Acta Cybernetica — *online first paper version* — pages 1–11.

On the Steps of Emil Post: from Normal Systems to the Correspondence Decision Problem*

Vesa Halava[a] and Tero Harju[b]

Abstract

In 1946 Emil Leon Post (*Bull. Amer. Math. Soc.* **52** (1946), 264–268) introduced his famous correspondence decision problem, nowadays known as the Post Correspondence Problem (PCP). Post proved the undecidability of the PCP by a reduction from his normal systems. In the present article we follow the steps of Post, and give another, somewhat simpler and more straightforward proof of the undecidability of the problem by using the same source of reductions as Post did. We investigate these, very different, techniques, and point out out some peculiarities in the approach taken by Post.

Keywords: normal systems, Post correspondence problem, undecidability, assertion problem

1 Introduction

The original formulation of the *Post correspondence problem* (or, as Post called it, the *correspondence decision problem* [8]), PCP for short, is stated as follows:

Problem 1 (Post Correspondence Problem)**.** *Let $A = \{a, b\}$ be a binary alphabet, and denote by A^* the set of all finite words over A. Given a finite set of pairs of words in A^*,*

$$W = \{(u_i, v_i) \mid u_i, v_i \in A^*,\ i = 1, 2, \ldots, n\},$$

does there exist a nonempty sequence $i_1, i_2, \ldots, i_k$ of indices, where $i_j \in \{1, 2, \ldots, n\}$ for $1 \leq j \leq k$, such that

$$u_{i_1} u_{i_2} \cdots u_{i_k} = v_{i_1} v_{i_2} \cdots v_{i_k}\,? \tag{1}$$

*The first author was supported by emmy.network foundation under the aegis of the Fondation de Luxembourg.

[a]Complex Systems Research Group, Department of Mathematics and Statistics, University of Turku, Turku, Finland. E-mail: vesa.halava@utu.fi, ORCID: https://orcid.org/0000-0003-3633-4902.

[b]Department of Mathematics and Statistics, University of Turku, Turku, Finland. E-mail: harju@utu.fi, ORCID: https://orcid.org/0000-0002-9640-6309.

DOI: 10.14232/actacyb.284625

In the history of computability, the Post correspondence problem and its many variants have played an important role as simply defined algorithmically undecidable problems that can be used to prove other undecidability results. Here we concentrate on the undecidability proofs of the PCP itself.

A standard textbook proof of the undecidability of the PCP employs the *halting problem of the Turing machines* as the base of the reduction; see e.g. [9], or the construction by Claus [2] from the *word problem of the semi-Thue systems*, which gives better undecidability bounds on the number of pairs in the sets. The integer $n = |W|$ in Problem 1 is said to be the *size* of the set W. The set W is called an *instance* of the PCP. Recently, Neary [5] showed that the PCP is undecidable for $|W| = 5$ using the (Post) *tag systems* that form a special class of Post normal systems; see [6].

In his article [8], Post proved that the PCP is unsolvable, i.e., undecidable, by a technical and nontrivial reduction from the *assertion problem of the Post normal systems*. We shall give another proof by utilizing the same source.

There are several proofs of the PCP. The standard reductions from the Turing machines, semi-Thue systems and tag systems to the PCP have a common leading idea: An instance of the PCP is constructed so that any solution to the instance is a (encoded) concatenation of the configurations required in the computation or derivation of the original machine or system. This is *not the case* in Post's original proof. Indeed, he relies simply on the words in the rules of a derivation in a normal system. A sequence of these words imply a required derivation in the normal system if and only if the sequence is a solution of a particular instance of the PCP. The new proof presented in this article is based on the idea in the standard type – a solution exists to the constructed instance of the PCP if and only if the solution is a concatenation of the full configurations required of the given Post normal system.

We note that, in Post's definition, the PCP is defined for binary words. Actually, the cardinality of the alphabet A is not relevant, since every instance of the PCP with any alphabet size has an equivalent one in terms of binary words using an injective encoding into binary alphabet $\{a, b\}^*$ from A^*. For example, if $A = \{a_1, a_2, \ldots, a_k\}$, then φ defined by $\varphi(a_i) = a^i b$ is such an encoding. Note, however, that the PCP is decidable for sets of pairs W over unary alphabet.

The structure of this article is the following: In Section 2, we present the basic notions, notations needed in this article. Especially, we introduce the normal systems and present preliminary results on them. In Section 3 we present Post's construction, following Post's original article, but also give some explanatory steps for readability. In Section 4 we present our main contribution: another proof for the undecidability of the Post Correspondence Problem using the same source of undecidability as in Section 3 in the Post's construction.

Short preliminary version of this article can be found in [3].

2 Normal systems

Let A be a finite alphabet and denote by A^* the set of all finite words over A including the empty word ε. The length of a word u, i.e., the number of occurrences of letters in u, is denoted by $|u|$. For words u and w, the word u is a *prefix* of w if there exists a word v such that $w = uv$. If u is a prefix of w with $w = uv$, we denote the *suffix* v also by $u^{-1}w$.

For a word $w \in A^*$, if $w = a_1 \cdots a_{n-1}a_n$ where $a_i \in A$ for all $i = 1, \ldots, n$, then the *reverse of* w is defined to be $w^{\mathrm{R}} = a_n a_{n-1} \cdots a_1$.

The words u and v are *(cyclic) conjugates* if there exist words x and y such that $u = xy$ and $v = yx$.

We give a formal definition of a normal system instead of the bit informal one used by Post in [8].

Let $A = \{a, b\}$ be a binary alphabet, and let X be a variable ranging over the words in A^*. A *normal system* $S = (w, P)$ consists of an *initial word* $w \in A^+$ and a finite set P of *rules* of the form $\alpha X \mapsto X\beta$, where $\alpha, \beta \in A^*$. We say that a word v is a *successor* of a word u, if there is a rule $\alpha X \mapsto X\beta$ in P such that $u = \alpha u'$ and $v = u'\beta$. We denote this by $u \to v$. Let $\to^*$ be the reflexive and transitive closure of $\to$. Then $u \to^* v$ holds if and only if $u = v$ or there is a finite sequence of words $u = v_1, v_2, \ldots, v_n = v$ such that $v_i \to v_{i+1}$ for $i = 1, 2, \ldots, n-1$. A normal system is a special case of the *Post canonical system* for which Post proved in 1943 the Normal-Form Theorem; see [6]. On the other hand, the tag systems mentioned in the introduction are a special class of the normal systems that have a constant length left for rule words α; see [6].

The *assertion* of a normal system $S = (w, P)$ is the set

$$\mathcal{A}_S = \{v \in A^* \mid w \to^* v\} \; . \tag{2}$$

Problem 2 (Assertion Problem)**.** *Given a normal system $S = (w, P)$ and a word u, does $u \in \mathcal{A}_S$ hold?*

The following result is crucial for the construction presented in this article, but the reference for it is a bit peculiar: in footnote 2 of [8], Post gives citation to his paper [7] for an informal proof and to Church [1] for a formal proof, but with a comment that a verification of the recursiveness of the reduction is needed and then he gives guidelines for missing details on the footnote.

Proposition 1. *The assertion problem for normal systems is undecidable.*

Actually, the problem remains undecidable even if we assume that in each rule $\alpha X \mapsto X\beta$ in P the words α and β are non-empty; see Post [8], footnote 3. A normal system with non-empty rule words is called a *standard* normal system in the literature. Therefore, we can assume in the following that the normal systems are standard. This assumption is indeed crucial when we construct instances of the PCP from the normal systems in Sections 3 and 4.

3 Undecidability of the PCP by Emil Post

In this section we present the original proof and construction of Emil Post in [8]. Occasionally we use more modern terminology instead of Post's original terms.

Let $u \in \mathcal{A}_S$, where $S = (w, P)$. As the assertion problem is trivial for the case $u = w$, we assume that $u \neq w$. Therefore, there exists a sequence

$$w = \alpha_1 x_1,\ x_1\beta_1 = \alpha_2 x_2, \ldots, x_{k-1}\beta_{k-1} = \alpha_k x_k,\ x_k\beta_k = u\,, \tag{3}$$

where $\alpha_i X \to X\beta_i$ is a rule for each i with $x_j \in A^*$ for all j. The idea in Post's proof is to present the set of equations in (3) as a single equation in the form of a word equality (1).

Consider the word $w\beta_1\beta_2\cdots\beta_k$ obtained from (3). Using the equations in (3), we obtain, for each $j = 1, 2, \ldots, k$,

$$w\beta_1\beta_2\cdots\beta_{j-1} = \alpha_1\alpha_2\cdots\alpha_j x_j\,, \tag{4}$$

and finally

$$w\beta_1\beta_2\cdots\beta_k = \alpha_1\alpha_2\cdots\alpha_k u\,. \tag{5}$$

By (4), we have, for each $j = 1, 2, \ldots, k$,

$$|w\beta_1\beta_2\cdots\beta_{j-1}| \geq |\alpha_1\alpha_2\cdots\alpha_j|\,. \tag{6}$$

So we have shown that the derivation sequence (3) implies (4), which further implies (5) together with the inequalities (6). Actually, $u \in \mathcal{A}_S$, that is, existence of a sequence (3) is indeed equivalent to the join of (5) and (6). For this we need to prove the above implications in the opposite direction.

First, we show that the join of the equalities (5) and (6) imply the equations in (4). For this it suffices to choose

$$x_j = (\alpha_1\alpha_2\cdots\alpha_j)^{-1}(w\beta_1\beta_2\cdots\beta_{j-1})$$

for all j.

Furthermore, for the equations in (3), we obtain, for all $1 \leq j \leq k$,

$$\begin{aligned}\alpha_{j+1}x_{j+1} &= \alpha_{j+1}(\alpha_1\alpha_2\cdots\alpha_{j+1})^{-1}(w\beta_1\beta_2\cdots\beta_j)\\ &= (\alpha_1\alpha_2\cdots\alpha_j)^{-1}(w\beta_1\beta_2\cdots\beta_{j-1})\beta_j = x_j\beta_j\,,\end{aligned}$$

and we have the equations in (3) except the first and the last ones. The first one is obtained directly from (4) by setting $j = 1$. Also, the last one will follow, since

$$\alpha_1\alpha_2\cdots\alpha_k x_k\beta_k = w\beta_1\beta_2\cdots\beta_{k-1}\beta_k = \alpha_1\alpha_2\cdots\alpha_k u\,.$$

We have proved that (5) and (6) are satisfied if and only if the equations in (3) are satisfied, i.e., (5) and (6) are equivalent to the condition $u \in \mathcal{A}_S$.

Finally, we need to get rid of the extra condition (6). This is done by constructing a new normal system S_1, where (5) implies (6), and $uc \in \mathcal{A}_{S_1}$ if and only if $u^{\mathrm{R}} \in \mathcal{A}_S$ holds, where c is a new letter introduced below.

For this, let first $S' = (w^{\mathrm{R}}, P')$, where

$$P' = \left\{ X\alpha^{\mathrm{R}} \mapsto \beta^{\mathrm{R}} X \;\mid\; \alpha X \mapsto X\beta \in P \right\}.$$

Strictly speaking the system S' is not normal. It is a 'dual' of a normal system. However, we can still write $u \in \mathcal{A}_s$ if and only if $u^{\mathrm{R}} \in \mathcal{A}_{S'}$. Next we design a system $S'' = (w^{\mathrm{R}}c, P'')$, where c is a new letter. Let

$$P'' = \left\{ X\alpha^{\mathrm{R}}c \mapsto \beta^{\mathrm{R}} Xc \;\mid\; \alpha X \mapsto X\beta \in P \right\}.$$

It is immediate that $u^{\mathrm{R}} \in \mathcal{A}_{S'}$ if and only if $u^{\mathrm{R}}c \in \mathcal{A}_{S''}$. Obviously, S'' is even less normal as the letter c is kept constantly in the end of the words of the derivations.

Finally, let $S_1 = (w^{\mathrm{R}}c, P_1)$ be the normal system, where

$$P_1 = \left\{ \alpha^{\mathrm{R}}cX \mapsto Xc\beta^{\mathrm{R}} \mid \alpha X \mapsto X\beta \in P \right\} \cup \{ yX \mapsto Xy \mid y \in \{a, b, c\} \}.$$

Notice that the rules $yX \mapsto Xy$ for $y \in \{a, b, c\}$ imply that any sequence can be transformed to its conjugates. Therefore, if a rule is applied in S'', then the corresponding rule can be applied in S_1, since in the rules in P_1 the left hand sides are conjugates of the left hand sides of the rules in P'' and the right hand sides are conjugates of the right hand sides of the corresponding rules in P''. Let

$$\mathcal{C}_v = \{ w \mid w \text{ is a conjugate of } v \}$$

called the *conjugacy class* of the word v.

Then we have

$$\mathcal{A}_{S_1} = \mathcal{A}_{S''} \cup \bigcup_{v \in \mathcal{A}_{S''}} \mathcal{C}_v = (\mathcal{A}_{S'})^{\mathrm{R}} c \cup \bigcup_{v \in \mathcal{A}_{S'}} \mathcal{C}_{v^{\mathrm{R}}c} = (\mathcal{A}_S)^{\mathrm{R}} c \cup \bigcup_{v \in \mathcal{A}_S} \mathcal{C}_{v^{\mathrm{R}}c} = \bigcup_{v \in \mathcal{A}_S} \mathcal{C}_{v^{\mathrm{R}}c}.$$

To verify the above equality of sets, assume $u \in \mathcal{A}_{S_1}$. Denote by $x \xrightarrow{\mathcal{C}} y$ if $y \in \mathcal{C}_x$. For u, there exist a sequence of words and successors

$$\begin{aligned} w^R c \xrightarrow{\mathcal{C}} \alpha_1^R c x_1, \quad x_1 c \beta_1^R \xrightarrow{\mathcal{C}} \alpha_2^R c x_2, \quad x_2 c \beta_2^R \xrightarrow{\mathcal{C}} \alpha_3^R c x_3, \ldots, \\ x_{n-1} c \beta_{n-1}^R \xrightarrow{\mathcal{C}} \alpha_n^R c x_n, \quad x_n c \beta_n^R \xrightarrow{\mathcal{C}} u\,, \end{aligned} \tag{7}$$

where $\alpha^R cX \mapsto Xc\beta_n^R$ are in P_1 and $\xrightarrow{\mathcal{C}}$ part are done with rules $yX \mapsto Xy$ in S_1. Using cyclic shifts for all words in sequence (7) so that the special symbol c is the right most symbol of the word makes $\xrightarrow{\mathcal{C}}$ to be equality, and we get that there exists a sequence

$$\begin{aligned} &w^R c = x_1 \alpha_1^R c, \quad \beta_1^R x_1 c = x_2 \alpha_2^R c, \quad \beta_2^R x_2 c = x_3 \alpha_3^R c, \ldots, \\ &\beta_{n-1}^R x_{n-1} c = x_n \alpha_n^R c, \quad \beta_n^R x_n c \xrightarrow{\mathcal{C}} u\,. \end{aligned}$$

Now cancelling the letters c, taking reverse of all words and using the equality $(xy)^R = y^R x^R$, we have the sequence

$$\begin{aligned} &w = (w^R)^R = (x_1 \alpha_1^R)^R = \alpha_1 x_1^R, \quad x_1^R \beta_1 = (\beta_1^R x_1)^R = (x_2 \alpha_2^R)^R = \alpha_2 x_2^R, \quad \ldots, \\ &x_{n-1}^R \beta_{n-1} = (\beta_{n-1}^R x_{n-1})^R = (x_n \alpha_n^R)^R = \alpha_n x_n^R, \quad x_n^R \beta_n = v, \end{aligned}$$

for a word v with $vc \xrightarrow{\mathcal{C}} u^R$. We have proved that $u \in \mathcal{A}_{S_1}$ implies $u^R \in \mathcal{C}_{vc}$ for a word $v \in \mathcal{A}_S$. We note that $u^R \in \mathcal{C}_{vc}$ is equivalent to $u \in \mathcal{C}_{v^{\mathrm{R}}c}$. As the above verification works also in the other direction, we have proved that

$$\mathcal{A}_{S_1} = \bigcup_{v \in \mathcal{A}_S} \mathcal{C}_{v^{\mathrm{R}}c}.$$

Denote the rules of P_1 in the form $\gamma X \mapsto X\delta$. As in the above for the system S, we obtain that if $u^{\mathrm{R}}c \in \mathcal{A}_{S_1}$ then

$$w^{\mathrm{R}}c\delta_1\delta_2\cdots\delta_k = \gamma_1\gamma_2\cdots\gamma_k u^{\mathrm{R}}c\,, \tag{8}$$

where $\gamma_i X \mapsto X\delta_i \in P_1$ for each i. We shall prove that in S_1 the condition (8) implies the condition

$$|w^{\mathrm{R}}c\delta_1\delta_2\cdots\delta_{j-1}| \geq |\gamma_1\gamma_2\cdots\gamma_j|\,, \tag{9}$$

for all $j = 1, 2, \ldots, k$.

Assume contrary to the claim that there is a solution to (8) such that for some s,

$$|w^{\mathrm{R}}c\delta_1\delta_2\cdots\delta_{s-1}| < |\gamma_1\gamma_2\cdots\gamma_s|\,,$$

and let v be a nonempty word in $\{a, b, c\}^*$ such that

$$w^{\mathrm{R}}c\delta_1\delta_2\cdots\delta_{s-1}v = \gamma_1\gamma_2\cdots\gamma_s\,. \tag{10}$$

Now for each rule $\gamma_i X \mapsto X\delta_i$ of P_1, either both sides contain one c or neither of them contains c. Therefore if γ_s contains no occurrences of c then the left hand side of (10) would have at least one more occurrence of c than the right hand side; a contradiction. If c occurs in γ_s, then c is necessarily the last letter of γ_s and v would also end with c, and again the left hand side has more occurrences of the letter c than the right hand side; again a contradiction.

Therefore we have shown that $u \in \mathcal{A}_s$ if and only if $u^{\mathrm{R}}c \in \mathcal{A}_{S_1}$, which holds if and only if there exist rules $\gamma_i X \mapsto X\delta_i \in P_1$ for $i = 1, \ldots, k$ with

$$w^{\mathrm{R}}c\delta_1\delta_2\cdots\delta_k = \gamma_1\gamma_2\cdots\gamma_k u^{\mathrm{R}}c\,. \tag{11}$$

We then begin by the equation (11), and use the technique which is nowadays called desynchronization. Let d be a new symbol absent in S_1 and define a mapping $\ell_d : \{a, b, c\} \to \{a, b, c, d\}$ which writes the letter d before (to the left hand side of) every letter in a word, and define r_d similarly writing d to the right hand side of every letter. The mappings ℓ_d and r_d extend to morphisms in the natural manner. They are called desynchronizing morphisms. Now from equation (11), we obtain

$$d\ell_d\big(w^{\mathrm{R}}c\delta_1\delta_2\cdots\delta_k\big)dd = ddr_d\big(\gamma_1\gamma_2\cdots\gamma_k u^{\mathrm{R}}c\big)d\,, \tag{12}$$

where both sides begin and end with a double dd, and elsewhere d is between all pairs of letters from $\{a, b, c\}$. We let

$$W = \{(\ell_d(\gamma), r_d(\delta)) \mid \gamma X \mapsto X\delta \in P_1\} \cup \big\{(d\ell_d(w^{\mathrm{R}}c), dd), (dd, r_d(u^{\mathrm{R}}c)d)\big\} \tag{13}$$

be an instance of the PCP. It is straightforward that if $u \in \mathcal{A}_S$, then the instance W has a solution. We note that the assumption that the normal system S is standard, i.e., the rule words are non-empty, is needed at this point to guarantee that the desynchronization works properly.

What is left is to show is the converse: if W has a solution, then $u \in \mathcal{A}_S$. For this, assume that W has a solution, and choose a minimal solution, i.e. a solution that does not contain any solutions as a proper prefix. It is immediate that if W has a solution, it has a minimal solution.

Obviously, a solution must begin with the pair $(d\ell_d(w^{\mathrm{R}}c), dd)$ as that is the only pair having a common nonempty prefix. Similarly, a solution must end with the pair $(dd, r_d(u^{\mathrm{R}}c)d)$, since that is the only pair in W with a common nonempty suffix. On the other hand, these two special pairs with occurrences of the word dd cannot appear in the middle of any minimal solution as dd can be covered only by these two pair. Therefore, if $i_1, \ldots, i_k$ is a minimal solution to the instance W, then

$$u_{i_1}u_{i_2}\cdots u_{i_k} = v_{i_1}v_{i_2}\cdots v_{i_k} \text{ with } (u_{i_j}, v_{i_j}) \in W \text{ for } j = 1, \ldots, k\,,$$

then $(u_{i_1}, v_{i_1}) = (d\ell_d(w^{\mathrm{R}}c), dd)$, $(u_{i_k}, v_{i_k}) = (dd, r_d(u^{\mathrm{R}}c)d)$ and

$$(u_{i_j}, v_{i_j}) = (\ell_d(\gamma_j), r_d(\delta_j)) \text{ and } \gamma_j X \mapsto X\delta_j \in P_1\,,$$

for $j = 2, \ldots, k-1$. It follows that the minimal solution corresponds to the equation (12) which implies that $u \in \mathcal{A}_s$.

By Proposition 1, we have

Theorem 3. *The PCP is undecidable.*

Recall that $\{a, b, c, d\}^*$ can be embedded into $\{a, b\}^*$ by an injective morphisms. In this way we obtain instances in the binary alphabet as originally considered by Post.

Actually, Post proved the undecidability of a special form of the PCP, called the *generalized PCP* in the literature; see for example [4].

Theorem 4. *It is undecidable for given set of pairs $\{(u_i, v_i) \mid 1 \le i \le n\}$ of words whether or not there exist a sequence $i_1, i_2, \ldots, i_k$ such that*

$$u_1u_{i_1}\cdots u_{i_k}ku_n = v_1v_{i_1}\cdots v_{i_k}v_n\,. \tag{14}$$

Note that in Theorem 4 the first pair and the last pair of the required solution are fixed in (14) to be (u_1, v_1) and (u_n, v_n), respectively. In W constructed in (13), $(u_1, v_1) = (d\ell_d(w^{\mathrm{R}}c), dd)$ and $(u_n, v_n) = (dd, r_d(u^{\mathrm{R}}c)d)$. Note also that in (14) we could assume that, $i_j \notin \{1, n\}$ for all $j = 1, \ldots, k$.

4 Another proof for the PCP from the normal systems

In this section we give a new proof for the undecidability of the PCP by a reduction to the assertion problem of the normal systems, i.e., we show that if the PCP

is decidable, then the assertion problem of the normal systems is also decidable, contradicting Proposition 1. For this, we take an arbitrary (non-trivial) instance of the assertion problem, that is, normal system $S = (w, P)$ and word u with $u \neq w$, and construct an instance $W_{S,u}$ of the PCP such that $W_{S,u}$ has a solution if and only if $u \in \mathcal{A}_S$.

The present proof takes a modern approach of connecting the configurations of a derivation in the normal systems to a solution of the PCP – instead of the rule words used by PCP as was done in the original proof by Post in Section 3.

Let $S = (w, P)$ be a normal system over the binary alphabet $\{a, b\}$ where $P = \{p_1, \ldots, p_t\}$ and $p_j = \alpha_j X \mapsto X\beta_j$ for $j = 1, \ldots, t$. As Post did, we begin with the sequence (3), but use different indices: we assume that there exists a sequence of equalities

$$w = \alpha_{i_1} x_1,\ x_1\beta_{i_1} = \alpha_{i_2} x_2, \ldots, x_{k-1}\beta_{i_{k-1}} = \alpha_{i_k} x_k,\ x_k\beta_{i_k} = u\,, \tag{15}$$

for the input word u where $\alpha_{i_j} X \mapsto X\beta_{i_j} \in P$ for $j = 1, \ldots, k$. Instead of the equations (5) and (6), we take

$$wx_1\beta_{i_1}x_2\beta_{i_2}\cdots x_k\beta_{i_k} = \alpha_{i_1}x_1\alpha_{i_2}x_2\cdots\alpha_{i_k}x_k u, \tag{16}$$

where the configurations in (15) are concatenated – left hand sides on the left and right hand sides on the right.

Let c and f be new letters. We split each rule $p_j \in P$ to two pairs p_j^α and p_j^β as follows:

$$p_j^\alpha = (\ell_d(c^j f), r_d(f\alpha_j)) \quad \text{and} \quad p_j^\beta = (\ell_d(\beta_j), r_d(c^j)),$$

where r_d and ℓ_d are the desynchronizing mappings for the letter d. The word $c^j f$ is a marker word that forces a solution of the (the below) instance of the PCP to choose the pairs jointly. Consider the following instance of the PCP:

$$\begin{aligned} W = &\{(d\ell_d(fw), dd), (dd, r_d(fu)d), (da, ad), (db, bd)\} \\ &\cup \{p_j^\alpha, p_j^\beta \mid j = 1, \ldots, t\}. \end{aligned} \tag{17}$$

To see the idea encoded in W, let us first assume that W has a solution. A solution must necessarily begin with the pair $(d\ell_d(fw), dd)$ that we now write in the form

$$\begin{aligned} &\mathrm{L}\colon d\ell_d(fw) \\ &\mathrm{R}\colon dd \end{aligned}$$

In order to produce $r_d(fw)$ to the right hand side, we need to use a pair which has f as a first symbol on the right hand side. As the pair $(dd, r_d(fu)d)$ produces dd to the end of left hand side, which then has to match with dd produced by the right hand side, we must have

$$d\ell_d(fw) \cdot dd = dd \cdot r_d(fu)d,$$

implying $w = u$. In a non-trivial case of the assertion problem, $u \neq w$, for producing $r_d(fw)$ to the right hand side, there must exist a pair $p^{\alpha}_{i_1} = (\ell_d(c^{i_1}f), r_d(f\alpha_{i_1}))$ in W with $w = \alpha_{i_1}x_1$. After this we have

$$\begin{aligned} &\mathrm{L}\colon d\ell_d(fwc^{i_1}f)\\ &\mathrm{R}\colon ddr_d(f\alpha_{i_1}) \end{aligned}$$

Now the first occurrence of the letter c forces the use of the pairs $(da, ad), (db, bd) \in W$ in order to have x_1 and cover the start word w of the left hand side. So now we have

$$\begin{aligned} &\mathrm{L}\colon d\ell_d(fwc^{i_1}fx_1)\\ &\mathrm{R}\colon ddr_d(f\alpha_{i_1}x_1) \end{aligned}$$

Next to match c^{i_1}, we must chose the other half of the rule P_{i_1}, i.e., the pair $p^{\beta}_{i_1} = (\ell_d(\beta_{i_1}), r_d(c^{i_1}))$. We then have

$$\begin{aligned} &\mathrm{L}\colon d\ell_d(fwc^{i_1}fx_1\beta_{i_1})\\ &\mathrm{R}\colon ddr_d(f\alpha_{i_1}x_1c^{i_1}). \end{aligned}$$

In other words, after forgetting the synchronizing letters d, the left hand side has the overflow $fx_1\beta_{i_1}$. As above, the occurrence of f forces to chose the rule $p^{\alpha}_{i_2}$, then write x_{i_2} and the other half of the rule p_{i_2}, the pair $p^{\beta}_{i_2}$ etc. Therefore, at some point we must have

$$\begin{aligned} &\mathrm{L}\colon d\ell_d(fwc^{i_1}fx_1\beta_{i_1}c^{i_2}f\cdots fx_{t-1}\beta_{i_{t-1}}c^{i_t}fx_t\beta_{i_t})\\ &\mathrm{R}\colon ddr_d(f\alpha_{i_1}x_1c^{i_1}f\alpha_{i_2}x_2c^{i_2}f\cdots f\alpha_{i_t}x_tc^{i_t}) \end{aligned} \tag{18}$$

with each w, u, x_{i_j}, α_{i_j} and β_{i_j} satisfying (15) up to index $t \geq 2$. Note that

$$d\ell_d(fwc^{i_1}fx_1\beta_{i_1}c^{i_2}f\cdots fx_{t-1}\beta_{i_{t-1}}c^{i_t})d = ddr_d(f\alpha_{i_1}x_1c^{i_1}f\alpha_{i_2}x_2c^{i_2}f\cdots f\alpha_{i_t}x_tc^{i_t}).$$

A minimal solution in W must end with pair $(dd, r_d(fu)d)$ in order to match the d's in the solution. As $r_d(fu)d$ begin with f and has no c's, and left hand side has one more f than the right hand side in (18), the pair $(dd, r_d(fu)d)$ has to match $fx_t\beta_{i_t}$ in (18). Therefore, at some point $t = k$ in (15) and (16) with $x_k\beta_{i_k} = u$ and

$$\begin{aligned} &d\ell_d(fwc^{i_1}fx_1\beta_{i_1}c^{i_2}f\cdots c^{i_k}fx_k\beta_{i_k})dd\\ =&ddr_d(f\alpha_{i_1}x_1c^{i_1}f\alpha_{i_2}x_2c^{i_2}f\cdots\alpha_{i_k}x_kc^{i_k}fu)d. \end{aligned}$$

The other direction is clear: Suppose $u \in \mathcal{A}_S$. Then there exists a sequence of equations (15) satisfying (16). We start with (16), and place symbols f and words c^i accordingly using the equations in (15). Then we get

$$fwc^{i_1}fx_1\beta_{i_1}c^{i_2}f\cdots c^{i_k}fx_k\beta_{i_k} = f\alpha_{i_1}x_1c^{i_1}f\alpha_{i_2}x_2c^{i_2}f\cdots\alpha_{i_k}x_kc^{i_k}fu.$$

Now placing dd to both ends and d between the letters a, b, c, f, we obtain

$$\begin{aligned} &d\ell_d(fwc^{i_1}fx_1\beta_{i_1}c^{i_2}f\cdots c^{i_k}fx_k\beta_{i_k})dd \\ &= ddr_d(f\alpha_{i_1}x_1c^{i_1}f\alpha_{i_2}x_2c^{i_2}f\cdots\alpha_{i_k}x_kc^{i_k}fu)d\,. \end{aligned} \tag{19}$$

What is left is to show that the words in (19) can be build with pairs of W, correspondingly. For this, for a word $x \in \{a,b\}^*$, $x = e_1\cdots e_n$ and $e_i \in \{a,b\}$, denote by $\bar{x}$ the sequence of pairs $(de_1, e_1d), \ldots, (de_n, e_nd)$ from W. The idea is that the sequence $\bar{x}$ writes $\ell_d(x)$ to the left hand side and $r_d(x)$ to the right hand side in a solution. Let Z be the following sequence of pairs of W,

$$(d\ell_d(fw), dd), p^{\alpha}_{i_1}, \bar{x_1}, p^{\beta}_{i_1}, p^{\alpha}_{i_2}, \bar{x_2}, p^{\beta}_{i_2}, \ldots, p^{\alpha}_{i_k}, \bar{x_k}, p^{\beta}_{i_k}, (dd, r_d(fu)d).$$

Now, Z is a solution of the PCP, as

$$\begin{aligned} &d\ell_d(fw)\ell_d(c^{i_1}f)\ell_d(x_1)\ell_d(\beta_{i_1})\ell_d(c^{i_2}f)\cdots\ell_d(c^{i_k}f)\ell_d(x_k)\ell_d(\beta_{i_k})dd \\ &= d\ell_d(fwc^{i_1}fx_1\beta_{i_1}c^{i_2}f\cdots c^{i_k}fx_k\beta_{i_k})dd \\ &= ddr_d(f\alpha_{i_1}x_1c^{i_1}f\alpha_{i_2}x_2c^{i_2}f\cdots\alpha_{i_k}x_kc^{i_k}fu)d \\ &= ddr_d(f\alpha_{i_1})r_d(x_1)r_d(c^{i_1})r_d(f\alpha_{i_2})r_d(x_2)\cdots r_d(f\alpha_{i_k})r_d(x_k)r_d(c^{i_k})r_d(fu)d\,, \end{aligned} \tag{20}$$

where the first and the last words are the left hand and the right hand sides (respectively) of words in pairs in Z catenated correspondingly. This implies the existence of a solution of the PCP for the set W when $u \in \mathcal{A}_S$. Therefore, the PCP is undecidable.

5 Conclusion

A shorter and bit simpler proof for the undecidability of the PCP was given using the same source of undecidability, the Post normal systems, as in the original proof by Post. We are in no doubt that the present proof could have been found by Emil Post as well, but as a true pioneer of the field of computability he immediately would have noticed the following deficiency of the construction: when considering the size of an instance as constructed in the proof, Post's original construction gives an instance of size $|P| + 5$, but our new construction gives an instance of size $2|P| + 4$. As the undecidable problem in the normal system, the cardinality of P must be at least two, we realize that Post's proof gives a better bound for the undecidability.

Acknowledgement

The authors thank the anonymous referees whose comments helped to improve the presentation of the article.

References

[1] Church, Alonzo. Review of [6]. *The Journal of Symbolic Logic*, 8(1):50–52, 1943. DOI: 10.2307/2268006.

[2] Claus, V. Some remarks on PCP(k) and related problems. *Bull. EATCS*, 12:54–61, 1980.

[3] Halava, Vesa. Another proof of undecidability for the correspondence decision problem - Had I been Emil Post. *CoRR*, abs/1411.5197, 2014.

[4] Harju, Tero and Karhumäki, Juhani. *Morphisms*, pages 439–510. Springer Berlin Heidelberg, Berlin, Heidelberg, 1997. DOI: 10.1007/978-3-642-59136-5_7.

[5] Neary, Turlough. Undecidability in binary tag systems and the Post correspondence problem for five pairs of words. In *32nd International Symposium on Theoretical Aspects of Computer Science*, volume 30 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages 649–661. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2015.

[6] Post, Emil Leon. Formal reductions of the general combinatorial decision problem. *Amer. J. Math.*, 65:197–215, 1943. DOI: 10.2307/2371809.

[7] Post, Emil Leon. Recursively enumerable sets of positive integers and their decision problems. *Bull. Amer. Math. Soc.*, 50:284–316, 1944. DOI: 10.1090/S0002-9904-1944-08111-1.

[8] Post, Emil Leon. A variant of a recursively unsolvable problem. *Bull. Amer. Math. Soc.*, 52:264–268, 1946. DOI: 10.1090/S0002-9904-1946-08555-9.

[9] Sipser, Michael. *Introduction to the theory of computation. 3rd. ed.* Cengage Learning, 3rd. ed. edition, 2013.