



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



CC creative commons
COMMONS DEED

Attribution-NonCommercial-NoDerivs 2.5

You are free:

- to copy, distribute, display, and perform the work

Under the following conditions:

BY: **Attribution.** You must attribute the work in the manner specified by the author or licensor.

Noncommercial. You may not use this work for commercial purposes.

No Derivative Works. You may not alter, transform, or build upon this work.

- For any reuse or distribution, you must make clear to others the license terms of this work.
- Any of these conditions can be waived if you get permission from the copyright holder.

Your fair use and other rights are in no way affected by the above.

This is a human-readable summary of the [Legal Code \(the full license\)](#).

[Disclaimer](#) 

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

The Unambiguity of Segmented Morphisms[☆]

Dominik D. Freydenberger^{*,a,1}, Daniel Reidenbach^b

^a*Institut für Informatik, Johann Wolfgang Goethe-Universität,
Postfach 111932, 60054 Frankfurt am Main, Germany*

^b*Department of Computer Science, Loughborough University, Loughborough, Leicestershire,
LE11 3TU, United Kingdom*

Abstract

This paper studies the ambiguity of morphisms in free monoids. A morphism σ is said to be ambiguous with respect to a string α if there exists a morphism τ which differs from σ for a symbol occurring in α , but nevertheless satisfies $\tau(\alpha) = \sigma(\alpha)$; if there is no such τ then σ is called unambiguous. Motivated by the recent initial paper on the ambiguity of morphisms, we introduce the definition of a so-called segmented morphism σ_n , which, for any $n \in \mathbb{N}$, maps every symbol in an infinite alphabet onto a word that consists of n distinct factors in $\mathbf{a}\mathbf{b}^+\mathbf{a}$, where \mathbf{a} and \mathbf{b} are different letters. For every n , we consider the set $U(\sigma_n)$ of those finite strings over an infinite alphabet with respect to which σ_n is unambiguous, and we comprehensively describe its relation to any $U(\sigma_m)$, $m \neq n$.

Thus, our work features the first approach to a characterisation of sets of strings with respect to which certain fixed morphisms are unambiguous, and it leads to fairly counter-intuitive insights into the relations between such sets. Furthermore, it shows that, among the widely used homogeneous morphisms, most segmented morphisms are optimal in terms of being unambiguous for a preferably large set of strings. Finally, our paper yields several major improvements of crucial techniques previously used for research on the ambiguity of morphisms.

Key words: Combinatorics on words, Morphisms, Ambiguity, Pattern languages

[☆]A preliminary version of this paper was presented at the conference DLT 2007.

*Corresponding author.

Email addresses: freydenberger@em.uni-frankfurt.de (Dominik D. Freydenberger),
D.Reidenbach@lboro.ac.uk (Daniel Reidenbach)

¹A part of this work was done during the author's stays at the Technische Universität Kaiserslautern and the Research Group on Mathematical Linguistics, URV, Tarragona, Spain.

1. Introduction

In the present paper, we examine a fundamental combinatorial property of morphisms in free monoids: their (potential) *ambiguity*. We say that a morphism σ is *ambiguous* with respect to a string α if there exists a morphism τ such that $\tau(\alpha) = \sigma(\alpha)$, but $\tau(x) \neq \sigma(x)$ for at least one symbol x occurring in α ; if there is no such morphism τ then we call σ *unambiguous* with respect to α . For instance, the morphism $\sigma : \{A, B, C\}^* \rightarrow \{a, b\}^*$ given by $\sigma(A) := a$, $\sigma(B) := b$ and $\sigma(C) := ab$ is ambiguous with respect to the string $\alpha := ABACBC$, since the morphism τ given by $\tau(A) := a$, $\tau(B) := ba$ and $\tau(C) := b$ satisfies

$$\tau(ABACBC) = abababab = \sigma(ABACBC).$$

In contrast, it can be easily verified that the morphism σ' given by $\sigma'(A) := a$, $\sigma'(B) := b$ and $\sigma'(C) := ba$ is unambiguous with respect to α .

While our paper, thus, deals with an elementary phenomenon related to a crucial concept in noncommutative discrete mathematics, the original motivation for research on the ambiguity of morphisms is mainly derived from an algorithmic topic in computer science, namely *pattern inference*, i. e., a notion of the problem of computing a pattern that is common to a given set of words (cf. Ng and Shinohara [15]). In this context, as demonstrated by Reidenbach [17, 18], it is particularly valuable to investigate the existence of *unambiguous* morphisms for arbitrary strings. In addition to this, our subject shows direct connections to many other fields that are based on finite strings and morphisms. This notably holds for those concepts that consider several morphic images of a single string, such as *pattern languages* (cf. Mateescu and Salomaa [13]) and *equality sets* (cf. Harju and Karhumäki [8]), a topic that comprises the famous undecidable *Post Correspondence Problem* (cf. Post [16]) and related aspects (see, e. g., Lipponen and Păun [12]).

In consideration of previous literature and the needs of pattern inference, we study the ambiguity of morphisms that map a string over an infinite alphabet Δ onto a string over a binary alphabet Σ ; for the sake of convenience we choose $\Delta := \mathbb{N}$ and $\Sigma := \{a, b\}$. For the remainder of this paper, we call any “symbol” in \mathbb{N} a *variable*, any symbol in $\{a, b\}$ a *letter*, any string in \mathbb{N}^* a *pattern* and any string in $\{a, b\}^*$ a *word*. Furthermore, we separate the variables in a pattern by a dot “.”, so as to guarantee that, e. g., the pattern $\alpha := 1 \cdot 2 \cdot 1$ is not confused with the pattern $\alpha' := 1 \cdot 21$.

As mentioned above, a vital question arising from the research on pattern inference is that of the existence of unambiguous morphisms. This problem is resolved by the main result in the initial systematic paper on the ambiguity of morphisms by Freydenberger, Reidenbach and Schneider [6]. It says that, for any pattern α , there exists a morphism $\sigma_\alpha^{\text{su}}$ that is unambiguous with respect to α if and only if α is *succinct*, i. e., there is a certain complex factorisation for α , which is also crucial within the scopes of pattern languages and *fixed points* of endomorphisms (cf. Levé and Richomme [11]). From a technical point of view, the morphism $\sigma_\alpha^{\text{su}}$ has two main properties:

1. Since there is no single morphism that is unambiguous with respect to all succinct patterns, $\sigma_\alpha^{\text{su}}$ has to be tailor-made for α . More precisely, $\sigma_\alpha^{\text{su}}$ is *heterogeneous* with respect to α , which means that there exist certain variables x, y in α such that the first (or, if appropriate, the last) letter of $\sigma_\alpha^{\text{su}}(x)$ differs from the first (or last, respectively) letter of $\sigma_\alpha^{\text{su}}(y)$; the choice of x and y depends on the structure of α .
2. The morphism $\sigma_\alpha^{\text{su}}$ maps each variable in α onto a word that consists of exactly *three* distinct *segments*, i. e., factors taken from $\mathbf{ab}^+\mathbf{a}$ (or, in order to guarantee heterogeneity, $\mathbf{ba}^+\mathbf{b}$).

A closer look at the approach by Freydenberger et al. [6] – which is mainly meant to prove the *existence* of an unambiguous morphism with respect to any succinct pattern – reveals that it is not optimal, as there exist numerous patterns with respect to which there is a significantly less complex unambiguous morphism. For instance, as demonstrated by Reidenbach [17], the standard injective morphism σ_0 given by $\sigma_0(x) := \mathbf{ab}^x$, $x \in \mathbb{N}$, is unambiguous with respect to *every* pattern α satisfying, for some $m \in \mathbb{N}$ and $e_1, e_2, \dots, e_m \geq 2$, $\alpha = 1^{e_1} \cdot 2^{e_2} \cdot \dots \cdot m^{e_m}$ (where the superscripts e_j refer to the concatenation). With regard to this result, it is noteworthy that, first, σ_0 maps each variable onto a much shorter word than $\sigma_\alpha^{\text{su}}$ and, second, σ_0 is *homogeneous*, i. e., for all variables $x, y \in \mathbb{N}$, $\sigma_0(x)$ and $\sigma_0(y)$ have the same first and the same last letter. Consequently, σ_0 is unambiguous with respect to each pattern in a reasonably rich set, although it does not show any of the two decisive properties of $\sigma_\alpha^{\text{su}}$.

In contrast to this, no homogeneous morphism σ is unambiguous on the pattern $\alpha := 1 \cdot 2 \cdot 1 \cdot 3 \cdot 3 \cdot 2$. If σ is homogeneous, there exist words $w_x \in \{\mathbf{a}, \mathbf{b}\}^*$ and a letter $a \in \{\mathbf{a}, \mathbf{b}\}$ such that $\sigma(x) = w_x a$ for every symbol x occurring in α . Then the morphism τ defined by $\tau(1) := w_1$, $\tau(2) := a \sigma(2)$ and $\tau(3) := a w_3$ proves the ambiguity of σ on α . On the other hand, the morphism σ' defined by $\sigma'(1) := \mathbf{a}$, $\sigma'(2) := \mathbf{a}$ and $\sigma'(3) := \mathbf{b}$ is unambiguous on α . Evidently, σ' is heterogeneous.

In the present paper, we wish to further develop the theory of unambiguous morphisms. In accordance with the structure of $\sigma_\alpha^{\text{su}}$, we focus on *segmented* morphisms σ_n , which map every variable onto n distinct segments. More precisely, for every $n \in \mathbb{N}$, we define the homogeneous morphism σ_n by

$$\sigma_n(x) := \mathbf{ab}^{nx-(n-1)}\mathbf{a} \mathbf{ab}^{nx-(n-2)}\mathbf{a} \dots \mathbf{ab}^{nx-1}\mathbf{a} \mathbf{ab}^{nx}\mathbf{a}$$

for every $x \in \mathbb{N}$. We illustrate this definition by three examples, namely the morphisms σ_3 , σ_4 and σ_5 , which are given by

$$\begin{aligned} \sigma_3(x) &= \mathbf{ab}^{3x-2}\mathbf{a} \mathbf{ab}^{3x-1}\mathbf{a} \mathbf{ab}^{3x}\mathbf{a}, \\ \sigma_4(x) &= \mathbf{ab}^{4x-3}\mathbf{a} \mathbf{ab}^{4x-2}\mathbf{a} \mathbf{ab}^{4x-1}\mathbf{a} \mathbf{ab}^{4x}\mathbf{a}, \\ \sigma_5(x) &= \mathbf{ab}^{5x-4}\mathbf{a} \mathbf{ab}^{5x-3}\mathbf{a} \mathbf{ab}^{5x-2}\mathbf{a} \mathbf{ab}^{5x-1}\mathbf{a} \mathbf{ab}^{5x}\mathbf{a} \end{aligned}$$

for every $x \in \mathbb{N}$. With regard to such morphisms, we introduce the set $U(\sigma_n) \subseteq \mathbb{N}^+$ of all patterns with respect to which σ_n is unambiguous, and we give a characterisation of $U(\sigma_m)$ for $m \geq 3$. Furthermore, for every $n \in \mathbb{N}$, we compare

$U(\sigma_n)$ with every $U(\sigma_m)$, $m \neq n$, and, since every σ_n is a biprefix code (cf. Harju and Karhumäki [8]), we complement our approach by additionally considering the set $U(\sigma_0)$ of the suffix code σ_0 as introduced above.

Our studies in the present paper are centred around the following hypothesis:

Hypothesis 1. For $0 \leq i < j$, $U(\sigma_i) \subseteq U(\sigma_j)$.

The results of our analysis shall demonstrate that, in contrast to Hypothesis 1, these sets do not form a real hierarchy. More precisely, $U(\sigma_m) = U(\sigma_3)$ for all $m \geq 3$, and, while the sets $U(\sigma_0)$, $U(\sigma_1)$ and $U(\sigma_3)$ are strictly contained in $U(\sigma_3)$, they are pairwise incomparable.

Our approach is largely motivated by the intrinsic interest involved in the examination of the unambiguity of *fixed* instead of tailor-made morphisms. Therefore, we face a task which gives less definitional leeway than the original setting studied by Freydenberger et al. [6], and hence our paper reveals new elementary phenomena related to the ambiguity of morphisms that have not been discovered by the previous approach.

Our choice of segmented morphisms as the main objects under consideration, in turn, is primarily derived from the observation that σ_3 is the homogeneous version of $\sigma_\alpha^{\text{su}}$ (apart from a minor technical detail). Hence, the insights gained into $U(\sigma_3)$ immediately yield a deeper understanding of the necessity of the heterogeneity of $\sigma_\alpha^{\text{su}}$ and, thus, of a crucial concept introduced in [6]. In addition to this, our partly surprising results on Hypothesis 1 and the respective comparison between segmented and general homogeneous morphisms suggest that – in a similar manner as the work by, e. g., Halava et al. [7] with respect to the Post Correspondence Problem – we deal with a vital type of morphisms that addresses some of the very foundations of the problem field of ambiguity. Finally, it is surely worth mentioning that the properties of segmented morphisms have also been studied in the context of pattern languages (see, e. g., Jiang et al. [10]), and, in particular, recent papers prove the substantial impact of the (un-)ambiguity of such morphisms on pattern inference (cf. Reidenbach [17, 18]). Thus, our results provide a worthwhile starting point for further considerations in a prominent algorithmic research field. In the present paper, however, we do not explicitly discuss this aspect of our work.

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

Let $\mathbb{N} := \{1, 2, 3, \dots\}$ and $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. For any two sets A, B let $A \setminus B := \{a \in A \mid a \notin B\}$. Let Σ be an *alphabet*, i. e., an enumerable set of symbols. We consider 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 Σ) is a finite sequence of symbols from Σ . For the *concatenation* of two

strings w_1, w_2 we write $w_1 \cdot w_2$ or simply $w_1 w_2$. The notation $|x|$ stands for the size of a set x or the length of a string x , respectively. We denote the *empty string* by λ , i. e., $|\lambda| = 0$. 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 , $V(\alpha)$ denotes the set of all variables occurring in α . For every alphabet Σ , Σ^* is the set of all (empty and nonempty) strings over Σ , and $\Sigma^+ := \Sigma^* \setminus \{\lambda\}$. Furthermore, we use the regular operations $+$, $*$ and \cdot on sets and letters in the usual way. For any $w \in \Sigma^*$ and any $n \in \mathbb{N}$, w^n describes the n -fold concatenation of w , and $w^0 := \lambda$. We say that a string $v \in \Sigma^*$ is a *factor* of a string $w \in \Sigma^*$ if and only if, for some $u_1, u_2 \in \Sigma^*$, $w = u_1 v u_2$. If $u_1 = \lambda$ (or $u_2 = \lambda$), then v is a *prefix* of w (or a *suffix*, respectively). We say that a pattern α is in *canonical form* if its variables are introduced in the natural order, i. e., there is an $n \in \mathbb{N}$ such that $V(\alpha) = \{1, \dots, n\}$, and for every $x \in V(\alpha)$ and every $\beta, \gamma \in \mathbb{N}^*$ with $\alpha = \beta \cdot x \cdot \gamma$, $y \in V(\beta)$ holds for every $y < x$.

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 $\sigma : \mathbb{N}^* \rightarrow \{\mathbf{a}, \mathbf{b}\}^*$ 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 explicitly declare a morphism only for those variables that are relevant in the respective context. A morphism $\sigma : \mathbb{N}^* \rightarrow \{\mathbf{a}, \mathbf{b}\}^*$ is called *nonerasing* if $\sigma(x) \neq \lambda$ for all $x \in \mathbb{N}$. If, for every $\alpha, \beta \in \mathbb{N}^*$, $\sigma(\alpha) = \sigma(\beta)$ holds if and only if $\alpha = \beta$, σ is called *injective*. Note that every injective morphism is nonerasing.

For morphisms σ, τ and any $\alpha \in \mathbb{N}^+$, we say that $\sigma \neq \tau$ *with respect to* α if there is an $x \in V(\alpha)$ such that $\sigma(x) \neq \tau(x)$. If it is clear which pattern α we refer to, we write this as $\sigma \neq \tau$. For any pattern $\alpha \in \mathbb{N}^+$ with $\sigma(\alpha) \neq \lambda$, we call $\sigma(\alpha)$ *unambiguous (with respect to* α *or on* α) if there is no morphism $\tau : \mathbb{N}^* \rightarrow \{\mathbf{a}, \mathbf{b}\}^*$ such that $\tau(\alpha) = \sigma(\alpha)$ and $\tau \neq \sigma$; otherwise, we call σ *ambiguous (with respect to* α *or on* α). For a given morphism σ , let $U(\sigma)$ denote the set of all $\alpha \in \mathbb{N}^+$ such that σ is unambiguous on α .

As explained in Section 1, the main result of the initial paper on the ambiguity of morphisms by Freydenberger et al. [6] characterises those patterns with respect to which there is an unambiguous nonerasing morphism. This insight is based on the notion of so-called *succinct* and *prolix* (i. e., non-succinct) patterns. Some of our subsequent results again refer to these patterns, but our reasoning does not require the technical details of their involved definition. Instead, we introduce the following elementary equivalent concept: We call a pattern α *morphically imprimitive* if and only if there are a pattern β and morphisms $\phi, \psi : \mathbb{N}^* \rightarrow \mathbb{N}^*$ such that $|\beta| < |\alpha|$, $\phi(\alpha) = \beta$ and $\psi(\beta) = \alpha$. Accordingly, if α is not morphically imprimitive then we say that it is *morphically primitive*.

Example 2. Let $\alpha := 1 \cdot 1 \cdot 2 \cdot 3 \cdot 4 \cdot 1 \cdot 1 \cdot 2 \cdot 4 \cdot 3$, $\beta := 1 \cdot 2 \cdot 3 \cdot 1 \cdot 3 \cdot 2$, let the morphisms ϕ be defined by $\phi(1) := \lambda$, $\phi(2) := 1$, $\phi(3) := 2$ and $\phi(4) := 3$ and let the morphism ψ be defined by $\psi(1) := 1 \cdot 1 \cdot 2$, $\psi(2) := 3$ and $\psi(3) := 4$. As

$|\beta| < |\alpha|$, $\phi(\alpha) = \beta$ and $\psi(\beta) = \alpha$, α is morphically imprimitive. With some straightforward effort, based on checking all patterns in canonical form that are shorter than β , it can be verified that β is morphically primitive.

As shown by Reidenbach and Schneider [19], the partition of \mathbb{N}^* into the set of morphically primitive patterns and the set of morphically imprimitive patterns is characteristic for various aspects related to finite words and morphisms. First, the set of morphically imprimitive patterns exactly corresponds to the set of *fixed points* of nontrivial morphisms, i. e., for every morphically imprimitive pattern α there is a morphism $\phi : \mathbb{N}^* \rightarrow \mathbb{N}^*$ satisfying $\phi(\alpha) = \alpha$ and, for an $x \in V(\alpha)$, $\phi(x) \neq x$. Second, those patterns designated by Mitchell [14] as “succinct”, namely the shortest generators of their *E-pattern languages*, equal the morphically primitive patterns. Finally, and most importantly for the present paper, fixed points as well as the non-shortest generators of E-pattern languages are known to possess a characteristic factorisation (shown by Head [9] and Reidenbach [18], respectively), which is used by Freydenberger et al. [6] to *define* their concept of prolix patterns. Consequently, we can rephrase the main result of the latter paper in terms of our definition of morphic primitivity as follows:

Theorem 3 (Freydenberger et al. [6]). *A pattern α is morphically primitive if and only if there is a nonerasing morphism $\sigma_\alpha^{\text{su}} : \mathbb{N}^* \rightarrow \{\mathbf{a}, \mathbf{b}\}^*$ such that $\sigma_\alpha^{\text{su}}$ is unambiguous on α .*

Furthermore, the morphism $\sigma_\alpha^{\text{su}}$ needs to be tailor-made for α :

Theorem 4 (Freydenberger et al. [6]). *There is no nonerasing morphism $\sigma : \mathbb{N}^* \rightarrow \{\mathbf{a}, \mathbf{b}\}^*$ such that σ is unambiguous on every morphically primitive pattern.*

These two insights serve as the main fundament of the present work.

Within the scope of the present paper, we call a morphism $\sigma : \mathbb{N}^* \rightarrow \{\mathbf{a}, \mathbf{b}\}^*$ *left-homogeneous* if all $\sigma(x)$ share a common nonempty prefix. Likewise, we call σ *right-homogeneous* if all $\sigma(x)$ share a common nonempty suffix. Furthermore, σ is *homogeneous* if it is both left-homogeneous and right-homogeneous. Otherwise, σ is *heterogeneous*.

For every $n \in \mathbb{N}$, we define the morphism σ_n by

$$\sigma_n(x) := \mathbf{ab}^{nx-(n-1)}\mathbf{a} \mathbf{ab}^{nx-(n-2)}\mathbf{a} \dots \mathbf{ab}^{nx-1}\mathbf{a} \mathbf{ab}^{nx}\mathbf{a}$$

for every $x \in \mathbb{N}$. We call σ_n the *segmented morphism with n segments* and we refer to factors in $\mathbf{ab}^+\mathbf{a}$ as *segments*. In this work, we mostly concentrate on the morphisms $\sigma_1, \sigma_2, \sigma_3$ given by $\sigma_1(x) := \mathbf{ab}^x\mathbf{a}$, $\sigma_2(x) := \mathbf{ab}^{2x-1}\mathbf{a} \mathbf{ab}^{2x}\mathbf{a}$ and $\sigma_3(x) := \mathbf{ab}^{3x-2}\mathbf{a} \mathbf{ab}^{3x-1}\mathbf{a} \mathbf{ab}^{3x}\mathbf{a}$. Although it is not a segmented morphism, we also study the morphism σ_0 given by $\sigma_0(x) := \mathbf{ab}^x$, as it is quite similar to σ_1 and often used to encode words over infinite alphabets using only two letters. Note that, by definition, every segmented morphism (as well as σ_0) is homogeneous.

There is an important property of all σ_n with $n \geq 3$ that can be derived from the proof of Lemma 28 in [6]:

Lemma 5. *Let $\alpha \in \mathbb{N}^+$ be a morphically primitive pattern, and let $n \geq 3$. Then, for every morphism $\tau : \mathbb{N}^* \rightarrow \{\mathbf{a}, \mathbf{b}\}^*$ with $\tau(\alpha) = \sigma_n(\alpha)$ and for every $x \in V(\alpha)$, $\tau(x)$ contains the factor $\mathbf{a} \mathbf{a} \mathbf{b}^{nx-(n-2)} \mathbf{a} \dots \mathbf{a} \mathbf{b}^{nx-1} \mathbf{a} \mathbf{a}$.*

This lemma is very useful in the next section.

3. Morphisms with Three or More Segments

Due to Freydenberger et al. [6], we know that there are characteristic regularities in morphically imprimitive patterns which render every nonerasing morphism ambiguous on these patterns. Although morphic primitivity prohibits those regularities, some other structures supporting ambiguity of segmented morphisms can occur. For example, as mentioned above, it is easy to see that σ_1 is ambiguous on the morphically primitive pattern $\alpha := 1 \cdot 2 \cdot 1 \cdot 3 \cdot 3 \cdot 2$, e. g. by considering morphisms τ_1 or τ_2 which are given by $\tau_1(1) := \mathbf{a} \mathbf{b}$, $\tau_1(2) := \mathbf{a} \mathbf{a} \mathbf{b}^2 \mathbf{a}$ and $\tau_1(3) := \mathbf{a} \mathbf{a} \mathbf{b}^3$ and $\tau_2(1) := \mathbf{a} \mathbf{b} \mathbf{a} \mathbf{a}$, $\tau_2(2) := \mathbf{b}^2 \mathbf{a}$ and $\tau_2(3) := \mathbf{b}^3 \mathbf{a} \mathbf{a}$. In both cases, the arising ambiguity can be understood (albeit rather metaphorically) as a kind of communication where occurrences of 1 decide which modification is applied to their image under σ_1 and communicate this change to occurrences of 2, where applicable using 3 as a carrier. Those patterns that show such a structure can be generalised as follows:

Definition 6. Let $\alpha \in \mathbb{N}^+$. An *SCRN-partition*² for α is a partition of $V(\alpha)$ into pairwise disjoint sets S, C, R and N such that $\alpha \in (N^* S C^* R)^+ N^*$.

As suggested by the above example, the existence of an SCRN-partition for a pattern α is a sufficient condition for the ambiguity of any segmented morphism (and σ_0 as well). In fact, this even holds for every *left-* or *right-homogeneous* morphism:

Proposition 7. *Let $\alpha \in \mathbb{N}^+$. If α has an SCRN-partition, then any left-homogeneous or right-homogeneous morphism σ is ambiguous on α .*

PROOF. If σ is left-homogeneous, there exists a $p \in \{\mathbf{a}, \mathbf{b}\}^+$ such that, for every $x \in \mathbb{N}$, there is an $s_x \in \{\mathbf{a}, \mathbf{b}\}^*$ with $\sigma(x) = p s_x$. Let S, C, R, N be an SCRN-partition for α . We define τ by

$$\tau(x) := \begin{cases} \sigma(x) p & \text{for } x \in S, \\ s_x & \text{for } x \in R, \\ s_x p & \text{for } x \in C, \\ \sigma(x) & \text{for } x \in N. \end{cases}$$

As we are using an SCRN-partition, $\alpha \notin N^*$; therefore, $\tau \neq \sigma$ holds. It is easy to see that $\tau(\alpha) = \sigma(\alpha)$. Thus, σ is ambiguous on α .

²The letters S, C, R and N stand for *sender, carrier, receiver* and *neutral*, respectively.

The above construction shows that, if σ is left-homogeneous, τ can be obtained from σ by removing the common prefix of the images of the variables in $R \cup C$ and appending it to the images of the variables in $S \cup C$. If σ is right-homogeneous, an appropriate morphism τ can be constructed by analogous application of this technique to the common suffixes of all variables in $S \cup C$. \square

We now wish to demonstrate that, for σ_n with $n \geq 3$, this condition is even characteristic. If σ_n is ambiguous on some morphically primitive $\alpha \in \mathbb{N}^+$ (i. e., there is some $\tau \neq \sigma_n$ with $\tau(\alpha) = \sigma_n(\alpha)$), every variable possessing different images under τ and σ_n still keeps all its characteristic inner segments under τ (cf. Lemma 5). Any change is therefore limited to some gain or loss of its (or its neighbours') outer segments and has to be communicated along factors corresponding to the SC^*R -factors induced by an SCRN-partition. This allows us to construct an SCRN-partition from τ and leads to the following theorem:

Theorem 8. *Let $\alpha \in \mathbb{N}^+$. Then, for every $n \geq 3$, σ_n is ambiguous on α if and only if α is morphically imprimitive or has an SCRN-partition.*

PROOF. Due to Theorem 3, we can safely restrict ourselves to morphically primitive α , since every nonerasing morphism is ambiguous on every morphically imprimitive $\alpha \in \mathbb{N}^+$.

We begin with the only-if-direction. Assume σ_n is ambiguous on some morphically primitive $\alpha \in \mathbb{N}^+$; then there exists some morphism $\tau \neq \sigma_n$ with $\tau(\alpha) = \sigma_n(\alpha)$. Lemma 5 guarantees that every $\tau(x)$ contains the $n - 2$ inner segments of $\sigma_n(x)$. This allows us to develop the notions of left-invariant and right-invariant variables in $V(\alpha)$. We say that a variable $x \in V(\alpha)$ is *left-invariant* if $\tau(x) = \mathbf{ab}^{n_x - (n-1)} \mathbf{a} w$ for some $w \in \{\mathbf{a}, \mathbf{b}\}^+$. In other words, when reading from the left to the right, $\tau(x)$ begins with the first segment of $\sigma_n(x)$; and together with Lemma 5, this implies that $\tau(x)$ begins with the first $n - 1$ segments of $\sigma(x)$. Analogously, we say that $x \in V(\alpha)$ is *right-invariant* if $\tau(x) = w \mathbf{ab}^{n_x} \mathbf{a}$ for some $w \in \{\mathbf{a}, \mathbf{b}\}^+$. Using Lemma 5 we observe:

Fact 1: Let $x, y \in V(\alpha)$ such that xy is a factor of α . Then x is right-invariant if and only if y is left-invariant.

Next, we define the SCRN-partition for α . For every $x \in V(\alpha)$ let:

- $x \in N$ if x is left-invariant and right-invariant,
- $x \in S$ if x is left-invariant, but not right-invariant,
- $x \in C$ if x is neither left-invariant, nor right-invariant,
- $x \in R$ if x is right-invariant, but not left-invariant.

To show that $\alpha \in (N^*SC^*R)^+ N^*$, we read α from the left to the right. As the first variable has no left neighbour, it has to be left-invariant; thus, it must belong to N or S . If it belongs to N , it is right-invariant, and we can argue in the same way for the next variable. As $\alpha \in N^+$ would contradict $\tau \neq \sigma_n$,

sooner or later some variable from S must occur. This variable is not right-invariant, thus, Fact 1 implies that its right neighbour is not left-invariant, and consequently, that variable must belong to C or R . If it is from C (and thus, not right-invariant) as well, we continue observing variables from C until a variable from R is encountered; so α has a prefix from N^*SC^*R . But as all variables from R are right-invariant, we now have the same situation as when we started. We conclude $\alpha \in (N^*SC^*R)^+ N^*$; therefore, α has an SCRN-partition. The if-direction follows from Proposition 7. \square

Thus, Hypothesis 1 is indeed correct with regard to morphisms with at least three segments, although (perhaps surprisingly) the sets $U(\sigma_n)$ are neither *proper* subsets nor *proper* supersets of each other.

In terms of our sets $U(\sigma)$ of patterns on which a morphism σ is unambiguous, we can paraphrase Theorem 8 as follows:

Corollary 9. *For every $n \geq 3$,*

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

As a sidenote, consider *generalised segmented morphisms with n segments* as morphisms $\sigma_G : \mathbb{N}^* \rightarrow \Sigma^*$ where $\sigma_G(x) \in (\mathbf{ab}^+\mathbf{a})^n$ for all $x \in \mathbb{N}$, and for every $w \in \mathbf{ab}^+\mathbf{a}$, there is at most one $x \in \mathbb{N}$ such that w is a factor of $\sigma_G(x)$. It can be shown that if $n \geq 3$, Lemma 5 holds for σ_G as well. Thus, for every generalised segmented morphism σ_G with at least three segments, $U(\sigma_G) = U(\sigma_3)$. Furthermore, as σ_3 is basically the homogeneous version of the heterogeneous unambiguous morphism $\sigma_\alpha^{\text{su}}$ constructed by Freydenberger et al. [6], Theorem 8 precisely separates the patterns for which there is an unambiguous *homogeneous* morphism from those patterns where an unambiguous morphism has to be *heterogeneous*. Thus, a characteristic criterion on the existence of homogeneous unambiguous morphisms for any pattern can directly be derived from the following insight:

Corollary 10. *Let $\sigma : \mathbb{N}^* \rightarrow \Sigma^*$ be a nonerasing morphism that is left-homogeneous or right-homogeneous. Then*

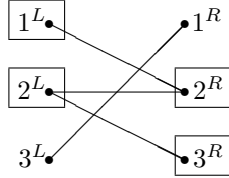
$$U(\sigma) \subseteq U(\sigma_3) = \{\alpha \in \mathbb{N}^+ \mid \alpha \text{ is morphically primitive and } \alpha \text{ has no SCRN-partition}\}.$$

Next, we present a surprising alternative characterisation of $U(\sigma_3)$. To this end, we need to go into greater detail and introduce some preparatory technical concepts. Theorem 8 demonstrates that, for σ_n with $n \geq 3$, ambiguity on morphically primitive patterns is inherently related to the occurrence of global regularities that depend on local interactions between *neighbouring variables* only. In fact, these regularities can be described by the equivalence classes L_i^\sim and R_i^\sim on $V(\alpha)$ introduced by Freydenberger et al. [6] as fundamental tools to construct tailor-made unambiguous morphisms $\sigma_\alpha^{\text{su}}$. In the present paper, we describe these equivalence classes using an equivalent definition. This new method is more convenient than the established one and, thus, significantly simplifies the approach in [6]. It is based on the adjacency graph of a pattern, a construction

that was first employed by Baker et al. [1] to simplify the Bean-Ehrenfeucht-McNulty-Zimin characterisation of avoidable patterns, cf. Cassaigne [2]. Like Baker et al., we associate a pattern $\alpha \in \mathbb{N}^+$ with a bipartite graph $\text{AG}(\alpha)$, the *adjacency graph* of α , defined as follows: The vertex set consists of two marked copies of $V(\alpha)$, V^L and V^R (for left and right, respectively), i. e., for each $x \in V(\alpha)$, there is an element $x^L \in V^L$ and an element $x^R \in V^R$. There is an edge $x^L - y^R$ for $x, y \in V(\alpha)$ if and only if xy is a factor of α .

In *contrast* to Baker et al., we consider a partition of $V^L \cup V^R$ into sets H_1, \dots, H_n such that each H_i is the set of vertices of a maximal and connected subgraph of $\text{AG}(\alpha)$. We call such a set H_i a *neighbourhood in α* and refer to the set of all neighbourhoods as $H(\alpha)$. For every neighbourhood H_i , the *left neighbourhood class* L_i^\sim denotes the set of all x such that x^L is in H_i and likewise the *right neighbourhood class* R_i^\sim the set of all x such that x^R is in H_i .

Example 11. Let $\alpha := 1 \cdot 2 \cdot 3 \cdot 1 \cdot 2 \cdot 2 \cdot 3$. We obtain $H_1 = \{1^L, 2^L, 2^R, 3^R\}$ and $H_2 = \{3^L, 1^R\}$ and therefore $L_1^\sim = \{1, 2\}$, $L_2^\sim = \{3\}$, $R_1^\sim = \{2, 3\}$ and $R_2^\sim = \{1\}$. In the following figure, we display the adjacency graph of α . Boxes mark the elements of H_1 :



As every nonerasing morphism is ambiguous on a morphically imprimitive pattern (cf. Theorem 3), we mainly deal with morphically primitive patterns. It is useful to note that, apart from patterns α of length 1 (such as $\alpha = 1$), no morphically primitive pattern contains variables that occur only once. Therefore, in morphically primitive patterns every neighbourhood contains elements from V^L and V^R , and every variable belongs to exactly one left and one right neighbourhood class.

Utilising our definition of neighbourhood classes, we now give the alternative characterisation of $U(\sigma_n)$, $n \geq 3$:

Theorem 12. Let $\alpha' \in \mathbb{N}^*$, $l, r \in \mathbb{N}$ and $\alpha := l\alpha'r$. For every $n \geq 3$, σ_n is ambiguous on α if and only if α is morphically imprimitive or there is a neighbourhood $H_i \in H(\alpha)$ such that $l \notin R_i^\sim$ and $r \notin L_i^\sim$.

PROOF. Again, Theorem 3 allows us to restrict our considerations to morphically primitive patterns. First, assume that, for some morphically primitive pattern $\alpha = l\alpha'r$ with $\alpha' \in \mathbb{N}^*$, there is some morphism $\tau \neq \sigma_n$ such that $\tau(\alpha) = \sigma_n(\alpha)$. Now we refer to the definition of left-invariant and right-invariant variables in the proof of Theorem 8, and we construct an SCRN-partition for α as in that proof, i. e., for every $x \in V(\alpha)$, let:

- $x \in N$ if x is left-invariant and right-invariant,

- $x \in S$ if x is left-invariant, but not right-invariant,
- $x \in C$ if x is neither left-invariant, nor right-invariant,
- $x \in R$ if x is right-invariant, but not left-invariant.

We observe that Fact 1 mentioned in the proof of Theorem 8 can be extended to neighbourhood classes, as Lemma 5 applies:

Fact 2: Let $H_i \in H(\alpha)$. If L_i^\sim contains a right-invariant variable or R_i^\sim contains a left-invariant variable, then all variables in L_i^\sim are right-invariant and all variables in R_i^\sim are left-invariant.

Proof of Fact 2: Assume that $x \in L_i^\sim$ and $y \in R_i^\sim$. Then there exist an $m \in \mathbb{N}$ and variables $z_0, \dots, z_m \in V(\alpha)$ such that $x = z_0$, $y = z_m$ and $z_0^L - z_1^R - \dots - z_{m-1}^L - z_m^R$ is a path in $\text{AG}(\alpha)$. By definition, $z_i^L - z_j^R$ is an edge of $\text{AG}(\alpha)$ if and only if $z_i z_j$ is a factor of α , and by Fact 1, this implies that z_i is right-invariant if and only if z_j is left-invariant. Thus, x is right-invariant if and only if y is left-invariant. This implies that, whenever one $x \in L_i^\sim$ is right-invariant, all variables in R_i^\sim are left-invariant; furthermore, whenever one $y \in R_i^\sim$ is left-invariant, all variables in L_i^\sim are right-invariant. These two observations combined prove Fact 2.

Now, we choose some $x \in S$. Let i be the (uniquely defined) value for which $x \in L_i^\sim$. As $x \in S$, by definition, x is not right-invariant. Thus, by Fact 2, no $y \in R_i^\sim$ is left-invariant, which implies that $l \notin R_i^\sim$ must hold, as $l \in R_i^\sim$ would imply that l is not left-invariant; but as l is the leftmost variable of α , this would contradict $\tau(\alpha) = \sigma(\alpha)$. Likewise, $r \notin L_i^\sim$ must hold: Indeed, if we assume $r \in L_i^\sim$, then Fact 2 implies that r is not right-invariant, but as r is the rightmost variable of α , this contradicts $\tau(\alpha) = \sigma(\alpha)$. Thus, H_i is a neighbourhood in $H(\alpha)$ with $l \notin R_i^\sim$ and $r \notin L_i^\sim$.

For the other direction, let $\alpha := la'r$ be morphically primitive with some neighbourhood H_i such that $l \notin R_i^\sim$ and $r \notin L_i^\sim$. Now define S, C, R, N by $S = L_i^\sim \setminus R_i^\sim$, $C = L_i^\sim \cap R_i^\sim$, $R = R_i^\sim \setminus L_i^\sim$ and $N = V(\alpha) \setminus (L_i^\sim \cup R_i^\sim)$. The four sets form a partition of $V(\alpha)$, so it merely remains to be shown that $\alpha \in (N^*SC^*R)^+ N^*$. First observe that, by definition, $l \in S \cup N$ and $r \in R \cup N$. Furthermore, for any factor xy of α , if $x \in S$ or $x \in C$, then $x \in L_i^\sim$. Therefore, $y \in R_i^\sim$ and thus $y \in C \cup R$. Likewise, $x \in N$ or $x \in R$ implies $x \notin L_i^\sim$ and $y \notin R_i^\sim$, which leads to $y \in N \cup S$ and $\alpha \in (N^*SC^*R)^+ N^*$. By Theorem 8, we conclude that σ_n is ambiguous on α . \square

Consequently, for any morphically primitive pattern, the neighbourhood classes of just the *first* and the *last* variable decide on the ambiguity of σ_n , $n \geq 3$. We consider it very counterintuitive that this largely “local” criterion is equivalent to the “global” one in Theorem 8, which asks for the existence of a particular factorisation of the full pattern α in question. Furthermore, we note that the construction of $H(\alpha)$ can be done efficiently, e. g. by using a Union-Find-algorithm (cf. Sedgewick [20]).

This theorem provides a useful corollary for a class of patterns first described by Baker et al. [1]. We call a pattern $\alpha \in \mathbb{N}^+$ *locked* if and only if $|H(\alpha)| = 1$ and thus $L_1^\sim = R_1^\sim = V(\alpha)$. We observe the following consequence:

Corollary 13. *Let $\alpha \in \mathbb{N}^+$. If α is morphically primitive and locked, then $\alpha \in U(\sigma_3)$.*

This corollary is of use in the next section, where we shall see that having less than three segments entails other types of ambiguity than the one described in the present section.

4. Morphisms with Less than Three Segments

In this section, we examine the effects caused by reducing the number of segments. As postulated in Hypothesis 1, one might expect no change in the corresponding sets U , or a small hierarchy that reflects the number of segments, but, as we shall see, neither is the case. In particular, we show that the sets $U(\sigma_0)$, $U(\sigma_1)$ and $U(\sigma_2)$ are pairwise incomparable. To this end, we construct the following five patterns:

Definition 14. We define α_0 , α_1 and α_2 as follows:

$$\begin{aligned}\alpha_0 &:= 1 \cdot 2 \cdot 3 \cdot 1 \cdot 3 \cdot 2, \\ \alpha_1 &:= 1 \cdot 2 \cdot 2 \cdot 3 \cdot 1 \cdot 1 \cdot 3 \cdot 1, \\ \alpha_2 &:= (1 \cdot 2 \cdot 3 \cdot 3 \cdot 4)^2 \cdot 5 \cdot 2 \cdot 6 \cdot 5 \cdot 7 \cdot (8 \cdot 6)^2 \cdot (9 \cdot 7)^2 \cdot 10 \cdot 4 \cdot 11 \cdot 4 \cdot \\ &\quad 10 \cdot 12 \cdot 11 \cdot 12 \cdot (3 \cdot 13)^2 \cdot (14 \cdot 3 \cdot 2 \cdot 15)^2.\end{aligned}$$

Furthermore, we define $\alpha_{0 \setminus 2}$ through

$$\alpha_{0 \setminus 2} := \beta_1 \cdot \beta_2 \cdot \dots \cdot \beta_9 \cdot 16 \cdot \beta_1 \cdot 16 \cdot \beta_2 \cdot 16 \cdot \dots \cdot 16 \cdot \beta_9 \cdot (16)^{17}$$

with

$$\begin{aligned}\beta_1 &:= 1 \cdot 2 \cdot 3, & \beta_2 &:= 4 \cdot 5 \cdot 4, \\ \beta_3 &:= 6 \cdot 7 \cdot 6 \cdot 8, & \beta_4 &:= 1 \cdot 7 \cdot 3, \\ \beta_5 &:= 9 \cdot 6 \cdot 6 \cdot 10, & \beta_6 &:= 11 \cdot 12, \\ \beta_7 &:= 13 \cdot 7 \cdot 7 \cdot 4 \cdot 14 \cdot 12, & \beta_8 &:= 15 \cdot 14, \\ \beta_9 &:= 9 \cdot 6.\end{aligned}$$

Finally, we define $\alpha_{1 \setminus 2}$ by $\alpha_{1 \setminus 2} := 1^2 \cdot \delta \cdot 1 \cdot p(\delta) \cdot 1$, where

$$\begin{aligned}\delta &:= \delta_1 \cdot 1 \cdot \delta_2 \cdot 1 \cdot \delta_3 \cdot 1 \cdot \delta_4 \cdot 1 \cdot \delta_5 \cdot 1 \cdot \gamma_1 \cdot \delta_6 \cdot 1 \cdot \delta_7 \cdot 1 \cdot \gamma_2 \cdot 1 \cdot \delta_8 \cdot \\ &\quad 1 \cdot \gamma_3 \cdot 1 \cdot \delta_9 \cdot 1 \cdot \delta_{10} \cdot 1 \cdot \delta_{11} \cdot 1 \cdot \delta_{12} \cdot 1 \cdot \delta_{13} \cdot 1 \cdot \delta_{14},\end{aligned}$$

and $p(1) := \lambda$, $p(x) := x$ for all $x \in \mathbb{N} \setminus \{1\}$, and furthermore

$$\begin{aligned}\delta_1 &:= 2 \cdot 3 \cdot 3 \cdot 4, & \delta_2 &:= 3 \cdot 2 \cdot 2 \cdot 5, \\ \delta_3 &:= 6 \cdot 7, & \delta_4 &:= 8 \cdot 9, \\ \delta_5 &:= 10 \cdot 11, & \delta_6 &:= 18 \cdot 19,\end{aligned}$$

$$\begin{aligned}
\delta_7 &:= 6 \cdot 20 \cdot 9, & \delta_8 &:= 6 \cdot 23 \cdot 11, \\
\delta_9 &:= 27 \cdot (2 \cdot 20)^2 \cdot 28 \cdot 2 \cdot 29, & \delta_{10} &:= 30 \cdot 2 \cdot (20 \cdot 2 \cdot 23 \cdot 2)^3 \cdot 20 \cdot (31)^4 \cdot 32, \\
\delta_{11} &:= 33 \cdot 3 \cdot 34 \cdot (23 \cdot 3)^2 \cdot 35, & \delta_{12} &:= 36 \cdot 20 \cdot 20 \cdot 28 \cdot 2 \cdot 29, \\
\delta_{13} &:= 33 \cdot 3 \cdot 34 \cdot 23 \cdot 23 \cdot 37, & \delta_{14} &:= 18 \cdot 31 \cdot 32, \\
\gamma_1 &:= 12 \cdot 1 \cdot 13 \cdot 1 \cdot \dots \cdot 17 \cdot 1, & \gamma_2 &:= 21 \cdot 1 \cdot 22, \\
\gamma_3 &:= 24 \cdot 1 \cdot 25 \cdot 1 \cdot 26.
\end{aligned}$$

We begin by establishing the relation between $U(\sigma_3)$ and the other sets:

Theorem 15. *The sets $U(\sigma_0)$, $U(\sigma_1)$ and $U(\sigma_2)$ are strictly included in $U(\sigma_3)$.*

PROOF. For all three languages, the inclusion directly follows from Corollary 10, since σ_0 , σ_1 and σ_2 are nonerasing and homogeneous. To prove strictness, we show that σ_0 , σ_1 , σ_2 are ambiguous on the patterns α_0 , α_1 , α_2 , respectively, as given by Definition 14. According to Reidenbach and Schneider [19], a pattern α is morphically imprimitive if and only if there exists a factorisation $\alpha = \beta_0 \gamma_1 \beta_1 \gamma_2 \beta_2 [\dots] \beta_{n-1} \gamma_n \beta_n$ 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' , $0 \leq k' \leq n$, $V(\gamma_k) \cap V(\beta_{k'}) = \emptyset$,
3. for every k , $1 \leq k \leq n$, there exists an $i_k \in V(\gamma_k)$ such that $|\gamma_k|_{i_k} = 1$ and, for every k' , $1 \leq k' \leq n$, if $i_k \in V(\gamma_{k'})$ then $\gamma_k = \gamma_{k'}$.

Using this characterisation, it is easy to see that all three patterns are morphically primitive. Furthermore, note that these patterns are locked, as can be shown through the corresponding adjacency graphs. This is quite trivial for α_0 and α_1 ; the edge set of $\text{AG}(\alpha_0)$ consists of the edges $1^L - 2^R$, $1^L - 3^R$, $2^L - 3^R$, $3^L - 1^R$ and $3^L - 2^R$, while the edge set of $\text{AG}(\alpha_1)$ consists of $1^L - 1^R$, $1^L - 2^R$, $1^L - 3^R$, $2^L - 2^R$, $2^L - 3^R$ and $3^L - 1^R$. In both cases, one can easily find a path between any two vertices. On the other hand, $\text{AG}(\alpha_2)$ is more complicated; its edges are $1^L - 2^R$, $2^L - 3^R$, $2^L - 6^R$, $2^L - 15^R$, $3^L - 2^R$, $3^L - 3^R$, $3^L - 4^R$, $3^L - 13^R$, $4^L - 1^R$, $4^L - 5^R$, $4^L - 10^R$, $4^L - 11^R$, $5^L - 2^R$, $5^L - 7^R$, $6^L - 5^R$, $6^L - 8^R$, $6^L - 9^R$, $7^L - 8^R$, $7^L - 9^R$, $7^L - 10^R$, $8^L - 6^R$, $9^L - 7^R$, $10^L - 4^R$, $10^L - 12^R$, $11^L - 4^R$, $11^L - 12^R$, $12^L - 3^R$, $12^L - 11^R$, $13^L - 3^R$, $13^L - 14^R$, $14^L - 3^R$ and $15^L - 14^R$. Still, with some straightforward effort, one can show that $\text{AG}(\alpha_2)$ is connected as well. Hence, all three patterns are locked, as each has exactly one neighbourhood, and due to Corollary 13, σ_3 is unambiguous on each of the patterns.

We start with α_0 and define τ by $\tau(1) := \sigma_0(1 \cdot 2)$, $\tau(2) := \sigma_0(2)$ and $\tau(3) := \mathbf{b}$. Then $\tau \neq \sigma_0$, but $\tau(\alpha_0) = \sigma_0(1 \cdot 2) \cdot \sigma_0(2) \cdot \mathbf{b} \cdot \sigma_0(1 \cdot 2) \cdot \mathbf{b} \cdot \sigma_0(2) = \sigma_0(\alpha_0)$. Therefore, σ_0 is ambiguous on α_0 . For α_1 , we set $\tau(1) := \mathbf{a}$, $\tau(2) := \mathbf{baab}$ and $\tau(3) := \mathbf{ba}\sigma_1(3)\mathbf{ab}$. It is easy to see that $\tau \neq \sigma_1$ and $\tau(\alpha_1) = \sigma_1(\alpha_1)$. With regard to σ_2 , we consider the morphism τ given by

$$\tau(1) := \sigma_2(1 \cdot 2 \cdot 3) \mathbf{ab}^5 \mathbf{a} \mathbf{ab}^3, \quad \tau(2) := \mathbf{b}^3 \mathbf{a} \mathbf{ab}^3,$$

$$\begin{aligned}
\tau(3) &:= \lambda, & \tau(4) &:= b^4 a ab^8 a, \\
\tau(5) &:= \sigma_2(5) a, & \tau(6) &:= ba \sigma_2(6), \\
\tau(7) &:= b^{13} a ab^{14} a, & \tau(8) &:= ab^{15} a ab^{15}, \\
\tau(9) &:= \sigma_2(9) a, & \tau(10) &:= \sigma_2(10) ab^3, \\
\tau(11) &:= \sigma_2(11) ab^3, & \tau(12) &:= b^{20} a ab^{24} a, \\
\tau(13) &:= \sigma_2(3 \cdot 13), & \tau(14) &:= ab^{27} a ab^{25}, \\
\tau(15) &:= b^2 a ab^6 a \sigma_2(2 \cdot 15).
\end{aligned}$$

Then $\tau \neq \sigma_2$. Proving $\tau(\alpha_2) = \sigma_2(\alpha_2)$ is less obvious, but straightforward:

$$\begin{aligned}
\sigma_2(\alpha_2) &= \underbrace{\overbrace{aba ab^2 a ab^3 a ab^4 a}^{\sigma_2(1)} \overbrace{ab^5 a ab^6 a ab^5 a ab^3}^{\sigma_2(2)} \overbrace{b^3 a ab^3}^{\sigma_2(3)} \overbrace{b^4 a ab^8 a}^{\sigma_2(4)}}_{\tau(1)} \\
&= \underbrace{\overbrace{aba ab^2 a ab^3 a ab^4 a}^{\sigma_2(1)} \overbrace{ab^5 a ab^6 a ab^5 a ab^3}^{\sigma_2(2)} \overbrace{b^3 a ab^3}^{\sigma_2(3)} \overbrace{b^4 a ab^8 a}^{\sigma_2(4)}}_{\tau(1)} \\
&= \underbrace{\overbrace{ab^9 a ab^{10} a a}^{\sigma_2(5)} \overbrace{b^3 a ab^3}^{\sigma_2(2)} \overbrace{ba ab^{11} a ab^{12} a}^{\sigma_2(6)} \overbrace{ab^9 a ab^{10} a a}^{\sigma_2(5)} \overbrace{b^{13} a ab^{14} a}^{\sigma_2(7)}}_{\tau(5)} \\
&= \underbrace{\overbrace{ab^{15} a ab^{15}}^{\sigma_2(8)} \overbrace{ba ab^{11} a ab^{12} a}^{\sigma_2(6)} \overbrace{ab^{15} a ab^{15}}^{\sigma_2(8)} \overbrace{ba ab^{11} a ab^{12} a}^{\sigma_2(6)}}_{\tau(8)} \\
&= \underbrace{\overbrace{ab^{17} a ab^{18} a a}^{\sigma_2(9)} \overbrace{b^{13} a ab^{14} a}^{\sigma_2(7)} \overbrace{ab^{17} a ab^{18} a a}^{\sigma_2(9)} \overbrace{b^{13} a ab^{14} a}^{\sigma_2(7)}}_{\tau(9)} \\
&= \underbrace{\overbrace{ab^{19} a ab^{20} a ab^3}^{\sigma_2(10)} \overbrace{b^4 a ab^8 a}^{\sigma_2(4)} \overbrace{ab^{21} a ab^{22} a ab^3}^{\sigma_2(11)} \overbrace{b^4 a ab^8 a}^{\sigma_2(4)}}_{\tau(10)} \\
&= \underbrace{\overbrace{ab^{19} a ab^{20} a ab^3}^{\sigma_2(10)} \overbrace{b^{20} a ab^{24} a}^{\sigma_2(12)} \overbrace{ab^{21} a ab^{22} a ab^3}^{\sigma_2(11)} \overbrace{b^{20} a ab^{24} a}^{\sigma_2(12)}}_{\tau(10)} \\
&= \underbrace{\overbrace{ab^5 a ab^6 a ab^{25} a ab^{26} a}^{\sigma_2(3)} \overbrace{ab^5 a ab^6 a ab^{25} a ab^{26} a}^{\sigma_2(13)}}_{\tau(13)} \\
&= \underbrace{\overbrace{ab^{27} a ab^{25}}^{\sigma_2(14)} \overbrace{b^3 a ab^3}^{\sigma_2(3)} \overbrace{b^2 a ab^6 a ab^3 a ab^4 a}^{\sigma_2(2)} \overbrace{ab^{29} a ab^{30} a}^{\sigma_2(15)}}_{\tau(14)}
\end{aligned}$$

$$\begin{aligned}
& \overbrace{\mathbf{ab}^{27}\mathbf{a}\mathbf{ab}^{25}}^{\sigma_2(14)} \overbrace{\mathbf{b}^3\mathbf{a}\mathbf{ab}^3}^{\sigma_2(3)} \overbrace{\mathbf{b}^2\mathbf{a}\mathbf{ab}^6\mathbf{a}}^{\sigma_2(2)} \overbrace{\mathbf{ab}^3\mathbf{a}\mathbf{ab}^4\mathbf{a}}^{\sigma_2(15)} \\
& \quad \tau(14) \quad \tau(2) \quad \tau(15) \\
& = \tau(\alpha_2).
\end{aligned}$$

This concludes the proof of Theorem 15. \square

Thus, with respect to the relationship between $U(\sigma_3)$ and each of $U(\sigma_0)$, $U(\sigma_1)$ and $U(\sigma_2)$, Hypothesis 1 is indeed correct.

The proof for Theorem 15 is of additional interest as Freydenberger et al. [6] propose to study a morphism $\sigma_\alpha^{2\text{-seg}}$ that maps each variable x in a morphically primitive pattern α onto a word that merely consists of the left and the right segment of $\sigma_\alpha^{\text{su}}(x)$ (recall that $\sigma_\alpha^{\text{su}}$ is a heterogeneous morphism which maps every variable x onto *three* segments and is tailored to α ; with regard to the details, see [6]). In [6] it is asked whether, for every morphically primitive pattern α , $\sigma_\alpha^{2\text{-seg}}$ is unambiguous on α , thus suggesting the chance for a major improvement of $\sigma_\alpha^{\text{su}}$. With regard to this question, we now consider the pattern α_2 . In the above proof, it is stated that $|H(\alpha_2)| = 1$, and hence α_2 is a locked pattern (by definition), which implies that $\sigma_{\alpha_2}^{2\text{-seg}}$ only maps the variable 1 onto a word $\mathbf{b}\dots\mathbf{b}$ and all other variables in α_2 onto words $\mathbf{a}\dots\mathbf{a}$. Consequently, for each $x \in V(\alpha_2) \setminus \{1\}$, $\sigma_{\alpha_2}^{2\text{-seg}}(x) = \sigma_2(x)$. Therefore – and since, for the corresponding τ introduced in the proof of Theorem 15, the word $\tau(1)$ completely contains $\sigma_2(1)$ – we can define a morphism τ' by $\tau'(1) := \sigma_{\alpha_2}^{2\text{-seg}}(1 \cdot 2 \cdot 3) \mathbf{ab}^5\mathbf{a}\mathbf{ab}^3$ and $\tau'(x) := \tau(x)$, $x \in V(\alpha_2) \setminus \{1\}$, and this definition yields $\tau'(\alpha_2) = \sigma_{\alpha_2}^{2\text{-seg}}(\alpha_2)$. So, there exists a morphically primitive pattern α (namely α_2) such that $\sigma_\alpha^{2\text{-seg}}$ is ambiguous on α . Thus, α_2 does not only prove $U(\sigma_2) \subset U(\sigma_3)$, but it also provides a negative answer to a question posed in [6].

Returning to the focus of the present paper, the examples in the proof of Theorem 15 demonstrate ambiguity phenomena that are intrinsic for their respective kind of morphisms and cause ambiguity on patterns that are neither morphically imprimitive nor have an SCRN-partition. With regard to σ_0 , the fact that for each x, y with $x < y$, $\sigma_0(x)$ is a prefix of $\sigma_0(y)$ can lead to ambiguity, as demonstrated by α_0 . Concerning σ_1 , a variable x can obtain an image $\tau(x) = \mathbf{a}$ both by “giving” \mathbf{ab}^x to the left or $\mathbf{b}^x\mathbf{a}$ to the right, like the variable 1 in α_1 .

The situation is less obvious and somewhat more complicated for σ_2 , as suggested by the fact that we do not know a shorter pattern serving the same purpose as α_2 . Here, a variable x can obtain an image $\tau(x) \in \mathbf{b}^*\mathbf{aab}^*$, which can be used both as a middle part of some $\sigma_2(y)$, and as the borderline between some $\sigma_2(y)$ and some $\sigma_2(z)$. In the proofs of Theorem 17 and Theorem 18 we utilise further examples for complicated cases of σ_2 -ambiguity.

Especially in the context of Hypothesis 1, it is natural to ask whether these phenomena can be used to find patterns where one of the three morphisms σ_0 , σ_1 , σ_2 is ambiguous, and another is not. We begin with a comparison of $U(\sigma_0)$ and $U(\sigma_1)$:

Theorem 16. *The sets $U(\sigma_0)$ and $U(\sigma_1)$ are incomparable.*

PROOF. The ambiguity of σ_0 on α_0 and of σ_1 on α_1 is established in the proof of Theorem 15. It remains to show that $\alpha_0 \in U(\sigma_1)$ and $\alpha_1 \in U(\sigma_0)$.

Claim 1: $\alpha_0 \in U(\sigma_1)$. *Proof:* Assume to the contrary that there is a morphism $\tau \neq \sigma_1$ with $\tau(\alpha_0) = \sigma_1(\alpha_0)$. Then $\tau(1) \neq \lambda$ must hold; otherwise, this would result in $\tau(\alpha_0) = \tau(2 \cdot 3 \cdot 3 \cdot 2) = \sigma_1(1 \cdot 2 \cdot 3 \cdot 1 \cdot 3 \cdot 2) = \mathbf{abaab^2aab^3aabaab^3aab^2a}$ – but such a τ cannot exist. Likewise, if $\tau(1) = \sigma_1(1)$, then $\tau(2 \cdot 3) = \sigma_1(2 \cdot 3)$ and $\tau(3 \cdot 2) = \sigma_1(3 \cdot 2)$. By exhausting all possible decompositions of $\sigma_1(2 \cdot 3)$, we can easily see that this implies $\tau(2) = \sigma_1(2)$ and $\tau(3) = \sigma_1(3)$, which contradicts $\tau \neq \sigma_1$. Furthermore, we observe that $\tau(1)$ must not contain the whole word $\sigma_1(1 \cdot 2) = \mathbf{abaabba}$, as this would contradict $\tau(\alpha_0) = \sigma_1(\alpha_0)$.

Thus, two possibilities remain: First, $\tau(1) = \mathbf{ab}^m$ with $m \in \{0, 1\}$, and second, $\tau(1) = \mathbf{abaab}^n$ with $n \in \{0, 1, 2\}$. Before we go into details, recall that in both cases $\tau(1 \cdot 2 \cdot 3) = \sigma_1(1 \cdot 2 \cdot 3)$ and $\tau(1 \cdot 3 \cdot 2) = \sigma_1(1 \cdot 3 \cdot 2)$ must hold. Thus, if $\tau(1) = \mathbf{ab}^m$, we obtain $\tau(2 \cdot 3) = \mathbf{b}^{1-n}\mathbf{a}\sigma_1(2 \cdot 3)$ and $\tau(3 \cdot 2) = \mathbf{b}^{1-n}\mathbf{a}\sigma_1(3 \cdot 2)$. However, as there is no cyclic permutation mapping $\tau(2 \cdot 3)$ on $\tau(3 \cdot 2)$, this is impossible. Likewise, $\tau(1) = \mathbf{abaab}^n$ leads to $\tau(2 \cdot 3) = \mathbf{b}^{2-n}\mathbf{aab}^3\mathbf{a}$ and $\tau(3 \cdot 2) = \mathbf{b}^{3-n}\mathbf{aab}^2\mathbf{a}$. Again, no appropriate cyclic permutation exists. Therefore, σ_1 must be unambiguous on α_0 .

Claim 2: $\alpha_1 \in U(\sigma_0)$. *Proof:* Assume to the contrary that there is a morphism $\tau \neq \sigma_0$ with $\tau(\alpha_1) = \sigma_0(\alpha_1)$. If $\tau(1) = \lambda$, then $\tau(\alpha_1) = \tau(2 \cdot 2 \cdot 3 \cdot 3)$; but now it is impossible to obtain $\tau(\alpha_1) = \sigma_0(\alpha_1)$, because $\sigma_0(\alpha_1)$ does not consist of two subsequent squares. As 1 is the first and the last variable of α_1 , either $\tau(1) = \mathbf{ab}$ or $\tau(1) = \mathbf{ab}w\mathbf{ab}$ for some $w \in \{\mathbf{a}, \mathbf{b}\}^*$.

If $\tau(1) = \mathbf{ab} = \sigma_0(1)$, we obtain $\tau(\alpha_1) = \mathbf{ab}\tau(2 \cdot 2 \cdot 3)\mathbf{abab}\tau(3)\mathbf{ab}$. Now, the first of the two inner occurrences of \mathbf{ab} has to correspond to the σ_0 -image of one of the two inner occurrences of 1. If it corresponds to the first, then $\tau(3) = \sigma_0(3)$, and $\tau(2) = \sigma_0(2)$ follows immediately, which contradicts $\tau \neq \sigma_0$. But if it corresponds to the second, then $\tau(3) = \mathbf{bb}$, which leads to $\tau(\alpha_1) = \mathbf{ab}\tau(2 \cdot 2)\mathbf{bbabab}^3\mathbf{ab}$. As $\sigma_0(\alpha_1)$ ends with $\mathbf{ababab}^3\mathbf{ab}$, this contradicts $\tau(\alpha_1) = \sigma_0(\alpha_1)$. Finally, if $\tau(1) = \mathbf{ab}w\mathbf{ab}$ for some $w \in \{\mathbf{a}, \mathbf{b}\}^*$, we obtain

$$\tau(\alpha_1) = \mathbf{ab}w\mathbf{ab}\tau(2 \cdot 2 \cdot 3)\mathbf{ab}w\mathbf{abab}w\mathbf{ab}\tau(3)\mathbf{ab}w\mathbf{ab}.$$

By comparing the number of occurrences of \mathbf{a} in this representation of $\tau(\alpha_1)$ and in $\sigma_0(\alpha_1)$, we observe that none of w , $\tau(2)$, $\tau(3)$ can contain any \mathbf{a} . It is easy to see that in this case, $\tau(\alpha_1) = \sigma_0(\alpha_1)$ is impossible. \square

Note that this is the first result that directly contradicts Hypothesis 1. Furthermore, the fact that $U(\sigma_0)$ and $U(\sigma_1)$ can be separated by two very short examples might be considered evidence that the two languages are by far not as similar as the two morphisms.

We proceed with a comparison of $U(\sigma_0)$ and $U(\sigma_2)$:

Theorem 17. *The sets $U(\sigma_0)$ and $U(\sigma_2)$ are incomparable.*

PROOF. We use the patterns $\alpha_0 \in U(\sigma_2) \setminus U(\sigma_0)$ and $\alpha_{0 \setminus 2} \in U(\sigma_0) \setminus U(\sigma_2)$.

Claim 1: $\alpha_0 \in U(\sigma_2) \setminus U(\sigma_0)$. *Proof:* The ambiguity of σ_0 on α_0 is established in the proof of Theorem 15; the proof for $\alpha_0 \in U(\sigma_2)$ is very similar to the one for $\alpha_0 \in U(\sigma_1)$ in the proof of Theorem 16. Assume there is a morphism $\tau \neq \sigma_2$ such that $\tau(\alpha_0) = \sigma_2(\alpha_0)$. As in the proof for $\alpha_0 \in U(\sigma_1)$, $\tau(1) = \lambda$ and $\tau(1) = \sigma_2(1)$ contradict $\tau(\alpha_0) = \sigma_2(\alpha_0)$ and $\tau \neq \sigma_2$, respectively. Following the same reasoning as before, we realise that $\tau(1)$ must be a proper prefix of $\sigma_2(1) \mathbf{ab}^3 \mathbf{a}$, which can be summarised using the following cases:

1. $\tau(1) = \sigma_2(1) \mathbf{ab}^n$ for some $n \in \{0, 1, 2, 3\}$,
2. $\tau(1) = \mathbf{aba} \mathbf{ab}^n$ for some $n \in \{0, 1, 2\}$,
3. $\tau(1) = \mathbf{aba}$,
4. $\tau(1) = \mathbf{ab}^n$ for some $n \in \{0, 1\}$.

Again, we know that $\tau(1 \cdot 2 \cdot 3) = \sigma_2(1 \cdot 2 \cdot 3)$ and $\tau(1 \cdot 3 \cdot 2) = \sigma_2(1 \cdot 3 \cdot 2)$. Thus, we obtain the following pairs of equations, numbered according to the cases they are derived from:

$$\begin{aligned}
(1) \quad \tau(2 \cdot 3) &= \mathbf{b}^{3-n} \mathbf{a} \mathbf{ab}^4 \mathbf{a} \sigma_2(3), & \tau(3 \cdot 2) &= \mathbf{b}^{5-n} \mathbf{a} \mathbf{ab}^6 \mathbf{a} \sigma_2(2), \\
(2) \quad \tau(2 \cdot 3) &= \mathbf{b}^{2-n} \mathbf{a} \sigma_2(2 \cdot 3), & \tau(3 \cdot 2) &= \mathbf{b}^{2-n} \mathbf{a} \sigma_2(3 \cdot 2), \\
(3) \quad \tau(2 \cdot 3) &= \mathbf{ab}^2 \mathbf{a} \sigma_2(2 \cdot 3), & \tau(3 \cdot 2) &= \mathbf{ab}^2 \mathbf{a} \sigma_2(3 \cdot 2), \\
(4) \quad \tau(2 \cdot 3) &= \mathbf{b}^{1-n} \mathbf{a} \mathbf{ab}^2 \mathbf{a} \sigma_2(2 \cdot 3), & \tau(3 \cdot 2) &= \mathbf{b}^{1-n} \mathbf{a} \mathbf{ab}^2 \mathbf{a} \sigma_2(3 \cdot 2).
\end{aligned}$$

In each of the four cases, there does not exist a cyclic permutation mapping $\tau(2 \cdot 3)$ to $\tau(3 \cdot 2)$. Therefore, σ_2 is unambiguous on α_0 .

Claim 2: $\alpha_{0 \setminus 2} \notin U(\sigma_2)$. *Proof:* We refer to the morphism τ given by

$$\begin{aligned}
\tau(1) &:= \sigma_2(1) \mathbf{ab}^2, & \tau(2) &:= \mathbf{ba} \mathbf{ab}^2, \\
\tau(3) &:= \mathbf{b}^2 \mathbf{a} \sigma_2(3), & \tau(4) &:= \lambda, \\
\tau(5) &:= \sigma_2(4 \cdot 5 \cdot 4), & \tau(6) &:= \mathbf{a}, \\
\tau(7) &:= \mathbf{b}^{11} \mathbf{a} \mathbf{ab}^{12}, & \tau(8) &:= \sigma_2(7 \cdot 6 \cdot 8), \\
\tau(9) &:= \sigma_2(9) \mathbf{ab}^{11} \mathbf{a} \mathbf{ab}^{12}, & \tau(10) &:= \mathbf{b}^{11} \mathbf{a} \mathbf{ab}^{12} \mathbf{a} \sigma_2(10), \\
\tau(11) &:= \sigma_2(11) \mathbf{ab}^{23} \mathbf{a} \mathbf{ab}^{12}, & \tau(12) &:= \mathbf{b}^{12} \mathbf{a}, \\
\tau(13) &:= \sigma_2(13 \cdot 7 \cdot 7 \cdot 4) \mathbf{ab}^{27} \mathbf{a} \mathbf{ab}^{13}, & \tau(14) &:= \lambda, \\
\tau(15) &:= \sigma_2(15 \cdot 14), & \tau(16) &:= \sigma_2(16).
\end{aligned}$$

Since the proof of $\tau(\alpha_{0 \setminus 2}) = \sigma_2(\alpha_{0 \setminus 2})$ is rather lengthy but straightforward, we do not give it as extensively as for $\tau(\alpha_2) = \sigma_2(\alpha_2)$ in the proof of Theorem 15, but leave it to the reader.

Claim 3: $\alpha_{0 \setminus 2} \in U(\sigma_0)$. *Proof:* Assume there exists a morphism $\tau \neq \sigma_0$ with $\tau(\alpha_{0 \setminus 2}) = \sigma_0(\alpha_{0 \setminus 2})$. First, we direct our attention to $\tau(16)$. If $\tau(16) = \lambda$, we obtain

$$\tau(\alpha_{0 \setminus 2}) = (\tau(\beta_1 \cdot \beta_2 \cdot \dots \cdot \beta_9))^2,$$

but as $\sigma_0(\alpha_{0 \setminus 2})$ is not a square, this contradicts $\tau(\alpha_{0 \setminus 2}) = \sigma_0(\alpha_{0 \setminus 2})$. Moreover, it is easily seen that $\tau(16) = \sigma_0(16)$ is the only possibility that does not contradict

$\tau(\alpha_{0\setminus 2}) = \sigma_0(\alpha_{0\setminus 2})$. But then every occurrence of 16 generates the same part of the whole image under both morphisms, which allows us to conclude $\tau(\beta_i) = \sigma_0(\beta_i)$ for all $i \in \{1, \dots, 9\}$.

Next, we focus on $\tau(6)$. Due to $\tau(\beta_9) = \sigma_0(\beta_9)$ and $\tau(\beta_5) = \sigma_0(\beta_5)$, we are able to infer

$$\tau(9 \cdot 6) = \sigma_0(9 \cdot 6) \quad (1)$$

and

$$\tau(6 \cdot 10) = \sigma_0(6 \cdot 10). \quad (2)$$

First, assume $\tau(6) = \lambda$. Together with $\tau(\beta_3) = \sigma_0(\beta_3)$, this leads to

$$\tau(7 \cdot 8) = \sigma_0(6 \cdot 7 \cdot 6 \cdot 8). \quad (3)$$

If $\tau(7) = \lambda$, then $\tau(\beta_4) = \sigma_0(\beta_4)$ enforces $\tau(1 \cdot 3) = \sigma_0(1 \cdot 7 \cdot 3)$. But then

$$|\tau(\beta_1)| \geq |\tau(1 \cdot 3)| = |\sigma_0(\beta_4)| > |\sigma_0(\beta_1)|,$$

which contradicts $\sigma_0(\beta_1) = \tau(\beta_1)$. Thus, $\tau(7) \neq \lambda$ must hold. Due to (3), $\tau(7) = \mathbf{a}w$ for some $w \in \{\mathbf{a}, \mathbf{b}\}^*$. But as 7 occurs squared in β_7 , and 6 does not occur in β_7 at all, $w = \mathbf{b}^n$ for some $n \leq 6$. Then $\tau(7 \cdot 7) = \mathbf{a}\mathbf{b}^n \mathbf{a}\mathbf{b}^n$, which leads to $n \in V(\beta_7)$, and therefore $n = 4$. Using this on $\tau(\beta_4) = \sigma_0(\beta_4)$ leads to

$$\tau(3) = \mathbf{b}^3 \sigma_0(3),$$

which contradicts $\tau(\beta_1) = \sigma_0(\beta_1)$. Therefore, $\tau(6) \neq \lambda$; and (1) and (2) leave no other possibility than $\tau(6) = \sigma_0(6)$, which implies $\tau(9) = \sigma_0(9)$ and $\tau(10) = \sigma_0(10)$.

Using $\tau(6) = \sigma_0(6)$, we can derive the following three possibilities for $\tau(7)$ and $\tau(8)$ from β_3 :

$$\begin{aligned} (1) \quad \tau(7) &= \lambda, & \tau(8) &= \mathbf{b} \sigma_0(6 \cdot 8), \\ (2) \quad \tau(7) &= \sigma_0(7 \cdot 6), & \tau(8) &= \mathbf{b}^2, \\ (3) \quad \tau(7) &= \sigma_0(7), & \tau(8) &= \sigma_0(8). \end{aligned}$$

In the first case, using $\tau(7) = \lambda$ on $\tau(\beta_4) = \sigma_0(\beta_4)$ leads to

$$\tau(1 \cdot 3) = \sigma_0(\beta_4).$$

By applying this to $\tau(\beta_1) = \sigma_0(\beta_1)$, we obtain the contradiction

$$|\sigma_0(\beta_1)| = |\tau(\beta_1)| \geq |\tau(1 \cdot 3)| = |\sigma_0(\beta_4)| > |\sigma_0(\beta_1)|.$$

In the second case, $\tau(7) = \sigma_0(7 \cdot 6)$ is inconsistent with $\tau(\beta_7) = \sigma_0(\beta_7)$. Thus, the third case holds. Using $\tau(7) = \sigma_0(7)$ on $\tau(\beta_7) = \sigma_0(\beta_7)$ provides us with $\tau(13) = \sigma_0(13)$ and

$$\tau(4 \cdot 14 \cdot 12) = \sigma_0(4 \cdot 14 \cdot 12). \quad (4)$$

If $\tau(4) = \lambda$, then $\tau(14 \cdot 12) = \sigma_0(4 \cdot 14 \cdot 12)$ must hold. But then τ cannot satisfy both $\tau(\beta_6) = \sigma_0(\beta_6)$ and $\tau(\beta_8) = \sigma_0(\beta_8)$. Moreover, since $\tau(\beta_2) = \sigma_0(\beta_2)$, we

deduce $\tau(4) = \sigma(4)$ and also $\tau(5) = \sigma_0(5)$. The former allows us to shorten (4) to $\tau(14 \cdot 12) = \sigma_0(14 \cdot 12)$. This, together with $\tau(\beta_6) = \sigma_0(\beta_6)$ and $\tau(\beta_8) = \sigma_0(\beta_8)$, leads to $\tau(x) = \sigma_0(x)$ for $x \in \{11, 12, 14, 15\}$.

Finally, $\tau(7) = \sigma_0(7)$ and $\tau(\beta_4) = \sigma_0(\beta_4)$ imply $\tau(1) = \sigma_0(1)$ and $\tau(3) = \sigma_0(3)$; and applying this to $\tau(\beta_1) = \sigma_0(\beta_1)$ gives $\tau(2) = \sigma_0(2)$. Therefore, $\tau(x) = \sigma_0(x)$ for all $x \in V(\alpha_{0 \setminus 2})$, which contradicts $\tau \neq \sigma_0$. Thus, σ_0 is unambiguous on $\alpha_{0 \setminus 2}$. \square

Note that $\alpha_{0 \setminus 2}$ is not the shortest known pattern in $U(\sigma_0) \setminus U(\sigma_2)$. The similar pattern $\alpha_{0 \setminus 2}$ described by Freydenberger and Reidenbach [5] (which has a length of 53 and 15 different variables) is also contained in this set, but the present version allows the proof to be more concise.

Although Theorem 16 and Theorem 17 have already demonstrated that Hypothesis 1 does not hold when comparing $U(\sigma_0)$ to $U(\sigma_1)$ or $U(\sigma_2)$, one might still conjecture that this phenomenon relies on the fact that σ_0 is not a segmented morphism, or that the larger number of segments in σ_2 make this morphism “less ambiguous” than σ_1 . But our most sophisticated example pattern $\alpha_{1 \setminus 2}$ shows that this assertion is not correct:

Theorem 18. *The sets $U(\sigma_1)$ and $U(\sigma_2)$ are incomparable.*

PROOF. For this proof we use the patterns α_1 and $\alpha_{1 \setminus 2}$.

Claim 1: $\alpha_1 \in U(\sigma_2) \setminus U(\sigma_1)$. *Proof:* Recall that σ_1 is ambiguous on α_1 (cf. proof of Theorem 15). Assume to the contrary that there is a morphism $\tau \neq \sigma_2$ such that $\tau(\alpha_1) = \sigma_2(\alpha_1)$. Then $\tau(1) \neq \lambda$; otherwise, $\tau(\alpha_1) = \tau(2 \cdot 2 \cdot 3 \cdot 3) \neq \sigma_2(\alpha_1)$ would lead to a contradiction, as $\sigma_2(\alpha_1)$ does not consist of two subsequent squares. As α_1 starts and ends with 1, $\tau(1)$ must start and end with **a**. But as α_1 additionally contains the factor $1 \cdot 1$, only the alternatives $\tau(1) = \mathbf{a}$ and $\tau(1) = \sigma_2(1)$ remain as possible solutions. If $\tau(1) = \sigma_2(1)$, the factor $1 \cdot 3 \cdot 1$ implies $\tau(3) = \sigma_2(3)$, and thus $\tau(2) = \sigma_2(2)$. This contradicts $\tau \neq \sigma_2$. Therefore, $\tau(1) = \mathbf{a}$ must hold.

There must be at least one variable $x \in \{2, 3\}$ such that $\tau(x) = u \sigma_2(x) v$ for some $u, v \in \{\mathbf{a}, \mathbf{b}\}^*$ with $uv \neq \lambda$. If $x = 2$, then $\tau(\alpha_1)$ contains the factor $\tau(2 \cdot 2) = u \sigma_2(x) v u \sigma_2(x) v$. Due to $uv \neq \lambda$, this contradicts $\tau(\alpha_1) = \sigma_2(\alpha_1)$. Thus, $\tau(3) = u \sigma_2(3) v$ for some $u, v \in \{\mathbf{a}, \mathbf{b}\}^*$ with $uv \neq \lambda$ and

$$\begin{aligned} \tau(\alpha_1) &= \tau(1 \cdot 2 \cdot 2 \cdot 3 \cdot 1 \cdot 1 \cdot 3 \cdot 1) \\ &= \mathbf{a} \tau(2 \cdot 2) u \sigma_2(3) v \mathbf{a} \mathbf{a} u \sigma_2(3) v \mathbf{a}. \end{aligned}$$

As $\tau(\alpha_1) = \sigma_2(\alpha_1)$, $v \mathbf{a} = \sigma_2(1)$ and consequently $\mathbf{a} u = \sigma_2(1)$. This leads to

$$\tau(\alpha_1) = \mathbf{a} \tau(2 \cdot 2) u \sigma_2(3) \sigma_2(1) \sigma_2(1) \sigma_2(3) \sigma_2(1)$$

and thus, $\mathbf{a} \tau(2 \cdot 2) u = \sigma_2(1 \cdot 2 \cdot 2)$. This contradicts $u = \mathbf{b} \mathbf{a} \mathbf{b}^2 \mathbf{a}$; thus, we know that $\alpha_1 \in U(\sigma_2) \setminus U(\sigma_1)$.

Claim 2: $\alpha_{1 \setminus 2} \in U(\sigma_1)$. *Proof:* Assume to the contrary that there is a morphism $\tau \neq \sigma_1$ such that $\tau(\alpha_{1 \setminus 2}) = \sigma_1(\alpha_{1 \setminus 2})$. First, observe the following two facts that hold for all $x \in V(\alpha_{1 \setminus 2})$:

Fact 1: If xx is a factor of $\alpha_{1\setminus 2}$, neither $\tau(x) \in \mathbf{ab}^+$, nor $\tau(x) \in \mathbf{b}^+\mathbf{a}$.

Fact 2: If xx is a factor of $\alpha_{1\setminus 2}$ and $\tau(x)$ contains $\sigma_1(x)$, then $\tau(x) = \sigma_1(x)$.

Otherwise, we immediately obtain a contradiction to $\tau(\alpha_{1\setminus 2}) = \sigma_1(\alpha_{1\setminus 2})$.

As those facts are of vital importance, take special note that the variables 1, 2, 3, 20 and 23 occur squared in $\alpha_{1\setminus 2}$.

Now we focus on $\tau(1)$. Assume $\tau(1) = \lambda$, then

$$\tau(\alpha_{1\setminus 2}) = \tau(1^2 \cdot \delta \cdot 1 \cdot p(\delta) \cdot 1) = \tau(p(\delta) \cdot p(\delta)) = \tau(p(\delta))\tau(p(\delta)).$$

As $\sigma_1(\alpha_{1\setminus 2})$ is obviously not a square, $\tau(1) \neq \lambda$. As 1 is the first and the last variable of $\alpha_{1\setminus 2}$, $\tau(\alpha_{1\setminus 2})$ must begin and start with the letter \mathbf{a} and, as 1 is also the second letter of $\alpha_{1\setminus 2}$, $|\tau(1)| > 1$. Due to Fact 1 and Fact 2, we obtain $\tau(1) = \sigma_1(1)$ and therefore, $\tau(x) = \sigma_1(x)$ for all $x \in V(\gamma_1) \cup V(\gamma_2) \cup V(\gamma_3)$ and $\tau(\delta_i) = \sigma_1(\delta_i)$ for all $i \in \{1, 2, \dots, 12\}$.

Next, we show that $\tau(2)$ must be empty. Assume the opposite; then, by the same reasoning as for 1, $\tau(2) = \sigma_1(2)$ must hold, since all other possibilities lead to contradictions by a careful inspection of δ_1 . Indeed, due to Fact 2, $\tau(2)$ cannot be longer than $\sigma_1(2)$, and due to Fact 1, it cannot be of the form \mathbf{ab}^+ . If $\tau(2) = \mathbf{a}$, a look at δ_2 and Facts 1 and 2 show that no consistent $\tau(3)$ can be found. Therefore, $\tau(2) = \sigma_1(2)$ must hold. But this immediately implies $\tau(x) = \sigma_1(x)$ for $x \in \{3, 20, 23\}$, as those variables are surrounded by occurrences of the variable 1 or 2 in at least one position in $\alpha_{1\setminus 2}$, and thus (with some straightforward effort) the same equality holds for all $x \in V(\alpha_{1\setminus 2}) \setminus \{18, 19, 31, 32\}$. If $\tau(31) = \sigma_1(31)$, then $\tau = \sigma_1$. Using δ_{10} , this leads to $\tau(31) = \lambda$ and $\tau(32) = \sigma_1(31^4 \cdot 32)$. This contradicts $\tau(\delta_{14}) = \sigma_1(\delta_{14})$. We conclude $\tau(2) = \lambda$.

Likewise, we show that $\tau(3) = \lambda$. If $\tau(3) \neq \lambda$, then $\tau(3) = \sigma_1(3)$. With δ_2 it follows that $\tau(2) = \sigma_1(2) \neq \lambda$, therefore $\tau(3) = \lambda$.

Finally, we consider the different possibilities for $\tau(6)$. With respect to δ_3 we observe the following possibilities:

1. $\tau(6) = \lambda$,
2. $\tau(6) = \mathbf{ab}^k$, where $0 \leq k \leq 6$,
3. $\tau(6) = \sigma_1(6)$,
4. $\tau(6) = \sigma_1(6) \mathbf{ab}^k$, where $0 \leq k \leq 7$.

As 6 occurs also in the factor $6 \cdot 20$ in δ_7 and in $6 \cdot 23$ in δ_8 , $\tau(6)$ cannot completely contain $\sigma_1(7)$. Analogously, 8 and 10 in δ_4 and δ_5 limit how far $\tau(9)$ and $\tau(11)$ can extend to the left; a fact that is useful in our reasoning below.

Case 1: If $\tau(6) = \lambda$, then $\tau(20)$ contains all of $\sigma_1(6)$, as due to δ_4 , $\tau(9)$ is limited by $\sigma_1(8)$ and therefore cannot acquire more than the nine rightmost letters of $\sigma_1(20)$. But this implies that $\sigma_1(6)$ occurs in $\tau(\alpha_{1\setminus 2})$ at least as often as 20 occurs in $\alpha_{1\setminus 2}$. This contradicts $\sigma_1(\alpha_{1\setminus 2}) = \tau(\alpha_{1\setminus 2})$.

Case 2: Assume $\tau(6) = \mathbf{ab}^k$, where $0 \leq k \leq 6$. Together with to Fact 2, δ_7 and δ_8 lead to $|\tau(9)| > |\sigma_1(9)|$ and $|\tau(11)| > |\sigma_1(11)|$, respectively. Due to δ_4 , $\tau(9)$ cannot contain all of $\sigma_1(20)$, and due to δ_5 , $\tau(11)$ cannot contain all of $\sigma_1(23)$. Therefore, $\tau(9) = \mathbf{b}^m \mathbf{a} \cdot \sigma_1(9)$ with $0 \leq m \leq 8$ and $\tau(11) = \mathbf{b}^n \mathbf{a} \cdot \sigma_1(11)$

with $0 \leq n \leq 10$; thus $\tau(20) = \mathbf{b}^{8-k} \mathbf{a} \mathbf{a} \mathbf{b}^{20-m}$ and $\tau(23) = \mathbf{b}^{8-k} \mathbf{a} \mathbf{a} \mathbf{b}^{23-n}$. We now consider $\tau(\delta_{10})$:

$$\begin{aligned} \tau(\delta_{10}) &= \tau(30 \cdot 2 \cdot (20 \cdot 2 \cdot 23 \cdot 2)^3 \cdot 20 \cdot (31)^4 \cdot 32) \\ &= \tau(30 \cdot (20 \cdot 23)^3 \cdot 20 \cdot (31)^4 \cdot 32) \\ &= \tau(30) \cdot (\mathbf{b}^{8-k} \mathbf{a} \mathbf{a} \mathbf{b}^{20-m} \cdot \mathbf{b}^{8-k} \mathbf{a} \mathbf{a} \mathbf{b}^{23-n})^3 \cdot \mathbf{b}^{8-k} \mathbf{a} \mathbf{a} \mathbf{b}^{20-m} \cdot \tau((31)^4 \cdot 32) \\ &= \tau(30) \cdot \mathbf{b}^{8-k} \mathbf{a} \cdot (\mathbf{a} \mathbf{b}^{28-k-m} \mathbf{a} \mathbf{a} \mathbf{b}^{31-k-n} \mathbf{a})^3 \cdot \mathbf{a} \mathbf{b}^{20-m} \cdot \tau((31)^4 \cdot 32). \end{aligned}$$

However, there are no $x, y \in V(\delta_{10})$ such that $(xy)^3$ is a factor of δ_{10} . This contradicts the assumption $\tau(\alpha_{1 \setminus 2}) = \sigma_1(\alpha_{1 \setminus 2})$.

Case 3: We consider the case $\tau(6) = \sigma_1(6)$ and look at the possibilities for $\tau(20)$ with special attention to δ_7 . Clearly, $\tau(20)$ cannot contain more than $\sigma_1(20)$, as this would contradict Fact 2. If $\tau(20) = \sigma_1(20)$, then the factor $20 \cdot 2 \cdot 20$ in δ_9 and $\tau(2) = \lambda$ is inconsistent with $\tau(\alpha_{1 \setminus 2}) = \sigma_1(\alpha_{1 \setminus 2})$. But $\tau(20) = \mathbf{a} \mathbf{b}^{20-m}$ for some $m \leq 8$ is prohibited by Fact 1, as 20 occurs as a square in δ_{12} . Finally, as already mentioned above, $\tau(20) = \lambda$ is impossible due to δ_3, δ_4 and δ_7 . Hence, $\tau(6) = \sigma_1(6)$ would not leave any possibility of defining $\tau(20)$, which contradicts the existence of τ .

Case 4: Assume $\tau(6) = \sigma_1(6) \cdot \mathbf{a} \mathbf{b}^k$ for some $0 \leq k \leq 7$. Now, considering Fact 1 together with δ_7 and δ_8 , respectively, we see that each of $\tau(20)$ and $\tau(23)$ is either of the form $\mathbf{b}^+ \mathbf{a} \mathbf{a} \mathbf{b}^*$ or \mathbf{b}^* . We first consider the case that $\tau(20) \in \mathbf{b}^+ \mathbf{a} \mathbf{a} \mathbf{b}^*$, i. e., $\tau(20) = \mathbf{b}^{20-k} \mathbf{a} \mathbf{a} \mathbf{b}^m$ and $\tau(9) = \mathbf{b}^{9-m} \mathbf{a}$ for some $m \leq 9$. Applying this to δ_9 , we obtain:

$$\begin{aligned} \tau(\delta_9) &= \tau(27 \cdot 2 \cdot 20 \cdot 2 \cdot 20 \cdot 28 \cdot 2 \cdot 29) \\ &= \tau(27) \cdot \tau(20 \cdot 20) \cdot \tau(28 \cdot 29) \\ &= \tau(27) \cdot \mathbf{b}^{20-k} \mathbf{a} \mathbf{a} \mathbf{b}^m \cdot \mathbf{b}^{20-k} \mathbf{a} \mathbf{a} \mathbf{b}^m \cdot \tau(28 \cdot 29) \\ &= \tau(27) \cdot \mathbf{b}^{20-k} \mathbf{a} \cdot \mathbf{a} \mathbf{b}^{20+m-k} \mathbf{a} \cdot \mathbf{a} \mathbf{b}^m \cdot \tau(28 \cdot 29). \end{aligned}$$

As $20 + m - k \geq 13$ and $20 - k \geq 13$, the full segment $\mathbf{a} \mathbf{b}^{20+m-k} \mathbf{a}$ in $\tau(20 \cdot 20)$ must be identical to $\sigma_1(28)$, as $\sigma_1(\delta_9)$ allows for no other match. Therefore $20 + m - k = 28$ and thus $m = k + 8$. But this implies $m \geq 8$ and contradicts the fact that $\mathbf{a} \mathbf{b}^m$ must be a factor of $\sigma_1(2)$. Therefore, $\tau(20) \in \mathbf{b}^*$ must hold, which leads to $\tau(20) = \mathbf{b}^{20-k-m}$ and $\tau(9) = \mathbf{b}^m \mathbf{a} \cdot \sigma_1(9)$ for some $m \leq 8$.

Now, assume $\tau(23) = \mathbf{b}^{23-k} \mathbf{a} \mathbf{a} \mathbf{b}^n$ and $\tau(11) = \mathbf{b}^{11-n} \mathbf{a}$ for some $n \leq 11$. Applying this to δ_{11} , we obtain:

$$\begin{aligned} \tau(\delta_{11}) &= \tau(33 \cdot 3 \cdot 34 \cdot 23 \cdot 3 \cdot 23 \cdot 3 \cdot 35) \\ &= \tau(33 \cdot 34) \cdot \mathbf{b}^{23-k} \mathbf{a} \cdot \mathbf{a} \mathbf{b}^{23+n-k} \mathbf{a} \cdot \mathbf{a} \mathbf{b}^n \cdot \tau(35). \end{aligned}$$

As both $23 - k$ and $23 + n - k$ are at least 16, the left part of $\tau(23 \cdot 23)$ must be a part of $\sigma_1(34)$, the right part a part of $\sigma_1(3)$ and the middle part must equal $\sigma_2(23)$. Therefore, $23 + n - k = 23$ and thus $n = k$ and $k \leq 3$. This leads to the following equation:

$$\tau(\delta_{10}) = \tau(30 \cdot 2 \cdot (20 \cdot 2 \cdot 23 \cdot 2)^3 \cdot 20 \cdot (31)^4 \cdot 32)$$

$$\begin{aligned}
&= \tau(30) \cdot \tau((20 \cdot 23)^3 \cdot 20) \cdot \tau((31)^4 \cdot 32) \\
&= \tau(30) \cdot (\mathbf{b}^{20-k-m} \cdot \mathbf{b}^{23-k} \mathbf{a} \mathbf{a} \mathbf{b}^k)^3 \cdot \mathbf{b}^{20-k-m} \cdot \tau((31)^4 \cdot 32) \\
&= \tau(30) \cdot \mathbf{b}^{43-2k-m} \mathbf{a} \mathbf{a} \mathbf{b}^k \cdot (\mathbf{b}^{43-2k-m} \mathbf{a} \mathbf{a} \mathbf{b}^k)^2 \cdot \mathbf{b}^{20-k-m} \cdot \tau((31)^4 \cdot 32) \\
&= \tau(30) \cdot \mathbf{b}^{43-2k-m} \mathbf{a} \cdot (\mathbf{a} \mathbf{b}^{43-k-m} \mathbf{a})^2 \cdot \mathbf{a} \mathbf{b}^{20-m} \cdot \tau((31)^4 \cdot 32).
\end{aligned}$$

As no other letter occurs squared in δ_{10} , we have $43 - k - m = 31$ and thus $k + m = 12$. As $m \leq 8$, we obtain $k \geq 4$, but due to the abovementioned considerations for δ_{11} we know that $k \leq 3$ must hold. This is a contradiction.

Finally, consider the case $\tau(23) = \mathbf{b}^{23-k-n}$ with $n \leq 10$. Now $\tau(\delta_{10})$ contains a factor $\tau((20 \cdot 23)^3 \cdot 20) = \mathbf{b}^l$ with $l = 4(20 - k - m) + 3(23 - k - n)$. Recall that $k \leq 7$, $m \leq 8$ and $n \leq 10$ and observe that $l \geq 4(20 - 15) + 3(23 - 17) = 38$. Thus, \mathbf{b}^{38} is a factor of $\tau(\delta_{10})$, but it is obviously not a factor of $\sigma_1(\delta_{10})$. This contradicts $\tau(\alpha_{1 \setminus 2}) = \sigma_1(\alpha_{1 \setminus 2})$. Therefore, σ_1 is unambiguous on $\alpha_{1 \setminus 2}$.

Claim 3: $\alpha_{1 \setminus 2} \notin U(\sigma_2)$. *Proof:* Let $\tau(x) := \lambda$ for $x \in \{2, 3, 28, 31, 34\}$ and $\tau(x) := \sigma_2(x)$ for $x \in V(\gamma_1) \cup V(\gamma_2) \cup V(\gamma_3)$. For all other $x \in V(\alpha_{1 \setminus 2})$, define $\tau(x)$ as follows:

$$\begin{aligned}
\tau(4) &:= \sigma_2(2 \cdot 3 \cdot 3 \cdot 4), & \tau(5) &:= \sigma_2(3 \cdot 2 \cdot 2 \cdot 5), \\
\tau(6) &:= \sigma_2(6) \mathbf{a} \mathbf{b}^{11}, & \tau(7) &:= \mathbf{b}^2 \mathbf{a} \mathbf{a} \mathbf{b}^{14} \mathbf{a}, \\
\tau(8) &:= \mathbf{a} \mathbf{b}^{15} \mathbf{a} \mathbf{a} \mathbf{b}^3, & \tau(9) &:= \mathbf{b}^{13} \mathbf{a} \sigma_2(9), \\
\tau(10) &:= \mathbf{a} \mathbf{b}^{19} \mathbf{a} \mathbf{a} \mathbf{b}^8, & \tau(11) &:= \mathbf{b}^{12} \mathbf{a} \sigma_2(11), \\
\tau(18) &:= \sigma_2(18) \mathbf{a} \mathbf{b}^{27}, & \tau(19) &:= \mathbf{b}^{10} \mathbf{a} \mathbf{a} \mathbf{b}^{38} \mathbf{a}, \\
\tau(20) &:= \mathbf{b}^{28} \mathbf{a} \mathbf{a} \mathbf{b}^{27}, & \tau(23) &:= \mathbf{b}^{34} \mathbf{a} \mathbf{a} \mathbf{b}^{34}, \\
\tau(27) &:= \sigma_2(27 \cdot 2 \cdot 20 \cdot 2) \mathbf{a} \mathbf{b}^{39} \mathbf{a} \mathbf{a} \mathbf{b}^{12}, & \tau(29) &:= \mathbf{b}^{29} \mathbf{a} \sigma_2(2 \cdot 29), \\
\tau(30) &:= \sigma_2(30 \cdot 2 \cdot (20 \cdot 2 \cdot 23 \cdot 2)^3) \mathbf{a} \mathbf{b}^{39} \mathbf{a} \mathbf{a} \mathbf{b}^{12}, & \tau(32) &:= \mathbf{b}^{34} \mathbf{a} \mathbf{a} \mathbf{b}^{62} \mathbf{a} \cdot \sigma_2(32), \\
\tau(33) &:= \sigma_2(33 \cdot 3) \mathbf{a} \mathbf{b}^{33}, & \tau(35) &:= w \sigma_2(3 \cdot 23 \cdot 3 \cdot 35), \\
\tau(36) &:= \sigma_2(36 \cdot 20) \mathbf{a} \mathbf{b}^{39} \mathbf{a} \mathbf{a} \mathbf{b}^{12}, & \tau(37) &:= w \sigma_2(23 \cdot 37),
\end{aligned}$$

where $w := \mathbf{b}^{11} \mathbf{a} \mathbf{a} \mathbf{b}^{46} \mathbf{a}$. Obviously $\tau \neq \sigma_2$. As $\tau(x) = \sigma_2(x)$ for $x \in V(\gamma_1) \cup V(\gamma_2) \cup V(\gamma_3)$, especially for $x = 1$, it suffices to show $\tau(\delta_i) = \sigma_2(\delta_i)$ for all $i \in \{1, 2, \dots, 14\}$. For δ_1 and δ_2 , the claim holds trivially. For the other δ_i , the process is straightforward, but somewhat lengthy.

Thus, $\alpha_{1 \setminus 2} \in U(\sigma_1) \setminus U(\sigma_2)$, and therefore $U(\sigma_1)$ and $U(\sigma_2)$ are incomparable. \square

Having a length of 177 and 37 different variables, $\alpha_{1 \setminus 2}$ is not the shortest example separating $U(\sigma_1)$ from $U(\sigma_2)$. First of all, one could remove the factors γ_i and their images under p without changing the ambiguity properties. These factors are only present to bring the whole pattern in canonical form and to demonstrate that the construction does not rely on some numbering tricks. Furthermore, the authors firmly believe that although its presence greatly simplifies the proof of the unambiguity of σ_1 , all occurrences of the variable 1 can

be removed from $\alpha_{1\setminus 2}$ without causing ambiguity of σ_1 or unambiguity of σ_2 . Thus, the authors conjecture that the shortest pattern in $U(\sigma_1) \setminus U(\sigma_2)$ has at most length 140 and 25 variables and do not see much room for improvement for any of these two parameters.

Note that we do not know any nontrivial characterisation of $U(\sigma_0)$, $U(\sigma_1)$ and $U(\sigma_2)$. Thus, and due to the NP-completeness of the underlying problem (cf. Ehrenfeucht and Rozenberg [4]) as well as the growth of the search space (cf. Reidenbach and Schneider [19]), we cannot refer to a computationally feasible method to successfully seek for any patterns in $U(\sigma_1) \setminus U(\sigma_2)$, $U(\sigma_0) \setminus U(\sigma_2)$ or $U(\sigma_3) \setminus U(\sigma_2)$. In particular, this means that the patterns α_2 , $\alpha_{0\setminus 2}$ and $\alpha_{1\setminus 2}$ are “handmade” and therefore we cannot answer the question of whether there exist shorter examples than α_2 , the pattern $\alpha_{0\setminus 2}$ from Freydenberger and Reidenbach [5] (cf. our remark below the proof of Theorem 17) and the modification of $\alpha_{1\setminus 2}$ mentioned above that are suitable for proving Theorems 15, 17 and 18, respectively. The intricacy of the ambiguity phenomena relevant for the construction of such patterns, however, suggests that our examples cannot be shortened significantly.

5. Conclusion and Open Problems

In the present paper, we have studied the unambiguity of an important family of injective morphisms. More precisely, we have examined the impact of the number n of segments of a segmented morphism σ_n on the set $U(\sigma_n)$ of patterns for which σ_n is unambiguous. Our main results show that a change of n , surprisingly, does not give rise to a “real” hierarchy of sets of patterns, as the three pairwise incomparable languages $U(\sigma_0)$, $U(\sigma_1)$ and $U(\sigma_2)$ are all contained in one common superset $U(\sigma_3)$, that is also the maximum any left- or right-homogeneous morphism can achieve. We have established the result on $U(\sigma_3)$ by several characteristic criteria on $U(\sigma_3)$, which additionally entail a substantial improvement of the main technique introduced in the initial paper [6] on the unambiguity of morphisms.

Contrary to this, a major part of our results on σ_0 , σ_1 and σ_2 are not based on criteria, but on example patterns. We regard it as a very interesting problem to find characterisations of $U(\sigma_0)$, $U(\sigma_1)$ and $U(\sigma_2)$. In consideration of the remarkable complexity of the patterns $\alpha_{0\setminus 2}$, $\alpha_{1\setminus 2}$ and α_2 , however, we expect this to be an extraordinarily cumbersome task.

Acknowledgements

We wish to thank the anonymous reviewers for their careful and valuable comments, which, in particular, have led to stronger statements in Proposition 7 and Corollary 10. Furthermore, we thank the referees of the conference version [5] of this paper for their helpful remarks.

References

- [1] K. A. Baker, G. F. McNulty, and W. Taylor. Growth problems for avoidable words. *Theoretical Computer Science*, 69:319–345, 1989.
- [2] J. Cassaigne. Unavoidable patterns. In M. Lothaire, editor, *Algebraic Combinatorics on Words*, volume 90 of *Encyclopedia of Mathematics and Its Applications*, pages 111–134. Cambridge Mathematical Library, 2002.
- [3] C. Choffrut and J. Karhumäki. Combinatorics of words. In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Languages*, volume 1, pages 329–438. Springer, 1997.
- [4] A. Ehrenfeucht and G. Rozenberg. Finding a homomorphism between two words is NP-complete. *Information Processing Letters*, 9:86–88, 1979.
- [5] D.D. Freydenberger and D. Reidenbach. The unambiguity of segmented morphisms. In *Proc. 11th International Conference on Developments in Language Theory, DLT 2007*, volume 4588 of *Lecture Notes in Computer Science*, pages 181–192, 2007.
- [6] D.D. Freydenberger, D. Reidenbach, and J.C. Schneider. Unambiguous morphic images of strings. *International Journal of Foundations of Computer Science*, 17:601–628, 2006.
- [7] V. Halava, T. Harju, J. Karhumäki, and M. Latteux. Extension of the decidability of the marked PCP to instances with unique blocks. *Theoretical Computer Science*, 380:355–362, 2007.
- [8] 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.
- [9] T. Head. Fixed languages and the adult languages of OL schemes. *International Journal of Computer Mathematics*, 10:103–107, 1981.
- [10] T. Jiang, A. Salomaa, K. Salomaa, and S. Yu. Decision problems for patterns. *Journal of Computer and System Sciences*, 50:53–63, 1995.
- [11] F. Levé and G. Richomme. On a conjecture about finite fixed points of morphisms. *Theoretical Computer Science*, 339:103–128, 2005.
- [12] M. Lipponen and G. Păun. Strongly prime PCP words. *Discrete Applied Mathematics*, 63:193–197, 1995.
- [13] A. Mateescu and A. Salomaa. Patterns. In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Languages*, volume 1, pages 230–242. Springer, 1997.

- [14] A.R. Mitchell. Learnability of a subclass of extended pattern languages. In *Proc. 11th Annual Conference on Computational Learning Theory, COLT 1998*, pages 64–71, 1998.
- [15] Y.K. Ng and T. Shinohara. Developments from enquiries into the learnability of the pattern languages from positive data. *Theoretical Computer Science*, 397:150–165, 2008.
- [16] E.L. Post. A variant of a recursively unsolvable problem. *Bulletin of the American Mathematical Society*, 52:264–268, 1946.
- [17] D. Reidenbach. A non-learnable class of E-pattern languages. *Theoretical Computer Science*, 350:91–102, 2006.
- [18] D. Reidenbach. Discontinuities in pattern inference. *Theoretical Computer Science*, 397:166–193, 2008.
- [19] D. Reidenbach and J.C. Schneider. Morphically primitive words. *Theoretical Computer Science*, 410:2148–2161, 2009.
- [20] R. Sedgewick. *Algorithms*. Addison-Wesley, 1983.