On the Dual Post Correspondence Problem*

Joel D. Day[1], Daniel Reidenbach**[1], and Johannes C. Schneider[2]

[1] Department of Computer Science, Loughborough University,
Loughborough, Leicestershire, LE11 3TU, UK
J.Day-10@student.lboro.ac.uk
D.Reidenbach@lboro.ac.uk

[2] Fachbereich Informatik, Technische Universität Kaiserslautern,
Postfach 3049, 67653 Kaiserslautern, Germany
jschneider@informatik.uni-kl.de

Abstract. The Dual Post Correspondence Problem asks whether, for a given word α, there exists a pair of distinct morphisms σ, τ, 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 α over larger than binary alphabets. 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 σ, τ is the set of all words α that satisfy $\sigma(\alpha) = \tau(\alpha)$. Equality sets were introduced by A. Salomaa [13] and Engelfriet and Rozenberg [4], and they 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. [10]). Furthermore, since the famous undecidable *Post Correspondence Problem* (PCP) by Post [11] asks whether, for given morphisms σ, τ, there *exists* a word α 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 α, there exist a pair of distinct morphisms

* This work was supported by the London Mathematical Society, grant SC7-1112-02.

** Corresponding author.

σ, τ (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 Post Correspondence Problem is of particular interest for the research on equality sets as it helps to identify words that can only occur in *trivial* equality sets (i. e., equality sets $E(\sigma, \tau)$ where σ or τ are periodic). The existence of these 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 α has a solution to the DPCP if and only there exists a non-periodic solution to the word equation $\alpha = \alpha'$, where α' is renaming of α. A further related concept is the *ambiguity of morphisms* (see, e. g., Freydenberger et al. [6, 5], Schneider [14]). Research on this topic mainly asks whether, for a given word α, there exists a morphism σ that is *unambiguous* for it, i. e., there is no other morphism τ satisfying $\sigma(\alpha) = \tau(\alpha)$. Using this terminology, a word does not have a solution to the DPCP 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, for larger alphabets, it is not even known whether the problem is nontrivial, i. e., whether there are periodicity forcing words, and if so, what they look like. It is the purpose of the present paper to study the DPCP for words over arbitrary alphabets. Our main results shall, firstly, establish an approach to the problem that reduces the complexity of the word equations involved, secondly, demonstrate that most words are not periodicity forcing and why that is the case and, thirdly, prove that the DPCP is nontrivial for all alphabet sizes.

Due to space constraints, all proofs have been omitted from this paper.

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 Σ, we call the former a *pattern*. Given a pattern $\alpha \in \mathbb{N}^*$, we call symbols occurring in α *variables* and denote the set of variables in α 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 α. Let $\alpha \in \{x_1, x_2, \ldots, x_n\}^*$ be a pattern. The *Parikh vector of* α, denoted by $\mathrm{P}(\alpha)$, is the vector $(|\alpha|_{x_1}, |\alpha|_{x_2}, \ldots, |\alpha|_{x_n})$.

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 $a \in \mathcal{A}$. For the *composition* of two morphisms $g, h : \mathcal{A}^* \to \mathcal{A}^*$, we write $g \circ h$, i. e., for every $w \in \mathcal{A}^*$, $g \circ h(w) = g(h(w))$. 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 α to its subpattern $\pi_N(\alpha)$ consisting of variables in N only. Let $\Delta \subset \mathbb{N}$ be a set of variables and Σ be an alphabet. Then two morphisms $\sigma, \tau : \Delta^* \to \Sigma^*$ are *distinct* if and only if there exists an $x \in \Delta$ such that $\sigma(x) \neq \tau(x)$.

Let $\alpha \in \mathbb{N}^+$. We call α *morphically imprimitive* if and only if there exist a pattern β with $|\beta| < |\alpha|$ and morphisms $\varphi, \psi : \mathbb{N}^* \to \mathbb{N}^*$ satisfying $\varphi(\alpha) = \beta$ and $\psi(\beta) = \alpha$. If α is not morphically imprimitive, we call α *morphically primitive.* As demonstrated by Reidenbach and Schneider [12], 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.

We now formally define the Dual PCP as a set:

Definition 1. *Let Σ 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 $x \in \mathrm{var}(\alpha)$.*

We wish to investigate what patterns $\alpha \in \mathbb{N}^+$ are contained in DPCP, and what patterns are not. Since all morphisms with unary target alphabets are periodic and since we can encode any Σ, $|\Sigma| \geq 2$, over $\{\mathtt{a}, \mathtt{b}\}$, we choose $\Sigma := \{\mathtt{a}, \mathtt{b}\}$ from now on.

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 (periodic) morphisms $\sigma, \tau : \mathbb{N}^* \to \{\mathtt{a}, \mathtt{b}\}^*$ satisfying $\sigma(\alpha) = \tau(\alpha)$.*

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

In the literature, patterns not in DPCP are called *periodicity forcing* since they force every pair of morphisms that agree on the pattern to be periodic. This notion can be extended to sets of patterns in a natural way: Let $\Delta \subset \mathbb{N}$ be a set of variables, and let β_1, β_2, ..., $\beta_n \in \Delta^+$ be patterns. The set $\{\beta_1, \beta_2, ..., \beta_n\}$ is *periodicity forcing* if, for every pair of distinct morphisms σ, $\tau : \Delta^* \to \{\mathtt{a}, \mathtt{b}\}^*$ which agree on every β_i for $1 \leq i \leq n$, σ and τ are periodic.

From Culik II and Karhumäki [2] it is known that DPCP is decidable. Furthermore, the following specific results on two-variable patterns that are or are not in DPCP can be derived from the literature on word equations and binary equality sets:

Proposition 3 **([2]).** *Every two-variable pattern of length 4 or less is in* DPCP*. Every renaming or mirrored version 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*.*

Furthermore, we have the following examples of longer patterns.

Proposition 4 **([7]).** *For any* $i \in \mathbb{N}$*,* $(1 \cdot 2)^i \cdot 1 \in \mathrm{DPCP}$*.*

Proposition 5 **([2]).** *For any* $i, j \in \mathbb{N}$*,* $1^i \cdot 2^j \in \mathrm{DPCP}$*.*

Proposition 6 **([8]).** *For any* $i \in \mathbb{N}$*,* $1 \cdot 2^i \cdot 1 \in \mathrm{DPCP}$*.*

Note that, for $i, j > 1$, the three propositions above give morphically primitive example patterns. Thus, the results are not trivially achievable by applying Corollary 18 in Section 4.

Proposition 7 **([3]).** $1^2 \cdot 2^3 \cdot 1^2 \notin \mathrm{DPCP}$*.*

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. A pattern $\alpha \in \{1, 2\}^+$ is called *ratio-primitive* if and only if, for every proper prefix β of α, it is $|\beta|_1/|\beta|_2 \neq |\alpha|_1/|\alpha|_2$. Otherwise, α is called *ratio-imprimitive.*

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

Proposition 8. *If* $\alpha \in \mathrm{DPCP}$*, then, for every* k*,* $\alpha^k \in \mathrm{DPCP}$*. If* $\alpha \notin \mathrm{DPCP}$*, then, for every* k*,* $\alpha^k \notin \mathrm{DPCP}$*.*

Proposition 9. *If* $\alpha, \beta \in \mathrm{DPCP}$ *with* $\mathrm{var}(\alpha) \cap \mathrm{var}(\beta) = \emptyset$*, then* $\alpha\beta \in \mathrm{DPCP}$*.*

If we apply Proposition 8 to existing examples, then we can state the following insight:

Corollary 10. *There are patterns of arbitrary length in* DPCP*. There are patterns of arbitrary length not in* DPCP*.*

The existing literature on the Dual PCP mainly studies two-variable patterns. In contrast to this, as mentioned in Section 1, we wish to investigate the structure of DPCP for patterns over any numbers of variables. In this regard, we can state a number of immediate observations:

Proposition 11. *Let* $\alpha \in \mathbb{N}^+$*,* $|\mathrm{var}(\alpha)| = 1$*. Then* $\alpha \notin \mathrm{DPCP}$*.*

It is easy to give example patterns with three or more variables that belong to DPCP. Proposition 19 in Section 4 gives a construction principle. Furthermore, as soon as a pattern α is projectable to a subpattern $\beta \in \mathrm{DPCP}$, also $\alpha \in \mathrm{DPCP}$.

Proposition 12. *Let* $\alpha \in \mathbb{N}^+$ *and* $V \subseteq \mathrm{var}(\alpha)$ *with* $\pi_V(\alpha) \in \mathrm{DPCP}$*. Then* $\alpha \in \mathrm{DPCP}$*.*

On the other hand, this implies that every $\alpha \notin \mathrm{DPCP}$ must not be projectable to a subpattern from DPCP.

Corollary 13. *Let* $\alpha \notin \mathrm{DPCP}$*. Then for every* $V \subseteq \mathrm{var}(\alpha)$*,* $\pi_V(\alpha) \notin \mathrm{DPCP}$*.*

Consequently, on the one hand, 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 α is in DPCP is to solve the word equation $\alpha = \alpha'$, where α' is a renaming of α 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 α. The pattern α 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. It is generally extremely difficult to determine the complete solution set to such an equation, and as a result, only a limited class of examples is known.

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 $\alpha \notin \mathrm{DPCP}$, and we identify conditions that, if satisfied, yield $\varphi(\alpha) \notin \mathrm{DPCP}$.

The main result of this section characterises such morphisms φ:

Theorem 14. *Let* $\alpha \in \mathbb{N}^+$ *be a pattern that is 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* $\sigma \circ \varphi(\alpha) = \rho(\alpha) = \tau \circ \varphi(\alpha)$*,* σ *and* τ *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* $\sigma \circ \varphi = \rho = \tau \circ \varphi$*,* $\sigma = \tau$*.*

As briefly mentioned above, Theorem 14 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 patterns that are not in DPCP, which mainly exists for patterns over two variables, and expand this knowledge by studying the existence of morphisms φ that *preserve non-periodicity* (i. e., if certain morphisms σ are non-periodic, then $\sigma \circ \varphi$ needs to be non-periodic; see Condition (i)) and *preserve distinctness* (i. e., if certain morphisms σ, τ are distinct, then $\sigma \circ \varphi$ and $\tau \circ \varphi$ need to be distinct; see Condition (ii)).

Theorem 14 can be used to characterise the patterns in DPCP, but it is mainly suitable as a tool to find patterns that are *not* in DPCP. We shall study this option in Section 5, where we, due to our focus on the *if* direction of Theorem 14, can drop the additional conditions on non-periodicity and distinctness preserving morphisms φ that are postulated by the Theorem. In addition to reducing the need for studying word equations, the use of morphisms to generate examples not in DPCP shall prove to have another key benefit; since morphisms can be applied to infinitely many pre-image patterns, the construction of a single morphism automatically produces an infinite set of examples. This process can be applied iteratively – with morphisms providing new examples of patterns which can then potentially be used as the pre-images for the same, or other morphisms. Before we study this in more details, 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 [12] to characterise the morphically primitive patterns.

Definition 15. *Let $\alpha \in \mathbb{N}^+$. An* ambiguity factorisation (of α) *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 $|\delta_2| = |\delta'_2|$.*

Using this concept, we now can give a strong sufficient condition for patterns in DPCP:

Theorem 16. *Let $\alpha \in \mathbb{N}^+$. If there exists an ambiguity factorisation of α, then $\alpha \in$ DPCP.*

The following example illustrates Definition 15 and Theorem 16:

Example 17. Let the pattern α 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 partition, as is implied by the marked γ parts and the variables in bold face, which stand for the x_i.

We now consider two distinct non-periodic morphisms σ and τ, 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 σ and τ agree on α. $\Diamond$

Since ambiguity partitions are more general than imprimitivity partitions, we can immediately conclude that a natural set of patterns is included in DPCP:

Corollary 18. *Let $\alpha \in \mathbb{N}^+$. If α is morphically imprimitive, then $\alpha \in$ DPCP.*

Since most patterns are morphically imprimitive (see Reidenbach and Schneider [12]), this implies that most patterns are in DPCP, which confirms our intuitive considerations at the beginning of Section 3.

While ambiguity partitions are a powerful tool, they are technically rather involved. In this respect, our next sufficient condition on patterns in DPCP is much simpler, since it merely asks whether a pattern can be split in two factors that do not have any variables in common:

Proposition 19. *Let $\alpha \in \mathbb{N}^+$, $|\operatorname{var}(\alpha)| \geq 3$. If, for some $\alpha_1, \alpha_2 \in \mathbb{N}^+$ with $\operatorname{var}(\alpha_1) \cap \operatorname{var}(\alpha_2) = \emptyset$, $\alpha = \alpha_1\, \alpha_2$, then $\alpha \in$ DPCP.*

Note that it is possible to extend Proposition 19 quite substantially since the same technique can be applied to, e g., the pattern $\alpha := \alpha_1\alpha_2\alpha_1\alpha_2$ and much more sophisticated types of patterns where certain factors have disjoint variable sets and can therefore be allocated to different periodic morphisms each. The following proposition is such an extension of Proposition 19.

Proposition 20. *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 $\operatorname{var}(\alpha_i) = \{x, y\}$, 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.

The following example pattern is covered by Proposition 20: $1 \cdot 1 \cdot 2 \cdot 2 \cdot 2 \cdot 3 \cdot 1 \cdot 2 \cdot 2 \cdot 1 \cdot 2 \cdot 3 \cdot 1 \cdot 1 \cdot 1 \cdot 1 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2$. Although Proposition 20 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 12.

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 with more variables which are also not in DPCP, ultimately allowing for the inductive proof of Theorem 33, which provides a strong positive answer.

As discussed in Section 3, this is accomplished by simplifying the conditions of Theorem 14, so that they ask the morphism φ to be (i) *non-periodicity preserving* and (ii) *distinctness-preserving*:

Lemma 21. *Let Δ_1, Δ_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 $y \in \Delta_1$ such that $x \in \mathrm{var}(\varphi(y))$, 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 σ, $\tau : {\Delta_2}^* \to \{\mathtt{a}, \mathtt{b}\}^*$, 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.

Remark 22. Condition (i) of Lemma 21 is identical to asking that $\sigma\circ\varphi$ is periodic if and only if σ is periodic, since if σ is periodic, then $\sigma\circ\varphi$ will always be periodic as well.

While Lemma 21 provides a clear proof technique for demonstrating that a given pattern is not in DPCP, the conditions are abstract, and it does not directly lead to any new examples. The next step, therefore, is to investigate the existence and nature of morphisms φ which satisfy both conditions.

Since the main focus of the following results is concerned with properties of compositions of morphisms, the following two facts are included formally.

Fact 23. Let Δ_1, Δ_2 be sets of variables. let $\varphi : {\Delta_1}^* \to {\Delta_2}^*$ and $\sigma : {\Delta_2}^* \to \{\mathtt{a}, \mathtt{b}\}^*$ be morphisms. The morphism $\sigma \circ \varphi$ is periodic if and only if there exists a (primitive) word $w \in \Sigma^*$ such that for each $i \in \Delta_1$, there exists an $n \in \mathbb{N}_0$ with $\sigma(\varphi(i)) = w^n$.

Fact 24. Let Δ_1, Δ_2 be sets of variables. let $\varphi : {\Delta_1}^* \to {\Delta_2}^*$ and $\sigma : {\Delta_2}^* \to \{\mathtt{a}, \mathtt{b}\}^*$ be morphisms. The morphisms $\sigma \circ \varphi$ and $\tau \circ \varphi$ are distinct if and only if there exists a variable $i \in \Delta_1$ such that $\sigma(\varphi(i)) \neq \tau(\varphi(i))$.

Facts 23 and 24 highlight how properties such as periodicity and distinctness of a composition of two morphisms can be determined by observing certain properties of specific sets of patterns. Since the conditions in Lemma 21 rely only on these properties, it is apparent that, further than requiring that $\alpha \notin$ DPCP, the structure of α is not relevant. It is instead dependent on $\mathrm{var}(\alpha)$.

Each condition from Lemma 21 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 23, and therefore Condition (i) of Lemma 21 relies on specific systems of word equations having only periodic solutions. The following proposition provides a tool for demonstrating exactly that.

Proposition 25. *(Lothaire [9]) 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 21 for a particular morphism $\varphi : {\Delta_1}^* \rightarrow {\Delta_2}^*$, it is necessary to identify which morphisms $\sigma : {\Delta_2}^* \rightarrow \{\mathtt{a}, \mathtt{b}\}^*$ result in the composition $\sigma \circ \varphi$ being periodic. The next proposition gives the required characteristic condition on σ for $\sigma \circ \varphi$ to be periodic. Each term $\sigma(\gamma_i)$ in equality (1) below 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, allowing the relationship between σ and the periodicity of $\sigma \circ \varphi$ to be expressed formally.

Proposition 26. *Let Δ_1 and Δ_2 be sets of variables and let $\varphi : {\Delta_1}^* \rightarrow {\Delta_2}^*$, $\sigma : {\Delta_2}^* \rightarrow \{\mathtt{a}, \mathtt{b}\}^*$ be morphisms. For every $i \in \Delta_1$, let $\varphi(i) := \beta_i$, and let $\{\gamma_1, \gamma_2, \ldots \gamma_n\}$ be the set of all patterns β_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 $k_1, k_2, \ldots k_n \in \mathbb{N}$ such that*

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

Corollary 27. *Let Δ_1 and Δ_2 be sets of variables, let $\varphi : {\Delta_1}^* \rightarrow {\Delta_2}^*$ be a morphism, and let $\varphi(i) := \beta_i$ for every $i \in \Delta_1$. The morphism φ satisfies Condition (i) of Lemma 21 if and only if, for every non-periodic morphism $\sigma : {\Delta_2}^* \rightarrow \{\mathtt{a}, \mathtt{b}\}^*$,*

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

$$\sigma(\gamma_1)^{k_1} = \sigma(\gamma_2)^{k_2} = \cdots = \sigma(\gamma_n)^{k_n} \tag{2}$$

where $\{\gamma_1, \gamma_2, \ldots, \gamma_n\}$ is the set of all patterns β_i such that $\sigma(\beta_i) \neq \varepsilon$.

Corollary 27 also provides a proof technique. Since there are finitely many combinations of $\beta_1, \beta_2, \ldots, \beta_m$, it is clear that the satisfaction of Condition (ii) of Corollary 27 will always rely on finitely many cases. By considering all possible sets $\{\gamma_1, \gamma_2, \ldots.\gamma_n\}$, infinitely many morphisms can be accounted for in a finite and often very concise manner. Thus, it becomes much simpler to demonstrate that there cannot exist a non-periodic morphism σ such that $\sigma \circ \varphi$ is periodic, and therefore that Condition (i) of Lemma 21 is satisfied. We now give an example of such an approach.

Example 28. Let $\Delta_1 := \{1, 2, 3, 4\}$ and let $\Delta_2 := \{5, 6, 7, 8\}^*$. Let $\varphi : {\Delta_1}^* \rightarrow {\Delta_2}^*$ 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 $\sigma : \{5, 6, 7, 8\}^* \rightarrow \{\mathtt{a}, \mathtt{b}\}^*$. Note that if $\sigma(5 \cdot 6) \neq \varepsilon$ then $\sigma(6 \cdot 5) \neq \varepsilon$ and vice-versa. Also note that since σ 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 β_j with $\sigma(\beta_j) \neq \varepsilon$. Therefore, for any non-periodic morphism σ, there exists a minimum of two patterns β_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 σ is non-periodic, $\sigma(7) \neq \varepsilon$ and $\sigma(8) \neq \varepsilon$. By Proposition 26, $\sigma \circ \varphi$ 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 Proposition 25, this is the case only if σ 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 26, 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

$$\sigma(5 \cdot 6)^{k_1} = \sigma(6 \cdot 5)^{k_2} = \sigma(6 \cdot 8 \cdot 8 \cdot 5)^{k_3} = \sigma(5 \cdot 6 \cdot 7 \cdot 7)^{k_4} \tag{3}$$

By Proposition 25, 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 (3) 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 Proposition 25, 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, σ is periodic, which is a contradiction, so the composition $\sigma \circ \varphi$ is non-periodic.

All possibilities for non-periodic morphisms σ 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 φ satisfies Condition (i) of Lemma 21. $\diamondsuit$

Condition (ii) of Lemma 21 is now considered. Fact 24 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 which satisfy the condition.

Proposition 29. *Let Δ_1, Δ_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 φ satisfies Condition (ii) of Lemma 21 if and only if $\{\beta_1, \beta_2, \ldots, \beta_n\}$ is a periodicity forcing set.*

Proposition 29 facilitates a formal comparison of the word equations involved in directly finding patterns not in DPCP and the word equations that need to be considered when using Lemma 21. Furthermore, it shows the impact of the choice of α on the complexity of applying the Lemma. However, it does not immediately provide a nontrivial morphism φ that satisfies Condition (ii) of Lemma 21. Therefore, we consider the following technical tool:

Proposition 30. *Let Δ_1, Δ_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 $x \in \Delta_2 \setminus \mathrm{var}(\beta_i)$, let there exist β_j and patterns γ_1, γ_2, such that $\beta_j = \gamma_1 \cdot \gamma_2$ and*

(i) *$x \in \mathrm{var}(\gamma_1)$, and for every $y \in \mathrm{var}(\gamma_1)$ with $y \neq x$, $y \in \mathrm{var}(\beta_i)$,*
(ii) *$\gamma_1 \notin \mathrm{DPCP}$ with $|\mathrm{var}(\gamma_1)| \geq |\mathrm{var}(\beta_i)|$,*

(iii) $\mathrm{P}(\gamma_2)$ *and* $\mathrm{P}(\beta_i)$ *are linearly dependent.*

Then φ *satisfies Condition (ii) of Lemma 21.*

The following example demonstrates the structure given in Proposition 30. It is chosen such that it also satisfies Corollary 27, allowing for the construction given in Proposition 32.

Example 31. 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 $\varphi(5) := \gamma_1 \cdot \gamma_2$ 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 β_4 and γ_1 are not in DPCP. Let σ, $\tau : \{1, 2, 3\}^* \to \{\mathtt{a}, \mathtt{b}\}^*$ be distinct morphisms, at least one of which is non-periodic, that agree on β_4. By definition of DPCP, this is only possible if σ and τ agree on, or are periodic over $\{1, 2\}$.

If σ and τ agree on $\{1, 2\}$, then they agree on γ_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 σ and τ 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 σ and τ agree on exactly one variable in $\{1, 2\}$, then they cannot agree on β_4. Consider the case that σ and τ do not agree on 1 or 2. Then they must be periodic over $\{1, 2\}$, so $\sigma(2 \cdot 1 \cdot 1 \cdot 2 \cdot 1) = \sigma(1 \cdot 2 \cdot 1 \cdot 1 \cdot 2)$ (and likewise for τ). It follows that $\sigma(2 \cdot 1 \cdot 1 \cdot 2 \cdot 1) = \tau(2 \cdot 1 \cdot 1 \cdot 2 \cdot 1)$ and, as a consequence, σ and τ 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 σ or τ, $\sigma(3)$ or $\tau(3)$ must have a different primitive root to $\sigma(1)$ or $\tau(1)$, respectively. This means that σ and τ are distinct over $\{1, 3\}$, and at least one of them must be non-periodic over $\{1, 3\}$. This implies that σ and τ cannot agree on γ_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 29, the morphism φ satisfies Condition (ii) of Lemma 21. ♢

The next proposition introduces a pattern over three variables which is not in DPCP. This not only demonstrates that this is possible for patterns over more than two variables, but provides the basis for the construction given in Theorem 33, which shows that there are patterns of arbitrarily many variables not in DPCP.

Proposition 32. *The pattern* $1 \cdot 2 \cdot 1 \cdot 1 \cdot 2 \cdot 1 \cdot 3 \cdot 1 \cdot 1 \cdot 3 \cdot 2 \cdot 1 \cdot 1 \cdot 2 \cdot 1 \cdot 1 \cdot 2 \cdot 1 \cdot 1 \cdot 2 \cdot 1 \cdot 2 \cdot 1 \cdot 1 \cdot 2 \cdot 1 \cdot 3 \cdot 1 \cdot 1 \cdot 3 \cdot 2 \cdot 1 \cdot 1 \cdot 2 \cdot 1$ *is not in* DPCP.

It is now possible to state the following theorem, the proof of which provides a construction for a pattern not in DPCP over an arbitrary number of variables. This is achieved by considering, for any $n \geq 2$, the morphism $\varphi_n : \{1, 2, \dots, n\}^* \to \{1, 2, \dots, n+1\}^*$, given by $\varphi_n(1) := 1 \cdot 2 \cdot 1 \cdot 1 \cdot 2$, and for $2 \leq x \leq n$, $\varphi_n(x) := 1 \cdot (x+1) \cdot 1 \cdot 1 \cdot (x+1) \cdot 2 \cdot 1 \cdot 1 \cdot 2 \cdot 1$. This morphisms satisfies the conditions for Lemma 21, i. e., it maps any n-variable pattern that is not in DPCP to an $n+1$-variable pattern that is also not in DPCP. It follows that if there exists a pattern with n variables not in DPCP, then there exists a pattern with n variables not in DPCP. Thus, by induction, there exist such patterns for any number of variables.

Theorem 33. *There are patterns of arbitrarily many variables not in* DPCP.

Hence, we may conclude that the Dual PCP is nontrivial for all alphabets with at least two variables, and we can show this in a constructive manner.

Acknowledgements. The authors wish to thank the anonymous referees 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. Engelfriet and G. Rozenberg. Equality languages and fixed point languages. *Information and Control*, 43:20–49, 1979.
5. D.D. Freydenberger, H. Nevisi, and D. Reidenbach. Weakly unambiguous morphisms. *Theoretical Computer Science*, 448:21–40, 2012.
6. D.D. Freydenberger, D. Reidenbach, and J.C. Schneider. Unambiguous morphic images of strings. *International Journal of Foundations of Computer Science*, 17:601–628, 2006.
7. J. Hadravová and Š. Holub. Large simple binary equality words. *International Journal of Foundations of Computer Science*, 23:1385–1403, 2012.
8. J. Karhumäki and E. Petre. On some special equations on words. Technical Report 583, Turku Centre for Computer Science, TUCS, 2003. `http://tucs.fi:8080/publications/insight.php?id=tKaPe03a`.
9. M. Lothaire. *Combinatorics on Words*. Addison-Wesley, Reading, MA, 1983.
10. A. Mateescu, A. Salomaa, K. Salomaa, and S. Yu. P, NP, and the Post Correspondence Problem. *Information and Computation*, 121:135–142, 1995.
11. E.L. Post. A variant of a recursively unsolvable problem. *Bulletin of the American Mathematical Society*, 52:264–268, 1946.
12. D. Reidenbach and J.C. Schneider. Morphically primitive words. *Theoretical Computer Science*, 410:2148–2161, 2009.
13. A. Salomaa. Equality sets for homomorphisms of free monoids. *Acta Cybernetica*, 4:127–139, 1978.
14. J.C. Schneider. Unambiguous erasing morphisms in free monoids. *Theoretical Informatics and Applications (RAIRO)*, 44:193–208, 2010.