Theoretical Computer Science 601 (2015) 15–20

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

On the n-permutation Post Correspondence Problem

Mari Ernvall [a,1], Vesa Halava [a,b,*], Tero Harju [a]

[a] Department of Mathematics and Statistics, University of Turku, FI-20014 Turku, Finland
[b] Department of Computer Science, University of Liverpool, UK

ARTICLE INFO

Article history:
Received 4 December 2013
Accepted 10 April 2015
Available online 14 July 2015

Keywords:
Permutation Post Correspondence Problem
Semi-Thue system
Word problem
Deterministic
Cyclic derivation

ABSTRACT

We give new and simpler proof for the undecidability of the n-permutation Post Correspondence Problem that was originally proved by K. Ruohonen [10]. Our proof uses a recent result on deterministic semi-Thue systems according to which it is undecidable for a given deterministic semi-Thue system T and a word u whether or not there exists a nonempty cyclic derivation $u \to_T^+ u$ in T.

1. Introduction

In the history of computation the Post Correspondence Problem and its variants have played a major role as a simply defined algorithmically undecidable problem that can be used to prove other undecidability results. For example, several problems in the formal language theory and in the theory of integer matrices are shown to be undecidable by reducing the Post Correspondence Problem to them.

The original formulation of the Post Correspondence Problem, or PCP for short, by Emil Post [8] is the following:

Problem 1 *(PCP).* Let B be an alphabet, and let B^* be the set of all finite words over B, including the empty word ε. Given an integer n and two finite ordered lists of words

$$(u_1, u_2, \ldots, u_n) \text{ and } (v_1, v_2, \ldots, v_n) \tag{1}$$

where $u_i, v_i \in B^*$ for all $i = 1, 2, \ldots, n$, does there exist a finite nonempty sequence $i_1, i_2, \ldots, i_k$ of indices such that

$$u_{i_1} u_{i_2} \cdots u_{i_k} = v_{i_1} v_{i_2} \cdots v_{i_k}? \tag{2}$$

The PCP is given an equivalent form in Problem 2. Let A and B be two alphabets. A mapping $h: A^* \to B^*$ is a *morphism*, if $h(uv) = h(u)h(v)$ holds for all words $u, v \in A^*$. Note that a morphism is uniquely defined by the images of the letters of the domain alphabet. For a given instance in (1) with $u_i, v_i \in B^*$, let $A = \{a_1, a_2, \ldots, a_n\}$ be an alphabet and define two morphisms $g, h: A^* \to B^*$ by

* Corresponding author at: Department of Mathematics and Statistics, University of Turku, FI-20014 Turku, Finland.
E-mail addresses: mari.huova@utu.fi (M. Ernvall), vesa.halava@utu.fi (V. Halava), harju@utu.fi (T. Harju).

[1] Supported by the Väisälä Foundation.

http://dx.doi.org/10.1016/j.tcs.2015.07.021
0304-3975/

$$g(a_i) = u_i \quad \text{and} \quad h(a_i) = v_i$$

for all $i = 1, 2, \ldots, n$. Then the original form of the PCP is equivalent to the following problem.

Problem 2 *(PCP).* Given two morphisms $g, h\colon A^* \to B^*$, does there exist a nonempty word $w \in A^+$ such that

$$g(w) = h(w)\,?$$

Now, a pair $I = (g, h)$ of morphisms is said to be an *instance* of the PCP, and a word w satisfying $h(w) = g(w)$ is called a *solution* of the instance I. The *size* of the instance (g, h) is the cardinality of the domain alphabet A.

Several variants of the PCP are known to be undecidable. By a variant we mean a restriction of the PCP to a specific type of instances. For example, it is known that the PCP is undecidable for instances of size 7; see [7]. It is also known that the PCP is undecidable for instances of injective morphisms; see [5,11], and also [4] for a more recent proof to this end.

In [10], K. Ruohonen proved that the following variants of the PCP are undecidable. For any natural number $n \geq 1$, define:

Problem 3 *(n-Permutation PCP (nPPCP)).* Let n be a positive integer. Given two morphisms $g, h\colon A^* \to B^*$, does there exist a word $w = w_1 w_2 \cdots w_n$ and a permutation σ of the set $\{1, 2, \ldots, n\}$ such that

$$g(w_1 \cdots w_n) = h(w_{\sigma(1)} \cdots w_{\sigma(n)}).$$

Note that the permutation σ and its existence is tied up to the existence of solution, where a solution consists of the component words w_i and a permutation for them. Indeed, if the problem were defined for a fixed permutation σ, then the problem would be undecidable simply because by choosing σ to be the identity permutation, the PCP would be a special case of the problem. On the other hand, if restricted to non-trivial permutations, then the question would be open and, actually, our construction for the undecidability of the nPPCP works for this problem also, as we shall note later.

Clearly, the 1PPCP is just the PCP in Problem 2. Also, the 2PPCP is of independent interest. This special case is also called the *circular PCP*; see [10]. Indeed, we may formulate the circular PCP in the following way: given two morphisms $g, h\colon A^* \to B^*$, does there exist words $u, v \in A^*$ with $uv \neq \varepsilon$ such that

$$g(uv) = h(vu).$$

We can omit the permutation from the definition of the 2PPCP as the identity permutation corresponds to the case where $u = \varepsilon$. Here the words $w_1 = uv$ and $w_2 = vu$ are called *conjugates* of each other. Hence, the circular PCP could be stated by asking whether there exist nonempty conjugates w_1 and w_2 such that $g(w_1) = h(w_2)$. The phrase 'circular PCP' refers to the problem setting where the words are considered to be cyclic, i.e., the last letter is followed by the first letter.

The original undecidability proofs of nPPCP and circular PCP by Ruohonen [10] employ an undecidable property of linearly bounded automata. These proofs are rather long and technical, and therefore, there is a request for simpler proofs for these problems. In [3], instead of linearly bounded automata, the authors employed a special variant of the word problem for semi-Thue systems while proving the undecidability of the circular PCP. Here we shall use the same techniques for the nPPCP for any $n \geq 1$.

Let us briefly discuss this special form of the word problem. A *semi-Thue system* T is a pair (Σ, R) where $\Sigma = \{a_1, a_2, \ldots, a_n\}$ is a finite alphabet, the elements of which are called *generators* of T, and $R \subseteq \Sigma^* \times \Sigma^*$ is a relation. The elements of R are called the *rules* of T. We write $u \to_T v$, if there exists a rule $(x, y) \in R$ such that $u = u_1 x u_2$ and $v = u_1 y u_2$ for some words u_1 and u_2. We denote by $\to_T^*$ the reflexive and transitive closure of $\to_T$, and by $\to_T^+$ the transitive closure of $\to_T$. Note that the index T is omitted from the notation, i.e., we shall write $\to$, when the semi-Thue system studied is clear from the context.

If the relation R is symmetric, then T is a *Thue system* and then T corresponds to a semigroup with generators Σ and relation R.

In the *word problem* for a semi-Thue system $T = (\Sigma, R)$ we are given two words $u, v \in \Sigma^*$ and the task is to determine whether or not there exists a *derivation* from u to v using the rules in R, i.e., $u \to_T^* y$. The first proofs for undecidability of the word problem of (semi-)Thue systems were given independently by Post [9] and Markov [6].

Let $T = (\Sigma, R)$ be a semi-Thue system such that $\Sigma = A \cup B$ with $A \cap B = \emptyset$. Then T is called *B-deterministic*, if

1. $R \subseteq A^*BA^* \times A^*BA^*$, namely, if the rules contain a unique letter from B on both sides, and
2. for all words $w \in A^*BA^*$, if there exists a rule in R giving $w \to_T w'$, then the rule is unique in T.

In [3] it was proved that the word problem is undecidable for B-deterministic semi-Thue systems even in the following strong form:

Theorem 1. *It is undecidable whether or not there exists a nonempty (cyclic) derivation $aSc \to_T^+ aSc$ for instances where $T = (\Sigma, R)$ is a B-deterministic semi-Thue system where $\Sigma = A \cup B$ and $A \cap B = \emptyset$, and $a, c \in A$ and $S \in B$.*