International Journal of Foundations of Computer Science


# On the Dual Post Correspondence Problem*†

Joel D. Day‡

*Department of Computer Science,*
*Loughborough University, Loughborough, UK*
*J.Day@lboro.ac.uk*

Daniel Reidenbach

*Department of Computer Science,*
*Loughborough University, Loughborough, UK*
*D.Reidenbach@lboro.ac.uk*

Johannes C. Schneider

*DIaLOGIKa GmbH, Pascalschacht 1,*
*66125 Saarbrücken, Germany*
*johannes.schneider@dialogika.de*

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

The Dual Post Correspondence Problem asks whether, for a given word $\alpha$, there exists a pair of distinct morphisms $\sigma, \tau$, one of which needs to be non-periodic, such that $\sigma(\alpha) = \tau(\alpha)$ is satisfied. This problem is important for the research on equality sets, which are a vital concept in the theory of computation, as it helps to identify words that are in trivial equality sets only.

Little is known about the Dual PCP for words $\alpha$ over larger than binary alphabets, especially for so-called ratio-primitive examples. In the present paper, we address this question in a way that simplifies the usual method, which means that we can reduce the intricacy of the word equations involved in dealing with the Dual PCP. Our approach yields large sets of words for which there exists a solution to the Dual PCP as well as examples of words over arbitrary alphabets for which such a solution does not exist.

*Keywords*: Morphisms; Equality sets; Dual Post Correspondence Problem; Periodicity forcing sets; Word equations; Ambiguity of morphisms

## 1. Introduction

The *equality set* $E(\sigma, \tau)$ of two morphisms $\sigma, \tau$ is the set of all words $\alpha$ that satisfy $\sigma(\alpha) = \tau(\alpha)$. Introduced by A. Salomaa [16] and Engelfriet and Rozenberg [6], they

*A preliminary version of this paper was presented at the conference DLT 2013.
†This work was supported by the London Mathematical Society, grant SC7-1112-02.
‡Corresponding author.

can be used to characterise crucial concepts in the theory of computation, such as the recursively enumerable set (see Culik II [1]) and the complexity classes P and NP (see Mateescu et al. [13]). Furthermore, since the famous undecidable *Post Correspondence Problem* (PCP) by Post [14] asks whether, for given morphisms $\sigma, \tau$, there *exists* a word $\alpha$ satisfying $\sigma(\alpha) = \tau(\alpha)$, it is simply the emptiness problem for equality sets.

Culik II and Karhumäki [2] study an alternative problem for equality sets, called the *Dual Post Correspondence Problem* (Dual PCP or DPCP for short): they ask whether, for any given word $\alpha$, there exist a pair of distinct morphisms $\sigma, \tau$ (called a *solution* to the DPCP) such that $\sigma(\alpha) = \tau(\alpha)$. Note that, in order for this problem to lead to a rich theory, at least one of the morphisms needs to be non-periodic. If a word does not have such a pair of morphisms, then it is called *periodicity forcing*, since the only solutions to the corresponding instance of the DPCP are periodic.

The Dual PCP is of particular interest for the research on equality sets as it identifies those words which belong to some non-trivial equality set (i.e., where at least one of the morphisms is periodic), and those which do not. The existence of the later words (namely the periodicity forcing ones) is a rather peculiar property of equality sets when compared to other types of formal languages, and it illustrates their combinatorial intricacy. In addition, the DPCP shows close connections to a special type of *word equations*, since a word $\alpha$ has a solution to the DPCP if and only if there exists a non-periodic solution to the word equation $\alpha = \alpha'$, where $\alpha'$ is renaming of $\alpha$. A further related concept is the *ambiguity of morphisms* (see, e. g., Freydenberger et al. [8, 7], Schneider [17]), since a word does not have a solution to the DPCP if and only if *every* non-periodic morphism is unambiguous for it.

Previous research on the DPCP has established its decidability and numerous insights into words over *binary* alphabets that do or do not have a solution. In contrast to this, there has been no work explicitly considering the case of larger alphabets, for which very little is known. The focus of the present paper is therefore to study the DPCP for words over arbitrarily large alphabets. Our main results, firstly, establish an approach to the problem that reduces the complexity of the word equations involved, secondly provide large classes of patterns (such as the morphically imprimitive ones [15]) that satisfy the DPCP, and, finally, show that the Dual PCP is non-trivial in the general case.

## 2. Definitions and Basic Observations

Let $\mathbb{N} := \{1, 2, \ldots\}$ be the set of natural numbers, and let $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. We often use $\mathbb{N}$ as an infinite alphabet of symbols. In order to distinguish between a word over $\mathbb{N}$ and a word over a (possibly finite) alphabet $\Sigma$, we call the former a *pattern*. Given a pattern $\alpha \in \mathbb{N}^*$, we call symbols occurring in $\alpha$ *variables* and denote the set of variables in $\alpha$ by $\mathrm{var}(\alpha)$. 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 $11 \cdot 2$. For a set $X$, the notation $|X|$ refers to the cardinality of $X$, and for a word $X$, $|X|$

stands for the length of $X$. By $|\alpha|_x$, we denote the number of occurrences of the variable $x$ in the pattern $\alpha$. Let $\alpha \in \{1, 2, \dots, n\}^*$ be a pattern. The *Parikh vector of* $\alpha$, denoted by $\mathrm{P}(\alpha)$, is the vector $(|\alpha|_1, |\alpha|_2, \dots, |\alpha|_n)$.

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 $\mathcal{A}$. Such a morphism $h$ is called *periodic* if and only if there exists a $v \in \mathcal{B}^*$ such that $h(a) \in v^*$ for every $a \in \mathcal{A}$. The morphisms $g$, $h : \mathcal{A}^* \to \mathcal{B}^*$ are *distinct* if and only if there exists an $a \in \mathcal{A}$ such that $g(a) \neq h(a)$. For the *composition* of two morphisms $g, h : \mathcal{A}^* \to \mathcal{A}^*$, we write $g \circ h$, i. e., for every $w \in \mathcal{A}^*$, $g \circ h(w) = g(h(w))$. Thus we have the following simple observations, which, to aid the exposition, are included formally. For alphabets $\mathcal{A}$, $\mathcal{B}$, let $f : \mathcal{A}^* \to \mathcal{B}^*$ and $g$, $h : \mathcal{B}^* \to \{\mathtt{a}, \mathtt{b}\}^*$ be morphisms. Then:

(Fact 1) *the morphism $g \circ f$ is periodic if and only if there exists a (primitive* [12]*) word $w \in \{\mathtt{a}, \mathtt{b}\}^*$ such that for each $a \in \mathcal{A}$, there exists an $n \in \mathbb{N}_0$ with $g(f(a)) = w^n$, and*

(Fact 2) *the morphisms $g \circ f$ and $h \circ f$ are distinct if and only if there exists an $a \in \mathcal{A}$ such that $g(f(a)) \neq h(f(a))$.*

In this paper, we usually consider morphisms $\sigma : \mathbb{N}^* \to \{\mathtt{a}, \mathtt{b}\}^*$ and morphisms $\varphi : \mathbb{N}^* \to \mathbb{N}^*$. For a set $N \subseteq \mathbb{N}$, the morphism $\pi_N : \mathbb{N}^* \to \mathbb{N}^*$ is defined by $\pi_N(x) := x$ if $x \in N$ and $\pi_N(x) := \varepsilon$ if $x \notin N$. Thus, for a pattern $\alpha \in \mathbb{N}^+$, $\pi_N(\alpha)$ is the *projection* of $\alpha$ to its subpattern $\pi_N(\alpha)$ consisting of variables in $N$ only.

Let $\alpha \in \mathbb{N}^+$. We call $\alpha$ *morphically imprimitive* if and only if there exist a pattern $\beta$ with $|\beta| < |\alpha|$ and morphisms $\varphi, \psi : \mathbb{N}^* \to \mathbb{N}^*$ satisfying $\varphi(\alpha) = \beta$ and $\psi(\beta) = \alpha$. If $\alpha$ is not morphically imprimitive, we call $\alpha$ *morphically primitive.* As demonstrated by Reidenbach and Schneider [15], the partition of the set of all patterns into morphically primitive and morphically imprimitive ones is vital in several branches of combinatorics on words and formal language theory, and some of our results in the main part of the present paper shall again be based on this notion. It is convenient to formally define the Dual PCP as a set:

**Definition 1.** *Let $\Sigma$ be an alphabet.* DPCP *is the set of all $\alpha \in \mathbb{N}^+$ such that there exist a non-periodic morphism $\sigma : \mathbb{N}^* \to \Sigma^*$ and an (arbitrary) morphism $\tau : \mathbb{N}^* \to \Sigma^*$ satisfying $\sigma(\alpha) = \tau(\alpha)$ and $\sigma(x) \neq \tau(x)$ for an $x \in \mathrm{var}(\alpha)$.*

Since all morphisms with unary target alphabets are periodic and since we can encode any $\Sigma$, $|\Sigma| \geq 2$, over a binary alphabet, we only consider $\Sigma := \{\mathtt{a}, \mathtt{b}\}$. The following proposition explains why in the definition of DPCP at least one morphism must be non-periodic.

**Proposition 2.** *For every $\alpha \in \mathbb{N}^+$ with $|\mathrm{var}(\alpha)| \geq 2$, there exist distinct (periodic) morphisms $\sigma, \tau : \mathbb{N}^* \to \{\mathtt{a}, \mathtt{b}\}^*$ satisfying $\sigma(\alpha) = \tau(\alpha)$.*

**Proof.** W. l. o. g., let $\mathrm{var}(\alpha) \supseteq \{1, 2\}$. Let $k := |\alpha|_1$, $l := |\alpha|_2$. Let the morphism $\sigma : \mathbb{N}^* \to \{\mathtt{a}, \mathtt{b}\}^*$ be defined by $\sigma(1) := \mathtt{a}^l$ and $\sigma(x) := \varepsilon$ for every variable $x \neq 1$. Similarly, let the morphism $\tau : \mathbb{N}^* \to \{\mathtt{a}, \mathtt{b}\}^*$ be defined by $\tau(2) := \mathtt{a}^k$ and $\tau(x) := \varepsilon$ for every variable $x \neq 2$. Thus, $\sigma(\alpha) = (\mathtt{a}^l)^k = \mathtt{a}^{lk} = (\mathtt{a}^k)^l = \tau(\alpha)$. □

Hence, allowing periodic morphisms would turn the Dual PCP into a trivial problem. Note also that for patterns $\alpha$ with $|\mathrm{var}(\alpha)| = 1$, every morphism is unambiguous so all unary patterns are periodicity forcing.

The Dual PCP can be extended to sets of patterns in a natural way: let $\beta_1$, $\beta_2$, ..., $\beta_n \in \mathbb{N}^*$ be patterns, and let $\Delta := \mathrm{var}(\beta_1) \cup \mathrm{var}(\beta_2) \cup \cdots \cup \mathrm{var}(\beta_n)$. The set $\{\beta_1, \beta_2, ..., \beta_n\}$ is *periodicity forcing* if, for every pair of distinct morphisms $\sigma$, $\tau : \Delta^* \to \{\mathtt{a}, \mathtt{b}\}^*$ which agree on every $\beta_i$ for $1 \leq i \leq n$, $\sigma$ and $\tau$ are periodic. The strong relationship between periodicity forcing words and periodicity forcing sets is immediately clear from the definitions, however this connection can be seen to be even stronger when considering ratio-imprimitive patterns.

In [2], ratio-imprimitivity is defined for binary patterns as follows: a pattern $\alpha \in \{1, 2\}^+$ is called *ratio-primitive* if and only if, for every proper prefix $\beta$ of $\alpha$, it is $|\beta|_1/|\beta|_2 \neq |\alpha|_1/|\alpha|_2$. Otherwise, $\alpha$ is called *ratio-imprimitive*. We extend this definition to arbitrarily large alphabets by considering the ratio of occurrences of all variables in the prefix $\beta$: a pattern $\alpha \in \mathbb{N}^+$ is ratio-primitive if and only if, for every proper prefix $\beta$ of $\alpha$, there does not exist a rational number $k$ such that $k\,\mathrm{P}(\beta) = \mathrm{P}(\alpha)$.

The significance of this definition is that if $\alpha = \beta \cdot \gamma$, where for some $k$, $\mathrm{P}(\alpha) = k\,\mathrm{P}(\beta)$, then for any two morphisms $\sigma$, $\tau$, the equality $\sigma(\alpha) = \tau(\alpha)$ holds if and only if $\sigma(\beta) = \tau(\beta)$ and $\sigma(\gamma) = \tau(\gamma)$. Therefore, for ratio-imprimitive patterns, periodicity forcing words are equivalent to non-unary periodicity forcing sets. It is often the 'truly unary' examples – the ratio-primitive words, which are the most challenging to classify when considering the Dual PCP – and these are therefore given particular attention.

From Culik II and Karhumäki [2] it is known that DPCP is decidable. From the literature on word equations and binary equality sets, it can be inferred that for any $i$, $j \in \mathbb{N}$, we have that $(1 \cdot 2)^i \cdot 1 \in$ DPCP [9], $1^i \cdot 2^j \in$ DPCP [2] and $1 \cdot 2^i \cdot 1 \in$ DPCP [11]. Note that, for $i, j > 1$, all these patterns are morphically primitive. Thus, the results are not trivially achievable by applying Corollary 10 in Section 4. Furthermore:

**Proposition 3 ([2])** *Every two-variable pattern of length 4 or less is in* DPCP. *Every renaming and/or mirror image of the patterns* $1 \cdot 2 \cdot 1 \cdot 1 \cdot 2$, $1 \cdot 2 \cdot 1 \cdot 2 \cdot 2$ *is not in* DPCP. *These are the only patterns of length 5 that are not in* DPCP. *In particular, the (morphically primitive) patterns* $1 \cdot 1 \cdot 2 \cdot 2 \cdot 2$, $1 \cdot 2 \cdot 1 \cdot 2 \cdot 1$, $1 \cdot 2 \cdot 2 \cdot 1 \cdot 1$ *and* $1 \cdot 2 \cdot 2 \cdot 2 \cdot 1$ *are in* DPCP.

**Proposition 4 ([3])** $1^2 \cdot 2^3 \cdot 1^2 \notin$ DPCP.

It is worth noting that the proof of the latter proposition takes about 9 pages. This illustrates how difficult it can be to show that certain example patterns do not belong to DPCP. In [2], Culik II and Karhumäki state without proof that any ratio-primitive pattern $\alpha \in (1^3 \cdot 1^* \cdot 2^3 \cdot 2^*)^2$ is not in DPCP.

While the above examples are partly hard to find, some general statements on DPCP and its complement can be obtained with little effort:

**Proposition 5.** *The following statements hold for any $\alpha$, $\beta \in \mathbb{N}^+$:*

*(1) Let $k \in \mathbb{N}$. Then $\alpha^k \in \mathrm{DPCP}$ if and only if $\alpha \in \mathrm{DPCP}$.*
*(2) If $\mathrm{var}(\alpha) \cap \mathrm{var}(\beta) = \emptyset$, then $\alpha\beta \in \mathrm{DPCP}$.*
*(3) Let $|\mathrm{var}(\alpha)| = 1$. Then $\alpha \notin \mathrm{DPCP}$.*
*(4) Let $V \subseteq \mathrm{var}(\alpha)$ with $\pi_V(\alpha) \in \mathrm{DPCP}$. Then $\alpha \in \mathrm{DPCP}$.*

**Proof.** Statements 1, 3, and 4 follow directly from the properties of morphisms. Consider statement 2. If $\alpha$ and $\beta$ are both unary, then the situation is covered by the example $1^i \cdot 2^j$ (see above). Otherwise, assume that $|\mathrm{var}(\alpha_1)| \geq 2$ (if not, then $|\mathrm{var}(\alpha_2)| \geq 2$, and the proof is analogous). W. l. o. g., let $\mathrm{var}(\alpha_1) \supseteq \{1, 2\}$ and $3 \in \mathrm{var}(\alpha_2)$. We choose the morphisms $\sigma$ and $\tau$ from the proof of Proposition 2 and define morphisms $\sigma', \tau' : \mathbb{N}^* \to \{\mathtt{a}, \mathtt{b}\}^*$ by $\sigma'(1) := \sigma(1)$, $\tau'(1) := \tau(1)$, $\sigma'(2) := \sigma(2)$, $\tau'(2) := \tau(2)$, $\sigma'(3) := \tau'(3) := \mathtt{b}$ and $\sigma'(x) := \tau'(x) := \varepsilon$ for every $x > 3$. Clearly, $\sigma'$ and $\tau'$ are non-periodic and distinct. Furthermore, $\sigma'(\alpha) = \mathtt{a}^{lk}\,\mathtt{b}^{|\alpha|_3} = \tau'(\alpha)$ and, thus, $\alpha \in \mathrm{DPCP}$. □

It follows immediately from the first statement that there are patterns of arbitrary length both in, and not in DPCP, and from the second statement, that patterns exist in DPCP over arbitrarily large alphabets. The equivalent statement for patterns not in DPCP is much harder to verify, and is addressed in Section 5. It is also worth noting that by the last statement, the discovery of one pattern not in DPCP directly leads to a multitude of patterns not in DPCP (namely, all of its subpatterns). On the other hand, this situation makes it very difficult to find such example patterns since arbitrary patterns easily contain subpatterns from DPCP.

## 3. A Characteristic Condition

The most direct way to decide on whether a pattern $\alpha$ is in DPCP is to solve the word equation $\alpha = \alpha'$, where $\alpha'$ is a renaming of $\alpha$ such that $\mathrm{var}(\alpha) \cap \mathrm{var}(\alpha') = \emptyset$. Indeed, the set of solutions corresponds exactly to the set of all pairs of morphisms which agree on $\alpha$. The pattern $\alpha$ is in DPCP if and only if there exists such a solution which is non-periodic. This explains why Culik II and Karhumäki [2] use Makanin's Algorithm for demonstrating the decidability of DPCP. Furthermore, it demonstrates why, in many respects, the more challenging questions often concern patterns *not* in DPCP. For such patterns, it is not enough to simply find a single non-periodic solution, but instead every single solution to the equation $\alpha = \alpha'$ must be accounted for. This is, in general, an extremely difficult and time consuming task.

This section presents an alternative approach which attempts to reduce the difficulties associated with such equations. To this end, we apply a morphism $\varphi : \mathbb{N}^* \to \mathbb{N}^*$ to a pattern $\alpha \notin$ DPCP, and we identify conditions that, if satisfied, yield $\varphi(\alpha) \notin$ DPCP. Our first main result characterises such morphisms $\varphi$:

**Theorem 6.** *Let $\alpha \in \mathbb{N}^+$ be a pattern not in* DPCP*, and let $\varphi : \mathrm{var}(\alpha)^* \to \mathbb{N}^*$ be a morphism. The pattern $\varphi(\alpha)$ is not in* DPCP *if and only if*

**(i)** *for every periodic morphism $\rho : \mathrm{var}(\alpha)^* \to \{\mathtt{a}, \mathtt{b}\}^*$ and for all distinct morphisms $\sigma, \tau : \mathrm{var}(\varphi(\alpha))^* \to \{\mathtt{a}, \mathtt{b}\}^*$ with $\sigma \circ \varphi(\alpha) = \rho(\alpha) = \tau \circ \varphi(\alpha)$, $\sigma$ and $\tau$ are periodic and*

**(ii)** *for every non-periodic morphism $\rho : \mathrm{var}(\alpha)^* \to \{\mathtt{a}, \mathtt{b}\}^*$ and for all morphisms $\sigma, \tau : \mathrm{var}(\varphi(\alpha))^* \to \{\mathtt{a}, \mathtt{b}\}^*$ with $\sigma \circ \varphi = \rho = \tau \circ \varphi$, $\sigma = \tau$.*

**Proof.** We define $\beta := \phi(\alpha)$, and we begin with the *if* direction. We assume to the contrary that $\beta \in$ DPCP. Hence, there exist distinct morphisms $\sigma', \tau' : \mathrm{var}(\beta)^* \to \{\mathtt{a}, \mathtt{b}\}^*$ such that $\sigma'(\beta) = \tau'(\beta)$, and at least one of these morphisms is non-periodic. According to Condition (i) of the Theorem, this implies that every morphism $\rho$ with $\sigma'(\beta) = \rho(\alpha) = \tau'(\beta)$ is non-periodic. This particularly holds for every morphism $\rho$ satisfying $\rho = \sigma' \circ \varphi$ or $\rho = \tau' \circ \varphi$. Thus, due to Condition (ii) of the Theorem, the fact that $\sigma'$ and $\tau'$ are distinct implies that $\sigma' \circ \varphi$ and $\tau' \circ \varphi$ are distinct. Consequently, $\sigma' \circ \varphi$ and $\tau' \circ \varphi$ are distinct non-periodic morphisms, and they satisfy $\sigma' \circ \varphi(\alpha) = \sigma'(\beta) = \tau'(\beta) = \tau' \circ \varphi(\alpha)$. This contradicts $\alpha \notin$ DPCP, and therefore $\beta \notin$ DPCP must be satisfied.

We continue with the *only if* direction. Let $\beta \notin$ DPCP, and let $\rho : \mathrm{var}(\alpha)^* \to \{\mathtt{a}, \mathtt{b}\}^*$ be any morphism. If $\rho$ is periodic, then we assume to the contrary that there exist distinct morphisms $\sigma', \tau' : \mathrm{var}(\beta)^* \to \{\mathtt{a}, \mathtt{b}\}^*$ such that at least one of these morphisms is non-periodic and $\sigma' \circ \varphi(\alpha) = \rho(\alpha) = \tau' \circ \varphi(\alpha)$. Hence, $\sigma'(\beta) = \tau'(\beta)$, which implies $\beta \in$ DPCP; this is a contradiction. Thus, all distinct morphisms $\sigma, \tau : \mathrm{var}(\beta)^* \to \{\mathtt{a}, \mathtt{b}\}^*$ with $\sigma \circ \varphi(\alpha) = \rho(\alpha) = \tau \circ \varphi(\alpha)$ must be periodic. This proves Condition (i) of the Theorem. If, on the other hand, $\rho$ is non-periodic, then we assume to the contrary that there are distinct morphisms $\sigma', \tau' : \mathrm{var}(\beta)^* \to \{\mathtt{a}, \mathtt{b}\}^*$ satisfying $\sigma' \circ \varphi = \rho = \tau' \circ \varphi$. Hence, $\sigma'(\beta) = \rho(\alpha) = \tau'(\beta)$. Furthermore, the non-periodicity of $\rho$ directly implies that $\sigma'$ and $\tau'$ are non-periodic. This again contradicts $\beta \notin$ DPCP. Consequently, for all $\sigma, \rho : \mathrm{var}(\beta)^* \to \{\mathtt{a}, \mathtt{b}\}^*$ with $\sigma \circ \varphi = \rho = \tau \circ \varphi$, $\sigma = \tau$ must hold true. This concludes the proof of Condition (ii) and, hence, of the Theorem. □

As briefly mentioned above, Theorem 6 shows that insights into the structure of DPCP can be gained in a manner that partly circumvents the solution of word equations. Instead, we can make use of prior knowledge on periodicity forcing words, which mainly exists for patterns over two variables, and expand this knowledge by studying the existence of morphisms $\varphi$ that *preserve non-periodicity* (i. e., if certain morphisms $\sigma$ are non-periodic, then $\sigma \circ \varphi$ needs to be non-periodic; see Condition (i)) and *preserve distinctness* (i. e., if certain morphisms $\sigma, \tau$ are distinct, then $\sigma \circ \varphi$ and $\tau \circ \varphi$ need to be distinct; see Condition (ii)).

While Theorem 6 provides a characterisation, it is mainly suitable when looking for periodicity forcing words. We make use of this in Section 5, where, due to our focus on the *if* direction of Theorem 6, the conditions may be simplified, allowing for example morphisms, and thus large classes of periodicity forcing words to be identified. Before we study this in more detail, we wish to consider patterns that are in DPCP in the next section.

## 4. On Patterns in DPCP

In the present section, we wish to establish major sets of patterns over arbitrarily many variables that are in DPCP. Our first criterion is based on so-called *ambiguity factorisations*, which are a generalisation of imprimitivity factorisations used by Reidenbach and Schneider [15] to characterise the morphically primitive patterns. Using this concept, we can give a strong sufficient condition for patterns in DPCP.

**Definition 7.** *Let $\alpha \in \mathbb{N}^+$. An* ambiguity factorisation (of $\alpha$) *is a mapping $f : \mathbb{N}^+ \to \mathbb{N}^n \times (\mathbb{N}^+)^n$, $n \in \mathbb{N}$, such that, for $f(\alpha) = (x_1, x_2, \dots, x_n; \gamma_1, \gamma_2, \dots, \gamma_n)$, there exist $\beta_0, \beta_1, \dots, \beta_n \in \mathbb{N}^*$ satisfying $\alpha = \beta_0\, \gamma_1\, \beta_1\, \gamma_2\, \beta_2 \dots \gamma_n\, \beta_n$ and*

**(i)** *for every $i \in \{1, 2, \dots, n\}$, $|\gamma_i| \geq 2$,*
**(ii)** *for every $i \in \{0, 1, \dots, n\}$ and for every $j \in \{1, 2, \dots, n\}$, $\mathrm{var}(\beta_i) \cap \mathrm{var}(\gamma_j) = \emptyset$,*
**(iii)** *for every $i \in \{1, 2, \dots, n\}$, $|\gamma_i|_{x_i} = 1$ and if $x_i \in \mathrm{var}(\gamma_{i'})$ for an $i' \in \{1, 2, \dots, n\}$, $\gamma_i = \delta_1\, x_i\, \delta_2$ and $\gamma_i' = \delta_1'\, x_i\, \delta_2'$, then $|\delta_1| = |\delta_1'|$ and $|\delta_2| = |\delta_2'|$.*

**Theorem 8.** *Let $\alpha \in \mathbb{N}^+$. If there exists an ambiguity factorisation of $\alpha$, then $\alpha \in$ DPCP.*

**Proof.** Let $f(\alpha) = (x_1, x_2, \dots, x_n;\, \gamma_1, \gamma_2, \dots, \gamma_n)$ be an ambiguity factorisation of $\alpha$. Hence, there exist $\beta_0, \beta_1, \dots, \beta_n \in \mathbb{N}^*$ satisfying $\alpha = \beta_0\, \gamma_1\, \beta_1\, \gamma_2\, \beta_2 \dots \gamma_n\, \beta_n$. We now consider the sets $B := \bigcup_{i \in \{0,1,\dots,n\}} \mathrm{var}(\beta_i)$, $\Gamma := \bigcup_{i \in \{1,2,\dots,n\}} \mathrm{var}(\gamma_i)$ and $X := \{x_1, x_2, \dots, x_n\}$. Furthermore, we define morphisms $\psi, \sigma, \tau : \mathbb{N}^* \to \{\mathtt{a}, \mathtt{b}\}^*$ in the following way:

- $\psi(x) := \mathtt{a}$ for $x \in X$, $\psi(x) := \mathtt{b}$ for $x \in \Gamma \setminus X$, and $\psi(x) := \varepsilon$ otherwise,
- $\sigma(x) := \mathtt{a}$ for $x \in X$, $\sigma(x) := \mathtt{bb}$ for $x \in \Gamma \setminus X$, and $\sigma(x) := \varepsilon$ otherwise, and
- $\tau(x_i) := \psi(\gamma_i)$ for $x_i \in \{x_1,\ x_2,\ ...,\ x_n\}$, $\tau(x) := \mathtt{b}$ for $x \in \Gamma \setminus X$, and $\tau(x) := \varepsilon$ otherwise.

Clearly, $\sigma$ and $\tau$ are non-periodic and distinct. Now, let $i \in \{1, 2, \dots, n\}$. It is $|\gamma_i|_x = 1$ for exactly one $x \in X$, namely $x = x_i$. Hence, $\gamma_i = \gamma_1\, x_i\, \gamma_2$ with $\gamma_1 \gamma_2 \in (\Gamma \setminus X)^+$. Thus, $\sigma(\gamma_i) = \mathtt{b}^{2|\gamma_1|}\, \mathtt{a}\, \mathtt{b}^{2|\gamma_2|} = \mathtt{b}^{|\gamma_1|}\, \mathtt{b}^{|\gamma_1|}\, \mathtt{a}\, \mathtt{b}^{|\gamma_2|}\, \mathtt{b}^{|\gamma_2|} = \mathtt{b}^{|\gamma_1|}\, \psi(\gamma_i)\, \mathtt{b}^{|\gamma_2|} = \tau(\gamma_i)$. Since this holds for every $i$ and since $\sigma(x) = \varepsilon = \tau(x)$ for every $x \in B$, we have $\sigma(\alpha) = \tau(\alpha)$, which proves $\alpha \in$ DPCP. $\square$

The following example illustrates Definition 7 and Theorem 8:

**Example 9.** *Let the pattern $\alpha$ be given by*

$$\alpha := \underbrace{\mathbf{1} \cdot 2 \cdot 2}_{\gamma_1} \cdot 3 \cdot \underbrace{2 \cdot \mathbf{4} \cdot 5 \cdot 2}_{\gamma_2} \cdot \underbrace{5 \cdot \mathbf{4} \cdot 2 \cdot 5}_{\gamma_3} \cdot 3 \cdot \underbrace{\mathbf{1} \cdot 2 \cdot 2}_{\gamma_4}$$

*This pattern has an ambiguity factorisation, as is implied by the marked $\gamma$ parts and the variables in bold face, which stand for the $x_i$.*

*We now consider two distinct non-periodic morphisms $\sigma$ and $\tau$, given by $\sigma(1) = \sigma(4) = \mathtt{a}$, $\sigma(2) = \sigma(5) = \mathtt{bb}$, $\sigma(3) = \varepsilon$ and $\tau(1) = \mathtt{abb}$, $\tau(4) = \mathtt{babb}$, $\tau(2) = \tau(5) = \mathtt{b}$, $\tau(3) = \varepsilon$. It can be verified with limited effort that $\sigma$ and $\tau$ agree on $\alpha$.* $\Diamond$

Since ambiguity factorisations are more general than the imprimitivity factorisations (see Reidenbach, Schneider [15]) used to characterise morphically imprimitive words, we can immediately conclude that this natural set of patterns is included in DPCP:

**Corollary 10.** *Let $\alpha \in \mathbb{N}^+$. If $\alpha$ is morphically imprimitive, then $\alpha \in$ DPCP.*

While ambiguity factorisations are a powerful tool, they are technically rather involved. In this respect, our next sufficient condition on patterns in DPCP is simpler, and expands on Statement 2 from Proposition 5.

**Proposition 11.** *Let $x, y, z \in \mathbb{N}$, and let $\alpha \in \{x, y, z\}^+$ be a pattern such that $\alpha = \alpha_0 z \alpha_1 z \dots \alpha_{n-1} z \alpha_n$, $n \in \mathbb{N}$. If,*

- *for every $i \in \{0, 1, \dots, n\}$, $\alpha_i = \varepsilon$ or $\mathrm{var}(\alpha_i) = \{x, y\}$, and*
- *for every $i, j \in \{0, 1, \dots, n\}$ with $\alpha_i \neq \varepsilon \neq \alpha_j$, $\frac{|\alpha_i|_x}{|\alpha_i|_y} = \frac{|\alpha_j|_x}{|\alpha_j|_y}$,*

*then $\alpha \in$ DPCP.*

**Proof.** W. l. o. g., let $\{x, y\} = \{1, 2\}$ and $z = 3$. Let $m \in \{0, 1, \dots, n\}$ such that $\alpha_m$ is of minimal length among all $\alpha_i \neq \varepsilon$. We choose the morphisms $\sigma$ and $\tau$ from the proof of Proposition 2 with $k := |\alpha_m|_1$ and $l := |\alpha_m|_2$ and set $\sigma(3) := \tau(3) := \mathtt{b}$. This makes $\sigma$ and $\tau$ non-periodic. With the same argumentation as in the proof of Proposition 2, it is $\sigma(\alpha_m) = \tau(\alpha_m)$. Since the second condition of Proposition 11 implies that, for every $i \in \{0, 1, \dots, n\}$ with $\alpha_i \neq \varepsilon$, $\frac{|\alpha_i|_x}{|\alpha_i|_y}$ is identical to $\frac{|\alpha_m|_x}{|\alpha_m|_y}$, it is $\sigma(\alpha_i) = \tau(\alpha_i)$ for every $i \in \{0, 1, \dots, n\}$ and, thus, $\sigma(\alpha) = \tau(\alpha)$, which proves $\alpha \in$ DPCP. $\square$

The following example pattern is covered by Proposition 11: $1 \cdot 1 \cdot 2 \cdot 2 \cdot 2 \cdot 3 \cdot 1 \cdot 2 \cdot 2 \cdot 1 \cdot 2 \cdot 3 \cdot 1 \cdot 1 \cdot 1 \cdot 1 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2$. Although Proposition 11 is restricted to three-variable patterns, it is worth mentioning that we can apply it to arbitrary patterns that have a three-variable subpattern of this structure. This is a direct consequence of Proposition 5 (Statement 4).

## 5. On Patterns Not in DPCP

As a result of the intensive research on binary equality sets, several examples of patterns over two variables are known not to be in DPCP (see Section 2). Hence, the most obvious question to ask is whether or not there exist such examples with more than two variables (and more generally, whether there exist examples for any given set of variables). The following results develop a structure for morphisms which map patterns not in DPCP to patterns over larger alphabets which are also not in DPCP, ultimately allowing for the inductive proof of Theorem 21, which provides a strong positive answer. A major advantage of the 'morphisms approach' presented in this section is that it facilitates the production of the more elusive ratio-primitive examples, since morphisms can easily preserve this property.

As discussed in Section 3, we simplify the conditions of Theorem 6, so that they ask the morphism $\varphi$ to be *non-periodicity preserving* and *distinctness preserving*:

**Lemma 12.** *Let $\Delta_1$, $\Delta_2$ be sets of variables. Let $\varphi : {\Delta_1}^* \to {\Delta_2}^*$ be a morphism such that, for every $x \in \Delta_2$, there exists a $y \in \Delta_1$ satisfying $x \in \mathrm{var}(\varphi(y))$, and*

**(i)** *for every non-periodic morphism $\sigma : {\Delta_2}^* \to \{\mathtt{a}, \mathtt{b}\}^*$, $\sigma \circ \varphi$ is non-periodic, and*
**(ii)** *for all distinct morphisms $\sigma$, $\tau : {\Delta_2}^* \to \{\mathtt{a}, \mathtt{b}\}^*$, where at least one is non-periodic, $\sigma \circ \varphi$ and $\tau \circ \varphi$ are distinct.*

*Then, for any $\alpha \notin$ DPCP with $\mathrm{var}(\alpha) = \Delta_1$, $\varphi(\alpha) \notin$ DPCP.*

**Proof.** Let $\alpha \notin$ DPCP be a pattern with $\mathrm{var}(\alpha) = \Delta_1$. Note that $\mathrm{var}(\varphi(\alpha)) = \Delta_2$. Assume to the contrary that $\varphi(\alpha) \in$ DPCP. Then there exist two distinct morphisms $\sigma$, $\tau : \mathrm{var}(\varphi(\alpha))^* \to \{\mathtt{a}, \mathtt{b}\}^*$ which agree on $\varphi(\alpha)$, and at least one of them is non-periodic. By Condition (i), at least one of $\sigma \circ \varphi$, $\tau \circ \varphi$ will be non-periodic and by Condition (ii), they are distinct. Clearly, if $\sigma$ and $\tau$ agree on $\varphi(\alpha)$, then $\sigma \circ \varphi$ and $\tau \circ \varphi$ agree on $\alpha$. Thus, $\sigma \circ \varphi$ and $\tau \circ \varphi$ are evidence that $\alpha \in$ DPCP. This is a contradiction, and so $\varphi(\alpha) \notin$ DPCP. □

Due to the nature of morphisms, it is apparent that, further than requiring that $\alpha \notin$ DPCP, the structure of $\alpha$ is not relevant.

**Remark 13.** *Condition (i) of Lemma 12 is identical to asking that $\sigma \circ \varphi$ is periodic if and only if $\sigma$ is periodic, since $\sigma \circ \varphi$ will always be periodic if $\sigma$ is periodic.*

We now investigate the existence and nature of morphisms $\varphi$ that satisfy both conditions. Each condition is relatively independent from the other, so it is appropriate to first establish classes of morphisms satisfying each one separately. Condition (i) is considered first. The satisfaction of Fact 1, and therefore Condition (i) of Lemma 12 relies on specific systems of word equations having only periodic solutions. The following lemma provides a tool for demonstrating exactly that.

**Lemma 14.** *(Lothaire* [12]*) All non-trivial, terminal-free word equations in two unknowns have only periodic solutions.*

In order to determine the satisfaction of Condition (i) of Lemma 12 for a particular morphism $\varphi : {\Delta_1}^* \to {\Delta_2}^*$, it is necessary to identify which morphisms $\sigma : {\Delta_2}^* \to \{\mathtt{a}, \mathtt{b}\}^*$ result in the composition $\sigma \circ \varphi$ being periodic. The next proposition gives the required characterisation.

**Proposition 15.** *Let $\Delta_1$ and $\Delta_2$ be sets of variables and let $\varphi : {\Delta_1}^* \to {\Delta_2}^*$, $\sigma : {\Delta_2}^* \to \{\mathtt{a}, \mathtt{b}\}^*$ be morphisms. For every $i \in \Delta_1$, let $\varphi(i) =: \beta_i$, and let $\{\gamma_1, \gamma_2, ..., \gamma_n\}$ be the set of all patterns $\beta_j$ such that $\sigma(\beta_j) \neq \varepsilon$. If $n < 2$, the composition $\sigma \circ \varphi$ is trivially periodic. For $n \geq 2$, $\sigma \circ \varphi$ is periodic if and only if there exist $k_1$, $k_2$, ... $k_n \in \mathbb{N}$ such that*

$$\sigma(\gamma_1)^{k_1} = \sigma(\gamma_2)^{k_2} = \cdots = \sigma(\gamma_n)^{k_n}. \tag{1}$$

**Proof.** Assume that $\sigma \circ \varphi$ is periodic. It follows from Fact 1 that there exists a word $w \in \Sigma^+$ such that for every $\gamma_i$, there exists an $m_i \in \mathbb{N}$ with $\sigma(\gamma_i) = w^{m_i}$. Hence there exist $k_1$, $k_2$, ..., $k_n \in \mathbb{N}$ such that equality 1 holds; simply choose $k_i := m_1 m_2 \cdots m_{i-1} m_{i+1} \cdots m_n$.

If there exist $k_1$, $k_2 \in \mathbb{N}$ such that $\sigma(\gamma_1)^{k_1} = \sigma(\gamma_2)^{k_2}$, then by Lemma 14 there exists a word $w \in \Sigma^+$ and numbers $m_1$, $m_2 \in \mathbb{N}_0$ such that $\sigma(\gamma_1) = w^{m_1}$ and $\sigma(\gamma_2) = w^{m_2}$. By continuing this argument, there must exist an $m_i \in \mathbb{N}_0$ for every $\gamma_i$, and therefore if equality 1 is satisfied, $\sigma \circ \varphi$ is periodic as required. □

Each term $\sigma(\gamma_i)$ in equality (1) corresponds directly to a word $\sigma \circ \varphi(j)$, for some $j \in \Delta_1$. The satisfaction of the system of equalities is identical to each word $\sigma \circ \varphi(i)$ sharing a primitive root, highlighting the nature of the relationship between $\sigma$ and the periodicity of $\sigma \circ \varphi$.

**Corollary 16.** *Let $\Delta_1$ and $\Delta_2$ be sets of variables, let $\varphi : {\Delta_1}^* \to {\Delta_2}^*$ be a morphism, and let $\varphi(i) =: \beta_i$ for every $i \in \Delta_1$. The morphism $\varphi$ satisfies Condition (i) of Lemma 12 if and only if, for every non-periodic morphism $\sigma : {\Delta_2}^* \to \{\mathtt{a}, \mathtt{b}\}^*$,*

**(i)** *there are at least two patterns $\beta_i$ such that $\sigma(\beta_i) \neq \varepsilon$, and*
**(ii)** *there do not exist $k_1$, $k_2$, ..., $k_n \in \mathbb{N}$ such that equality (1) is satisfied.*

**Proof.** Assume that $\varphi : {\Delta_1}^* \to {\Delta_2}^*$ is a morphism satisfying both conditions for the corollary. Then by Proposition 15, for every non-periodic morphism $\sigma : {\Delta_2}^* \to \{\mathtt{a}, \mathtt{b}\}^*$, the composition $\sigma \circ \varphi$ is also non-periodic. It follows directly that $\varphi$ satisfies Condition (i) of Lemma 12.

Assume that $\varphi$ does not satisfy Condition (i) of the Corollary. Then there exists a non-periodic morphism $\sigma : {\Delta_2}^* \to \{\mathtt{a}, \mathtt{b}\}^*$, such that there exists only one $i \in \Delta_1$ with $\sigma(\beta_i) \neq \varepsilon$. Therefore there exists only one $i \in \Delta_1$ such that $\sigma \circ \varphi(i) \neq \varepsilon$, so the composition $\sigma \circ \varphi$ is periodic. Thus there exists non-periodic $\sigma$ such that $\sigma \circ \varphi$ is periodic, and Condition (i) of Lemma 12 is not satisfied.

Assume that $\varphi$ does not satisfy Condition (ii) of the Corollary. Then it follows directly from Proposition 15 that there exists a non-periodic morphism $\sigma : {\Delta_2}^* \to$

$\{\mathtt{a},\mathtt{b}\}^*$ such that the composition $\sigma \circ \varphi$ is periodic and Condition (i) of Lemma 12 does not hold. □

Corollary 16 also provides a proof technique. Since any morphism $\varphi$ will erase exactly one of finitely many combinations of the patterns $\beta_1, \beta_2, ..., \beta_m$, it is clear that the satisfaction of Condition (ii) of Corollary 16 will always rely on finitely many cases. By considering all possible sets $\{\gamma_1, \gamma_2, ..., \gamma_n\}$, infinitely many morphisms $\sigma$ can be accounted for in a finite, often very concise manner. Thus, it becomes much simpler to demonstrate that there cannot exist a non-periodic morphism $\sigma$ such that $\sigma \circ \varphi$ is periodic, and thus that Condition (i) of Lemma 12 is satisfied. We now give an example of such an approach.

**Example 17.** *Let* $\Delta_1 := \{1,2,3,4\}$ *and let* $\Delta_2 := \{5,6,7,8\}^*$. *Let* $\varphi : {\Delta_1}^* \to {\Delta_2}^*$ *be the morphism given by* $\varphi(1) := 5 \cdot 6$, $\varphi(2) := 6 \cdot 5$, $\varphi(3) := 5 \cdot 6 \cdot 7 \cdot 7$ *and* $\varphi(4) := 6 \cdot 8 \cdot 8 \cdot 5$. *Consider all non-periodic morphisms* $\sigma : \{5,6,7,8\}^* \to \{\mathtt{a},\mathtt{b}\}^*$. *Note that if* $\sigma(5 \cdot 6) \neq \varepsilon$ *then* $\sigma(6 \cdot 5) \neq \varepsilon$ *and vice-versa. Also note that since* $\sigma$ *is non-periodic, there must be at least two variables* $x$ *such that* $\sigma(x) \neq \varepsilon$. *So if either* $\sigma(5 \cdot 6 \cdot 7 \cdot 7) \neq \varepsilon$, *or* $\sigma(6 \cdot 8 \cdot 8 \cdot 5) \neq \varepsilon$, *there must be at least one other pattern* $\beta_j$ *with* $\sigma(\beta_j) \neq \varepsilon$. *Therefore, for any non-periodic morphism* $\sigma$, *there exists a minimum of two patterns* $\beta_i$ *such that* $\sigma(\beta_i) \neq \varepsilon$. *Now consider all possible cases.*

*Assume first that* $\sigma(5 \cdot 6) = \varepsilon$. *Clearly* $\sigma(5) = \sigma(6) = \varepsilon$, *so* $\sigma(6 \cdot 5) = \varepsilon$. *Since* $\sigma$ *is non-periodic,* $\sigma(7) \neq \varepsilon$ *and* $\sigma(8) \neq \varepsilon$. *By Proposition 15,* $\sigma \circ \varphi$ *is periodic if and only if there exist* $k_1$, $k_2 \in \mathbb{N}$ *such that* $\sigma(7 \cdot 7)^{k_1} = \sigma(8 \cdot 8)^{k_2}$. *By Lemma 14, this is the case only if* $\sigma$ *is periodic and this is a contradiction, so* $\sigma \circ \varphi$ *is non-periodic.*

*Assume* $\sigma(5 \cdot 6) \neq \varepsilon$ *(so* $\sigma(6 \cdot 5) \neq \varepsilon$, $\sigma(6 \cdot 8 \cdot 8 \cdot 5) \neq \varepsilon$, *and* $\sigma(5 \cdot 6 \cdot 7 \cdot 7) \neq \varepsilon$*), then by Proposition 15, the composition* $\sigma \circ \varphi$ *is periodic if and only if there exist* $k_1$, $k_2$, $k_3$, $k_4 \in \mathbb{N}$ *such that*

$$\sigma(5 \cdot 6)^{k_1} = \sigma(6 \cdot 5)^{k_2} = \sigma(6 \cdot 8 \cdot 8 \cdot 5)^{k_3} = \sigma(5 \cdot 6 \cdot 7 \cdot 7)^{k_4} \tag{2}$$

*By Lemma 14, the first equality only holds if there exist a word* $w \in \{\mathtt{a},\mathtt{b}\}^*$ *and numbers* $p$, $q \in \mathbb{N}_0$ *such that* $\sigma(5) = w^p$ *and* $\sigma(6) = w^q$. *Thus, equality (2) is satisfied if and only if* $w^{k_1(p+q)} = (w^q \cdot \sigma(8 \cdot 8) \cdot w^p)^{k_3}$ *and* $w^{k_1(p+q)} = (w^{p+q} \cdot \sigma(7 \cdot 7))^{k_4}$. *By Lemma 14, this is only the case if there exist* $r$, $s \in \mathbb{N}$ *such that* $\sigma(7) = w^s$ *and* $\sigma(8) = w^r$. *Thus,* $\sigma$ *is periodic, which is a contradiction, so the composition* $\sigma \circ \varphi$ *is non-periodic.*

*All possibilities for non-periodic morphisms* $\sigma$ *have been exhausted, so for any non-periodic morphism* $\sigma : \{5, 6, 7, 8\}^* \to \{\mathtt{a},\mathtt{b}\}^*$, *the composition* $\sigma \circ \varphi$ *is also non-periodic and* $\varphi$ *satisfies Condition (i) of Lemma 12.* ◊

Condition (ii) of Lemma 12 is now considered. Fact 2 shows that it relies on the (non-)existence of distinct, non-periodic morphisms which agree on a set of patterns (more precisely, the set of morphic images of single variables). The following proposition provides a characterisation for morphisms that satisfy the condition.

**Proposition 18.** *Let $\Delta_1$, $\Delta_2$ be sets of variables, and let $\varphi : {\Delta_1}^* \to {\Delta_2}^*$ be a morphism. For every $i \in \Delta_1$, let $\varphi(i) =: \beta_i$. The morphism $\varphi$ satisfies Condition (ii) of Lemma 12 if and only if $\{\beta_1, \beta_2, \dots, \beta_n\}$ is a periodicity forcing set.*

**Proof.** Let $\Xi_i$ be the set of all pairs of distinct morphisms $\sigma$, $\tau : {\Delta_2}^* \to \{\mathtt{a},\mathtt{b}\}^*$, where at least one is non-periodic, satisfying $\sigma(\beta_i) = \tau(\beta_i)$. The set $\{\beta_1, \beta_2, ..., \beta_n\}$ is periodicity forcing if and only if $\Xi_1 \cap \Xi_2 \cap \cdots \cap \Xi_n = \emptyset$. Assume that $\Xi_1 \cap \Xi_2 \cap \cdots \cap \Xi_n$ is non-empty. Then, by definition, there exist distinct morphisms $\sigma$, $\tau : {\Delta_2}^* \to \{\mathtt{a},\mathtt{b}\}^*$, where at least one is non-periodic, such that $\sigma(\beta_i) = \tau(\beta_i)$ for every $i \in \Delta_1$. It follows from Fact 2 that $\sigma \circ \varphi$ and $\tau \circ \varphi$ are not distinct, so Condition (ii) does not hold. If, instead, $\Xi_1 \cap \Xi_2 \cap \cdots \cap \Xi_n$ is empty, then for any two non-periodic morphisms $\sigma$, $\tau : {\Delta_2}^* \to \{\mathtt{a},\mathtt{b}\}^*$, there exists a $j \in \Delta_1$ such that $\sigma(\beta_j) \neq \tau(\beta_j)$. Therefore, by Fact 2, $\sigma \circ \varphi$ and $\tau \circ \varphi$ must be distinct, and Condition (ii) holds as required. $\square$

Proposition 18 demonstrates the difference between the word equations involved in directly finding patterns not in DPCP and the word equations that need to be considered when using Lemma 12. Furthermore, it shows the impact of the choice of $\alpha$ on the complexity of applying the Lemma. However, it does not immediately provide a nontrivial morphism $\varphi$ that satisfies Condition (ii) of Lemma 12. Therefore, we consider the following technical tool:

**Proposition 19.** *Let $\Delta_1$, $\Delta_2$ be sets of variables, and let $\varphi : {\Delta_1}^* \to {\Delta_2}^*$ be a morphism. For every $k \in \Delta_1$, let $\varphi(k) =: \beta_k$ and let $\beta_i \notin \mathrm{DPCP}$ for some $i \in \Delta_1$. For every $x \in \Delta_2 \backslash \mathrm{var}(\beta_i)$, let there exist $\beta_j$ and patterns $\gamma_1$, $\gamma_2$, such that $\beta_j = \gamma_1 \cdot \gamma_2$ and*

*(1) $x \in \mathrm{var}(\gamma_1)$, and for every $y \in \mathrm{var}(\gamma_1)$ with $y \neq x$, $y \in \mathrm{var}(\beta_i)$,*
*(2) $\gamma_1 \notin \mathrm{DPCP}$ with $|\mathrm{var}(\gamma_1)| \geq |\mathrm{var}(\beta_i)|$,*
*(3) $\mathrm{P}(\gamma_2)$ and $\mathrm{P}(\beta_i)$ are linearly dependent.*

*Then $\varphi$ satisfies Condition (ii) of Lemma 12.*

**Proof.** Firstly, note that by Condition (2), if any of the $\gamma_1$ are unary, then $\beta_i$ must be unary. This results in a trivial situation, so we assume that $\beta_i$ (and thus also any pattern $\gamma_1$) contains at least two distinct variables.

W. l. o. g. let $\Delta_1 := \{1, 2, ..., n\}$. Assume that $\varphi$ is a morphism satisfying the conditions for the Proposition, and assume to the contrary that $\varphi$ does not satisfy Condition (ii) of Lemma 12. Hence by Proposition 18, the set $\{\beta_k \mid k \in \Delta_1\}$ is not periodicity forcing. This is equivalent to the statement

$$\Xi_1 \cap \Xi_2 \cap \cdots \cap \Xi_n \neq \emptyset$$

where, for any $k \in \Delta_1$, $\Xi_k$ is the set of all pairs of distinct morphisms $\sigma$, $\tau : {\Delta_2}^* \to \{\mathtt{a},\mathtt{b}\}^*$ such that at least one is non-periodic and $\sigma(\beta_k) = \tau(\beta_k)$. Thus we have two distinct morphisms $\sigma$, $\tau : {\Delta_2}^* \to \{\mathtt{a},\mathtt{b}\}^*$, such that at least one is non-periodic, and $\sigma(\beta_k) = \tau(\beta_k)$ for every $k \in \Delta_1$. Since $\beta_i \notin$ DPCP, this is only possible if either

(1) $\sigma(x) = \tau(x)$ for every $x \in \mathrm{var}(\beta_i)$, or
(2) there exists an $x \in \mathrm{var}(\beta_i)$ such that $\sigma(x) \neq \tau(x)$, and $\sigma$ and $\tau$ are periodic over $\mathrm{var}(\beta_i)$.

Let $\beta_j \neq \beta_i$ be arbitrary. Let $\gamma_1$, $\gamma_2$ be defined accordingly. Note that, by Condition (3), because $\sigma(\beta_i)$ and $\tau(\beta_i)$ have the same length, so do $\sigma(\gamma_2)$ and $\tau(\gamma_2)$. Similarly, since $\sigma(\beta_j)$ and $\tau(\beta_j)$ share the same length, then so do $\sigma(\gamma_1)$ and $\tau(\gamma_1)$. Moreover, because $\sigma$ and $\tau$ agree on $\beta_i$, they also agree on $\gamma_2$. It follows that $\sigma(\gamma_1) = \tau(\gamma_1)$.

Note that, by Condition (2), there is at most one variable in $\mathrm{var}(\beta_i)$ which does not occur in $\gamma_1$. Furthermore there is exactly one variable in $\mathrm{var}(\gamma_1)$ which does not occur in $\beta_i$ by definition. Consider Case (1): that $\sigma$ and $\tau$ are identical over $\mathrm{var}(\beta_i)$. Then there is exactly one variable over which they are not identical in $\mathrm{var}(\gamma_1)$. This implies $\sigma(\gamma_1) \neq \tau(\gamma_1)$ which is a contradiction.

On the other hand, if $\sigma$ and $\tau$ do not agree on all variables in $\mathrm{var}(\beta_i)$, then by the same reasoning there must be at least two variables on which they disagree. By Condition (2), this implies that at least one of those variables occurs in $\gamma_1$. Furthermore, since we now consider Case (2), there exists some word $w$ such that $\sigma(x), \tau(x) \in \{w\}^*$ for all $x \in \mathrm{var}(\beta_i)$. However, at least one morphism (w. l. o. g. let it be $\sigma$) is non-periodic, so there exists a variable $y \notin \mathrm{var}(\beta_i)$ such that $\sigma(y) \notin \{w\}^*$.

Since we chose $\beta_j$ arbitrarily, we can assume w. l. o. g. that $y$ occurs in $\beta_j$ and therefore in $\gamma_1$ (it cannot occur in $\gamma_2$ due to Condition (3)). This implies that $\sigma$ is non-periodic over $\gamma_1$. Since $\sigma$ and $\tau$ are also distinct over $\gamma_1$, and due to the fact that $\gamma_1 \notin$ DPCP, it follows that $\sigma(\gamma_1) \neq \tau(\gamma_1)$. Hence $\sigma(\beta_j) \neq \tau(\beta_j)$, which is a contradiction.

Therefore no pair of morphisms in $\Xi_i$ can be in $\Xi_1 \cap \Xi_2 \cap ...\Xi_{i-1} \cap \Xi_{i+1} \cap ...\Xi_n$, so $\Xi_1 \cap \Xi_2 \cap \cdots \cap \Xi_n = \emptyset$ and the set $\{\beta_i \mid i \in \Delta_1\}$ is periodicity forcing. Hence $\varphi$ satisfies Condition (ii) of Lemma 12 as required. $\square$

The following example demonstrates the structure given in Proposition 19. It is chosen such that it also satisfies Corollary 16, providing a basis for the construction given in Theorem 21.

**Example 20.** *Let* $\Delta_1 := \{4, 5\}$*, and let* $\Delta_2 := \{1, 2, 3\}$*. Let* $\varphi : {\Delta_1}^* \to {\Delta_2}^*$ *be the morphism given by* $\varphi(4) = \beta_4 := 1 \cdot 2 \cdot 1 \cdot 1 \cdot 2$ *and* $\varphi(5) = \beta_5 := \gamma_1 \cdot \gamma_2$ *where* $\gamma_1 := 1 \cdot 3 \cdot 1 \cdot 1 \cdot 3$ *and* $\gamma_2 := 2 \cdot 1 \cdot 1 \cdot 2 \cdot 1$*. Notice that* $\beta_4$ *and* $\gamma_1$ *are not in* DPCP*. Let* $\sigma$, $\tau : \{1, 2, 3\}^* \to \{\mathtt{a},\mathtt{b}\}^*$ *be distinct morphisms, at least one of which is non-periodic, that agree on* $\beta_4$*. By definition of* DPCP*, this is only possible if* $\sigma$ *and* $\tau$ *agree on, or are periodic over* $\{1, 2\}$*.*

*If* $\sigma$ *and* $\tau$ *agree on* $\{1, 2\}$*, then they agree on* $\gamma_2$*. This means that* $\sigma(\gamma_1 \cdot \gamma_2) = \tau(\gamma_1 \cdot \gamma_2)$ *if and only if* $\sigma(1 \cdot 3 \cdot 1 \cdot 1 \cdot 3) = \tau(1 \cdot 3 \cdot 1 \cdot 1 \cdot 3)$*. Furthermore* $\sigma$ *and* $\tau$

*are distinct, so cannot agree on* 3*. However, since* $\sigma(1) = \tau(1)$ *but* $\sigma(3) \neq \tau(3)$*, this cannot be the case, therefore* $\sigma \circ \varphi$ *and* $\tau \circ \varphi$ *are distinct.*

*Note that if* $\sigma$ *and* $\tau$ *agree on exactly one variable in* $\{1, 2\}$*, then they cannot agree on* $\beta_4$*. Consider the case that* $\sigma$ *and* $\tau$ *do not agree on* 1 *or* 2*. Then they must be periodic over* $\{1, 2\}$*, so* $\sigma(2 \cdot 1 \cdot 1 \cdot 2 \cdot 1) = \tau(2 \cdot 1 \cdot 1 \cdot 2 \cdot 1)$ *and, as a consequence,* $\sigma$ *and* $\tau$ *agree on* $\gamma_1 \cdot \gamma_2$ *if and only if they agree on* $\gamma_1 = 1 \cdot 3 \cdot 1 \cdot 1 \cdot 3$*. However, due to the non-periodicity of* $\sigma$ *or* $\tau$*,* $\sigma(3)$ *or* $\tau(3)$ *must have a different primitive root (see Lothaire* [12]*) to* $\sigma(1)$ *or* $\tau(1)$*, respectively. This means that* $\sigma$ *and* $\tau$ *are distinct over* $\{1, 3\}$*, and at least one of them must be non-periodic over* $\{1, 3\}$*. This implies that* $\sigma$ *and* $\tau$ *cannot agree on* $\gamma_1$*, and therefore* $\sigma \circ \varphi$ *and* $\tau \circ \varphi$ *are distinct.*

*Hence, there do not exist two distinct morphisms, at least one of which is non-periodic, that agree on* $1 \cdot 2 \cdot 1 \cdot 1 \cdot 2$ *and* $1 \cdot 3 \cdot 1 \cdot 1 \cdot 3 \cdot 2 \cdot 1 \cdot 1 \cdot 2 \cdot 1$*. These patterns, thus, form a periodicity forcing set, and, by Proposition 18, the morphism* $\varphi$ *satisfies Condition (ii) of Lemma 12.* $\Diamond$

It is now possible to state the following theorem, the proof of which provides a construction for ratio-primitive patterns not in DPCP over an arbitrary number of variables. It is worth noting that for ratio-imprimitive examples, the result is not as difficult to obtain. Results by Holub and Kortelainen [10] can be extended to produce examples over arbitrary alphabets by showing that so-called test sets of the set of permutations of $x_1 \cdot x_2 \cdots x_n$ can modified to produce periodicity forcing words (see [5]). These examples, however, are both highly ratio-imprimitive and highly restricted. It is a major advantage of the following construction is that by using a ratio-primitive pre-image (e.g., $1 \cdot 2 \cdot 1 \cdot 1 \cdot 2$), one can obtain ratio-primitive examples over any alphabet: a much stronger statement that the Dual PCP is non-trivial in the general case. Another is that by using morphisms such as the one constructed in the following proof, it is possible to obtain large and varied classes of examples over any alphabet.

**Theorem 21.** *There are ratio-primitive patterns of arbitrarily many variables not in* DPCP*.*

**Proof.** It is already established that there exist patterns over two variables which are not in DPCP. Proceed by induction, and assume that there exists a pattern over $n$ variables not in DPCP with $n \geq 2$, and let $\alpha \in \{1,\ 2,\ ...,\ n\}^*$ be such a pattern. Let $\varphi_n : \{1, 2, ..., n\}^* \to \{1, 2, ...,\ n+1\}^*$ be the morphism such that $\varphi_n(1) := 1 \cdot 2 \cdot 1 \cdot 1 \cdot 2$, and for $2 \leq x \leq n$, $\varphi_n(x) := 1 \cdot (x+1) \cdot 1 \cdot 1 \cdot (x+1) \cdot 2 \cdot 1 \cdot 1 \cdot 2 \cdot 1$. It is shown first that $\varphi_n$ satisfies Condition (i) of Lemma 12.

Let $\sigma : \{1, 2, 3, ..., n+1\}^* \to \{\mathtt{a}, \mathtt{b}\}^*$ be an arbitrary non-periodic morphism. Let $\gamma_1, \gamma_2, ..., \gamma_m$ be the patterns $\varphi_n(i)$ such that $\sigma(\varphi_n(i)) \neq \varepsilon$ with $1 \leq i \leq n$. Note that if $\sigma(1 \cdot 2 \cdot 1 \cdot 1 \cdot 2) \neq \varepsilon$, then $\sigma(\varphi_n(i)) \neq \varepsilon$ for every $i \in \{1, 2, \ldots, n\}$. If $\sigma(1 \cdot 2 \cdot 1 \cdot 1 \cdot 2) = \varepsilon$, then $\sigma(1) = \sigma(2) = \varepsilon$. Since $\sigma$ is non-periodic, there must exist $x,\ y \in \{3, 4, ...,\ n\}$ with $x \neq y$ such that $\sigma(x) \neq \varepsilon$ and $\sigma(y) \neq \varepsilon$. It follows that $\sigma(\varphi(x-1)) \neq \varepsilon$, and $\sigma(\varphi(y-1)) \neq \varepsilon$. In either case $m \geq 2$ whenever $\sigma$ is non-periodic, so $\varphi$ satisfies Condition (i) of Corollary 16.

Condition (ii) of Corollary 16, and therefore Condition (i) of Lemma 12 is satisfied by $\varphi$ if and only if there do not exist $k_1$, $k_2$, ..., $k_m \in \mathbb{N}$, such that

$$\sigma(\gamma_1)^{k_1} = \sigma(\gamma_2)^{k_2} = \cdots = \sigma(\gamma_m)^{k_m}. \tag{3}$$

Assume to the contrary that equality 3 holds. If $\sigma(1 \cdot 2 \cdot 1 \cdot 1 \cdot 2) = \varepsilon$, then $\sigma(1) = \sigma(2) = \varepsilon$. Since every term in equality 3 will have at most one other unknown, namely $\sigma(x)$ with $3 \leq x \leq n+1$, each individual equality $\sigma(\gamma_j)^{k_j} = \sigma(\gamma_{j+1})^{k_{j+1}}$ will be in exactly two unknowns. It follows by Lemma 14 that $\sigma$ must be periodic, which is a contradiction.

If, on the other hand, $\sigma(1 \cdot 2 \cdot 1 \cdot 1 \cdot 2) \neq \varepsilon$, then equality 3 can be expressed as a series of equalities of the form:

$$\sigma(1 \cdot 2 \cdot 1 \cdot 1 \cdot 2)^{k_i} = \sigma(1 \cdot x \cdot 1 \cdot 1 \cdot x \cdot 2 \cdot 1 \cdot 1 \cdot 2 \cdot 1)^{k_j} \tag{4}$$

where $3 \leq x \leq n+1$. By comparing the suffix of length $|\sigma(1)|+|\sigma(2)|$ on either side, $\sigma(1)\sigma(2) = \sigma(2)\sigma(1)$, so, by Lemma 14, there exists $w \in \{\mathtt{a},\mathtt{b}\}^+$ and $n$, $m \in \mathbb{N}_0$ such that $\sigma(1) = w^n$ and $\sigma(2) = w^m$. It follows that equality 4 can be expressed as

$$w^{k_i(3n+2m)} = (w^n \cdot \sigma(x) \cdot w^{2n} \cdot \sigma(x) \cdot w^{3n+2m})^{k_j}$$

which is an equation in two unknowns. Thus, there exists an $l \in \mathbb{N}_0$ such that $\sigma(x) = w^l$ for $3 \leq x \leq n+1$, and $\sigma$ is periodic, which is a contradiction. Consequently, equality 3 cannot be satisfied by a non-periodic morphism, hence $\varphi$ satisfies Condition (ii) of Corollary 16, and Condition (i) of Lemma 12.

Condition (ii) of Lemma 12 is now considered. Note that $\varphi(1) = 1 \cdot 2 \cdot 1 \cdot 1 \cdot 2 \notin$ DPCP. For every $x \in \{3, 4, ..., n+1\}$, there exists an $i \in \{1, 2, ..., n\}$ such that $\varphi(i) = \gamma_1 \cdot \gamma_2$ with $\gamma_1 = 1 \cdot x \cdot 1 \cdot 1 \cdot x$ and $\gamma_2 = 2 \cdot 1 \cdot 1 \cdot 2 \cdot 1$. Thus, $x \in \mathrm{var}(\gamma_1)$, and for every $y \in \mathrm{var}(\gamma_1)$ with $y \neq x$, $y \in \mathrm{var}(\varphi(1))$. Also, $1 \cdot x \cdot 1 \cdot 1 \cdot x = \gamma_1 \notin$ DPCP, and $|\mathrm{var}(\gamma_1)| = |\mathrm{var}(\varphi(1))|$. Furthermore, $|\gamma_2|_z = |\varphi(1)|_z$ for every $z \in \{1, 2, ..., n+1\}$. Therefore, by Proposition 19, $\varphi$ also satisfies Condition (ii) of Lemma 12.

Consequently, $\varphi(\alpha) \notin$ DPCP by Lemma 12, and $|\mathrm{var}(\varphi(\alpha))| = n+1$. So if there exists a pattern with $n$ variables not in DPCP, there exists a pattern with $n+1$ variables also not in DPCP. By induction, there exist patterns of arbitrarily many variables not in DPCP as required. It is not difficult to see that all the patterns obtained in this sequence are ratio-primitive. □

Note that while the above theorem focuses on ratio-primitive examples, corresponding classes of ratio-imprimitive periodicity forcing words can be obtained simply by substituting a ratio-imprimitive periodicity forcing word, such as $1{\cdot}2{\cdot}1{\cdot}1{\cdot}2{\cdot}2$, for $\alpha$. Thus, the above method also yields large and varied classes of ratio-imprimitive periodicity forcing words over arbitrarily large alphabets.

*Acknowledgements*

The authors wish to thank the anonymous referees of both the present and the conference version [4] of this paper for their helpful remarks and suggestions. Furthermore, numerous enlightening discussions with Paul C. Bell on an earlier version of this paper are gratefully acknowledged.

## References

[1] K. Culik II. A purely homomorphic characterization of recursively enumerable sets. *Journal of the ACM*, 26:345–350, 1979.

[2] K. Culik II and J. Karhumäki. On the equality sets for homomorphisms on free monoids with two generators. *Theoretical Informatics and Applications (RAIRO)*, 14:349–369, 1980.

[3] E. Czeizler, Š. Holub, J. Karhumäki, and M. Laine. Intricacies of simple word equations: An example. *International Journal of Foundations of Computer Science*, 18:1167–1175, 2007.

[4] J. D. Day, J. C. Schneider, and D. Reidenbach. On the dual post correspondence problem. In *Proc. 17th International Conference on Developments in Language Theory, DLT 2013*, volume 7907 of *Lecture Notes in Computer Science*, pages 167–178, 2013.

[5] J. D. Day, J. C. Schneider, and D. Reidenbach. Periodicity forcing words. In *Proc. 9th International Conference on Words, WORDS 2013*, volume 8079 of *Lecture Notes in Computer Science*, pages 107–118, 2013.

[6] J. Engelfriet and G. Rozenberg. Equality languages and fixed point languages. *Information and Control*, 43:20–49, 1979.

[7] D.D. Freydenberger, H. Nevisi, and D. Reidenbach. Weakly unambiguous morphisms. *Theoretical Computer Science*, 448:21–40, 2012.

[8] 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.

[9] J. Hadravová and Š. Holub. Large simple binary equality words. In *Proc. 12th International Conference on Developments in Language Theory, DLT 2008*, volume 5257 of *Lecture Notes in Computer Science*, pages 396–407, 2008.

[10] Š. Holub and J. Kortelainen. Linear size test sets for certain commutative languages. *Theoretical Informatics and Applications (RAIRO)*, 35:453–475, 2001.

[11] J. Karhumäki and E. Petre. On some special equations on words. Technical Report 583, Turku Centre for Computer Science, TUCS, 2003.

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

[13] A. Mateescu, A. Salomaa, K. Salomaa, and S. Yu. P, NP, and the Post Correspondence Problem. *Information and Computation*, 121:135–142, 1995.

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

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

[16] A. Salomaa. Equality sets for homomorphisms of free monoids. *Acta Cybernetica*, 4:127–139, 1978.

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