Restricted ambiguity of erasing morphisms

Restricted Ambiguity of Erasing Morphisms[☆]

Daniel Reidenbach[a], Johannes C. Schneider[b,*]

[a] *Department of Computer Science, Loughborough University, Loughborough, LE11 3TU, United Kingdom*
[b] *Fachbereich Informatik, Technische Universität Kaiserslautern, Postfach 3049, 67653 Kaiserslautern, Germany*

Abstract

A morphism h is called ambiguous for a string s if there is another morphism that maps s to the same image as h; otherwise, it is called unambiguous. In this paper, we examine some fundamental problems on the ambiguity of erasing morphisms. We provide a detailed analysis of so-called ambiguity partitions, and our main result uses this concept to characterise those strings that have a morphism of strongly restricted ambiguity. Furthermore, we demonstrate that there are strings for which the set of unambiguous morphisms, depending on the size of the target alphabet of these morphisms, is empty, finite or infinite. Finally, we show that the problem of the existence of unambiguous erasing morphisms is equivalent to some basic decision problems for nonerasing multi-pattern languages.

Keywords: Erasing Morphisms, Ambiguity, Pattern Languages

---

1. Introduction

The research on the ambiguity of morphisms is based on the following, elementary questions: Given a string[1] s and a morphism h, do there exist morphisms g with $g(s) = h(s)$, but $g(x) \neq h(x)$ for a symbol x in s? If so, what properties do these morphisms g have? For example, let $s := \mathtt{AABBCC}$, and let the morphism $h : \{\mathtt{A}, \mathtt{B}, \mathtt{C}\}^* \to \{\mathtt{a}, \mathtt{b}\}^*$ be given by $h(\mathtt{A}) := h(\mathtt{C}) := \mathtt{a}$ and $h(\mathtt{B}) := \mathtt{b}$. Then it can be easily verified that there is no morphism g satisfying $g(s) = \mathtt{aabbaa} = h(s)$ and $g(x) \neq h(x)$ for an $x \in \{\mathtt{A}, \mathtt{B}, \mathtt{C}\}$. Therefore, we call h *unambiguous* for s. On the other hand, if we consider the morphism $h' : \{\mathtt{A}, \mathtt{B}, \mathtt{C}\}^* \to \{\mathtt{a}, \mathtt{b}\}^*$, defined by $h'(\mathtt{A}) := h'(\mathtt{B}) := h'(\mathtt{C}) := (\mathtt{ab})^{10}$, then there are various other morphisms g that map s to $h'(s) = (\mathtt{ab})^{60}$. Hence, h' is *ambiguous* for s. Furthermore, for every n with $0 \leq n \leq 30$ and for every symbol $x \in \{\mathtt{A}, \mathtt{B}, \mathtt{C}\}$, there exists at least one morphism g satisfying $g(s) = h'(s)$ and $g(x) = (\mathtt{ab})^n$. Thus, the ambiguity of h' for s is largely unrestricted. In the present paper, we wish to investigate this phenomenon, and we shall mainly focus on the question of whether, for any string, there *exists* a morphism with a restricted ambiguity. To this end, we distinguish between two types of restrictions: maximally restricted ambiguity (i. e., unambiguity) and so-called *moderate ambiguity*, a sophisticated yet natural concept to be introduced below.

The existence of unambiguous and moderately ambiguous *nonerasing* morphisms has already been intensively studied (see, e. g., Freydenberger et al. [1], Reidenbach [9]), and characteristic criteria have been provided. These criteria reveal that the existence of such morphisms is alphabet-independent, i. e., for any string s over some alphabet $\mathcal{A}$ and for any alphabets Σ, Σ' with at least two letters each, s has an unambiguous or moderately ambiguous nonerasing morphism $h : \mathcal{A}^* \to \Sigma^*$ if and only if there is a morphism $h' : \mathcal{A}^* \to \Sigma'^*$ with the equivalent property. In the present work, we study the ambiguity of all morphisms, including *erasing morphisms*, which map a symbol in s to the empty string. As pointed out by Schneider [14], here the existence of an unambiguous erasing morphism does not only depend on the structure of the string, but also on the size of the target alphabet of the morphism, which turns the search for characteristic conditions into a rather intricate problem.

[☆] A preliminary version [12] of this paper was presented at the conference DLT 2010.
[*] Corresponding Author.
Email addresses: `D.Reidenbach@lboro.ac.uk` (Daniel Reidenbach), `jschneider@informatik.uni-kl.de` (Johannes C. Schneider)

[1] In recent literature, the term “word” is normally used for what we call a “string”. We use this dated terminology since we wish to restrict the term “word” to strings over a particular alphabet (see Section 2).

The examination of the ambiguity of morphisms is not only of intrinsic interest, but, due to the simplicity of the concept, also shows various connections to other topics in theoretical computer science and discrete mathematics. This primarily holds for those approaches where several morphisms are applied to one finite string, including pattern languages (see, e. g., Mateescu and Salomaa [8]) as well as equality sets (and, thus, the Post Correspondence Problem, cf. Harju and Karhumäki [2]). Particularly well understood are the relations to pattern languages, where several prominent problems have been solved using insights into the ambiguity of morphisms (see, e. g., Reidenbach [10]). Moreover, there are further connections of the ambiguity of morphisms to various concepts that involve morphisms such as fixed points of morphisms, avoidable patterns and word equations.

Our work is organised as follows: After giving some definitions and basic results, we provide a detailed analysis of *ambiguity partitions* (as introduced by Schneider [14]), which are a vital concept when investigating the ambiguity of erasing morphisms. In Section 4, we introduce and study moderate ambiguity, i. e., an important type of strongly restricted ambiguity. We characterise those strings for which there exist moderately ambiguous erasing morphisms, and this is the main result of our paper. In Section 5, we deal with unambiguous morphisms, and we study the number of such morphisms for certain strings. Finally, in Section 6, we reveal that the existence of unambiguous erasing morphisms can be characterised using basic decision problems for so-called nonerasing multi-pattern languages. This insight might be a worthwhile starting point for future research.

2. Definitions and Basic Notes

In the present section we give some basic definitions and results. For notations not explained explicitly, we refer the reader to Rozenberg and Salomaa [13].

Let $\mathbb{N} := \{1, 2, \ldots\}$ be the set of natural numbers. The *power set* of a set S is denoted by $\mathcal{P}(S)$. An *alphabet* $\mathcal{A}$ is an enumerable set of symbols. A *string (over $\mathcal{A}$)* is a finite sequence of symbols taken from $\mathcal{A}$. By $|X|$ we denote the cardinality of a set X or the length of a string X. The *empty* string ε is the unique sequence of symbols of length 0. For the *concatenation* of strings s, t we write $s \cdot t$ (or st for short). The string that results from the n-fold concatenation of a string s is denoted by s^n. The notation $\mathcal{A}^*$ refers to the set of all strings over $\mathcal{A}$, i. e., more precisely, the *free monoid* generated by $\mathcal{A}$; furthermore, $\mathcal{A}^+ := \mathcal{A}^* \setminus \{\varepsilon\}$. The number of occurrences of a symbol $x \in \mathcal{A}$ in a string $s \in \mathcal{A}^*$ is written as $|s|_x$. With regard to arbitrary strings $s, t \in \mathcal{A}^*$, we write $s = t \ldots$ if there exists an $u \in \mathcal{A}^*$ such that $s = tu$, we write $s = \ldots t$ if there exists an $u \in \mathcal{A}^*$ such that $s = ut$, and, finally, $s = \ldots t \ldots$ if there exist $u, v \in \mathcal{A}^*$ such that $s = utv$. We call t a *prefix*, *suffix* and *factor* of s, respectively. In contrast to this notation, if we omit some parts of a canonically given string, then we henceforth use the symbol $[\ldots]$; e. g., $s = \ldots a\, b\, [\ldots]\, f$ means that s ends with the string $a\, b\, c\, d\, e\, f$.

We often use $\mathbb{N}$ as an infinite alphabet of symbols. In order to distinguish between a string over $\mathbb{N}$ and a string over a (possibly finite) alphabet Σ, we call the former a *pattern* and the latter a *word*. Given a pattern $\alpha \in \mathbb{N}^*$, we call symbols occurring in α *variables* and denote the set of variables in α with $\mathrm{var}(\alpha)$. Hence, $\mathrm{var}(\alpha) \subseteq \mathbb{N}$. We use the symbol $\cdot$ to separate the variables in a pattern, so that, for instance, $1 \cdot 1 \cdot 2$ is not confused with $11 \cdot 2$.

Given arbitrary alphabets $\mathcal{A}, \mathcal{B}$, a *morphism* is a mapping $h : \mathcal{A}^* \to \mathcal{B}^*$ that is compatible with the concatenation, i. e., for all $v, w \in \mathcal{A}^*$, $h(vw) = h(v)h(w)$. Hence, h is fully defined for all $v \in \mathcal{A}^*$ as soon as it is defined for all symbols in $\mathcal{A}$. We call h *erasing* if and only if $h(a) = \varepsilon$ for an $a \in \mathcal{A}$; otherwise, h is called *nonerasing*. If we call a morphism h (non)erasing with a certain input string s in mind, we only demand h to be (non)erasing for the symbols occurring in s.

A pattern $\alpha \in \mathbb{N}^+$ is called a *fixed point (of a morphism h)* if $h(\alpha) = \alpha$. A morphism $h : \mathbb{N}^* \to \mathbb{N}^*$ is said to be *nontrivial* if $h(x) \neq x$ for an $x \in \mathbb{N}$. Let $V \subseteq \mathbb{N}$. We call $h : \mathbb{N}^* \to \mathbb{N}^*$ *nontrivial for V* if $h(x) \neq x$ for an $x \in V$. The morphism $\pi_V : \mathbb{N}^* \to \mathbb{N}^*$ is given by $\pi_V(x) := x$ if $x \in V$ and $\pi_V(x) := \varepsilon$ if $x \notin V$.

For any alphabet Σ, for any morphism $\sigma : \mathbb{N}^* \to \Sigma^*$ and for any pattern $\alpha \in \mathbb{N}^+$ with $\sigma(\alpha) \neq \varepsilon$, we call σ *unambiguous (for α)* if and only if there is no morphism $\tau : \mathbb{N}^* \to \Sigma^*$ satisfying $\tau(\alpha) = \sigma(\alpha)$ and, for some $x \in \mathrm{var}(\alpha)$, $\tau(x) \neq \sigma(x)$. If σ is not unambiguous for α, it is called *ambiguous (for α)*. We extend this definition to any word $w \in \Sigma^*$ in the natural way, i. e., w is said to be *unambiguous (for α)* if there is an unambiguous morphism $\sigma : \mathbb{N}^* \to \Sigma^*$ with $\sigma(\alpha) = w$, and w is called *ambiguous (for α)* if there is an ambiguous morphism $\sigma : \mathbb{N}^* \to \Sigma^*$ satisfying $\sigma(\alpha) = w$. Furthermore, with regard to the E-pattern language of α to be introduced in the subsequent paragraph, we say that a word $w \in L_{\mathrm{E},\Sigma}(\alpha)$ is (un-)ambiguous if w is (un-)ambiguous for α.

Basically, the set of all images of a pattern $\alpha \in \mathbb{N}^+$ under morphisms $\sigma : \mathbb{N}^* \to \Sigma^*$, where Σ is an arbitrary alphabet of so-called *terminal-symbols*, is called the *pattern language (generated by α)*. Formally, two main types of pattern languages of α are considered: its *E-pattern language* $L_{\mathrm{E},\Sigma}(\alpha) := \{\sigma(\alpha) \mid \sigma : \mathbb{N}^* \to \Sigma^* \text{ is a morphism}\}$ and its *NE-pattern language* $L_{\mathrm{NE},\Sigma}(\alpha) := \{\sigma(\alpha) \mid \sigma : \mathbb{N}^* \to \Sigma^* \text{ is a nonerasing morphism}\}$. Note that, in literature, pattern languages as defined above are usually called *terminal-free*, since, in a more general understanding of the concept, a pattern may additionally contain terminal symbols. The morphisms $\sigma : (\mathbb{N} \cup \Sigma)^* \to \Sigma^*$ applied to such a pattern $\alpha \in (\mathbb{N} \cup \Sigma)^+$ when generating its pattern language must then be *terminal-preserving*, i. e., for any $a \in \Sigma$, $\sigma(a) = a$ must be satisfied.

The following result characterises the inclusion of erasing pattern languages.

Theorem 1 (Jiang et al. [6]). *Let $\alpha, \beta \in \mathbb{N}^+$, and let Σ be an alphabet with $|\Sigma| \geq 2$. Then $L_{\mathrm{E},\Sigma}(\alpha) \subseteq L_{\mathrm{E},\Sigma}(\beta)$ if and only if there exists a morphism $h : \mathbb{N}^* \to \mathbb{N}^*$ satisfying $h(\beta) = \alpha$.*

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

Definition 1. Let $\alpha \in \mathbb{N}^+$. We call α *prolix* if and only if there exists a factorisation $\alpha = \beta_0\, \gamma_1\, \beta_1\, \gamma_2\, \beta_2 \dots \gamma_n\, \beta_n$ with $n \geq 1, \beta_i \in \mathbb{N}^*$, $0 \leq i \leq n$, and $\gamma_i \in \mathbb{N}^+$, $1 \leq i \leq n$, such that

1. for every $i \in \{1, 2, \dots, n\}$, $|\gamma_i| \geq 2$,
2. for every $i \in \{0, 1, \dots, n\}$, for every $j \in \{1, 2, \dots, n\}$, $\mathrm{var}(\beta_i) \cap \mathrm{var}(\gamma_j) = \emptyset$,
3. for every $i \in \{1, 2, \dots, n\}$, there exists an $y_i \in \mathrm{var}(\gamma_i)$ such that y_i occurs exactly once in γ_i and, for every $i' \in \{1, 2, \dots, n\}$, if $y_i \in \gamma_{i'}$ then $\gamma_i = \gamma_{i'}$.

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

A succinct pattern is the shortest generator of its respective E-pattern language, i. e., for any Σ, $|\Sigma| \geq 2$, and any succinct pattern α, there is no pattern β with $|\beta| < |\alpha|$ and $L_{\mathrm{E},\Sigma}(\beta) = L_{\mathrm{E},\Sigma}(\alpha)$. Furthermore, the set of prolix patterns exactly corresponds to the class of finite fixed points of nontrivial morphisms:

Theorem 2 (Head [3]). *A pattern $\alpha \in \mathbb{N}^+$ is prolix if and only if it is a fixed point of a nontrivial morphism $h : \mathbb{N}^* \to \mathbb{N}^*$.*

Hence, for every prolix pattern α, there exists a morphism $h : \mathbb{N}^* \to \mathbb{N}^*$ satisfying $h(\alpha) = \alpha$ and $h(x) \neq x$ for an $x \in \mathrm{var}(\alpha)$. Note that set of succinct patterns is also equivalent to the set of *morphically primitive* words (as introduced by Reidenbach and Schneider [11]).

Regarding the unambiguity of *nonerasing* morphisms, the classification of patterns into succinct and prolix patterns is vital:

Theorem 3 (Freydenberger, Reidenbach, and Schneider [1]). *Let $\alpha \in \mathbb{N}^+$, let Σ be an alphabet, $|\Sigma| \geq 2$. There exists an unambiguous nonerasing morphism $\sigma : \mathbb{N}^* \to \Sigma^*$ for α if and only if α is succinct.*

According to this result, for any prolix pattern α, every nonerasing morphism is ambiguous. In contrast to this negative insight, there are prolix patterns that have unambiguous *erasing* morphisms (as pointed out by Schneider [14]). However, this is not a universal property of prolix patterns; thus, certain prolix patterns do not have any unambiguous morphism at all. This phenomenon is the main topic of our paper.

3. Ambiguity Partitions

Previous results show that *ambiguity partitions* as introduced by Schneider [14] are a crucial notion when investigating the ambiguity of erasing morphisms, and the main result of our paper, given in Section 4, further illustrates their importance. In the present section, we therefore study some fundamental properties of this concept.

Definition 2. We inductively define an *ambiguity partition (for any $\alpha \in \mathbb{N}^+$)*:

(i) $(\emptyset, \mathrm{var}(\alpha))$ is an ambiguity partition for α.

(ii) If (E, N) is an ambiguity partition for α and there exists a morphism $h : \mathbb{N}^* \to \mathbb{N}^*$ that is nontrivial for N and satisfies $h(\alpha) = \pi_N(\alpha)$, then (E', N') is an ambiguity partition with $E' := E \cup \{x \in N \mid h(x) = \varepsilon\}$, $N' := \{x \in N \mid h(x) \neq \varepsilon\}$.

We illustrate this definition by the following example.

Example 1. We define two example patterns as follows:

$$\begin{aligned} \alpha_1 &:= 1 \cdot 2 \cdot 1 \cdot 2 \cdot 3 \cdot 3 \cdot 3 \cdot 3, \\ \alpha_2 &:= 1 \cdot 2 \cdot 1 \cdot 2 \cdot 3 \cdot 3 \cdot 3. \end{aligned}$$

Note that α_1 and α_2 only differ in the number of occurences of variable 3.

We first consider α_1. By Definition 2, point (i), $(E, N) := (\emptyset, \{1, 2, 3\})$ is an ambiguity partition for α_1. The morphism $h : \mathbb{N}^* \to \mathbb{N}^*$, defined by $h(1) := \varepsilon$, $h(2) := 1 \cdot 2$, $h(3) := 3$, satisfies $h(\alpha_1) = \alpha_1 = \pi_N(\alpha_1)$. Note that h is nontrivial for N. Thus, according to Definition 2, point (ii), $(E', N') := (\{1\}, \{2, 3\})$ is an ambiguity partition for α_1, too. Furthermore, the nontrivial morphism $h : \mathbb{N}^* \to \mathbb{N}^*$, defined by $h(1) := 1 \cdot 2$, $h(2) := \varepsilon$, $h(3) := 3$, also satisfies $h(\alpha_1) = \alpha_1 = \pi_N(\alpha_1)$. Thus, according to Definition 2, point (ii), $(E', N') := (\{2\}, \{1, 3\})$ is an ambiguity partition for α_1 as well. We continue with the ambiguity partition $(E, N) := (\{1\}, \{2, 3\})$ and define the morphism $h : \mathbb{N}^* \to \mathbb{N}^*$ by $h(1) := 2$, $h(2) := \varepsilon$, $h(3) := 3$. Hence, h is nontrivial for N and satisfies $h(\alpha_1) = 2 \cdot 2 \cdot 3 \cdot 3 \cdot 3 \cdot 3 = \pi_N(\alpha_1)$. Thus, we can again apply Definition 2, point (ii) and get another ambiguity partition $(E', N') := (\{1, 2\}, \{3\})$ for α_1. Let $(E, N) := (E', N')$. Finally, the morphism $h : \mathbb{N}^* \to \mathbb{N}^*$, defined by $h(1) := \varepsilon$, $h(2) := 3 \cdot 3$, $h(3) := \varepsilon$, is nontrivial for N and satisfies $h(\alpha_1) = 3 \cdot 3 \cdot 3 \cdot 3 = \pi_N(\alpha_1)$. Consequently, $(E', N') := (\{1, 2, 3\}, \emptyset) = (\mathrm{var}(\alpha_1), \emptyset)$ is an ambiguity partition for α_1, too.

Concerning α_2, we can use a similar reasoning to show that all $(E, N) \in A := \{(\emptyset, \{1, 2, 3\}), (\{1\}, \{2, 3\}), (\{2\}, \{1, 3\}), (\{1, 2\}, \{3\})\}$ are ambiguity partitions for α_2. Furthermore, for any $(E, N) \in A$, there is no morphism that satisfies $h(3) = \varepsilon$ and point (ii) of Definition 2. Hence, the ambiguity partitions in A are the only ambiguity partitions for α_2.

According to [14], Definition 2 permits a number of fundamental insights into the ambiguity of erasing morphisms to be established. They directly or indirectly result from the following, slightly technical fact:

Theorem 4 (Schneider [14]). *Let Σ be an alphabet. Let $\alpha \in \mathbb{N}^+$ and let (E, N) be an ambiguity partition for α. Then every morphism $\sigma : \mathbb{N}^* \to \Sigma^*$ satisfying $\sigma(x) \neq \varepsilon$ for an $x \in E$ is ambiguous for α.*

Consequently, for any pattern α, an ambiguity partition (E, N) for α gives us valuable information on the set S of variables in α which must be erased by unambiguous morphisms, since $S \supseteq E$. Thus, the larger the set E becomes, the more information we get. Therefore, we name ambiguity partitions with a set E of maximal size in the following definition:

Definition 3. Let $\alpha \in \mathbb{N}^+$. An ambiguity partition (E, N) for α is called *maximal* if and only if every ambiguity partition (E', N') for α satisfies $|E'| \leq |E|$ and $|N'| \geq |N|$.

This definition supports some of our proofs, and we can use it to express vital statements on the (non-)existence of morphisms with a restricted ambiguity. In Example 1, $(\{1, 2, 3\}, \emptyset)$ is a maximal ambiguity partition for α_1 and $(\{1, 2, \}, \{3\})$ is a maximal ambiguity partition for α_2.

From Definition 2, it is not obvious whether or not a maximal ambiguity partition for a pattern α is unique. In order to answer this question, the following technical lemma is useful:

Lemma 1. *Let $\alpha \in \mathbb{N}^+$ and (E_1, N_1), (E_2, N_2) be ambiguity partitions for an α. Then $(E_1 \cup E_2, N_1 \cap N_2)$ is an ambiguity partition for α.*

Proof. To begin with, we note that, since (E_1, N_1) and (E_2, N_2) are partitions of $\mathrm{var}(\alpha)$, $N_1 \cap N_2 = \mathrm{var}(\alpha) \setminus (E_1 \cup E_2)$ and, thus, $(E_1 \cup E_2, N_1 \cap N_2)$ is a partition of $\mathrm{var}(\alpha)$, too.

If $(E_2, N_2) = (\emptyset, \mathrm{var}(\alpha))$, the statement is obviously true. Hence, let $(E_2, N_2) \neq (\emptyset, \mathrm{var}(\alpha))$. Then, according to condition (ii) of Definition 2, there exist ambiguity partitions $(E^{(0)}, N^{(0)}) := (\emptyset, \mathrm{var}(\alpha))$, $(E^{(1)}, N^{(1)})$, $(E^{(2)}, N^{(2)})$, ..., $(E^{(m)}, N^{(m)}) := (E_2, N_2)$, $m \in \mathbb{N}$ and, for every $k \in \{0, 1, \ldots, m\}$, a morphism $h^{(k)} : \mathbb{N}^* \to \mathbb{N}^*$ satisfying

(1) $h^{(k)}$ is nontrivial for $N^{(k)}$,
(2) $h^{(k)}(\alpha) = \pi_{N^{(k)}}(\alpha)$,
(3) $E^{(k+1)} = E^{(k)} \cup \{x \in N^{(k)} \mid h^{(k)}(x) = \varepsilon\}$, and
(4) $N^{(k+1)} = \{x \in N^{(k)} \mid h^{(k)}(x) \neq \varepsilon\}$.

We now give a procedure that, starting with (E_1, N_1), successively constructs ambiguity partitions (E', N') with growing sets $E' \supseteq E_1$ until $E' = E_1 \cup E_2$.

> $E' := E_1$, $N' := N_1$.
> `while` $(E' \neq E_1 \cup E_2)$ `do`
>   Let k be maximal with $E^{(k)} \subseteq E'$ and $E^{(k+1)} \nsubseteq E'$. $(\star)$
>   Let $h := \pi_{N'} \circ h^{(k)}$.
> $E'_{new} := E' \cup \{x \in N' \mid h(x) = \varepsilon\}$.
> $N'_{new} := \{x \in N' \mid h(x) \neq \varepsilon\}$.
> $E' := E'_{new}$, $N' := N'_{new}$.
> `od`

We show the following:

(a) Every (E', N') constructed by the algorithm is an ambiguity partition for α.
(b) The algorithm terminates.

ad (a). For $(E', N') = (E_1, N_1)$, the statement trivially holds. Hence, we show that, in every `while` loop, (E'_{new}, N'_{new}) is an ambiguity partition for α. Since $E^{(k)} \subseteq E'$ and $E^{(k+1)} \nsubseteq E'$, there is an $x \in E^{(k+1)} \setminus E'$ with $h^{(k)}(x) = \varepsilon$ (cf. point (3)). Furthermore, $x \notin E'$ implies $x \in N'$. Thus, h is nontrivial for N' since $h(x) = \pi_{N'}(h^{(k)}(x)) = \varepsilon \neq x$. Moreover, $h(\alpha) = \pi_{N'}(h^{(k)}(\alpha)) = \pi_{N'}(\pi_{N^{(k)}}(\alpha)) = \pi_{N'}(\alpha)$ since $h^{(k)}(\alpha) = \pi_{N^{(k)}}(\alpha)$ (cf. point (2)) and $N' \subseteq N^{(k)}$ (due to $E^{(k)} \subseteq E'$). Thus, (E'_{new}, N'_{new}) is an ambiguity partition for α according to condition (ii) of Definition 2.

ad (b). At first, we show that $\{x \in N' \mid h(x) = \varepsilon\} = E^{(k+1)} \setminus E'$. If $x \in N'$ and $h(x) = \pi_{N'}(h^{(k)}(x)) = \varepsilon$, then $h^{(k)}(x) = \varepsilon$. Due to $N' \subseteq N^{(k)}$, this implies $x \in E^{(k+1)}$ (cf. point (3)). Hence, $\{x \in N' \mid h(x) = \varepsilon\} \subseteq E^{(k+1)} \setminus E'$. Now let $x \in E^{(k+1)} \setminus E'$.

Since $E^{(k)} \subseteq E'$, $x \in E^{(k+1)} \setminus E^{(k)}$ and, thus, $h^{(k)}(x) = \varepsilon$ (cf. point (3)). Hence, $h(x) = \pi_{N'}(h^{(k)}(x)) = \varepsilon$. Furthermore, $x \notin E'$ directly implies $x \in N'$. This shows $\{x \in N' \mid h(x) = \varepsilon\} \supseteq E^{(k+1)} \setminus E'$, which proves the equality of both of the sets.

Consequently, E' is only extended by variables in some $E^{(k+1)} \subseteq E_2$, which implies $E' \subseteq E_1 \cup E_2$ for every E'. Moreover, $(\star)$ makes sure that all variables from $E_2 \setminus E_1$ are added to some E' such that, finally, $E' = E_1 \cup E_2$ and the `while` loop ends.

The statements (a) and (b) imply that there exists an algorithm that constructs the ambiguity partition $(E', N') = (E_1 \cup E_2, N_1 \cap N_2)$, which proves the lemma. □

From Lemma 1, we can conclude that, for any pattern, there is exactly one maximal ambiguity partition:

Theorem 5. *Let $\alpha \in \mathbb{N}^+$ and (E, N) be a maximal ambiguity partition for α. Then (E, N) is unique.*

Proof. Assume to the contrary that there exists a maximal ambiguity partition (E', N') for α such that $E' \neq E$. Then, according to Lemma 1, $(E \cup E', N \cap N')$ is an ambiguity partition for α, but $|E \cup E'| > |E|$. Consequently, (E, N) is not maximal, which contradicts the assumption. □

Evidently, the uniqueness of the maximal ambiguity partition (E, N) of a pattern α is a nontrivial property only if $E \neq \mathrm{var}(\alpha)$. On the other hand, if $(\mathrm{var}(\alpha), \emptyset)$ is the maximal ambiguity partition of α, then Theorem 4 directly implies that the following statement on the existence of unambiguous morphisms holds true:

Corollary 1 (Schneider [14]). *Let Σ be an alphabet, and let $\alpha \in \mathbb{N}^+$. If $(\mathrm{var}(\alpha), \emptyset)$ is an ambiguity partition for α, then every morphism $\sigma : \mathbb{N}^* \to \Sigma^*$ is ambiguous for α.*

Referring to Example 1, it is obvious that Corollary 1 can be applied to α_1, but not to α_2.

Corollary 1, in the case of arbitrary alphabets Σ, uses ambiguity partitions $(\mathrm{var}(\alpha), \emptyset)$ to establish a *sufficient* criterion on the nonexistence of unambiguous morphisms. However, in general, this criterion is not characteristic:

Example 2. Let

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

With a little effort (cf. the proof of Theorem 5.1 in [14], where the above pattern α is called α_2), we can show that $(\mathrm{var}(\alpha), \emptyset)$ is not an ambiguity partition for α. Hence, Corollary 1 does not apply to α. However, if we consider the morphism $h : \mathbb{N}^* \to \mathbb{N}^*$, defined by $h(1) := 1 \cdot 4 \cdot 5$, $h(2) := 2 \cdot 4 \cdot 6$, $h(3) := 3 \cdot 5 \cdot 6$, $h(4 \cdot 5 \cdot 6) := \varepsilon$, then point (ii) of Definition 2 implies that $(E', N') := (\{4, 5, 6\}, \{1, 2, 3\})$ is an ambiguity partition for α.

We now wish to demonstrate that no morphism $\sigma : \mathbb{N}^* \to \{\mathtt{a}, \mathtt{b}\}^*$ is unambiguous for α. We assume to the contrary that there exists an unambiguous morphism $\sigma : \mathbb{N}^* \to \{\mathtt{a}, \mathtt{b}\}^*$. Then, according to Theorem 4, $\sigma(4) = \sigma(5) = \sigma(6) = \varepsilon$. If $\sigma(x) = \varepsilon$ for an $x \in \{1, 2, 3\}$, σ is ambiguous. For instance, if $x = 1$, then the morphism $\tau : \mathbb{N}^* \to \{\mathtt{a}, \mathtt{b}\}^*$, defined by $\tau(1) := \sigma(2)$, $\tau(2) := \varepsilon$, $\tau(3) := \sigma(3)$, $\tau(4 \cdot 5 \cdot 6) := \varepsilon$, contradicts σ being unambiguous for α. The cases $x = 2$ and $x = 3$ are analogous. Thus, σ is nonerasing for the variables $1, 2, 3$. Since the target alphabet $\{\mathtt{a}, \mathtt{b}\}$ of σ consists of 2 letters and σ maps 3 variables onto a nonempty word, $\sigma(1)$ and $\sigma(2)$, $\sigma(1)$ and $\sigma(3)$ or $\sigma(2)$ and $\sigma(3)$ must end with the same letter $c \in \{\mathtt{a}, \mathtt{b}\}$. Assume $\sigma(1) = w_1 c$ and $\sigma(2) = w_2 c$ with $w_1, w_2 \in \{\mathtt{a}, \mathtt{b}\}^*$ (the other cases are analogous), then τ, defined by $\tau(1) := w_1$, $\tau(2) := w_2$, $\tau(4) := c$, $\tau(3) := \sigma(3)$, $\tau(y) := \varepsilon$ for all other variables y, satisfies $\tau(\alpha) = \sigma(\alpha)$ and, thus, contradicts σ being unambiguous.

Consequently, no morphism $\sigma : \mathbb{N}^* \to \{\mathtt{a}, \mathtt{b}\}^*$ is unambiguous for α.

Note that α also demonstrates that the ambiguity of erasing morphisms strongly depends on the size of the target alphabet since the morphism $\sigma : \mathbb{N}^* \to \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}^*$, defined by $\sigma(1) := \mathtt{a}$, $\sigma(2) := \mathtt{b}$, $\sigma(3) := \mathtt{c}$ and $\sigma(y) := \varepsilon$ for all other variables y, is unambiguous for α. Example patterns α with this property do not only exist for target alphabet sizes 2 and 3, but for any pair of finite target alphabets (see Theorem 5.1 in [14]).

Nevertheless, for infinite target alphabets, a result even stronger than Corollary 1 is known:

Theorem 6 (Schneider [14]). *Let Σ be an infinite alphabet, and let $\alpha \in \mathbb{N}^+$. Then $(\mathrm{var}(\alpha), \emptyset)$ is an ambiguity partition for α if and only if every morphism $\sigma : \mathbb{N}^* \to \Sigma^*$ is ambiguous for α.*

Thus, when investigating the existence of unambiguous erasing morphisms, the question of whether or not $(\mathrm{var}(\alpha), \emptyset)$ is an ambiguity partition for α leads to an important (and sometimes even characteristic) partition of $\mathbb{N}^+$. Therefore, we now introduce a new terminology reflecting this question:

Definition 4. Let $\alpha \in \mathbb{N}^+$. We call α *morphically erasable* if and only if $(\mathrm{var}(\alpha), \emptyset)$ is an ambiguity partition for α. Otherwise, α is called *morphically unerasable*.

The pattern α_1 from Example 1 is morphically erasable, whereas α_2 is morphically unerasable.

Referring to Definition 4, Corollary 1 demonstrates that, for finite alphabets Σ, the search for patterns with unambiguous morphisms can be narrowed down to the morphically unerasable ones. Therefore, and since our main result in Section 4 again is based on this property, we now give a nontrivial characterisation of such patterns. To this end, we use a condition that is based on the inclusion of E-pattern languages, which is a well-investigated problem (see Jiang et al. [6]).

Condition 1. A pattern $\alpha \in \mathbb{N}^+$ satisfies Condition 1 if and only if there exists a set $N \subseteq \mathrm{var}(\alpha)$ such that, for every $M \subseteq \mathrm{var}(\alpha)$ with $M \not\supseteq N$ and for any alphabet Σ with $|\Sigma| \geq 2$, $L_{\mathrm{E},\Sigma}(\pi_M(\alpha)) \not\supseteq L_{\mathrm{E},\Sigma}(\pi_N(\alpha))$.

Lemma 2. *A pattern $\alpha \in \mathbb{N}^+$ satisfies Condition 1 if and only if α is morphically unerasable.*

Proof. We first show the *if* direction of Lemma 2. If α is unerasable, then there exists an ambiguity partition (E, N) for α such that

(i) $N \neq \emptyset$ and

(ii) there is no morphism h with $h(\alpha) = \pi_N(\alpha)$ and $h(x) \neq x$ for an $x \in N$.

Note that (ii) is true since otherwise there would – due to (i) – exist a $y \in N$ with $h(y) = \varepsilon$, and therefore, by definition, $E' := E \cup \{x \in N \mid h(x) = \varepsilon\}$ and $N' := \{x \in N \mid h(x) \neq \varepsilon\}$ would form an ambiguity partition for α. Furthermore, the ambiguity partition (E', N') would satisfy $E' \supset E$ and $N' \subset N$. Thus, if there was no ambiguity partition (E, N) with properties (i) and (ii), then eventually $(\mathrm{var}(\alpha), \emptyset)$ would be an ambiguity partition for α.

We now demonstrate that, for every $M \subseteq \mathrm{var}(\alpha)$ with $M \not\supseteq N$ and for any alphabet Σ with $|\Sigma| \geq 2$, $L_{\mathrm{E},\Sigma}(\pi_M(\alpha)) \not\supseteq L_{\mathrm{E},\Sigma}(\pi_N(\alpha))$, i. e., α satisfies Condition 1. To this end, assume to the contrary that there exists an $M' \subseteq \mathrm{var}(\alpha)$ with

(a) $M' \not\supseteq N$ and

(b) $L_{\mathrm{E},\Sigma}(\pi_{M'}(\alpha)) \supseteq L_{\mathrm{E},\Sigma}(\pi_N(\alpha))$.

Due to Theorem 1, (b) implies that there is a morphism g with $g(\pi_{M'}(\alpha)) = \pi_N(\alpha)$. We now define a morphism $h : \mathbb{N}^* \to \mathbb{N}^*$ as follows: for every $x \in \mathbb{N}$, let

$$h(x) := \begin{cases} g(x) & , x \in M', \\ \varepsilon & , \text{else}. \end{cases}$$

Thus, $h(\pi_{M'}(\alpha)) = \pi_N(\alpha)$. Furthermore, due to (a), there is an $x \in N$ with $h(x) = \varepsilon$. However, this directly implies that $h(\alpha) = \pi_N(\alpha)$ and $h(x) \neq x$ for an $x \in N$. This contradicts property (ii) of (E, N). Thus, the set M' does not exist, which means that Condition 1 is satisfied.

We proceed with the *only if* direction. If α satisfies Condition 1, then there exists a set $N \subseteq \mathrm{var}(\alpha)$ such that, for every $M \subseteq \mathrm{var}(\alpha)$ with $M \not\supseteq N$ and for an arbitrary alphabet Σ with $|\Sigma| \geq 2$, $L_{\mathrm{E},\Sigma}(\pi_M(\alpha)) \not\supseteq L_{\mathrm{E},\Sigma}(\pi_N(\alpha))$. Due to Theorem 1, this means that, for every $M \subseteq \mathrm{var}(\alpha)$ with $M \not\supseteq N$, there is no morphism g with $g(\pi_M(\alpha)) = \pi_N(\alpha)$. Thus, if a morphism h satisfies $h(\alpha) = \pi_N(\alpha)$, then h is trivial for N. Hence, for every $N' \supseteq N$ and for every morphism h' with $h'(\alpha) = \pi_{N'}(\alpha)$, h' is also trivial for N. This statement directly implies that $(\mathrm{var}(\alpha), \emptyset)$ is not an ambiguity partition for α, since any procedure for finding the ambiguity partition $(\mathrm{var}(\alpha), \emptyset)$ starts from the ambiguity partition $(\emptyset, \mathrm{var}(\alpha))$ and needs to eventually reach an intermediate stage where there is a morphism h', a set $N' \supseteq N$ and an $x \in N$ such that $h'(\alpha) = \pi_{N'}(\alpha)$ and $h'(x) = \varepsilon$. This concludes the proof of the *only if* direction. □

Summarising the above statements, we can note the following sufficient condition on the nonexistence of unambiguous erasing morphisms, that is equivalent to Corollary 1:

Theorem 7. *Let Σ be an alphabet. If an $\alpha \in \mathbb{N}^+$ does not satisfy Condition 1, then every morphism $\sigma : \mathbb{N}^* \to \Sigma^*$ with $\sigma(\alpha) \neq \varepsilon$ is ambiguous for α.*

Proof. Directly from Corollary 1 and Lemma 2. □

The original motivation for investigating the ambiguity of morphisms is derived from inductive inference of E-pattern languages – i. e., the problem of computing a pattern from the words in its pattern languages –, which strongly depends on the inclusion relation between E-pattern languages. In this context, certain morphisms with a restricted ambiguity are known to generate words that contain reliable and algorithmically usable information about their generating pattern (cf. Reidenbach [10]) and, thus, are a vital input to any inference procedure. Theorem 7 further illustrates this close connection between the two topics.

The techniques used in [10] are based on the notion of an ambiguity of specific *nonerasing* morphisms that is restricted in a particular manner. We now introduce and study an equivalent concept for *erasing* morphisms.

4. Moderate Ambiguity

Theorem 6 shows that, in case of an infinite alphabet Σ, the property of a pattern α being morphically unerasable is characteristic for the existence of a morphism $\sigma : \mathbb{N}^* \to \Sigma^*$ that is unambiguous for α. However, concerning finite target alphabets Σ, there are morphically unerasable patterns for which there exists no unambiguous morphism (see Example 2). Although we are, hence, not able to achieve unambiguity for every morphically unerasable pattern, we shall demonstrate below that a certain restricted ambiguity is possible, which can be interpreted as unambiguity of a morphism with regard to particular factors of $\sigma(\alpha)$. As briefly mentioned above, a similar property of nonerasing morphisms is used for many fundamental results on inductive inference of E-pattern languages, and an extensive analysis of this phenomenon is provided by Reidenbach [9].

In accordance with [9], we call the said type of ambiguity *moderate ambiguity*. Intuitively, it can be understood as follows: A morphism $\sigma : \mathbb{N}^* \to \Sigma^*$ is called *moderately ambiguous* for a pattern α if, for every variable position j of a variable x in α with $\sigma(x) \neq \varepsilon$, there exists a certain factor w_j of $\sigma(\alpha)$ at a certain position (between the l_jth and r_jth letter in $\sigma(\alpha)$) such that *every* morphism $\tau : \mathbb{N}^* \to \Sigma^*$ with $\tau(\alpha) = \sigma(\alpha)$ maps the variable x at position j to a word which covers at least the factor w_j at this particular position. We illustrate this type of ambiguity in the following example:

Example 3. Let $\Sigma := \{\mathtt{a}, \mathtt{b}\}$ and

$$\begin{array}{rcccccccccccccccccccc} \alpha & := & 1 & \cdot & 2 & \cdot & 1 & \cdot & 1 & \cdot & 2 & \cdot & 1 & \cdot & 1 & \cdot & 3 & \cdot & 1 & \cdot & 3 \\ & = & i_1 & \cdot & i_2 & \cdot & i_3 & \cdot & i_4 & \cdot & i_5 & \cdot & i_6 & \cdot & i_7 & \cdot & i_8 & \cdot & i_9 & \cdot & i_{10} \end{array}$$

Let $\sigma : \mathbb{N}^* \to \Sigma^*$ be a morphism defined by $\sigma(1) := \varepsilon$, $\sigma(2) := \mathtt{aba}$, $\sigma(3) := \mathtt{abb}$. The morphism σ is ambiguous for α since $\tau : \mathbb{N}^* \to \Sigma^*$, defined by $\tau(1) := \mathtt{a}$, $\tau(2) := \mathtt{b}$, $\tau(3) := \mathtt{bb}$, satisfies $\tau(\alpha) = \sigma(\alpha)$. Hence, the situation looks as follows:

$$\sigma(\alpha) \;=\; \overbrace{\underbrace{\mathtt{a}}_{\tau(1)}\,\underbrace{\mathtt{b}}_{\tau(2)}\,\underbrace{\mathtt{a}}_{\tau(1)}}^{\sigma(2)}\;\overbrace{\underbrace{\mathtt{a}}_{\tau(1)}\,\underbrace{\mathtt{b}}_{\tau(2)}\,\underbrace{\mathtt{a}}_{\tau(1)}}^{\sigma(2)}\;\overbrace{\underbrace{\mathtt{a}}_{\tau(1)}\,\underbrace{\mathtt{bb}}_{\tau(3)}}^{\sigma(3)}\;\overbrace{\underbrace{\mathtt{a}}_{\tau(1)}\,\underbrace{\mathtt{bb}}_{\tau(3)}}^{\sigma(3)} \;=\; \tau(\alpha)$$

However, we call σ moderately ambiguous since all morphisms τ' with $\tau'(\alpha) = \sigma(\alpha)$ map every variable i_k with $\sigma(i_k) \neq \varepsilon$ to a certain factor w_k of $\sigma(i_k)$ at a particular position. In this example, we have $w_2 = w_5 = \mathtt{b}$ and $w_8 = w_{10} = \mathtt{bb}$. We can verify that the only morphisms τ' with $\tau'(\alpha) = \sigma(\alpha)$ are σ itself and τ, and, as explained above, these two morphisms satisfy $\sigma(i_k) = \ldots w_k \ldots = \tau(i_k)$ for $k = 2, 5, 8, 10$.

We now formalise moderate ambiguity. As explained above, we consider this a very natural way of slightly relaxing the requirement of unambiguity, and the relevance of this concept has been demonstrated in the context of inductive inference of pattern languages. Nevertheless, our definition is quite involved, since we do not only postulate that, for a given pattern α and for every $x \in \mathrm{var}(\alpha)$, there exists a string w_x such that, for every morphism τ with $\tau(\alpha) = \sigma(\alpha)$, $\tau(x)$ contains w_x as a factor (which could be called *factor-preserving* ambiguity), but we also demand that these factors are located at fixed positions for all τ. This means that we need to identify and mark the positions of the factors.

Definition 5. Let Σ be an alphabet, let $\alpha = i_1 \cdot i_2 \cdot [\ldots] \cdot i_n$ with $n, i_1, i_2, \ldots, i_n \in \mathbb{N}$, and let $\sigma : \mathbb{N}^* \to \Sigma^*$ be a morphism satisfying $\sigma(\alpha) \neq \varepsilon$. Then σ is called *moderately ambiguous (for α)* provided that there exist $l_2, l_3, \ldots, l_n, r_1, r_2, \ldots, r_{n-1} \in \mathbb{N} \cup \{0\}$ such that, for every morphism $\tau : \mathbb{N}^* \to \Sigma^*$ with $\tau(\alpha) = \sigma(\alpha)$,

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

We call σ *strongly ambiguous (for α)* if and only if it is not moderately ambiguous (for α).

In the definition, for any pattern α and any moderately ambiguous morphism σ for α, a pair (l_k, r_k) for some $i_k \in \mathrm{var}(\alpha)$ with $\sigma(i_k) \neq \varepsilon$ "marks" the factor w_k from position l_k to r_k in $\sigma(\alpha)$. This factor must be covered by the image of i_k under every morphism τ with $\tau(\alpha) = \sigma(\alpha)$ – this is guaranteed by the conditions (iv) and (v). Considering Example 3, we choose the following markers l_i, r_k: Let $r_1 := 0$, $(l_2, r_2) := (2, 2)$, $(l_5, r_5) := (5, 5)$, $(l_8, r_8) := (8, 9)$, $l_{10} := 11$ and finally $(l_k, r_k) := (|\sigma(\alpha)| + 1, 0)$ for $k \in \{1, 3, 4, 6, 7, 9\}$ since, for these k, $\sigma(i_k) = \varepsilon$, and, thus, no factor has to be marked. It can be verified that these values of l_j, $2 \leq j \leq n$, and r_k, $1 \leq k \leq n-1$ meet the requirements (i)–(v) of Definition 5.

The following lemma is useful when studying moderate ambiguity since, in certain cases, it circumvents a check of the minutiae of Definition 5.

Lemma 3. *Let Σ be an alphabet, $\alpha \in \mathbb{N}^+$ and $\sigma : \mathbb{N}^* \to \Sigma^*$ be a morphism. If there exists a morphism $\tau : \mathbb{N}^* \to \Sigma^*$ such that $\tau(\alpha) = \sigma(\alpha)$, but $\tau(x) = \varepsilon \neq \sigma(x)$ for an $x \in \mathrm{var}(\alpha)$, then σ is not moderately ambiguous for α.*

Proof. Let $\alpha = i_1 \cdot i_2 \cdot [\ldots] \cdot i_n$ with $n, i_1, i_2, \ldots, i_n \in \mathbb{N}$, and let k be minimal such that $x = i_k$. Assume to the contrary that σ is moderately ambiguous for α. Let $l_2, l_3, \ldots, l_n, r_1, r_2, \ldots, r_{n-1} \in \mathbb{N}$ as defined in Definition 5.

Case 1: $k = 1$. Then $r_1 \geq 1$ since $\sigma(i_1) \neq \varepsilon$, but $|\tau(i_1)| = |\varepsilon| = 0 < r_1$. This contradicts condition (v) of Definition 5.

Case 2: $k = n$. Then $l_n \leq |\sigma(\alpha)|$ due to $\sigma(i_n) \neq \varepsilon$. But since $\tau(i_n) = \varepsilon$, $\tau(\alpha) = \tau(i_1 \cdot i_2 \cdot [\ldots] \cdot i_{n-1})$ and, thus, $|\tau(i_1 \cdot i_2 \cdot [\ldots] \cdot i_{n-1})| = |\tau(\alpha)| = |\sigma(\alpha)| \geq l_n$. This contradicts condition (iv) of Definition 5.

Case 3: $1 < k < n$. Thus, τ must satisfy $|\tau(i_1 \cdot i_2 \cdot [\ldots] \cdot i_{k-1})| < l_k$ and $|\tau(i_1 \cdot i_2 \cdot [\ldots] \cdot i_k)| \geq r_k$. However, since $\tau(i_1 \cdot i_2 \cdot [\ldots] \cdot i_{k-1}) = \tau(i_1 \cdot i_2 \cdot [\ldots] \cdot i_k)$, it follows that $l_k > r_k$, which contradicts σ being moderately ambiguous. □

As suggested by the definitions and further substantiated by Example 3, for any given morphism, the requirement of being moderately ambiguous is less strict than that of being unambiguous:

Proposition 1. *Let Σ be an alphabet, let $\sigma : \mathbb{N}^* \to \Sigma^*$ be a morphism, and let $\alpha \in \mathbb{N}^+$. If σ is unambiguous for α, then σ is moderately ambiguous for α. In general, the converse does not hold.*

Proof. We begin with the first statement in Proposition 1. Let $\alpha = i_1 \cdot i_2 \cdot [\ldots] \cdot i_n$ with $n, i_1, i_2, \ldots, i_n \in \mathbb{N}$, and let $\sigma : \mathbb{N}^* \to \Sigma^*$ be a morphism that is unambiguous for α. We define $r_k := |\sigma(i_1 \cdot i_2 \cdot [\ldots] \cdot i_k)|$ and $l_{k+1} := |\sigma(i_1 \cdot i_2 \cdot [\ldots] \cdot i_k)| + 1$ for every $1 \leq k \leq n - 1$. Since σ is unambiguous and, thus, every morphism τ with $\tau(\alpha) = \sigma(\alpha)$ for every $x \in \mathrm{var}(\alpha)$ necessarily satisfies $\tau(x) = \sigma(x)$, the correctness of conditions (i)–(v) of Definition 5 can be verified easily.

Regarding the second statement in Proposition 1, Example 3 gives a morphism σ and a pattern α such that σ is moderately ambiguous, but not unambiguous for α. □

This directly implies that if there exists no moderately ambiguous morphism for a pattern α, then there exists no unambiguous morphism for α and, thus, every morphism is strongly ambiguous for α.

With these new terms of ambiguity, we can give a stronger version of Theorem 4:

Theorem 8. *Let Σ be an alphabet. Let $\alpha \in \mathbb{N}^+$ and let (E, N) be an ambiguity partition for α. Then every morphism $\sigma : \mathbb{N}^* \to \Sigma^*$ satisfying $\sigma(x) \neq \varepsilon$ for an $x \in E$ is strongly ambiguous for α.*

Proof. We prove Theorem 8 by induction.

For $(E, N) = (\emptyset, \mathrm{var}(\alpha))$, the statement is obviously true.

Now, let (E', N') be an ambiguity partition derived from an ambiguity partition (E, N) using condition (ii) of Definition 2. Then there exists a nontrivial morphism $h : \mathbb{N}^* \to \mathbb{N}^*$ satisfying $(\star)\, h(\alpha) = \pi_N(\alpha)$. Furthermore, $E' = E \cup \{x \in N \mid h(x) = \varepsilon\}$ and $N' = \{x \in N \mid h(x) \neq \varepsilon\}$. Consider a morphism $\sigma : \mathbb{N}^* \to \Sigma^*$ satisfying $\sigma(x) \neq \varepsilon$ for an $x \in E'$. If $x \in E$, it follows by induction that σ is strongly ambiguous. Now assume that $(\star\star)\, \sigma(x) = \varepsilon$ for all $x \in E$ and $\sigma(n) \neq \varepsilon$ for an $n \in N$ with $h(n) = \varepsilon$ and, thus, $n \in E'$. Let $\tau : \mathbb{N}^* \to \Sigma^*$ be the morphism defined by $\tau(x) = \sigma(h(x))$ for all $x \in \mathrm{var}(\alpha)$. Due to $(\star)$ and $(\star\star)$, $\tau(\alpha) = \sigma(h(\alpha)) = \sigma(\pi_N(\alpha)) = \sigma(\alpha)$, but $\tau(n) = \varepsilon \neq \sigma(n)$. With the help of Lemma 3, we can conclude that σ is strongly ambiguous for α. □

The main result of our paper characterises those patterns that have a moderately ambiguous morphism. More precisely, it states that moderate ambiguity can be achieved if and only if the pattern is morphically unerasable:

Theorem 9. *Let Σ be an alphabet, $|\Sigma| \geq 2$, let $\alpha \in \mathbb{N}^+$. There exists a morphism $\sigma : \mathbb{N}^* \to \Sigma^*$ that is moderately ambiguous for α if and only if α is morphically unerasable.*

Proof. We show the *only if* direction by contraposition: Let α be morphically erasable. Hence, there is an ambiguity partition $(\mathrm{var}(\alpha), \emptyset)$ for α. Then it follows from Theorem 8 that no morphism is moderately ambiguous for α.

We continue with the *if* direction. Let α be morphically unerasable. Hence, for the maximal ambiguity partition (E, N) for α, it is $N \neq \emptyset$. Furthermore, let $\alpha = i_1 \cdot i_2 \cdot [\ldots] \cdot i_n$ with $n, i_1, i_2, \ldots, i_n \in \mathbb{N}$, and let $\sigma : \mathbb{N}^* \to \{\mathtt{a}, \mathtt{b}\}^*$ be a morphism defined by

$$\sigma(i) := \begin{cases} \mathtt{ab}^{(i-1)(2n+1)+1}\mathtt{aab}^{(i-1)(2n+1)+2}\mathtt{a}[\ldots]\mathtt{ab}^{i(2n+1)}\mathtt{a}, & \text{if } i \in N, \\ \varepsilon, & \text{else,} \end{cases}$$

for every $i \in \mathbb{N}$. Hence, every $\sigma(i)$, $i \in N$, consists of exactly $2n + 1$ segments of the form $\mathtt{ab}^+\mathtt{a}$. Note that, for variables in N, σ is similar to the morphism $\tau_{k,a,b}$ as introduced by Jiang et. al. [6].

The idea now is to show that, for all k with $i_k \in N$, the factor $\mathtt{a\,a\,b}^{(i_k-1)(2n+1)+n+1}\,\mathtt{a\,a}$, which comprises the middle segment of $\sigma(i_k)$, is always contained in the image of i_k under any morphism τ with $\tau(\alpha) = \sigma(\alpha)$. Thus, we first give $l_2, l_3, \ldots, l_n$, $r_1, r_2, \ldots, r_{n-1} \in \mathbb{N} \cup \{0\}$ as required by Definition 5, according to the factors $\mathtt{a\,a\,b}^{(i_k-1)(2n+1)+n+1}\,\mathtt{a\,a}$ in $\sigma(\alpha)$. To this end, for all $k \in \{1, 2, \ldots, n\}$ with $i_k \in N$, let $v_k, w_k \in \{\mathtt{a}, \mathtt{b}\}^*$ such that $\sigma(i_k) = v_k\,\mathtt{a\,a\,b}^{(i_k-1)(2n+1)+n+1}\,\mathtt{a\,a}\,w_k$. We define, for every $k \in \{2, 3, \ldots, n\}$,

$$l_k := \begin{cases} |\sigma(i_1 \cdot i_2 \cdot [\ldots] \cdot i_{k-1})\, v_k\, \mathtt{a}|, & \text{if } i_k \in N \\ |\sigma(\alpha)| + 1, & \text{else,} \end{cases}$$

and, for every $m \in \{1, 2, \ldots, n-1\}$,

$$r_m := \begin{cases} |\sigma(i_1 \cdot i_2 \cdot [\ldots] \cdot i_{m-1})\, v_m\, \mathtt{a\,a\,b}^{(i_m-1)(2n+1)+n+1}\,\mathtt{a\,a}|, & \text{if } i_m \in N \\ 0, & \text{else.} \end{cases}$$

It can be verified with little effort that the l_k, r_m satisfy points (i)–(iii) of Definition 5. In the following, we verify points (iv) and (v). To this purpose, we introduce a new notion and prove some claims.

Auxiliary Definition. Let $\tau : \mathbb{N}^* \to \{\mathtt{a}, \mathtt{b}\}^*$ be a morphism with $\tau(\alpha) = \sigma(\alpha)$. Then a segment $\mathtt{ab}^x\mathtt{a}$, $x \in \mathbb{N}$, is called *preserved by τ at position j* if and only if, for $\tau(\alpha) = u_1\,\mathtt{ab}^x\mathtt{a}\,u_2$ and $|u_1| = j - 1$, there exists an l, $0 \leq l \leq n - 1$, such that

- $|\tau(i_1 \cdot i_2 \cdot [\ldots] \cdot i_l)| \leq |u_1|$ and
- $|\tau(i_1 \cdot i_2 \cdot [\ldots] \cdot i_l \cdot i_{l+1})| \geq |u_1\mathtt{ab}^x\mathtt{a}|$;

otherwise, it is called *split by τ at position j.*
A segment $\mathtt{ab}^x\mathtt{a}$ is called *preserved by τ* if and only if it is preserved by τ at all its positions in $\sigma(\alpha)$; otherwise, it is called *split by τ*.

Claim 1. For every morphism $\tau : \mathbb{N}^* \to \{\mathtt{a}, \mathtt{b}\}^*$ with $\tau(\alpha) = \sigma(\alpha)$ and every $k \in \{1, 2, \ldots, n\}$, there exist at least $n + 2$ different segments $\mathtt{ab}^x\mathtt{a}$, $x \in I := \{(i_k - 1)(2n + 1) + 1, (i_k - 1)(2n + 1) + 2, \ldots, i_k(2n + 1)\}$, that are preserved by τ.

Proof (Claim 1). Let τ be a morphism with $\tau(\alpha) = \sigma(\alpha)$ and let $k \in \{1, 2, \ldots, n\}$. Let $l := |\alpha|_{i_k}$. Thus, there exist exactly $l(2n + 1)$ positions $p_1, p_2, \ldots, p_{l(2n+1)} \in \mathbb{N}$ in $\sigma(\alpha)$ where segments of the form $\mathtt{ab}^x\mathtt{a}$, $x \in I$, begin. Since there are n variables in α, at most at $n - 1$ positions such a segment $\mathtt{ab}^x\mathtt{a}$ can be split by τ. Thus, there are at least $l(2n+1)-(n-1)$ positions in α where a segment $\mathtt{ab}^x\mathtt{a}$ with $x \in I$ is preserved by τ. It is a simple combinatorial insight that if there are $l(2n + 1)$ coloured balls of which exactly l balls have the same colour, one can choose at maximum $(l-1)(2n+1)$ balls without having all l balls of one colour. Every ball more than $(l-1)(2n+1)$ gives another complete set of equally coloured balls. We can transfer these considerations to our setting by identifying balls with positions of segments having the same colour if they mark the same segment. We have at least $l(2n + 1) - (n - 1)$ positions in $\sigma(\alpha)$ where a segment $\mathtt{ab}^x\mathtt{a}$ with $x \in I$ is preserved by τ, and $l(2n + 1) - (n - 1) = (l - 1)(2n + 1) + (n + 2)$. Consequently, there are $n + 2$ segments $\mathtt{ab}^x\mathtt{a}$ with $x \in I$ that are preserved by τ, since they are preserved at each of their l positions. *q.e.d.(Claim 1).*

Claim 2. If there exists a morphism $\tau : \mathbb{N}^* \to \{\mathtt{a}, \mathtt{b}\}^*$ with $\tau(\alpha) = \sigma(\alpha)$ that does not satisfy (iv) or (v) (cf. Definition 5), then there exist $j_1, j_2 \in \mathrm{var}(\alpha)$, $j_1 \neq j_2$, with $\sigma(j_2) = \ldots\mathtt{ab}^{(j_1-1)(2n+1)+s}\mathtt{a}\ldots$ for an $s \in \{1, 2, \ldots, 2n + 1\}$. Furthermore, the segment $\mathtt{ab}^{(j_1-1)(2n+1)+s}\mathtt{a}$ is preserved by τ.

Proof (Claim 2). If (iv) is not satisfied, then there exists a $k \in \{1, 2, \ldots, n - 1\}$ with $|\tau(i_1 \cdot i_2 \cdot [\ldots] \cdot i_k)| \geq l_{k+1}$. It follows from the definition of l_{k+1} that $i_{k+1} \in N$. Hence, $\tau(i_1 \cdot i_2 \cdot [\ldots] \cdot i_k) = \sigma(i_1 \cdot i_2 \cdot [\ldots] \cdot i_k)\, v_{k+1}\, \mathtt{a}\ldots$ with

$$v_{k+1} = \mathtt{ab}^{(i_{k+1}-1)(2n+1)+1}\mathtt{aab}^{(i_{k+1}-1)(2n+1)+2}\mathtt{a}[\ldots]\mathtt{ab}^{(i_{k+1}-1)(2n+1)+n}.$$

From Claim 1, we know that $n+2$ segments $\mathtt{ab}^x\mathtt{a}$ with $x \in \{(i_{k+1}-1)(2n+1)+1, (i_{k+1}-1)(2n+1)+2, \ldots, i_{k+1}(2n+1)\}$ are preserved by τ and, hence, also at least one segment $\mathtt{ab}^{(i_{k+1}-1)(2n+1)+s}\mathtt{a}\ldots$ for an $s \in \{1, 2, \ldots, n\}$. Thus, there is a $j \in \{i_1, i_2, \ldots, i_k\}$ with $\tau(j) = \ldots\mathtt{ab}^{(i_{k+1}-1)(2n+1)+s}\mathtt{a}\ldots$. If we assume $j = i_{k+1}$, then

- $\tau(i_{k+1}) = \ldots \mathtt{ab}^{(i_{k+1}-1)(2n+1)+s}\mathtt{a} \ldots \mathtt{ab}^{(i_{k+1}-1)(2n+1)+s}\mathtt{a} \ldots$ or
- $\tau(i_{k+1}) = \ldots \mathtt{ab}^{(i_{k+1}-1)(2n+1)+s}\mathtt{a} \ldots$ and $\tau(j') = \ldots \mathtt{ab}^{(i_{k+1}-1)(2n+1)+s}\mathtt{a} \ldots$ for a $j' \neq i_{k+1}$,

because the number of occurrences of j in $i_1 \cdot i_2 \cdot [\ldots] \cdot i_k$ is strictly smaller than the number of occurrences of $\mathtt{ab}^{(i_{k+1}-1)(2n+1)+s}\mathtt{a}$ in $\tau(i_1 \cdot i_2 \cdot [\ldots] \cdot i_k)$. However, then the number of occurrences of $\mathtt{ab}^{(i_{k+1}-1)(2n+1)+s}\mathtt{a}$ in $\tau(\alpha)$ is greater than the number of occurrences of $\mathtt{ab}^{(i_{k+1}-1)(2n+1)+s}\mathtt{a}$ in $\sigma(\alpha)$, which contradicts $\tau(\alpha) = \sigma(\alpha)$. Thus, $j \neq i_{k+1}$.

If (v) is not satisfied, then there exists a $k \in \{1, 2, \ldots, n-1\}$ with $|\tau(i_1 \cdot i_2 \cdot [\ldots] \cdot i_k)| < r_k$. It follows from the definition of r_k that $i_k \in N$. Hence, $\tau(i_{k+1} \cdot i_{k+2} \cdot [\ldots] \cdot i_n) = \ldots \mathtt{a}\, w_k\, \sigma(i_{k+1} \cdot i_{k+2} \cdot [\ldots] \cdot i_n)$ with

$$w_k = \mathtt{b}^{(i_k-1)(2n+1)+1}\mathtt{aab}^{(i_k-1)(2n+1)+2}\mathtt{a}[\ldots]\mathtt{ab}^{(i_k-1)(2n+1)+n}\mathtt{a}.$$

Using an analogous reasoning to the one given in the case of (iv) not being satisfied, we can conclude that there is a $j \in \{i_1, i_2, \ldots, i_k\}$ with $j \neq i_k$ and $\tau(j) = \ldots \mathtt{ab}^{(i_k-1)(2n+1)+s}\mathtt{a} \ldots$ for an $s \in \{n+2, n+3, \ldots, 2n+1\}$ and a segment $\mathtt{ab}^{(i_k-1)(2n+1)+s}\mathtt{a}$ which is preserved by τ. This proves the claim. *q.e.d.(Claim 2).*

Claim 3. If there exist a morphism $\tau : \mathbb{N}^* \to \{\mathtt{a}, \mathtt{b}\}^*$ with $\tau(\alpha) = \sigma(\alpha)$ and $j_1, j_2 \in \mathrm{var}(\alpha)$, $j_1 \neq j_2$, with $\tau(j_2) = \ldots \mathtt{ab}^{(j_1-1)(2n+1)+s}\mathtt{a} \ldots$ for an $s \in \{1, 2, \ldots, 2n+1\}$ and a segment $\mathtt{ab}^{(j_1-1)(2n+1)+s}\mathtt{a}$ that is preserved by τ, then the ambiguity partition (E, N) is not maximal.

Proof (Claim 3). For $i = j_2$, let $x_i := (j_1 - 1)(2n+1) + s$, and for $i \in N \setminus \{j_2\}$, we choose an $x_i \in \{(i-1)(2n+1) + 1, (i-1)(2n+1) + 2, \ldots, i(2n+1)\}$ such that the segment $\mathtt{ab}^{x_i}\mathtt{a}$ is preserved by τ. These x_i exist due to Claim 1.

Moreover, we define a morphism $h : \mathbb{N}^* \to \mathbb{N}^*$ for every $y \in \mathrm{var}(\alpha)$ as follows:

$$h(y) := \begin{cases} t_1 \cdot t_2 \cdot \ldots t_k, & \text{if } \tau(y) = w_0\, \mathtt{ab}^{x_{t_1}}\, \mathtt{a}\, w_1\, \mathtt{ab}^{x_{t_2}}\, \mathtt{a}\, w_2 \ldots \mathtt{ab}^{x_{t_k}}\, w_k, k \in \mathbb{N}, \\ & \text{satisfying } w_i \in \{\mathtt{a}, \mathtt{b}\}^* \text{ and } w_i \neq \ldots \mathtt{ab}^{x_j}\mathtt{a} \ldots \\ & \text{for all } i \in \{0, 1, \ldots, k\} \text{ and all } j \in N, \\ \varepsilon, & \text{else.} \end{cases}$$

If there exists a j_2 with $\tau(j_2) = \ldots \mathtt{ab}^{(j_1-1)(2n+1)+s}\mathtt{a} \ldots$, then h, by definition, is nontrivial for N. Furthermore, $h(\alpha) = \pi_N(\alpha)$ since, for every $i \in N$, there exists exactly one corresponding x_i. However, according to condition (ii) of Definition 2, (E', N') as defined when applying the above morphism h to (E, N), is an ambiguity partition satisfying $|E'| > |E|$ and $|N'| < |N|$. This contradicts the assumption of (E, N) being maximal (cf. Definition 3). *q.e.d.(Claim 3).*

Since Claims 2 and 3 imply that any violation of (iv) or (v) of Definition 5 would lead to a contradiction, σ is moderately ambiguous for α. □

In addition to the facts that Theorem 9 provides an algorithmically verifiable characteristic condition on a vital problem regarding the existence of morphisms with a restricted ambiguity and, furthermore, implies the equivalent result for the weaker requirement of factor-preserving ambiguity, we consider two other aspects of it quite remarkable. Firstly, it confirms that ambiguity partitions are indeed a crucial tool when investigating the ambiguity of erasing morphisms, since they cannot only be used to give sufficient criteria on the subject (cf. Corollary 1) and characteristic criteria for special cases (cf. Theorem 6), but are also capable of expressing a key phenomenon in this field of study.

Secondly, it establishes a quite remarkable and counter-intuitive difference between the ambiguity of erasing and nonerasing morphisms. As demonstrated by Freydenberger et al. [1], the existence of a moderately ambiguous nonerasing morphism σ for a pattern implies the existence of an unambiguous nonerasing morphism σ'. More technically, it can be shown that σ can be turned into σ' by applying some minor yet sophisticated changes that depend on the structure of the pattern in question (see Reidenbach [9] for a detailed discussion of this topic). It is also important to note that the morphisms σ and σ' both use a binary target alphabet; hence, the existence of such morphisms – which characterises the succinct patterns, cf. Theorem 3 – exclusively depends on the pattern and not on the size of Σ (provided that Σ contains at least two letters). In contrast to these observations, Theorem 9 demonstrates that the existence of moderately ambiguous erasing morphisms does not imply the existence of unambiguous erasing morphisms:

Corollary 2. *Let Σ be an alphabet. There exists an $\alpha \in \mathbb{N}^+$ and a morphism $\sigma : \mathbb{N}^* \to \Sigma^*$ such that σ is moderately ambiguous for α, but no morphism is unambiguous for α.*

Proof. Directly from Schneider [14] (Theorem 5.1) and Theorem 9 since the example pattern in the proof of Theorem 5.1 in [14] is morphically unerasable. For the case $|\Sigma| = 2$, Example 2 can be consulted. □

Hence, the main result of our paper also shows that the technical concepts used by Freydenberger et al. [1] to turn a moderately ambiguous morphism into an unambiguous one necessarily fail for erasing morphisms. Since this insight is rather unexpected, it is also surprising that Theorem 9 is alphabet-independent, whereas any characterisation of the set of those patterns that have an unambiguous erasing morphism must incorporate the size of Σ (as demonstrated by Example 2 and to be further addressed by Section 5).

We wish to conclude this section with an insight into the complexity of the problem of deciding on the existence of moderately ambiguous morphisms:

Corollary 3. *Let Σ be an alphabet, $|\Sigma| \geq 2$. The problem of deciding, for any given $\alpha \in \mathbb{N}^+$, on whether there is an erasing morphism $\sigma : \mathbb{N}^* \to \Sigma^*$ that is moderately ambiguous for α, is NP-complete.*

Proof. Corollary 3 directly follows from Theorems 4.1 and 4.3 by Schneider [14] and Theorem 9. □

This nicely contrasts with the recent result by Holub [4], which implies that there is a polynomial-time procedure deciding on the existence of unambiguous nonerasing morphisms.

As briefly mentioned above, we now study another fundamental property of those patterns that can be used to prove Corollary 2.

5. Patterns with Finitely Many Unambiguous Morphisms

Once the existence of morphisms with a restricted ambiguity has been established for a given pattern, it is a natural problem to investigate the number of such morphisms. Since the existence of one moderately ambiguous morphism for a given pattern immediately implies an infinite number of such morphisms (the morphism used to prove Theorem 9 can easily be generalised), we now study the above-mentioned topic with regard to a maximal restriction of ambiguity, i. e. unambiguity. To this end, we introduce the following notation:

Definition 6. Let Σ be an alphabet and $\alpha \in \mathbb{N}^+$. Then $\mathrm{UNAMB}_\Sigma(\alpha)$ is the set of all $\sigma(\alpha)$, where $\sigma : \mathbb{N}^* \to \Sigma^*$ is a morphism that is unambiguous for α, and $\mathrm{UNAMB}_{\mathrm{NE},\Sigma}(\alpha)$ is the set of all $\sigma(\alpha)$, where $\sigma : \mathbb{N}^* \to \Sigma^*$ is a morphism that is nonerasing and unambiguous for α.

We wish to point out that the sets $\mathrm{UNAMB}_\Sigma(\alpha)$ and $\mathrm{UNAMB}_{\mathrm{NE},\Sigma}(\alpha)$ do not consist of morphisms, but of morphic images. This makes sure that all unambiguous morphisms indirectly collected by these sets necessarily differ on variables that are contained in $\mathrm{var}(\alpha)$.

We first consider the case of *nonerasing* morphisms.

Theorem 10. *Let $\alpha \in \mathbb{N}^+$. Then either, for all alphabets Σ with $|\Sigma| \geq 2$, $\mathrm{UNAMB}_{\mathrm{NE},\Sigma}(\alpha)$ is empty or, for all alphabets Σ with $|\Sigma| \geq 2$, $\mathrm{UNAMB}_{\mathrm{NE},\Sigma}(\alpha)$ is infinite.*

Proof. Theorem 10 by Freydenberger et al. [1] states that $\mathrm{UNAMB}_{\mathrm{NE},\Sigma}(\alpha) = \emptyset$ for prolix patterns α. For any succinct pattern α, Definition 21 in [1] introduces a morphism $\sigma^{\mathrm{su}}_\alpha : \mathbb{N}^* \to \{\mathtt{a}, \mathtt{b}\}^*$ by, for every $k \in \mathbb{N}$,

$$\sigma^{\mathrm{su}}_\alpha(k) := \begin{cases} \mathtt{ab}^{3k}\mathtt{a}\ \mathtt{ab}^{3k+1}\mathtt{a}\ \mathtt{ab}^{3k+2}\mathtt{a}\ , & \forall\, i :\ k \neq \min L^\sim_i\ \wedge\ \forall\, i' :\ k \neq \min R^\sim_{i'}, \\ \mathtt{ba}^{3k}\mathtt{b}\ \mathtt{ab}^{3k+1}\mathtt{a}\ \mathtt{ab}^{3k+2}\mathtt{a}\ , & \forall\, i :\ k \neq \min L^\sim_i\ \wedge\ \exists\, i' :\ k = \min R^\sim_{i'}, \\ \mathtt{ab}^{3k}\mathtt{a}\ \mathtt{ab}^{3k+1}\mathtt{a}\ \mathtt{ba}^{3k+2}\mathtt{b}\ , & \exists\, i :\ k = \min L^\sim_i\ \wedge\ \forall\, i' :\ k \neq \min R^\sim_{i'}, \\ \mathtt{ba}^{3k}\mathtt{b}\ \mathtt{ab}^{3k+1}\mathtt{a}\ \mathtt{ba}^{3k+2}\mathtt{b}\ , & \exists\, i :\ k = \min L^\sim_i\ \wedge\ \exists\, i' :\ k = \min R^\sim_{i'}, \end{cases}$$

where the $L^\sim_i$ and $R^\sim_{i'}$ are equivalence classes over $\mathrm{var}(\alpha)$ and depend on the structure of α. This morphism is unambiguous for α (cf. Theorem 16 in [1]). However, the proof for Theorem 16 in [1] does not make use of the actual values $3k$, $3k + 1$ and $3k + 2$, but it is only required that, for every k, these three values are unique. Hence, we can modify $\sigma^{\mathrm{su}}_\alpha$ in infinitely many ways by substituting $3nk$ for $3k$, $3nk + 1$ for $3k + 1$, $3nk + 2$ for $3k + 2$, where $n \in \mathbb{N}$ is arbitrarily chosen. The resulting morphism is then still unambiguous for α. □

If we study the equivalent question for the ambiguity of *erasing* morphisms, we can observe a novel phenomenon that establishes a further difference to the case of nonerasing morphisms. More precisely, for certain patterns α, the cardinality of $\mathrm{UNAMB}_\Sigma(\alpha)$ can be finite, and this essentially depends on the size of Σ:

Theorem 11. *Let $k \in \mathbb{N}$. Let $\Sigma_k, \Sigma_{k+1}, \Sigma_{k+2}$ be alphabets with $k, k+1, k+2$ letters, respectively. There exists an $\alpha_k \in \mathbb{N}^+$ such that*

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

Proof. We define α_k as follows:

$$\alpha_k := \beta_1\beta_2 \ldots \beta_{k+1}\beta_{k+1}\beta_k \ldots \beta_1$$

with

$$\beta_i := \prod_{\substack{(j,m)\in\\ \{1,2,\ldots,k+1\}\times\{\mathrm{le},\mathrm{ri}\}\setminus\{(i,\mathrm{le})\}}} x_{\{(i,\mathrm{le}),(j,m)\}} \cdot i \cdot \prod_{\substack{(j,m)\in\\ \{1,2,\ldots,k+1\}\times\{\mathrm{le},\mathrm{ri}\}\setminus\{(i,\mathrm{ri})\}}} x_{\{(i,\mathrm{ri}),(j,m)\}},$$ [2]

where, for all indices S, the x_S are distinct variables taken from $\mathbb{N} \setminus \{1, 2, \ldots, k+1\}$.

For instance,

$$\begin{aligned}
\alpha_2 \;=\; & x_{\{(1,\mathrm{le}),(1,\mathrm{ri})\}} \cdot x_{\{(1,\mathrm{le}),(2,\mathrm{le})\}} \cdot x_{\{(1,\mathrm{le}),(2,\mathrm{ri})\}} \cdot x_{\{(1,\mathrm{le}),(3,\mathrm{le})\}} \cdot x_{\{(1,\mathrm{le}),(3,\mathrm{ri})\}} \cdot 1 \cdot \\
& x_{\{(1,\mathrm{ri}),(1,\mathrm{le})\}} \cdot x_{\{(1,\mathrm{ri}),(2,\mathrm{le})\}} \cdot x_{\{(1,\mathrm{ri}),(2,\mathrm{ri})\}} \cdot x_{\{(1,\mathrm{ri}),(3,\mathrm{le})\}} \cdot x_{\{(1,\mathrm{ri}),(3,\mathrm{ri})\}} \cdot \\
& x_{\{(2,\mathrm{le}),(1,\mathrm{le})\}} \cdot x_{\{(2,\mathrm{le}),(1,\mathrm{ri})\}} \cdot x_{\{(2,\mathrm{le}),(2,\mathrm{ri})\}} \cdot x_{\{(2,\mathrm{le}),(3,\mathrm{le})\}} \cdot x_{\{(2,\mathrm{le}),(3,\mathrm{ri})\}} \cdot 2 \cdot \\
& x_{\{(2,\mathrm{ri}),(1,\mathrm{le})\}} \cdot x_{\{(2,\mathrm{ri}),(1,\mathrm{ri})\}} \cdot x_{\{(2,\mathrm{ri}),(2,\mathrm{le})\}} \cdot x_{\{(2,\mathrm{ri}),(3,\mathrm{le})\}} \cdot x_{\{(2,\mathrm{ri}),(3,\mathrm{ri})\}} \cdot \\
& x_{\{(3,\mathrm{le}),(1,\mathrm{le})\}} \cdot x_{\{(3,\mathrm{le}),(1,\mathrm{ri})\}} \cdot x_{\{(3,\mathrm{le}),(2,\mathrm{le})\}} \cdot x_{\{(3,\mathrm{le}),(2,\mathrm{ri})\}} \cdot x_{\{(3,\mathrm{le}),(3,\mathrm{ri})\}} \cdot 3 \cdot \\
& x_{\{(3,\mathrm{ri}),(1,\mathrm{le})\}} \cdot x_{\{(3,\mathrm{ri}),(1,\mathrm{ri})\}} \cdot x_{\{(3,\mathrm{ri}),(2,\mathrm{le})\}} \cdot x_{\{(3,\mathrm{ri}),(2,\mathrm{ri})\}} \cdot x_{\{(3,\mathrm{ri}),(3,\mathrm{le})\}} \cdot \\
& x_{\{(3,\mathrm{le}),(1,\mathrm{le})\}} \cdot x_{\{(3,\mathrm{le}),(1,\mathrm{ri})\}} \cdot x_{\{(3,\mathrm{le}),(2,\mathrm{le})\}} \cdot x_{\{(3,\mathrm{le}),(2,\mathrm{ri})\}} \cdot x_{\{(3,\mathrm{le}),(3,\mathrm{ri})\}} \cdot 3 \cdot \\
& x_{\{(3,\mathrm{ri}),(1,\mathrm{le})\}} \cdot x_{\{(3,\mathrm{ri}),(1,\mathrm{ri})\}} \cdot x_{\{(3,\mathrm{ri}),(2,\mathrm{le})\}} \cdot x_{\{(3,\mathrm{ri}),(2,\mathrm{ri})\}} \cdot x_{\{(3,\mathrm{ri}),(3,\mathrm{le})\}} \cdot \\
& x_{\{(2,\mathrm{le}),(1,\mathrm{le})\}} \cdot x_{\{(2,\mathrm{le}),(1,\mathrm{ri})\}} \cdot x_{\{(2,\mathrm{le}),(2,\mathrm{ri})\}} \cdot x_{\{(2,\mathrm{le}),(3,\mathrm{le})\}} \cdot x_{\{(2,\mathrm{le}),(3,\mathrm{ri})\}} \cdot 2 \cdot \\
& x_{\{(2,\mathrm{ri}),(1,\mathrm{le})\}} \cdot x_{\{(2,\mathrm{ri}),(1,\mathrm{ri})\}} \cdot x_{\{(2,\mathrm{ri}),(2,\mathrm{le})\}} \cdot x_{\{(2,\mathrm{ri}),(3,\mathrm{le})\}} \cdot x_{\{(2,\mathrm{ri}),(3,\mathrm{ri})\}} \cdot \\
& x_{\{(1,\mathrm{le}),(1,\mathrm{ri})\}} \cdot x_{\{(1,\mathrm{le}),(2,\mathrm{le})\}} \cdot x_{\{(1,\mathrm{le}),(2,\mathrm{ri})\}} \cdot x_{\{(1,\mathrm{le}),(3,\mathrm{le})\}} \cdot x_{\{(1,\mathrm{le}),(3,\mathrm{ri})\}} \cdot 1 \cdot \\
& x_{\{(1,\mathrm{ri}),(1,\mathrm{le})\}} \cdot x_{\{(1,\mathrm{ri}),(2,\mathrm{le})\}} \cdot x_{\{(1,\mathrm{ri}),(2,\mathrm{ri})\}} \cdot x_{\{(1,\mathrm{ri}),(3,\mathrm{le})\}} \cdot x_{\{(1,\mathrm{ri}),(3,\mathrm{ri})\}}.
\end{aligned}$$

Note that, e. g., $x_{\{(1,\mathrm{le}),(2,\mathrm{ri})\}} = x_{\{(2,\mathrm{ri}),(1,\mathrm{le})\}}$ since $\{(1,\mathrm{le}),(2,\mathrm{ri})\} = \{(2,\mathrm{ri}),(1,\mathrm{le})\}$. Hence, $|\mathrm{var}(\alpha_2)| = 18$. With

$$\begin{array}{lll}
x_{\{(1,\mathrm{le}),(1,\mathrm{ri})\}} := 4, & x_{\{(1,\mathrm{le}),(2,\mathrm{le})\}} := 5, & x_{\{(1,\mathrm{le}),(2,\mathrm{ri})\}} := 6, \\
x_{\{(1,\mathrm{le}),(3,\mathrm{le})\}} := 7, & x_{\{(1,\mathrm{le}),(3,\mathrm{ri})\}} := 8, & x_{\{(1,\mathrm{ri}),(2,\mathrm{le})\}} := 9, \\
x_{\{(1,\mathrm{ri}),(2,\mathrm{ri})\}} := 10, & x_{\{(1,\mathrm{ri}),(3,\mathrm{le})\}} := 11, & x_{\{(1,\mathrm{ri}),(3,\mathrm{ri})\}} := 12, \\
x_{\{(2,\mathrm{le}),(2,\mathrm{ri})\}} := 13, & x_{\{(2,\mathrm{le}),(3,\mathrm{le})\}} := 14, & x_{\{(2,\mathrm{le}),(3,\mathrm{ri})\}} := 15, \\
x_{\{(2,\mathrm{ri}),(3,\mathrm{le})\}} := 16, & x_{\{(2,\mathrm{ri}),(3,\mathrm{ri})\}} := 17, & x_{\{(3,\mathrm{le}),(3,\mathrm{ri})\}} := 18,
\end{array}$$

α_2 looks as follows:

$$\begin{aligned}
\alpha_2 \;=\; & 4\cdot 5\cdot 6\cdot 7\cdot 8\cdot\ 1\ \cdot 4\cdot 9\cdot 10\cdot 11\cdot 12\ \cdot\ 5\cdot 9\cdot 13\cdot 14\cdot 15\cdot\ 2\ \cdot 6\cdot 10\cdot 13\cdot 16\cdot 17\cdot \\
& 7\cdot 11\cdot 14\cdot 16\cdot 18\cdot\ 3\ \cdot 8\cdot 12\cdot 15\cdot 17\cdot 18\ \cdot\ 7\cdot 11\cdot 14\cdot 16\cdot 18\cdot 3\cdot 8\cdot 12\cdot 15\cdot 17\cdot 18\cdot \\
& 5\cdot 9\cdot 13\cdot 14\cdot 15\cdot\ 2\ \cdot 6\cdot 10\cdot 13\cdot 16\cdot 17\ \cdot\ 4\cdot 5\cdot 6\cdot 7\cdot 8\cdot\ 1\ \cdot 4\cdot 9\cdot 10\cdot 11\cdot 12.
\end{aligned}$$

[2]Note that the order of the pairs $(j,m) \in \{1,2,\ldots,k+1\} \times \{\mathrm{le},\mathrm{ri}\} \setminus \{(i,\mathrm{le})\}$ can be arbitrarily chosen when composing β_i. The same holds for the order of the pairs $(j,m) \in \{1,2,\ldots,k+1\} \times \{\mathrm{le},\mathrm{ri}\} \setminus \{(i,\mathrm{ri})\}$.

This example may be consulted for a better understanding of the proof although the subsequent argumentation deals with the general pattern α_k.

Now, let $N := \{1, 2, \ldots, k+1\}$ and $E := \mathrm{var}(\alpha_k) \setminus N$. The morphism $h : \mathbb{N}^* \to \mathbb{N}^*$, defined by $h(i) := \beta_i$ for every $i \in N$, $h(i); = \varepsilon$ for every $i \in E$, is nontrivial and satisfies $h(\alpha_k) = \alpha_k$. Thus, according to Definition 2, (E, N) is an ambiguity partition for α_k.

ad (i). W. l. o. g., let $\Sigma_k := \{1, 2, \ldots, k\}$. Assume to the contrary that there exists an unambiguous morphism $\sigma : \mathbb{N}^* \to \Sigma_k^*$ for α_k. Then, according to Theorem 8, $\sigma(e) = \varepsilon$ for every $e \in E$. Thus, one of the following cases must occur:

Case 1: For every $n \in N$, $\sigma(n) \neq \varepsilon$. Since N contains $k+1$ variables, but Σ_k consists of k letters only, there must be $i, j \in N$, $i \neq j$, and a $y \in \Sigma_k$ such that $\sigma(i) = yw_i$ and $\sigma(j) = yw_j$ with $w_i, w_j \in \Sigma_k^*$. But then the morphism $\tau : \mathbb{N}^* \to \Sigma_k^*$, defined by $\tau(x_{\{(i,\mathrm{le}),(j,\mathrm{le})\}}) := y$, $\tau(i) := w_i$, $\tau(j) := w_j$, $\tau(n) := \sigma(n)$ for every $n \in N \setminus \{i, j\}$ and $\tau(e) := \varepsilon$ for every $e \in E \setminus \{x_{\{(i,\mathrm{le}),(j,\mathrm{le})\}}\}$, satisfies $\tau(\alpha_k) = \sigma(\alpha_k)$ and, thus, contradicts σ being unambiguous for α_k.

Case 2: There exists an $n \in N$ with $\sigma(n) = \varepsilon$. If $n = 1$, then the morphism $\tau : \mathbb{N}^* \to \Sigma_k^*$, defined by $\tau(1) = \sigma(2)$, $\tau(2) = \varepsilon$, $\tau(n') = \sigma(n')$ for all $n' \in N \setminus \{1, 2\}$, $\tau(e) := \varepsilon$ for all $e \in E$, satisfies $\tau(\alpha_k) = \sigma(\alpha_k)$ and, thus, contradicts σ being unambiguous for α_k. If $n > 1$, then the morphism $\tau : \mathbb{N}^* \to \Sigma_k^*$, defined by $\tau(n) := \sigma(n-1)$, $\tau(n-1) = \varepsilon$, $\tau(n') := \sigma(n')$ for all $n' \in N \setminus \{n-1, n\}$, $\tau(e) := \varepsilon$ for all $e \in E$, satisfies $\tau(\alpha_k) = \sigma(\alpha_k)$ and, thus, contradicts σ being unambiguous for α_k.

Thus, there is no unambiguous morphism $\sigma : \mathbb{N}^* \to \Sigma_k^*$ for α_k.

ad (ii). W. l. o. g., let $\Sigma_{k+1} := \{1, 2, \ldots, k+1\}$. Then $\pi_N(\alpha_k) \in \mathrm{UNAMB}_{\Sigma_{k+1}}(\alpha_k)$, since every $e \in E$ occurs four times in α_k, whereas every $n \in N$ occurs only two times, and since $\pi_N(\alpha_k)$ is succinct and, thus, the only morphism satisfying $h(\pi_N(\alpha_k)) = \pi_N(\alpha_k)$ is the trivial one (cf. Theorem 2). Now, let $\sigma : \mathbb{N}^* \to \Sigma_{k+1}^*$ be a morphism. If $\sigma(e) \neq \varepsilon$ for an $e \in E$, σ is ambiguous for α_k (cf. Theorem 8). Hence, let $\sigma(e) = \varepsilon$ for every $e \in E$. Assume that $|\sigma(n)| > 1$ for some $n \in N$. Then, for every $i \in N \setminus \{n\}$, there exist $a_i \in \Sigma_{k+1}$ and $w_i \in \Sigma_{k+1}^*$ such that $\sigma(i) = a_i w_i$, and there exist $a_{\mathrm{le}}, a_{\mathrm{ri}} \in \Sigma_{k+1}$, $w_n \in \Sigma_{k+1}^*$ such that $\sigma(n) = a_{\mathrm{le}} w_n a_{\mathrm{ri}}$. Since $a_{\mathrm{le}}, a_{\mathrm{ri}}$ and the a_i, $i \in N \setminus \{n\}$, stand for $k+2$ letters, but $|\Sigma_{k+1}| = k+1$, one of the following cases must occur:

Case 1: $a_{\mathrm{le}} = a_{\mathrm{ri}}$. Then, the morphism $\tau : \mathbb{N}^* \to \Sigma^*$, defined by $\tau(x_{\{(n,\mathrm{le}),(n,\mathrm{ri})\}}) := a_{\mathrm{le}}$, $\tau(n) := w_n$ and $\tau(i) := \sigma(i)$ for all $i \in \mathrm{var}(\alpha_k) \setminus \{n, x_{\{(n,\mathrm{le}),(n,\mathrm{ri})\}}\}$, contradicts σ being unambiguous for α_k, since $\tau(\alpha_k) = \sigma(\alpha_k)$, but $\tau(n) \neq \sigma(n)$.

Case 2: $a_{\mathrm{le}} = a_j$ for some $j \in N \setminus \{n\}$. Then, the morphism $\tau : \mathbb{N}^* \to \Sigma^*$, defined by $\tau(x_{\{(n,\mathrm{le}),(j,\mathrm{le})\}}) := a_{\mathrm{le}}$, $\tau(n) := w_n a_{\mathrm{ri}}$, $\tau(j) := w_j$ and $\tau(i) := \sigma(i)$ for all $i \in \mathrm{var}(\alpha_k) \setminus \{n, j, x_{\{(n,\mathrm{le}),(j,\mathrm{le})\}}\}$, contradicts σ being unambiguous for α_k, since $\tau(\alpha_k) = \sigma(\alpha_k)$, but $\tau(j) \neq \sigma(j)$.

Case 3: $a_{\mathrm{ri}} = a_j$ for some $j \in N \setminus \{n\}$. Then, the morphism $\tau : \mathbb{N}^* \to \Sigma^*$, defined by $\tau(x_{\{(n,\mathrm{ri}),(j,\mathrm{le})\}}) := a_{\mathrm{ri}}$, $\tau(n) := a_{\mathrm{le}} w_n$, $\tau(j) := w_j$ and $\tau(i) := \sigma(i)$ for all $i \in \mathrm{var}(\alpha_k) \setminus \{n, j, x_{\{(n,\mathrm{ri}),(j,\mathrm{le})\}}\}$, contradicts σ being unambiguous for α_k, since $\tau(\alpha_k) = \sigma(\alpha_k)$, but $\tau(j) \neq \sigma(j)$.

Case 4: $a_j = a_m$ for some $j, m \in N \setminus \{n\}$, $j \neq m$. Then, the morphism $\tau : \mathbb{N}^* \to \Sigma^*$, defined by $\tau(x_{\{(j,\mathrm{le}),(m,\mathrm{le})\}}) := a_j$, $\tau(j) := w_j$, $\tau(m) := w_m$ and $\tau(i) := \sigma(i)$ for all $i \in \mathrm{var}(\alpha_k) \setminus \{j, m, x_{\{(j,\mathrm{le}),(m,\mathrm{le})\}}\}$, contradicts σ being unambiguous for α_k, since $\tau(\alpha_k) = \sigma(\alpha_k)$, but $\tau(j) \neq \sigma(j)$.

Consequently, only morphisms $\sigma : \mathbb{N}^* \to \Sigma_{k+1}^*$ with $\sigma(e) = \varepsilon$ and for every $e \in E$ and $\sigma(n) \in \Sigma_{k+1}$ for every $n \in N$ can be unambiguous for α_k. Since there are only finitely many $\sigma(\alpha)$ for such morphisms σ and since $\pi_N(\alpha_k) \in \mathrm{UNAMB}_{\Sigma_{k+1}}(\alpha_k)$, (ii) follows.

ad (iii): W. l. o. g., let $\Sigma_{k+2} := \{1, 2, \ldots, k+2\}$. For every n, let $\sigma_n : \mathbb{N}^* \to \Sigma_{k+2}^*$ be a morphism defined by $\sigma_n(i) := i$, $1 \leq i \leq k$, and $\sigma_n(k+1) := (k+1) \cdot k^n \cdot (k+2)$. For instance, $\sigma_3(\alpha_2) = 1 \cdot 2 \;\cdot\; 3 \cdot 2 \cdot 2 \cdot 2 \cdot 4 \;\cdot\; 3 \cdot 2 \cdot 2 \cdot 2 \cdot 4 \;\cdot\; 2 \cdot 1$.

We show that, for every $n \in \mathbb{N}$, $\sigma_n(\alpha_k) \in \mathrm{UNAMB}_{\Sigma_{k+2}}(\alpha_k)$. This implies (iii). Let $n \in \mathbb{N}$. Assume to the contrary that there exists a morphism $\tau : \mathbb{N}^* \to \Sigma_{k+2}^*$ with $\tau(\alpha_k) = \sigma_n(\alpha_k)$ and $\tau(j) \neq \sigma_n(j)$ for a $j \in \mathrm{var}(\alpha_k)$.

Claim 1. For every $i \in \{1, 2, \ldots, k-1\}$, $\tau(i) = \sigma_n(i)$. (Since $\alpha_k = 1 \ldots 2 \ldots [\ldots] k-1 \ldots k-1 \ldots k-2 \ldots [\ldots] 1$.)

Claim 2. For every $e \in E$ and $n \in N \setminus \{k\}$, $\tau(e) \neq \ldots n \ldots$. (Since every $e \in E$ occurs four times in α_k, but n occurs only two times in $\sigma_n(\alpha_k)$.)

Due to Claim 1, $\tau(\beta_k \beta_{k+1} \beta_{k+1} \beta_k) = \sigma_n(\beta_k \beta_{k+1} \beta_{k+1} \beta_k) = k \cdot (k+1) \cdot k^n \cdot (k+2) \cdot (k+1) \cdot k^n \cdot (k+2) \cdot k$ must be satisfied. Because of Claim 2, for every $e \in E$ and $m \in \{k+1, k+2\}$, $\tau(e) \neq \ldots m \ldots$. Therefore, it can be verified by straightforward considerations on the structure of α that $\tau(k)$ or $\tau(k+1)$ must equal $\sigma_n(k+1)$. In both cases, all occurrences of k except for two are covered by $\tau(\pi_{\{1,2,\ldots,k\}}(\alpha))$ or $\tau(\pi_{\{1,2,\ldots,k-1,k+1\}}(\alpha))$. Thus, only $\tau(k) = \sigma_n(k)$ and $\tau(k+1) = \sigma_n(k+1)$ is possible, since every $e \in E$ occurs four times in α_k, which implies $\tau(e) \neq k$. This contradicts σ_n being ambiguous for α_k, and therefore $\sigma_n(\alpha_k) \in \mathrm{UNAMB}_{\Sigma_{k+2}}(\alpha_k)$. □

6. Connections to NE-pattern Languages

In this final main section of our paper we wish to study a topic that, after the particularly strong result in Theorem 9, remains as the most fundamental open problem on erasing morphisms with a restricted ambiguity, namely a characterisation of those patterns that have an unambiguous erasing morphism. As a matter of fact, the main result of the present section can be understood as such a characterisation, but the immediate usefulness of the result is limited. Nevertheless, our examinations reveal some enlightening and rather counter-intuitive insights that might be useful for further investigations.

While the existence of a relation between the ambiguity of erasing morphisms and certain properties of E-pattern languages (as, e. g., demonstrated by Condition 1 and Theorem 7) is by no means surprising, our characterisation shall demonstrate likewise deep connections between the main subject of our paper and vital properties of *NE*-pattern languages. It reads as follows:

Theorem 12. *Let Σ be an alphabet, and let $\alpha \in \mathbb{N}^+$. For any partition (U, V) of $\mathcal{P}(\mathrm{var}(\alpha)) \setminus \{\emptyset\}$, let*

$$L_{\alpha,U,V} := \bigcup_{u \in U} L_{\mathrm{NE},\Sigma}(\pi_u(\alpha)) \cap \bigcup_{v \in V} L_{\mathrm{NE},\Sigma}(\pi_v(\alpha)).$$

There is no unambiguous word in $L_{\mathrm{E},\Sigma}(\alpha) \setminus \{\varepsilon\}$ if and only if there is no unambiguous word in $L_{\mathrm{E},\Sigma}(\alpha) \setminus (\{\varepsilon\} \cup L_{\alpha,U,V})$.

Proof. It is sufficient to show that every word in $L_{\alpha,U,V}$ is ambiguous for α. Hence, for any $u \in U$, let w be any word in $L_{\mathrm{NE},\Sigma}(\pi_u(\alpha))$. Thus, there is a nonerasing morphism $\sigma : \mathbb{N}^* \to \Sigma^*$ with $\sigma(\pi_u(\alpha)) = w$. If, for a $v \in V$, $w \in L_{\mathrm{NE},\Sigma}(\pi_v(\alpha))$, then there additionally is a nonerasing morphism $\tau : \mathbb{N}^* \to \Sigma^*$ with $\tau(\pi_v(\alpha)) = w$. We define a morphism $\sigma' : \mathbb{N}^* \to \Sigma^*$ by, for every $x \in \mathbb{N}$,

$$\sigma'(x) := \begin{cases} \sigma(x) & , x \in \mathrm{var}(\pi_u(\alpha)), \\ \varepsilon & , \text{else}, \end{cases}$$

and a morphism $\tau' : \mathbb{N}^* \to \Sigma^*$ by, again for every $x \in \mathbb{N}$,

$$\tau'(x) := \begin{cases} \tau(x) & , x \in \mathrm{var}(\pi_v(\alpha)), \\ \varepsilon & , \text{else}. \end{cases}$$

Then $\sigma'(\alpha) = \sigma(\pi_u(\alpha)) = w = \tau(\pi_v(\alpha)) = \tau'(\alpha)$. Furthermore, because of the fact that (U, V) is a partition of $\mathcal{P}(\mathrm{var}(\alpha)) \setminus \{\emptyset\}$, it directly follows that $u \neq v$ and, thus, $\{x \in \mathrm{var}(\alpha) \mid \sigma'(x) \neq \varepsilon\} \neq \{x \in \mathrm{var}(\alpha) \mid \tau'(x) \neq \varepsilon\}$. Consequently, w is ambiguous for α. Since w, u and v were arbitrarily chosen, this directly implies that every word $w \in L_{\alpha,U,V}$ is ambiguous for α. □

It is a noteworthy property of Theorem 12 that it covers the ambiguity of both erasing and nonerasing morphisms and, hence, allows a unified view on both topics. However, for the latter case, Theorem 3 already gives a definite answer, indirectly stating that, for every *succinct* pattern α, there is *no* partition (U, V) of $\mathcal{P}(\mathrm{var}(\alpha)) \setminus \{\emptyset\}$ such that every word in $L_{\mathrm{E},\Sigma}(\alpha) \setminus (\{\varepsilon\} \cup L_{\alpha,U,V})$ is ambiguous for α. Thus, we can completely concentrate on prolix patterns when investigating applicability and consequences of Theorem 12.

From a practical point of view, Theorem 12 is not too helpful yet, as it merely reduces the number of words that need to be examined with regard to their ambiguity. Thus, it cannot be seen as an applicable characterisation of those patterns that have an unambiguous erasing morphism. On the other hand, it constitutes a promising starting point for further research on that topic, asking how U and V have to be be chosen such that $L_{\alpha,U,V}$ has maximal size and what a maximal $L_{\alpha,U,V}$ looks like for a given α. In this regard, it is worth mentioning that example patterns α and sets $U, V \subseteq \mathcal{P}(\mathrm{var}(\alpha)) \setminus \{\emptyset\}$ can be given where $L_{\alpha,U,V}$ is a nonempty subset of $L_{\mathrm{E},\Sigma}(\alpha)$ or even equals $L_{\mathrm{E},\Sigma}(\alpha) \setminus \{\varepsilon\}$.

Since, for any pattern α, $L_{\mathrm{E},\Sigma}(\alpha)$ is equivalent to a finite union of NE-pattern languages (see Theorem 2.1 by Jiang et al. [5]), Theorem 12 shows that the existence of unambiguous *erasing* morphisms strongly depends on equivalence and inclusion of certain finite unions of NE-pattern languages (or *nonerasing multi-pattern languages*, as they are called by Kari et al. [7]). This is not only a rather counter-intuitive insight, but it also gives an idea of how difficult the problem of the existence of unambiguous erasing morphisms might be. More precisely, even the decidability of the inclusion problem for ordinary terminal-free NE-pattern languages is open and includes some prominent open problems on pattern avoidability (cf. [5]). The inclusion of terminal-free NE-pattern languages is also known to depend on the size of the target alphabet, which fits very well with what is known for the subject of our paper (see, e. g., Theorem 11).

The following sufficient condition illustrates how Theorem 12 can be used to find criteria on the nonexistence of unambiguous erasing morphisms:

Corollary 4. *Let Σ be an alphabet, and let $\alpha \in \mathbb{N}^+$. If there exists a partition (U, V) of $\mathcal{P}(\mathrm{var}(\alpha)) \setminus \{\emptyset\}$ with*

$$L_{\mathrm{E},\Sigma}(\alpha) \setminus \{\varepsilon\} = \bigcup_{u \in U} L_{\mathrm{NE},\Sigma}(\pi_u(\alpha)) = \bigcup_{v \in V} L_{\mathrm{NE},\Sigma}(\pi_v(\alpha)),$$

then there is no unambiguous word in $L_{\mathrm{E},\Sigma}(\alpha) \setminus \{\varepsilon\}$.

Proof. Directly from Theorem 12. □

We finally wish to mention that Theorem 12 and Corollary 4 do not need to be based on a *partition* (U, V) of $\mathcal{P}(\mathrm{var}(\alpha)) \setminus \{\emptyset\}$. Alternatively, they could refer to *arbitrary* disjoint subsets U and V of $\mathcal{P}(\mathrm{var}(\alpha)) \setminus \{\emptyset\}$.

7. Conclusion and Open Problems

Concerning the ambiguity of erasing morphisms, the partition of patterns into morphically unerasable and erasable patterns (introduced and studied in Section 3) has similar importance as the partition into succinct and prolix patterns regarding the ambiguity of nonerasing morphisms: Both partitions characterise a vital property of strings, namely the (non)existence of moderately ambiguous morphisms (cf. Theorem 9 and Reidenbach [9]). While, in the case of nonerasing morphisms, this restricted ambiguity can additionally be turned into unambiguity, this does not hold for erasing morphisms since their ambiguity essentially depends on the size of the target alphabet (cf. Corollary 2 and, featuring a rather unexpected insight, Theorem 11).

A characterisation of those patterns that have an unambiguous erasing morphism is the main remaining open problem on the subject of the present paper, and even its mere decidability is still unresolved. Due to the insights summarised above, it seems evident that any solution to it requires concepts that significantly differ from the techniques used regarding moderate ambiguity. Section 6 reveals fundamental and quite surprising connections between the ambiguity of erasing morphisms and decision problems for nonerasing multi-pattern languages. An examination of these topics might be a helpful starting point for future studies.

References

[1] D. Freydenberger, D. Reidenbach, J. Schneider, Unambiguous morphic images of strings, International Journal of Foundations of Computer Science 17 (2006) 601–628.

[2] T. Harju, J. Karhumäki, Morphisms, in: G. Rozenberg, A. Salomaa (Eds.), Handbook of Formal Languages, volume 1, Springer, 1997, pp. 439–510.

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

[4] Š. Holub, Polynomial-time algorithm for fixed points of nontrivial morphisms, Discrete Mathematics 309 (2009) 5069–5076.

[5] T. Jiang, E. Kinber, A. Salomaa, K. Salomaa, S. Yu, Pattern languages with and without erasing, International Journal of Computer Mathematics 50 (1994) 147–163.

[6] T. Jiang, A. Salomaa, K. Salomaa, S. Yu, Decision problems for patterns, Journal of Computer and System Sciences 50 (1995) 53–63.

[7] L. Kari, A. Mateescu, G. Păun, A. Salomaa, Multi-pattern languages, Theoretical Computer Science 141 (1995) 253–268.

[8] A. Mateescu, A. Salomaa, Patterns, in: G. Rozenberg, A. Salomaa (Eds.), Handbook of Formal Languages, volume 1, Springer, 1997, pp. 230–242.

[9] D. Reidenbach, The Ambiguity of Morphisms in Free Monoids and its Impact on Algorithmic Properties of Pattern Languages, Ph.D. thesis, Fachbereich Informatik, Technische Universität Kaiserslautern, 2006. Logos Verlag, Berlin.

[10] D. Reidenbach, Discontinuities in pattern inference, Theoretical Computer Science 397 (2008) 166–193.
[11] D. Reidenbach, J. Schneider, Morphically primitive words, Theoretical Computer Science 410 (2009) 2148–2161.
[12] D. Reidenbach, J. Schneider, Restricted ambiguity of erasing morphisms, in: Proc. 14th International Conference on Developments in Language Theory, DLT 2010, volume 6224 of *Lecture Notes in Computer Science*, pp. 387–398.
[13] G. Rozenberg, A. Salomaa, Handbook of Formal Languages, volume 1, Springer, Berlin, 1997.
[14] J. Schneider, Unambiguous erasing morphisms in free monoids, Theoretical Informatics and Applications (RAIRO) 44 (2010) 193–208.