \neq \emptyset$ and
 - (2) $k \in L_{j_1}$ and $k' \in L_{j_s}$.
- $k \sim_r k'$ if and only if there are $j_1, j_2, \dots, j_t \in \text{var}(\alpha)$, $t \geq 1$, such that
 - (1) $R_{j_1} \cap R_{j_2} \neq \emptyset$, $R_{j_2} \cap R_{j_3} \neq \emptyset$, \dots , $R_{j_{t-1}} \cap R_{j_t} \neq \emptyset$ and
 - (2) $k \in R_{j_1}$ and $k' \in R_{j_t}$.

Then, given in arbitrary order, let $L_1^\sim, L_2^\sim, \dots, L_p^\sim \subseteq \text{var}(\alpha)$ be all equivalence classes resulting from \sim_l and $R_1^\sim, R_2^\sim, \dots, R_q^\sim \subseteq \text{var}(\alpha)$ all equivalence classes resulting from \sim_r .

Consequently, in particular, two variables k, k' belong to the same equivalence class L_i^\sim if they are in the same set L_j (i. e., in terms of Definition 20, $L_j = L_{j_1} = L_{j_s}$) or if they are in two different, but non-disjoint sets $L_j, L_{j'}$ (i. e. $s = 2$, $j = j_1$, $j' = j_2$ and $L_{j_1} \cap L_{j_2} \neq \emptyset$). Obviously, the same holds for the sets R_j and the equivalence classes R_i^\sim .

Since all L_i^\sim (as well as all R_i^\sim) are pairwise disjoint we can avoid those problems described in Example 19. Thus, we can use Definition 20 for introducing a well-defined morphism $\sigma_\alpha^{\text{su}}$ which depends on the structure of a pattern α :

Definition 21. Let $\alpha \in \mathbb{N}^+$. Let $L_1^\sim, L_2^\sim, \dots, L_p^\sim$ and $R_1^\sim, R_2^\sim, \dots, R_q^\sim$ be the equivalence classes on $\text{var}(\alpha)$ given in Definition 20. Then, for $i \in \{1, 2, \dots, p\}$, $i' \in \{1, 2, \dots, q\}$ and for every $k \in \mathbb{N}$, the morphism $\sigma_\alpha^{\text{su}}$ is defined by

$$\sigma_\alpha^{\text{su}}(k) := \begin{cases} \mathbf{a} \mathbf{b}^{3k} \mathbf{a} \mathbf{a} \mathbf{b}^{3k+1} \mathbf{a} \mathbf{a} \mathbf{b}^{3k+2} \mathbf{a}, & \nexists i : k = \min L_i^\sim \wedge \nexists i' : k = \min R_{i'}^\sim, \\ \mathbf{b} \mathbf{a}^{3k} \mathbf{b} \mathbf{a} \mathbf{b}^{3k+1} \mathbf{a} \mathbf{a} \mathbf{b}^{3k+2} \mathbf{a}, & \nexists i : k = \min L_i^\sim \wedge \exists i' : k = \min R_{i'}^\sim, \\ \mathbf{a} \mathbf{b}^{3k} \mathbf{a} \mathbf{a} \mathbf{b}^{3k+1} \mathbf{a} \mathbf{b} \mathbf{a}^{3k+2} \mathbf{b}, & \exists i : k = \min L_i^\sim \wedge \nexists i' : k = \min R_{i'}^\sim, \\ \mathbf{b} \mathbf{a}^{3k} \mathbf{b} \mathbf{a} \mathbf{b}^{3k+1} \mathbf{a} \mathbf{b} \mathbf{a}^{3k+2} \mathbf{b}, & \exists i : k = \min L_i^\sim \wedge \exists i' : k = \min R_{i'}^\sim. \end{cases}$$

Obviously, for every $\alpha \in \mathbb{N}^+$, $\sigma_\alpha^{\text{su}}$ is injective. Moreover note that, in the conditions in the definition of $\sigma_\alpha^{\text{su}}$, we choose the *minimum* variables in the L_i^\sim and R_i^\sim merely for the sake of convenience: Actually, the claim for s-heterogeneity of the L_i^\sim and p-heterogeneity of the R_i^\sim only requires that in each equivalence class with more than one element there is at least one variable k_1 matching the first case (i. e. $\sigma_\alpha^{\text{su}}(k_1) = \mathbf{a}[\dots]\mathbf{a}$) and at least one variable k_2 (our definition changes this to *exactly* one variable k_2 , namely the minimum one) matching an appropriate different case (i. e. $\sigma_\alpha^{\text{su}}(k_2) \neq \mathbf{a}[\dots]\mathbf{a}$).

We now consider an example pattern so as to illustrate Definition 18, Definition 20 and Definition 21:

Example 22. Let $\alpha := 1 \cdot 2 \cdot 3 \cdot 1 \cdot 2 \cdot 4 \cdot 3 \cdot 1 \cdot 2 \cdot 3 \cdot 2 \cdot 4$. Note that α is succinct. Then $L_1 = \{3\}$, $L_2 = \{1, 3\}$, $L_3 = \{2, 4\}$, $L_4 = \{2\}$ and $R_1 = \{2\}$, $R_2 = \{3, 4\}$, $R_3 = \{1, 2\}$, $R_4 = \{3\}$ (see Definition 18). According to Definition 20, this leads to $L_1^\sim = \{1, 3\}$, $L_2^\sim = \{2, 4\}$ and $R_1^\sim = \{1, 2\}$, $R_2^\sim = \{3, 4\}$, and, thus, by Definition 21

$$\begin{aligned} \sigma_\alpha^{\text{su}}(1) &= \mathbf{b} \mathbf{a}^3 \mathbf{b} \mathbf{a} \mathbf{b}^4 \mathbf{a} \mathbf{b} \mathbf{a}^5 \mathbf{b}, \\ \sigma_\alpha^{\text{su}}(2) &= \mathbf{a} \mathbf{b}^6 \mathbf{a} \mathbf{a} \mathbf{b}^7 \mathbf{a} \mathbf{b} \mathbf{a}^8 \mathbf{b}, \\ \sigma_\alpha^{\text{su}}(3) &= \mathbf{b} \mathbf{a}^9 \mathbf{b} \mathbf{a} \mathbf{b}^{10} \mathbf{a} \mathbf{a} \mathbf{b}^{11} \mathbf{a} \text{ and} \\ \sigma_\alpha^{\text{su}}(4) &= \mathbf{a} \mathbf{b}^{12} \mathbf{a} \mathbf{a} \mathbf{b}^{13} \mathbf{a} \mathbf{a} \mathbf{b}^{14} \mathbf{a}. \end{aligned}$$

It is evident, that $\sigma_\alpha^{\text{su}}$ is much more complicated than it is needed for assuring heterogeneity since this property solely depends on the first and the last letter of the morphic images of the variables. In this regard, the subsequent examples explain that, in general, mere heterogeneity is not sufficient for guaranteeing unambiguity:

Example 23. Let $\alpha_1 := 1 \cdot 1 \cdot 2 \cdot 2 \cdot 3 \cdot 3$. This pattern is succinct, and $L_1^\sim = \text{var}(\alpha_1)$, $R_1^\sim = \text{var}(\alpha_1)$ (cf. Definition 20). Then, for L_1^\sim and R_1^\sim , the non-injective morphism σ given by $\sigma(1) := \mathbf{b}$ and $\sigma(i) := \mathbf{a}$, $i \in \text{var}(\alpha) \setminus \{1\}$ leads to the desired heterogeneity. Nevertheless, there is a morphism τ with $\tau(\alpha_1) = \sigma(\alpha_1)$ and $\tau \neq \sigma$, namely $\tau(1) := \mathbf{b}$, $\tau(2) := \varepsilon$, and $\tau(3) := \mathbf{a} \mathbf{a}$.

Our initial succinct example $\alpha' = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 1 \cdot 4 \cdot 3 \cdot 2$ shows that, contrary to what might be suggested by α_1 , there are square-free patterns with the same property.

With regard to α' , $R_1^\sim = L_1^\sim = \{1, 3\}$ and $R_2^\sim = L_2^\sim = \{2, 4\}$. Thus, the morphism σ with $\sigma(1) = \sigma(2) := \mathbf{a}$ and $\sigma(3) = \sigma(4) := \mathbf{b}$ entails heterogeneity, but, with the morphism $\tau(1) = \tau(2) := \mathbf{a}$, $\tau(3) := \varepsilon$, $\tau(4) := \mathbf{b}\mathbf{b}$, $\tau(\alpha') = \sigma(\alpha')$.

Example 23 reveals that there is a second reason for ambiguity not covered by our considerations on the pattern $\alpha = j \cdot k \cdot j \cdot k' \cdot j' \cdot k \cdot j' \cdot k'$ introduced in the proof of Proposition 8 and discussed in the beginning of the present section. For the patterns α_1 and α' in Example 23 and the regarded morphisms σ , the ambiguity of $\sigma(\alpha_1)$ and $\sigma(\alpha')$ results from the fact that there are variables $i, i' \in \text{var}(\alpha_1)$ and $i, i' \in \text{var}(\alpha')$, the images of which under the respective σ commute, i. e. $\sigma(i \cdot i') = \sigma(i' \cdot i)$. It is a classical insight in combinatorics on words that such a phenomenon can be avoided by choosing a morphism where the primitive root of $\sigma(i)$ does not equal the primitive root of $\sigma(i')$ (cf. Lothaire [11]), i. e., for some $m, n \in \mathbb{N}$, if $\rho(\sigma(i))$ is the shortest word such that $\rho(\sigma(i))^m = \sigma(i)$ and $\rho(\sigma(i'))$ is the shortest word such that $\rho(\sigma(i'))^n = \sigma(i')$ then $\rho(\sigma(i)) \neq \rho(\sigma(i'))$. Note that the unambiguous image $w_3 = \sigma''(\alpha') = \mathbf{a}\mathbf{a}\mathbf{b}\mathbf{b}\mathbf{b}\mathbf{a}\mathbf{a}\mathbf{b}\mathbf{a}\mathbf{b}\mathbf{a}\mathbf{b}$ of α' given in Section 1 is exactly based on these two requirements: the corresponding morphism σ'' leads to heterogeneity of all L_i^\sim and R_i^\sim , and, for each pair of variables $i, i' \in \text{var}(\alpha')$ with $i \neq i'$, $\rho(\sigma''(i)) \neq \rho(\sigma''(i'))$.

As shown by the morphism σ'' and suggested by some simple additional considerations, there are morphisms with much shorter images than $\sigma_\alpha^{\text{su}}$ which guarantee heterogeneity of the equivalence classes introduced in Definition 20 and non-commutativity of the morphic images of variables (or, more general: of subpatterns). In fact, there are two particular reasons for $\sigma_\alpha^{\text{su}}$ to map every variable $j \in \mathbb{N}$ onto three distinct segments $c \mathbf{d}^{3j+p} c$, $p \in \{0, 1, 2\}$, with $c, \mathbf{d} \in \{\mathbf{a}, \mathbf{b}\}$, $c \neq \mathbf{d}$. First, this obviously leads to the injectivity of $\sigma_\alpha^{\text{su}}$ such that the morphism conforms with our requirements explained in Section 1. Second, we can show that the chosen shape of $\sigma_\alpha^{\text{su}}$ “restricts” the ambiguity of $\sigma_\alpha^{\text{su}}(\alpha)$ (the exact meaning of this vague term can be derived from Lemma 28) simply by assigning the said three segments to each variable and, thus, without considering heterogeneity. Note that a very similar proof technique is used in [15].

Formalising these considerations we now can finally prove that, for every succinct $\alpha \in \mathbb{N}^+$, $\sigma_\alpha^{\text{su}}(\alpha)$ is strongly unambiguous, a statement which directly implies the correctness of Theorem 16. We begin with a list of simple combinatorial observations which impose a number of strong restrictions on a morphism τ with $\tau(\alpha) = \sigma_\alpha^{\text{su}}(\alpha)$. These first formal remarks exclusively deal with those variables j in α for which $\tau(j)$ contains any segment of $\sigma_\alpha^{\text{su}}(j)$ (note that a straightforward reasoning proves the existence of such variables for every morphism τ with $\tau(\alpha) = \sigma_\alpha^{\text{su}}(\alpha)$). Since the subsequent lemmata are needed for an unobstructed understanding of our main reasoning in the Lemmata 28 and 29, it is important to keep them in mind.

Our first lemma says that, for any morphism τ with $\tau(\alpha) = \sigma_\alpha^{\text{su}}(\alpha)$ and for any variable j in α , $\tau(j)$ does not contain any segment of $\sigma_\alpha^{\text{su}}(j)$ twice (or even more than twice). As the examined phenomenon does not require heterogeneity of the equivalence classes introduced in Definition 20, we can apply, for any pattern β ,

the morphism σ_β^{su} to any pattern α (note that every σ_β^{su} is defined explicitly for all $k \in \mathbb{N}$, cf. Definition 21):

Lemma 24. *Let $\alpha, \beta \in \mathbb{N}^+$. Let $\tau : \mathbb{N}^+ \rightarrow \{\mathbf{a}, \mathbf{b}\}^*$ be any morphism with $\tau(\alpha) = \sigma_\beta^{\text{su}}(\alpha)$. Then, for every $j \in \text{var}(\alpha)$, for every $p \in \{0, 1, 2\}$ and for every $\mathbf{c}, \mathbf{d} \in \{\mathbf{a}, \mathbf{b}\}$, $\mathbf{c} \neq \mathbf{d}$, $\tau(j) \neq \dots \mathbf{c} \mathbf{d}^{3j+p} \mathbf{c} \dots \mathbf{c} \mathbf{d}^{3j+p} \mathbf{c} \dots$.*

Proof. Assume to the contrary that there is a $j \in \text{var}(\alpha)$ with $|\tau(j)|_{\mathbf{c} \mathbf{d}^{3j+p} \mathbf{c}} \geq 2$. Then $|\tau(\alpha)|_{\mathbf{c} \mathbf{d}^{3j+p} \mathbf{c}} \geq 2|\alpha|_j > |\alpha|_j = |\sigma_\beta^{\text{su}}(\alpha)|_{\mathbf{c} \mathbf{d}^{3j+p} \mathbf{c}}$. This contradicts the condition $\tau(\alpha) = \sigma_\beta^{\text{su}}(\alpha)$. \square

In order to address the segments more precisely we henceforth use the following terms: For any variable $j \in \mathbb{N}$ and $\mathbf{c}, \mathbf{d}, \mathbf{e}, \mathbf{f} \in \{\mathbf{a}, \mathbf{b}\}$, $\mathbf{c} \neq \mathbf{d}$, $\mathbf{e} \neq \mathbf{f}$, we call $\mathbf{c} \mathbf{d}^{3j} \mathbf{c}$ the *left segment*, $\mathbf{a} \mathbf{b}^{3j+1} \mathbf{a}$ the *inner segment* and $\mathbf{e} \mathbf{f}^{3j+2} \mathbf{e}$ the *right segment* of $\sigma_\alpha^{\text{su}}(j)$. Additionally, we extend this terminology to segments of the image of any (sub-)pattern $\gamma \in \mathbb{N}^+$: A word $w \in \{\mathbf{a}, \mathbf{b}\}^+$ is a segment of $\sigma_\alpha^{\text{su}}(\gamma)$ if and only if there exists a variable $j \in \text{var}(\gamma)$ such that w is a segment of $\sigma_\alpha^{\text{su}}(j)$. The next lemma says that if, for any morphism τ with $\tau(\alpha) = \sigma_\alpha^{\text{su}}(\alpha)$ and for any variable j in α , $\tau(j)$ contains the left and the inner segment (or the inner and the right segment) of $\sigma_\alpha^{\text{su}}(j)$ then these segments must occur in the “natural” order:

Lemma 25. *Let $\alpha, \beta \in \mathbb{N}^+$. Let $\tau : \mathbb{N}^+ \rightarrow \{\mathbf{a}, \mathbf{b}\}^*$ be any morphism with $\tau(\alpha) = \sigma_\beta^{\text{su}}(\alpha)$. For every $j \in \text{var}(\alpha)$, for every $p \in \{0, 1\}$ and for every $\mathbf{c}, \mathbf{d}, \mathbf{e}, \mathbf{f} \in \{\mathbf{a}, \mathbf{b}\}$, $\mathbf{c} \neq \mathbf{d}$, $\mathbf{e} \neq \mathbf{f}$, if $\tau(j) = \dots \mathbf{c} \mathbf{d}^{3j+p} \mathbf{c} \dots$ and $\tau(j) = \dots \mathbf{e} \mathbf{f}^{3j+p+1} \mathbf{e} \dots$ then $\tau(j) = \dots \mathbf{c} \mathbf{d}^{3j+p} \mathbf{c} \mathbf{e} \mathbf{f}^{3j+p+1} \mathbf{e} \dots$.*

Proof. Because of $\tau(\alpha) = \sigma(\alpha)$ we know that $|\tau(\alpha)|_{\mathbf{c} \mathbf{d}^{3j+p} \mathbf{c}} = |\tau(\alpha)|_{\mathbf{e} \mathbf{f}^{3j+p+1} \mathbf{e}} = |\tau(\alpha)|_{\mathbf{c} \mathbf{d}^{3j+p} \mathbf{c} \mathbf{e} \mathbf{f}^{3j+p+1} \mathbf{e}} = |\alpha|_j$. Thus, for every occurrence of $\mathbf{c} \mathbf{d}^{3j+p} \mathbf{c}$ and $\mathbf{e} \mathbf{f}^{3j+p+1} \mathbf{e}$ in $\tau(\alpha)$, $\tau(\alpha) = \dots \mathbf{c} \mathbf{d}^{3j+p} \mathbf{c} \mathbf{e} \mathbf{f}^{3j+p+1} \mathbf{e} \dots$. Since $\tau(j) = \dots \mathbf{c} \mathbf{d}^{3j+p} \mathbf{c} \dots$ and $\tau(j) = \dots \mathbf{e} \mathbf{f}^{3j+p+1} \mathbf{e} \dots$, Lemma 25 follows immediately. \square

Furthermore, it follows similarly that if, for any morphism τ with $\tau(\alpha) = \sigma_\alpha^{\text{su}}(\alpha)$ and for any variable j in α , $\tau(j)$ contains the left and the right segment of $\sigma_\alpha^{\text{su}}(j)$ then it must also contain the inner segment of $\sigma_\alpha^{\text{su}}(j)$ and, again, these segments must occur in the canonical order:

Lemma 26. *Let $\alpha, \beta \in \mathbb{N}^+$. Let $\tau : \mathbb{N}^+ \rightarrow \{\mathbf{a}, \mathbf{b}\}^*$ be any morphism with $\tau(\alpha) = \sigma_\beta^{\text{su}}(\alpha)$. For every $j \in \text{var}(\alpha)$ and for every $\mathbf{c}, \mathbf{d}, \mathbf{e}, \mathbf{f} \in \{\mathbf{a}, \mathbf{b}\}$, $\mathbf{c} \neq \mathbf{d}$, $\mathbf{e} \neq \mathbf{f}$, if $\tau(j) = \dots \mathbf{c} \mathbf{d}^{3j} \mathbf{c} \dots$ and $\tau(j) = \dots \mathbf{e} \mathbf{f}^{3j+2} \mathbf{e} \dots$ then $\tau(j) = \dots \mathbf{c} \mathbf{d}^{3j} \mathbf{c} \mathbf{a} \mathbf{b}^{3j+1} \mathbf{a} \mathbf{e} \mathbf{f}^{3j+2} \mathbf{e} \dots$.*

We conclude our list of preliminary observations on the impact of some $j \in \text{var}(\alpha)$ for which $\tau(j)$ contains any segment $s = \mathbf{c} \mathbf{d}^{3j+p} \mathbf{c}$, $p \in \{0, 1, 2\}$, with an immediate consequence of Lemma 24. It is based on the fact that, for any n , $1 \leq n \leq |\alpha|_j$, the n th occurrence (counted from the left) of j in α under both σ and τ

necessarily generates the n th occurrence of s in $\sigma_\alpha^{\text{su}}(\alpha) = \tau(\alpha)$. Thus, in addition to what is stated in Lemmata 24, 25, 26, we can also use a variable j with the said property for drawing conclusions about the images under τ of neighbours of j in α . In anticipation of the requirements of the subsequent main Lemmata 28, 29 and for the sake of a more concise presentation, we focus on a variable $j \in \text{var}(\alpha)$ for which $\tau(j)$ contains the *inner* segment of $\sigma_\alpha^{\text{su}}(j)$:

Lemma 27. *Let $\alpha, \beta \in \mathbb{N}^+$. Let, for some variable $j \in \text{var}(\alpha)$ and $\alpha_1, \alpha_2 \in \mathbb{N}^*$, $\alpha = \alpha_1 \cdot j \cdot \alpha_2$. Let $\tau : \mathbb{N}^+ \rightarrow \{\mathbf{a}, \mathbf{b}\}^*$ be any morphism with $\tau(\alpha) = \sigma_\beta^{\text{su}}(\alpha)$. If, for some $w_1, w_2 \in \{\mathbf{a}, \mathbf{b}\}^*$, $\tau(j) = w_1 \mathbf{a} \mathbf{b}^{3j+1} \mathbf{a} w_2$ then, for some $\mathbf{c}, \mathbf{d}, \mathbf{e}, \mathbf{f} \in \{\mathbf{a}, \mathbf{b}\}$, $\mathbf{c} \neq \mathbf{d}$, $\mathbf{e} \neq \mathbf{f}$,*

- $\tau(\alpha_1) w_1 = \sigma(\alpha_1) \mathbf{c} \mathbf{d}^{3j} \mathbf{c}$ and
- $w_2 \tau(\alpha_2) = \mathbf{e} \mathbf{f}^{3j+2} \mathbf{e} \sigma(\alpha_2)$.

We now proceed with a discussion of the consequences of $\sigma_\alpha^{\text{su}}$ mapping every variable onto a word that consists of three distinct segments. The subsequent crucial lemma shows that, for *every* morphism τ with $\tau(\alpha) = \sigma_\alpha^{\text{su}}(\alpha)$ and for *every* variable $j \in \text{var}(\alpha)$, $\tau(j)$ must contain the inner segment of $\sigma_\alpha^{\text{su}}(j)$ and the letter to the left and the letter to the right of this segment—this is what above has been referred to as a “restricted” ambiguity of $\sigma_\alpha^{\text{su}}(\alpha)$. Recall that the examined property does not depend on the heterogeneity of the equivalence classes L_i^\sim and R_i^\sim , and therefore we again give the lemma in terms of the image of a succinct pattern α under any morphism σ_β^{su} , $\beta \in \mathbb{N}^+$:

Lemma 28. *Let $\alpha, \beta \in \mathbb{N}^+$. Then α is succinct if and only if, for every morphism $\tau : \mathbb{N}^+ \rightarrow \{\mathbf{a}, \mathbf{b}\}^*$ with $\tau(\alpha) = \sigma_\beta^{\text{su}}(\alpha)$ and for every $i \in \text{var}(\alpha)$, $\tau(i) = \dots \mathbf{g} \mathbf{a} \mathbf{b}^{3i+1} \mathbf{a} \mathbf{h} \dots$, $\mathbf{g}, \mathbf{h} \in \{\mathbf{a}, \mathbf{b}\}$.*

Proof. We begin with the *if*-direction and we prove it by contraposition. Hence, for any prolix α , we show that there is a morphism τ and an $i \in \text{var}(\alpha)$ such that $\tau(\alpha) = \sigma_\beta^{\text{su}}(\alpha)$, but $\tau(i) \neq \dots \mathbf{g} \mathbf{a} \mathbf{b}^{3i+1} \mathbf{a} \mathbf{h} \dots$:

Let α be prolix. Then there is a decomposition $\alpha = \beta_0 \gamma_1 \beta_1 \gamma_2 \beta_2 [\dots] \beta_{n-1} \gamma_n \beta_n$ satisfying the conditions of Definition 1. With regard to this decomposition and for every k , $1 \leq k \leq n$, let i_k be the smallest $i \in \text{var}(\gamma_k)$ such that $|\gamma_k|_{i_k} = 1$ and, for every k' , $1 \leq k' \leq n$, if $i_k \in \text{var}(\gamma_{k'})$, then $\gamma_k = \gamma_{k'}$. By definition (i. e., more precisely, the third condition of Definition 1), for every γ_k , $1 \leq k \leq n$, such an i_k exists, and, for every $\beta_{k'}$, $0 \leq k' \leq n$, $i_k \notin \text{var}(\beta_{k'})$. Now we define τ as follows: For all k , $1 \leq k \leq n$, $\tau(i_k) := \sigma_\beta^{\text{su}}(\gamma_k)$, for all $i \in \text{var}(\gamma_k) \setminus \{i_k\}$, $\tau(i) := \varepsilon$, and, for all k , $0 \leq k \leq n$ and for all $i \in \text{var}(\beta_k)$, $\tau(i) := \sigma_\beta^{\text{su}}(i)$. As a consequence of the second and the third condition of Definition 1, τ is well-defined. Then $\sigma_\beta^{\text{su}}(\alpha) = \tau(\alpha)$, but, for all $i \in \text{var}(\gamma_k) \setminus \{i_k\}$ (which exist because of the first condition of Definition 1), $\tau(i) = \varepsilon \neq \dots \mathbf{g} \mathbf{a} \mathbf{b}^{3i+1} \mathbf{a} \mathbf{h} \dots$.

We proceed with the *only if*-direction and, again, we argue by contraposition.

In other words, we show that if there exists a morphism τ with $\tau(\alpha) = \sigma_{\beta}^{\text{su}}(\alpha)$ and, for some $i \in \text{var}(\alpha)$, $\tau(i) \neq \dots \mathbf{g a b}^{3i+1} \mathbf{a h} \dots$ then α is prolix:

We start with a partition of $\text{var}(\alpha)$ into subsets X_1, X_2, X_3 depending on any morphism τ satisfying the said conditions. From an informal point of view, this partition is given as follows: First, let X_1 be the set of all variables i in α such that $\tau(i)$ contains the inner segment of $\sigma_{\beta}^{\text{su}}(i)$, of the left and right segment of $\sigma_{\beta}^{\text{su}}(i)$ at least one letter and at least one segment of some $\sigma_{\beta}^{\text{su}}(j)$, $j \neq i$. Second, let X_2 be the set of all variables i in α such that $\tau(i)$ does not contain any letter of at least one segment of $\sigma_{\beta}^{\text{su}}(i)$. Third (and last), let X_3 be the set of all variables i in α such that $\tau(i)$ contains the inner segment of $\sigma_{\beta}^{\text{su}}(i)$ and at least one letter of the left and right segment of $\sigma_{\beta}^{\text{su}}(i)$, but no segment of some $\sigma_{\beta}^{\text{su}}(j)$, $j \neq i$.

Since $\tau(\alpha) = \sigma_{\beta}^{\text{su}}(\alpha)$ and thus, for every $i \in \text{var}(\alpha)$, $\tau(i)$ is a subword of $\sigma_{\beta}^{\text{su}}(\alpha)$ this vague definition of X_1, X_2 and X_3 results in several evident restrictions on the images under τ of the variables in α (cf. Lemmata 24, 25, 26, 27) such that the introduced subsets of $\text{var}(\alpha)$ read formally:

$$\begin{aligned} X_1 &:= \{i \in \text{var}(\alpha) \mid \tau(i) = \dots \mathbf{c d}^{3j+2} \mathbf{c e f}^{3i} \mathbf{e a b}^{3i+1} \mathbf{a g} \dots \text{ or} \\ &\quad \tau(i) = \dots \mathbf{h a b}^{3i+1} \mathbf{a c d}^{3i+2} \mathbf{c e f}^{3j} \mathbf{e} \dots, \\ &\quad \mathbf{c, d, e, f, g, h} \in \{\mathbf{a, b}\}, \mathbf{c} \neq \mathbf{d}, \mathbf{e} \neq \mathbf{f}, j \in \mathbb{N}\}, \\ X_2 &:= \{i \in \text{var}(\alpha) \mid \tau(i) \neq \dots \mathbf{g a b}^{3i+1} \mathbf{a h} \dots, \mathbf{g, h} \in \{\mathbf{a, b}\}\}, \\ X_3 &:= \{i \in \text{var}(\alpha) \mid \tau(i) = \dots \mathbf{g a b}^{3i+1} \mathbf{a h} \dots \text{ and} \\ &\quad \tau(i) \neq \dots \mathbf{c d}^{3j+2} \mathbf{c e f}^{3i} \mathbf{e} \dots \text{ and} \\ &\quad \tau(i) \neq \dots \mathbf{c d}^{3i+2} \mathbf{c e f}^{3j} \mathbf{e} \dots, \\ &\quad \mathbf{c, d, e, f, g, h} \in \{\mathbf{a, b}\}, \mathbf{c} \neq \mathbf{d}, \mathbf{e} \neq \mathbf{f}, j \in \mathbb{N}\}. \end{aligned}$$

Directly from the definition, it can be verified that $X_1 \cap X_2 = X_1 \cap X_3 = X_2 \cap X_3 = \emptyset$ and $X_1 \cup X_2 \cup X_3 = \text{var}(\alpha)$. According to the condition on our proof for the *only if*-direction, there is a variable $i \in \text{var}(\alpha)$ with $\tau(i) \neq \dots \mathbf{g a b}^{3i+1} \mathbf{a h} \dots$ and therefore $X_2 \neq \emptyset$. Note that our subsequent argumentation in Claim 3 shows that this leads to $X_1 \neq \emptyset$.

As we now wish to show that $X_2 \neq \emptyset$ implies α being prolix, we need to find an appropriate decomposition of α satisfying the three conditions of Definition 1. We start our argumentation with the following one:

$$\alpha = \bar{\beta}_0 \bar{\gamma}_1 \bar{\beta}_1 \bar{\gamma}_2 \bar{\beta}_2 [\dots] \bar{\beta}_{\bar{m}-1} \bar{\gamma}_{\bar{m}} \bar{\beta}_{\bar{m}}$$

with $\bar{m} \geq 1$ and

- $\bar{\beta}_0, \bar{\beta}_{\bar{m}} \in X_3^*$ and $\bar{\beta}_k \in X_3^+$, $1 \leq k \leq \bar{m} - 1$, and
- $\bar{\gamma}_k \in (X_1 \cup X_2)^+$, $1 \leq k \leq \bar{m}$.

Note that $\bar{m} \geq 1$ is granted because of $X_2 \neq \emptyset$.

Obviously, this decomposition is unique, and it satisfies the second condition of Definition 1 since X_1, X_2 and X_3 are disjoint:

Claim 1. For every $k, k', 1 \leq k, k' \leq \bar{m}$, $\text{var}(\bar{\gamma}_k) \cap \text{var}(\bar{\beta}_{k'}) = \emptyset$.

Concerning the first condition of Definition 1 we need to examine the given decomposition of α in a bit more detail. The subsequent claim says that, for every $\bar{\gamma}_k$, $\tau(\bar{\gamma}_k)$ “almost” corresponds to $\sigma_\beta^{\text{su}}(\bar{\gamma}_k)$, i. e. $\tau(\bar{\gamma}_k)$ contains *at least* $3|\bar{\gamma}_k| - 2$ complete segments of $\sigma_\beta^{\text{su}}(\bar{\gamma}_k)$ (and potentially some letters of two other segments) and *at most* $3|\bar{\gamma}_k|$ complete segments of $\sigma_\beta^{\text{su}}(\bar{\gamma}_k)$ and, moreover, for every variable $i \in \text{var}(\bar{\gamma}_k)$ the inner segment of $\sigma_\beta^{\text{su}}(i)$ as often as i is contained in $\bar{\gamma}_k$, and it does not contain any segment of $\sigma_\beta^{\text{su}}(i')$ if $i' \notin \text{var}(\bar{\gamma}_k)$:

Claim 2. For every $\bar{\gamma}_k = i_1 \cdot i_2 \cdot [\dots] \cdot i_s$, $s \in \mathbb{N}$, and for every $i' \in \text{var}(\alpha)$,

$$\begin{aligned} \tau(\bar{\gamma}_k) &= \dots \mathbf{g a b}^{3i_1+1} \mathbf{a c d}^{3i_1+2} \mathbf{c e f}^{3i_2} \mathbf{e} [\dots] \mathbf{a b}^{3i_s+1} \mathbf{a o} \dots \text{ \underline{and}} \\ \tau(\bar{\gamma}_k) &\neq \dots \mathbf{o a b}^{3i_1+1} \mathbf{a c d}^{3i_1+2} \mathbf{c} [\dots] \mathbf{e f}^{3i_s+2} \mathbf{e g h}^{3i'} \mathbf{g} \dots \text{ \underline{and}} \\ \tau(\bar{\gamma}_k) &\neq \dots \mathbf{g h}^{3i'+2} \mathbf{g c d}^{3i_1} \mathbf{c a b}^{3i_1+1} \mathbf{a} [\dots] \mathbf{a b}^{3i_s+1} \mathbf{a o} \dots, \end{aligned}$$

$\mathbf{c, d, e, f, g, h, o} \in \{\mathbf{a, b}\}$ with $\mathbf{c} \neq \mathbf{d}$, $\mathbf{e} \neq \mathbf{f}$, $\mathbf{g} \neq \mathbf{h}$.

For any subpattern δ of α satisfying the statement in Claim 2 we say that $\tau(\delta)$ *corresponds to* $\sigma_\beta^{\text{su}}(\delta)$ (apart from a negligible prefix and suffix).

Claim 2 follows from the fact that every $\bar{\gamma}_k$ is surrounded by $\bar{\beta}_{k-1} \in X_3^*$ and $\bar{\beta}_k \in X_3^*$. Thus, with Lemma 27 applied to the variables in X_3 , $\tau(\bar{\gamma}_k)$ is fixed by $\tau(\bar{\beta}_{k-1})$ and $\tau(\bar{\beta}_k)$: as these two subwords of $\tau(\alpha)$ by definition correspond to $\sigma_\beta^{\text{su}}(\bar{\beta}_{k-1})$ and $\sigma_\beta^{\text{su}}(\bar{\beta}_k)$, respectively, $\tau(\bar{\gamma}_k)$ must also correspond to $\sigma_\beta^{\text{su}}(\bar{\gamma}_k)$. Consequently—and since by definition, for no $\delta \in X_1^+$, $\tau(\delta)$ corresponds to $\sigma_\beta^{\text{su}}(\delta)$ — $\bar{\gamma}_k \notin X_1^+$.

We proceed our argumentation on the first condition of Definition 1 being satisfied for the regarded decomposition by a closer look at the images under τ of those subpatterns δ of α which exclusively consist of variables in X_2 . In this regard we can see that $\tau(\delta)$ necessarily does not correspond to $\sigma_\beta^{\text{su}}(\delta)$:

Claim 3. For every $\delta = i_1 \cdot i_2 \cdot [\dots] \cdot i_t$ with $t \in \mathbb{N}$, $\delta \in X_2^+$,

$$\tau(\delta) \neq \dots \mathbf{g a b}^{3i_1+1} \mathbf{a c d}^{3i_1+2} \mathbf{c e f}^{3i_2} \mathbf{e} [\dots] \mathbf{a b}^{3i_t+1} \mathbf{a h} \dots,$$

$\mathbf{c, d, e, f, g, h} \in \{\mathbf{a, b}\}$ with $\mathbf{c} \neq \mathbf{d}$, $\mathbf{e} \neq \mathbf{f}$.

The correctness of Claim 3 follows from a straightforward combinatorial examination of the definition of X_2 . Thus, and from Claim 2, it is $\bar{\gamma}_k \notin X_2^+$ and therefore $\bar{\gamma}_k$ must consist of variables in X_1 and of variables in X_2 :

Claim 4. For every k , $1 \leq k \leq \bar{m}$, $|\bar{\gamma}_k| \geq 2$.

Hence, the first condition of Definition 1 is satisfied for the above decomposition.

With regard to the third condition, however, the decomposition possibly requires some modifications. We wish to have a decomposition where there is *exactly* one occurrence of an $i \in X_1$ in each $\bar{\gamma}_k$ since this variable is meant to serve as the variable i_k referred to in the third condition of Definition 1. For the given decomposition, however, we can only conclude that there is *at least* one occurrence of an $i \in X_1$

in each $\bar{\gamma}_k$. Therefore we transform it into a specific decomposition where every $\bar{\gamma}_k$ contains exactly one $i \in X_1$. To this end, we apply two different types of operations, namely a *splitting* of certain $\bar{\gamma}_k$ and a *redefinition* of X_1 and X_3 .

We first split every $\bar{\gamma}_k$ that contains more than one occurrence of a variable from X_1 , and we do so by identifying all so-called *splitting points* in $\bar{\gamma}_k$. Intuitively, these splitting points lead to a maximum $s \in \mathbb{N}$ for which there exists a decomposition $\bar{\gamma}_k = \bar{\gamma}_{k,1} \bar{\gamma}_{k,2} [\dots] \bar{\gamma}_{k,s}$ such that, for every k' , $1 \leq k' \leq s$, $\tau(\bar{\gamma}_{k,k'})$ corresponds to $\sigma_\beta^{\text{su}}(\bar{\gamma}_{k,k'})$ “as far as possible”. Formally, a splitting point is an inner substring δ of $\bar{\gamma}_k$, i. e. $\bar{\gamma}_k = \bar{\gamma}_{k,l} \delta \bar{\gamma}_{k,r}$ with $\bar{\gamma}_{k,l} = i_1 \cdot i_2 \cdot [\dots] \cdot i_p$ and $\bar{\gamma}_{k,r} = i_{p+1} \cdot i_{p+2} \cdot [\dots] \cdot i_{p+q}$, $p, q \in \mathbb{N}$, $i_1, i_2, \dots, i_{p+q} \in X_1 \cup X_2$, that satisfies one of the following conditions:

(1) $\delta = \varepsilon$ and

$$\begin{aligned} \tau(\bar{\gamma}_{k,l}) &= \dots g a b^{3i_1+1} a c d^{3i_1+2} c [\dots] e f^{3i_p} e a b^{3i_p+1} a h \dots, \\ & c, d, e, f, g, h \in \{a, b\}, c \neq d, e \neq f, \text{ and} \\ \tau(\bar{\gamma}_{k,r}) &= \dots g a b^{3i_{p+1}+1} a c d^{3i_{p+1}+2} c [\dots] e f^{3i_{p+q}} e a b^{3i_{p+q}+1} a h \dots, \\ & c, d, e, f, g, h \in \{a, b\}, c \neq d, e \neq f, \text{ or} \end{aligned}$$

(2) $\delta = i'$, $i' \in \mathbb{N}$, and

$$\begin{aligned} \tau(\bar{\gamma}_{k,l}) &= \dots o a b^{3i_1+1} a c d^{3i_1+2} c [\dots] a b^{3i_p+1} a e f^{3i_p+2} e g h^{3i'} g \dots, \\ & c, d, e, f, g, h, o \in \{a, b\}, c \neq d, e \neq f, g \neq h, \text{ and} \\ \tau(\bar{\gamma}_{k,r}) &= \dots g h^{3i'+2} g c d^{3i_{p+1}} c a b^{3i_{p+1}+1} a [\dots] e f^{3i_{p+q}} e a b^{3i_{p+q}+1} a o \dots, \\ & c, d, e, f, g, h, o \in \{a, b\}, c \neq d, e \neq f, g \neq h. \end{aligned}$$

For a better understanding of the definition of a splitting point, recall Claim 2 and Claim 3. Furthermore, these claims are sufficient for verifying the following facts:

- A $\bar{\gamma}_k$ with only one occurrence of a variable from X_1 does not contain any splitting point.
- For every splitting point δ of type 2, i. e. $\delta = i' \in \mathbb{N}$, necessarily $i' \in X_2$.
- For two splitting points δ, δ' , necessarily $\bar{\gamma}_k \neq \dots \delta \delta' \dots$.

After *all* of the splitting points have been identified in $\bar{\gamma}_k$, for each of them the following *splitting operation* is performed:

- (1) If $|\delta| = 0$ then δ is renamed to $\hat{\beta}$.
- (2) If $|\delta| = 1$ then a $\hat{\beta} = \varepsilon$ is inserted to the right of δ , i. e. $\bar{\gamma}_k := \bar{\gamma}_{k,l} \delta \hat{\beta} \bar{\gamma}_{k,r}$.

Note that, in case 2, we can arbitrarily choose to insert $\hat{\beta}$ to the left or to the right of δ , but it is essential to do this for all splitting points in the same way. This will be relevant for our argumentation on the crucial Claim 6.

When this has been accomplished for all splitting points then we regard the following decomposition of α :

$$\alpha = \hat{\beta}_0 \hat{\gamma}_1 \hat{\beta}_1 \hat{\gamma}_2 \hat{\beta}_2 [\dots] \hat{\beta}_{\hat{m}-1} \hat{\gamma}_{\hat{m}} \hat{\beta}_{\hat{m}}$$

with $\hat{m} \geq 1$ and

- $\hat{\beta}_k \in X_3^*$, $0 \leq k \leq \hat{m}$, where, for every $1 \leq k' \leq \hat{m} - 1$, $\hat{\beta}_{k'} = \varepsilon$ if and only if at exactly the position of $\hat{\beta}_{k'}$ a $\hat{\beta}$ has been inserted by a splitting operation, i. e. in this case $\hat{\beta}_{k'}$ simply is a renaming of a $\hat{\beta}$, and
- $\hat{\gamma}_k \in (X_1 \cup X_2)^+$, $1 \leq k \leq \hat{m}$.

Consequently, if in some $\bar{\gamma}_k$ there is, e. g., exactly one splitting point, i. e. $\bar{\gamma}_k = \bar{\gamma}_{k,l} \delta \bar{\gamma}_{k,r}$, then, for some $h < \hat{m}$, the splitting operation leads to $\hat{\gamma}_h = \bar{\gamma}_{k,l}$ and $\hat{\gamma}_{h+1} = \bar{\gamma}_{k,r}$ (in case of $|\delta| = 0$) or $\hat{\gamma}_h = \bar{\gamma}_{k,l} \delta$ and $\hat{\gamma}_{h+1} = \bar{\gamma}_{k,r}$ (in case of $|\delta| = 1$). Additionally note that $\hat{m} \geq 1$ again follows from $X_2 \neq \emptyset$.

After the splitting operations we can record:

Claim 5. For every k , $1 \leq k \leq \hat{m}$, $\hat{\gamma}_k$ contains exactly one occurrence of an $i \in X_1$.

Claim 5 follows from Claim 2, Claim 3 and the definition of the splitting points.

Moreover, the resulting decomposition has a second crucial property:

Claim 6. For every k, k' , $1 \leq k, k' \leq \hat{m}$, if $\text{var}(\hat{\gamma}_k) \cap \text{var}(\hat{\gamma}_{k'}) \cap X_1 \neq \emptyset$ then $\hat{\gamma}_k = \hat{\gamma}_{k'}$.

Proof (Claim 6). If $|\hat{\gamma}_k| = |\hat{\gamma}_{k'}| = 1$ then Claim 6 trivially holds true. Therefore we restrict ourselves to the case $|\hat{\gamma}_k| \geq 2$ or $|\hat{\gamma}_{k'}| \geq 2$. Now assume to the contrary that there are k, k' , $1 \leq k, k' \leq \hat{m}$, with $\text{var}(\hat{\gamma}_k) \cap \text{var}(\hat{\gamma}_{k'}) \cap X_1 \neq \emptyset$ and $\hat{\gamma}_k \neq \hat{\gamma}_{k'}$. Because of Claim 5 we can write $\hat{\gamma}_k$ as $\hat{\gamma}_k = i_1 \cdot i_2 \cdot [\dots] \cdot i_p \cdot i_x \cdot i_{p+1} \cdot i_{p+2} \cdot [\dots] \cdot i_{p+q}$ with $p, q \in \mathbb{N}_0$, $i_x \in X_1$, $i_1, i_2, \dots, i_{p+q} \in X_2$ and $\hat{\gamma}_{k'}$ as $\hat{\gamma}_{k'} = i'_1 \cdot i'_2 \cdot [\dots] \cdot i'_r \cdot i_x \cdot i'_{r+1} \cdot i'_{r+2} \cdot [\dots] \cdot i'_{r+s}$ with $r, s \in \mathbb{N}_0$, $i'_1, i'_2, \dots, i'_{r+s} \in X_2$. Note that our condition $|\hat{\gamma}_k| \geq 2$ or $|\hat{\gamma}_{k'}| \geq 2$ implies $p + q + r + s \geq 1$.

We now assume, first, that $p = r$ and $q = s$ (we shall examine the case where there is $p \neq r$ or $q \neq s$ later) and, second, w. l. o. g. that $t \in \mathbb{N}$ is the largest number with $i_t \neq i'_t$ and $t \leq p$. The latter assumption does not restrict our reasoning since, for the case that the only different variables in $\hat{\gamma}_k, \hat{\gamma}_{k'}$ are to the right of i_x , an analogous argumentation can be applied. Under these two assumptions, we now examine Claim 3, which says that, for every y , $0 \leq y \leq p - t$, $\tau(i_t \cdot i_{t+1} \cdot [\dots] \cdot i_{t+y})$ does not correspond to $\sigma_{\beta}^{\text{su}}(i_t \cdot i_{t+1} \cdot [\dots] \cdot i_{t+y})$ (and of course $\tau(i'_t \cdot i'_{t+1} \cdot [\dots] \cdot i'_{t+y})$ does not correspond to $\sigma_{\beta}^{\text{su}}(i'_t \cdot i'_{t+1} \cdot [\dots] \cdot i'_{t+y})$) as all of the variables under consideration are in X_2 . More precisely, we may conclude that, again for every y , $0 \leq y \leq p - t$, $\tau(i_{t+y+1} \cdot i_{t+y+2} \cdot [\dots] \cdot i_x)$ contains the right segment of $\sigma_{\beta}^{\text{su}}(i_{t+y})$ (and, additionally, $\tau(i'_{t+y+1} \cdot i'_{t+y+2} \cdot [\dots] \cdot i_x)$ contains the right segment of $\sigma_{\beta}^{\text{su}}(i'_{t+y})$) since, otherwise, there would have been a splitting point somewhere between i_t and i_x (and between i'_t and i_x)—this statement can be verified by a closer look at the definition of both types of splitting points, where the condition for the left subpattern $\bar{\gamma}_{k,l}$ in the case of $\tau(\bar{\gamma}_{k,l})$ containing the right segment of $\sigma_{\beta}^{\text{su}}(i_{t+y})$ would have led to the insertion of the said splitting point. Thus, with $y := 0$, this implies for some $\mathbf{c}, \mathbf{d}, \mathbf{e}, \mathbf{f}, \mathbf{o} \in \{\mathbf{a}, \mathbf{b}\}$,

22 *D. D. Freydenberger, D. Reidenbach, J. C. Schneider*

$c \neq d, e \neq f,$

$$\begin{aligned} \tau(i_{t+1} \cdot i_{t+2} \cdot [\dots] \cdot i_x) &= \dots c d^{3i_t+2} c e f^{3i_{t+1}} e [\dots] a b^{3i_x+1} a o \dots \\ &= \tau(i'_{t+1} \cdot i'_{t+2} \cdot [\dots] \cdot i_x). \end{aligned}$$

On the other hand we know that, for the same e, f, o as above and some $g, h \in \{a, b\}, g \neq h,$

$$\sigma_\beta^{\text{su}}(i'_t \cdot i'_{t+1} \cdot i'_{t+2} \cdot [\dots] \cdot i_x) = \dots g h^{3i'_t+2} g e f^{3i_{t+1}} e [\dots] a b^{3i_x+1} a o \dots$$

However—since $i_x \in X_1$ and, hence, $\tau(i_x)$ contains the inner segment of $\sigma_\beta^{\text{su}}(i_x)$ —we know that $\tau(i_x)$ and $\sigma_\beta^{\text{su}}(i_x)$ generate the *same occurrence* of the subword $a b^{3i_x+1} a$ in $\tau(\alpha) = \sigma_\beta^{\text{su}}(\alpha)$. Thus, roughly speaking, if we compare $\tau(i'_t \cdot i'_{t+1} \cdot [\dots] \cdot i_x)$ with $\sigma_\beta^{\text{su}}(i'_t \cdot i'_{t+1} \cdot [\dots] \cdot i_x)$ then we deal with the same part of $\sigma_\beta^{\text{su}}(\alpha)$ (cf. our remarks introducing Lemma 27). Consequently, we can conclude from Lemma 27 that the conditions $\tau(\alpha) = \sigma_\beta^{\text{su}}(\alpha)$ and $i_x \in X_1$ imply $c d^{3i_t+2} c = g h^{3i'_t+2} g$ and, more precisely, $c = g, d = h$ and $i_t = i'_t$. This contradicts our assumption $i_t \neq i'_t$.

We proceed with the remaining case $p \neq r$ or $q \neq s$. Due to the same reason as given above and, thus, w. l. o. g., we focus on $p \neq r$. Additionally and again w. l. o. g., we assume $p < r$. If there is a $t \in \mathbb{N}_0, t < p$, such that $i_{p-t} \neq i'_{r-t}$ then we can apply exactly the same argument as above. Thus, $i_1 \cdot i_2 \cdot [\dots] \cdot i_p$ must be a suffix of $i'_1 \cdot i'_2 \cdot [\dots] \cdot i'_r$. If $\alpha = \hat{\gamma}_k \dots$ or $\alpha = \dots i_\# \hat{\gamma}_k \dots$ with $i_\# \neq i'_{r-p}$ then our argumentation again is equivalent to that on the case $p = r$. Hence, $i_\# = i'_{r-p}$. Since $i'_{r-p} \in X_2, \hat{\beta}_{k-1}$ must have been a splitting point separating i'_{r-p} and i_1 , whereas there has not been any splitting point between i'_{r-p} and i'_{r-p+1} . Since $i_1 \cdot i_2 \cdot [\dots] \cdot i_p$ is a suffix of $i'_1 \cdot i'_2 \cdot [\dots] \cdot i'_r$ this contradicts the definition of splitting points. \square (*Claim 6*)

In a final step, we now remove all $\hat{\gamma}_k$ with $|\hat{\gamma}_k| = 1$; this type of $\hat{\gamma}$ can occur, e. g., for $\hat{\gamma}_{k'} = i_1 \cdot i_2 \cdot i_3$ with $i_1, i_3 \in X_1$ and $i_2 \in X_2$. Consequently, for every $\hat{\gamma}_k = i, i \in X_1$, we shift i to X_3 , or, more precisely, we introduce $X'_3 := X_3 \cup \{i \mid \exists k : \hat{\gamma}_k = i\}$ and $X'_1 := X_1 \setminus \{i \mid \exists k : \hat{\gamma}_k = i\}$. Note that because of Claim 5 and Claim 6 this *redefinition* operation does not affect any $\hat{\gamma}_k$ with $|\hat{\gamma}_k| \geq 2$.

This leads to the final decomposition of α :

$$\alpha = \beta_0 \gamma_1 \beta_1 \gamma_2 \beta_2 [\dots] \beta_{m-1} \gamma_m \beta_m$$

with $m \geq 1$ and

- $\beta_k \in X'_3, 0 \leq k \leq m$, where, for every $1 \leq k' \leq m-1, \beta_{k'} = \varepsilon$ if and only if the variables to the right and to the left of $\beta_{k'}$ have been split by a splitting operation and none of the resulting neighbouring $\hat{\gamma}_{k''}$ has been removed by a shifting operation, and
- $\gamma_k \in (X'_1 \cup X_2)^+, 1 \leq k \leq m$.

Again, this decomposition is unique, and $m \geq 1$ is granted since X_2 is not redefined and (according to our assumption) $X_2 \neq \emptyset$.

We conclude the proof of Lemma 28 with the verification of the conditions in Definition 1:

Claim 7. For every k , $1 \leq k \leq m$, $|\gamma_k| \geq 2$.

Claim 7 is evident since the redefinition operation does not shorten or split any $\hat{\gamma}_k$ with $|\hat{\gamma}_k| \geq 2$. Consequently, the above decomposition conforms with the first condition of Definition 1. The next claim follows directly from the fact that X'_1 , X_2 and X'_3 are disjoint:

Claim 8. For every k , $1 \leq k \leq m$, and for every k' , $0 \leq k' \leq m$, $\text{var}(\gamma_k) \cap \text{var}(\beta_{k'}) = \emptyset$.

Thus, the second condition of Definition 1 is satisfied as well. Since, according to the notes on Claim 7, the splitting operation does not modify any $\hat{\gamma}_k$ with $|\hat{\gamma}_k| \geq 2$ we can easily conclude from Claim 6:

Claim 9. For every k , $1 \leq k \leq m$, γ_k contains exactly one $i \in X'_1$ and, for every k' , $1 \leq k' \leq m$, if $\text{var}(\gamma_k) \cap \text{var}(\gamma_{k'}) \cap X'_1 \neq \emptyset$ then $\gamma_k = \gamma_{k'}$.

This proves that the third condition of Definition 1 is satisfied.

Consequently, if there is an $i \in \text{var}(\alpha)$ such that $\tau(i) \neq \dots \mathbf{g a b}^{3i+1} \mathbf{a h} \dots$ then α is prolix. This proves the *only if*-direction and, hence, the lemma. \square

We are uncertain to which extent $\sigma_\alpha^{\text{su}}$ can be shortened without allowing alternative morphisms τ as given in Example 23, but we expect that strong unambiguity can also be ensured by a morphism which, for every variable $i \in \mathbb{N}$, assigns only the left and the right segment of $\sigma_\alpha^{\text{su}}(i)$ to i .

Referring to Lemma 28 we now conclude the proof of our main result with the detailed examination of the use of morphic heterogeneity of $\sigma_\alpha^{\text{su}}$:

Lemma 29. *Let $\alpha \in \mathbb{N}^+$. If, for every morphism $\tau : \mathbb{N}^+ \rightarrow \{\mathbf{a}, \mathbf{b}\}^*$ with $\tau(\alpha) = \sigma_\alpha^{\text{su}}(\alpha)$ and for every $i \in \text{var}(\alpha)$, $\tau(i) = \dots \mathbf{a b}^{3i+1} \mathbf{a} \dots$ then $\sigma_\alpha^{\text{su}}(\alpha)$ is strongly unambiguous.*

Proof. If $|\text{var}(\alpha)| = 1$ then every morphic image of α is strongly unambiguous, and therefore, in this case, Lemma 29 holds trivially. Hence, let $|\text{var}(\alpha)| \geq 2$. Additionally note that, because of Lemma 28 and the condition of Lemma 29, the pattern α must be succinct.

Now assume to the contrary that there is a morphism τ with $\tau(\alpha) = \sigma_\alpha^{\text{su}}(\alpha)$ such that, for every $i \in \text{var}(\alpha)$, $\tau(i) = \dots \mathbf{a b}^{3i+1} \mathbf{a} \dots$ and, for some $i' \in \text{var}(\alpha)$, $\tau(i') \neq \sigma_\alpha^{\text{su}}(i')$. W.l.o.g., we may assume that $|\tau(i')| \neq |\sigma_\alpha^{\text{su}}(i')|$ since the assumption that, for all $i \in \text{var}(\alpha)$, $|\tau(i)| = |\sigma_\alpha^{\text{su}}(i)|$ would imply, again for all $i \in \text{var}(\alpha)$, $\tau(i) = \sigma_\alpha^{\text{su}}(i)$. Consequently, because of $|\tau(\alpha)| = |\sigma_\alpha^{\text{su}}(\alpha)|$, there is a $j \in \text{var}(\alpha)$ such that $|\tau(j)| > |\sigma_\alpha^{\text{su}}(j)|$ and, thus, due to the condition $\tau(j) = \dots \mathbf{a b}^{3j+1} \mathbf{a} \dots$, for some $\mathbf{c}, \mathbf{d}, \mathbf{e} \in \{\mathbf{a}, \mathbf{b}\}$, $\mathbf{c} \neq \mathbf{d}$,

- (a) $\tau(j) = \dots \mathbf{e c d}^{3j} \mathbf{c a b}^{3j+1} \mathbf{a} \dots$ or
- (b) $\tau(j) = \dots \mathbf{a b}^{3j+1} \mathbf{a c d}^{3j+2} \mathbf{c e} \dots$

We restrict the following reasoning to case (a) as an analogous argumentation can be applied to case (b) (using R_j and \sim_r instead of L_j and \sim_l): Note that, due to $|\text{var}(\alpha)| \geq 2$ and the succinctness of α , $|\alpha|_j \geq 2$, and therefore the set L_j (given in Definition 21) of left neighbours of j in α is not empty. With regard to L_j we can conclude immediately that, for no $k \in L_j$, $\tau(k)$ completely contains the right segment of $\sigma_\alpha^{\text{su}}(k)$:

Claim 1. For every $k \in L_j$ and any $e, f \in \{a, b\}$, $e \neq f$, $\tau(k) \neq \dots e f^{3k+2} e \dots$.

Proof (Claim 1). As a result of the conditions $\tau(i) = \dots a b^{3i+1} a \dots$, $i \in \text{var}(\alpha)$, and $\tau(\alpha) = \sigma_\alpha^{\text{su}}(\alpha)$ we can apply Lemma 27 which says that, for some $c, d, e, f \in \{a, b\}$, $c \neq d$, $e \neq f$,

$$\tau(k \cdot j) = \dots a b^{3k+1} a e f^{3k+2} \boxed{e} c d^{3j} c a b^{3j+1} a \dots$$

Since $\tau(j) = \dots \boxed{e} c d^{3j} c a b^{3j+1} a \dots$ (cf. case (a)) and $\tau(k) = \dots a b^{3k+1} a \dots$ it follows immediately from Lemma 24 that $\tau(k) \neq \dots e f^{3k+2} e \dots$. \square (Claim 1)

Concerning j and any $k \in L_j$ we now examine the following cases:

Case 1: $\alpha = j \dots$

Then we can directly conclude from Lemma 24: $\sigma_\alpha^{\text{su}}(\alpha) = e f^{3j} e \dots \neq \tau(\alpha) = \dots g c d^{3j} c \dots$, with $c, d, e, f, g \in \{a, b\}$, $c \neq d$, $e \neq f$. This contradicts the condition $\sigma_\alpha^{\text{su}}(\alpha) = \tau(\alpha)$.

Case 2: $\alpha = \dots k$.

Since, from Claim 1 and for $c, d \in \{a, b\}$, $c \neq d$, $\tau(k) \neq \dots c d^{3k+2} c$ we may conclude using Lemma 24 that $\tau(\alpha) \neq \dots c d^{3k+2} c = \sigma_\alpha^{\text{su}}(\alpha)$. This again contradicts the condition $\sigma_\alpha^{\text{su}}(\alpha) = \tau(\alpha)$.

Case 3: $\alpha \neq j \dots$ and $\alpha \neq \dots k$.

For the equivalence classes $L_1^\sim, L_2^\sim, \dots, L_p^\sim$ on $\text{var}(\alpha)$ —which are explained in Definition 20—let $\iota \in \{1, 2, \dots, p\}$ with $L_j \subseteq L_\iota^\sim$. Then, since all $L_1^\sim, L_2^\sim, \dots, L_p^\sim$ are pairwise disjoint, this ι is unique. Now we can collect a number of facts that facilitate the argumentation in Case 3. The first holds as α is succinct:

Claim 2. If $\alpha \neq j \dots$ and $\alpha \neq \dots k$ then $|L_\iota^\sim| \geq 2$.

Proof (Claim 2). If $|L_j| \geq 2$ then Claim 2 holds by definition. Hence, let $L_j := \{k\}$. Then, for every occurrence of j in α , the conditions $\alpha \neq j \dots$ and $|\alpha|_j \geq 2$ lead to $\alpha = \dots k \cdot j \dots$. Thus, due to the succinctness of α , there are some $j_1, j_2, \dots, j_m \in \text{var}(\alpha)$, $m \geq 2$, with $\alpha = \dots k \cdot j_r \dots$, $1 \leq r \leq m$, since otherwise we could give a decomposition $\alpha = \beta_0 \gamma_1 \beta_1 \gamma_2 \beta_2 [\dots] \beta_{|\alpha|_k-1} \gamma_{|\alpha|_k} \beta_{|\alpha|_k}$ satisfying the conditions of Definition 1 with, for every h , $1 \leq h \leq |\alpha|_k$, $\gamma_h = k \cdot j$. Additionally, because of $\alpha \neq \dots k$, there must be an $s \in \{1, 2, \dots, m\}$ and a $\bar{k} \in \text{var}(\alpha)$, $\bar{k} \neq k$, with $\alpha = \dots \bar{k} \cdot j_s \dots$, since, otherwise, α would either be prolix—with a decomposition regarding each subpattern $k \cdot j_r$ as a γ_h —or start with a j_r , $r \in \{1, 2, \dots, m\}$, leading to the same argumentation as in Case 1. Consequently, $L_{j_s} \supseteq \{k, \bar{k}\}$ and therefore $L_j \subset \{k, \bar{k}\} \subseteq L_{j_s} \subseteq L_\iota^\sim$. \square (Claim 2)

Now, for every L^\sim among $L_1^\sim, L_2^\sim, \dots, L_p^\sim$, the next fact follows by definition since these equivalence classes are composed by union of non-disjoint sets (cf. Definition 21 and, e.g., Example 22):

Claim 3. If $|L^\sim| \geq 2$ then, for every $\hat{k} \in L^\sim$, there is an $L_j \subseteq L^\sim$ with $|L_j| \geq 2$ and $\hat{k} \in L_j$.

We conclude the list of preliminary claims with the following one, that deals with a crucial phenomenon which is reflected in the transitivity of \sim_l , namely the insight that, for no $\hat{k} \in L^\sim$, $\tau(\hat{k})$ completely contains the right segment of $\sigma_\alpha^{\text{su}}(\hat{k})$. Thus, the subsequent claim extends Claim 1:

Claim 4. For every $\hat{k} \in L^\sim$ and any $\mathbf{e}, \mathbf{f} \in \{\mathbf{a}, \mathbf{b}\}$, $\mathbf{e} \neq \mathbf{f}$, $\tau(\hat{k}) \neq \dots \mathbf{e} \mathbf{f}^{3\hat{k}+2} \mathbf{e} \dots$.

Proof (Claim 4). We stepwise expand the argumentation on Claim 1 to all variables in L^\sim : With regard to any $\hat{k}' \in L_j \subseteq L^\sim$, Claim 4 directly holds according to Claim 1. We now regard all $\hat{k}'' \in L^\sim$ for which there is an $L_{j'}$ with $\hat{k}', \hat{k}'' \in L_{j'}$ (recall that $\hat{k}' \in L_j$). Then—since Claim 4 is satisfied for \hat{k}' , i.e. $\tau(\hat{k}') \neq \dots \mathbf{e} \mathbf{f}^{3\hat{k}'+2} \mathbf{e} \dots$, and on the other hand, for some $\mathbf{c}, \mathbf{d}, \mathbf{e}, \mathbf{f} \in \{\mathbf{a}, \mathbf{b}\}$, $\mathbf{c} \neq \mathbf{d}$, $\mathbf{e} \neq \mathbf{f}$,

$$\tau(\hat{k}' \cdot j') = \dots \mathbf{a} \mathbf{b}^{3\hat{k}'+1} \mathbf{a} \mathbf{e} \mathbf{f}^{3\hat{k}'+2} \boxed{\mathbf{e}} \mathbf{c} \mathbf{d}^{3j'} \mathbf{c} \mathbf{a} \mathbf{b}^{3j'+1} \mathbf{a} \dots$$

while we have the condition $\tau(i) = \dots \mathbf{a} \mathbf{b}^{3i+1} \mathbf{a} \dots$, $i \in \text{var}(\alpha)$ —Lemma 24 allows the conclusion that $\tau(j') = \dots \boxed{\mathbf{e}} \mathbf{c} \mathbf{d}^{3j'} \mathbf{c} \mathbf{a} \mathbf{b}^{3j'+1} \mathbf{a} \dots$. Thus, with regard to the variables $\hat{k}'' \in L_{j'}$ and all subpatterns $\hat{k}'' \cdot j'$ of α , our argumentation is verbatim the same as the reasoning on the variables $k \in L_j$ and the subpatterns $k \cdot j$ given in Claim 1. Therefore Claim 4 holds for each $\hat{k}'' \in L_{j'}$ as well. Now we proceed to all $\hat{k}''' \in L^\sim$ for which there is an $L_{j''}$ with $\hat{k}'', \hat{k}''' \in L_{j''}$ (recall that $\hat{k}'' \in L_{j'}$). Then, as Claim 4 is satisfied for \hat{k}'' , Claim 4 holds for all \hat{k}''' and so on. Consequently, according to the construction of L^\sim (cf. Definition 21, which indirectly composes the equivalence classes $L_1^\sim, L_2^\sim, \dots, L_p^\sim$ by a union operation on non-disjoint sets) Claim 4 holds for every $\hat{k} \in L^\sim$. □ (Claim 4)

We now can conclude our argumentation on Case 3: According to Claim 2, $|L^\sim| \geq 2$. Let $k_\# := \min L^\sim$; then, due to Claim 3, there is a $j_\# \in \text{var}(\alpha)$ with $k_\# \in L_{j_\#}$ and $|L_{j_\#}| \geq 2$. Consequently, let $\bar{k}_\# \in \text{var}(\alpha)$ with $k_\# \neq \bar{k}_\#$ and $\{k_\#, \bar{k}_\#\} \subseteq L_{j_\#}$. Then, because of $k_\# = \min L^\sim$ and $\bar{k}_\# \in L^\sim$, $\sigma_\alpha^{\text{su}}(k_\#) = \dots \mathbf{b}$ and $\sigma_\alpha^{\text{su}}(\bar{k}_\#) = \dots \mathbf{a}$ (cf. Definition 21). Referring to the condition $\tau(i) = \dots \mathbf{a} \mathbf{b}^{3i+1} \mathbf{a} \dots$, $i \in \text{var}(\alpha)$, to Lemma 27 and to Claim 4, we can see that $\tau(j_\#)$ contains the left segment of $\sigma_\alpha^{\text{su}}(j_\#)$ and additionally, immediately to the left of this segment, the last letter of $\sigma_\alpha^{\text{su}}(k_\#)$ and $\sigma_\alpha^{\text{su}}(\bar{k}_\#)$. Therefore, the different rightmost letters of $\sigma_\alpha^{\text{su}}(k_\#)$ and $\sigma_\alpha^{\text{su}}(\bar{k}_\#)$ imply

$$\dots \boxed{\mathbf{b}} \mathbf{c} \mathbf{d}^{3j_\#} \mathbf{c} \mathbf{a} \mathbf{b}^{3j_\#+1} \mathbf{a} \dots = \tau(j_\#) = \dots \boxed{\mathbf{a}} \mathbf{c} \mathbf{d}^{3j_\#} \mathbf{c} \mathbf{a} \mathbf{b}^{3j_\#+1} \mathbf{a} \dots,$$

for some $\mathbf{c}, \mathbf{d} \in \{\mathbf{a}, \mathbf{b}\}$, $\mathbf{c} \neq \mathbf{d}$. This contradicts $\mathbf{a} \neq \mathbf{b}$. □

By Lemma 28 and Lemma 29, Theorem 16 follows immediately.

4.2. Additional Remarks on the Proof of Theorem 16

Although the proof of our main result is based on a procedure which gives a *concrete* morphism $\sigma_\alpha^{\text{su}}$ leading to the desired unambiguous word, it nevertheless shows some unsatisfactory traits. For instance, with regard to the pattern $\alpha := 1 \cdot 2 \cdot 3 \cdot 1 \cdot 2 \cdot 4 \cdot 3 \cdot 1 \cdot 2 \cdot 3 \cdot 2 \cdot 4$ introduced in Example 22, the much simpler injective morphism $\sigma'(i) = \mathbf{a} \mathbf{b}^i$, $i \in \mathbb{N}$, generates an unambiguous word as well, and even the non-injective morphism σ'' given by $\sigma''(2) := \mathbf{b}$, $\sigma''(1) = \sigma''(3) = \sigma''(4) := \mathbf{a}$ has this property. Additionally, for every pattern α satisfying, for some $n \geq 1$ and $p_1, p_2, \dots, p_n \geq 2$, $\alpha = 1^{p_1} \cdot 2^{p_2} \cdot [\dots] \cdot n^{p_n}$, $\sigma'(\alpha)$ is known to be strongly unambiguous (cf. Reidenbach [16]). Thus, $\sigma_\alpha^{\text{su}}$ is “sufficient” for generating an unambiguous word, but, in general, it is not “necessary” since there can be significantly shorter words with the desired property. Contrary to this, for our initial example $\alpha' = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 1 \cdot 4 \cdot 3 \cdot 2$, it is explained in Section 1 that a morphism is needed which utilises heterogeneity and, thus, which is more elaborate than σ' . Consequently, we expect that the given proof can be improved considerably, though any major progress might require a lot of effort—this holds, e. g., for the potential combinatorial interplay between the two decisive properties of $\sigma_\alpha^{\text{su}}$, namely restricted ambiguity caused by the assignment of some distinct segments to each variable and heterogeneity of the equivalence classes L_i^\sim and R_i^\sim , which we merely discuss separately in Lemma 28 and Lemma 29.

In fact, we do not know any pattern α for which there is demonstrably no shorter unambiguous image than $\sigma_\alpha^{\text{su}}(\alpha)$. Experience suggests, however, that corresponding example patterns might be extremely long and intricate so that we, with our current state of knowledge, do not dare to give any conjecture on the (non-)existence of such patterns.

5. Conclusion and Further Research Directions

In the present paper we have systematically examined the fundamental question for which strings $\alpha \in \mathbb{N}^+$ there exists a non-empty string w in $\{\mathbf{a}, \mathbf{b}\}^*$ such that w is unambiguous for α , i. e. there is exactly one morphism σ with $\sigma(\alpha) = w$. Our first major result has shown that there is no single nonerasing morphism σ such that, for every $\alpha \in \mathbb{N}^+$, $\sigma(\alpha)$ is unambiguous. A partition of \mathbb{N}^+ according to a sophisticated criterion—which has previously been discussed within the scope of pattern languages and fixed points of morphisms—has led to a class of strings (called “prolix”) for which every image under an arbitrary morphism is not unambiguous. Finally our main result has demonstrated that, for every string $\alpha \in \mathbb{N}^+$ which is not prolix (referred to as “succinct”), there is a tailor-made injective morphism $\sigma_\alpha^{\text{su}}$ such that $\sigma_\alpha^{\text{su}}(\alpha)$ is unambiguous. This statement can be easily rephrased in terms of pattern languages, equality sets, and finite fixed points of morphisms. All of our results hold for morphic images in arbitrary finitely generated free monoids instead of $\{\mathbf{a}, \mathbf{b}\}^*$ as well.

Although we have made significant first steps which partly require a cumbersome reasoning, a lot of questions on the unambiguity of morphic images are still open. We

conclude this paper with a short list of those problems (in parts already suggested in previous sections) we consider particularly worth dealing with: Can we shorten $\sigma_\alpha^{\text{su}}$ for every $\alpha \in \mathbb{N}^+$? If so, which is the minimum length required? For which strings $\alpha \in \mathbb{N}^+$ is there an unambiguous morphic image $w \in \{\mathbf{a}, \mathbf{b}\}^*$ with $|w| \leq |\alpha|$? More generally, what happens if we do not restrict ourselves to the ambiguity of images under injective morphisms σ ? For which major classes $\Pi \subset \mathbb{N}^+$ is there a single respective morphism σ_Π such that, for every $\alpha \in \Pi$, $\sigma_\Pi(\alpha)$ is unambiguous? How do the insights presented in our paper change as soon as we do not consider strings over \mathbb{N} , but only over some $\Delta \subset \mathbb{N}$? To which extent does a larger target alphabet $\Sigma \supset \{\mathbf{a}, \mathbf{b}\}$ allow a different reasoning and more compact unambiguous morphic images? For which prolix strings in \mathbb{N}^+ is there an unambiguous morphic image? Which additional phenomena can be observed within a restricted view solely dealing with morphic images in $\{\mathbf{a}, \mathbf{b}\}^+$?

Acknowledgements

The authors wish to thank the anonymous referees for their numerous, careful and helpful remarks and suggestions.

Daniel Reidenbach has been partly supported by the Deutsche Forschungsgemeinschaft (DFG) under Grant Wi 1638/1-3.

References

- [1] D. Angluin. Finding patterns common to a set of strings. *Journal of Computer and System Sciences*, 21:46–62, 1980.
- [2] J. Berstel and D. Perrin. *Theory of Codes*. Academic Press, Orlando, 1985.
- [3] 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.
- [4] E.M. Gold. Language identification in the limit. *Information and Control*, 10:447–474, 1967.
- [5] D. Hamm and J. Shallit. Characterization of finite and one-sided infinite fixed points of morphisms on free monoids. Technical Report CS-99-17, Dep. of Computer Science, University of Waterloo, 1999. <http://www.cs.uwaterloo.ca/~shallit/papers.html>.
- [6] T. Harju and J. Karhumäki. Morphisms. In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Languages*, volume 1, chapter 7, pages 439–510. Springer, 1997.
- [7] T. Head. Fixed languages and the adult languages of 0L schemes. *International Journal of Computer Mathematics*, 10:103–107, 1981.
- [8] S. Lange and R. Wiehagen. Polynomial-time inference of arbitrary pattern languages. *New Generation Computing*, 8:361–370, 1991.
- [9] F. Levé and G. Richomme. On a conjecture about finite fixed points of morphisms. *Theoretical Computer Science*, 339:103–128, 2005.
- [10] M. Lipponen and G. Păun. Strongly prime PCP words. *Discrete Applied Mathematics*, 63:193–197, 1995.
- [11] M. Lothaire. *Combinatorics on Words*. Addison-Wesley, Reading, MA, 1983.
- [12] A. Mateescu and A. Salomaa. Finite degrees of ambiguity in pattern languages. *RAIRO Informatique théorique et Applications*, 28:233–253, 1994.

- [13] 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.
- [14] E.L. Post. A variant of a recursively unsolvable problem. *Bulletin of the American Mathematical Society*, 52:264–268, 1946.
- [15] D. Reidenbach. A discontinuity in pattern inference. In *Proc. 21st Annual Symposium on Theoretical Aspects of Computer Science, STACS 2004*, volume 2996 of *Lecture Notes in Computer Science*, pages 129–140, 2004.
- [16] D. Reidenbach. A non-learnable class of E-pattern languages. *Theoretical Computer Science*, 350:91–102, 2006.