THEORY OF COMPUTATION
Post' Correspondence Problem - 24

Prof. Dan A. Simovici

UMB

The Post Correspondence Problem can be viewed as a single player game played with a special set of dominoes. There is an infinite supply of dominoes of each kind.
To win a game is to reach a situation where the same word appears on the top halves as one the bottom halves when we read across from left to right.

Each domino has a word that appears on each half. A typical domino is shown below;

aba
baab

A Post correspondence problem is a finite set of dominoes. Each move consists in placing one of the dominoes of the system to the right of the dominoes placed in the previous moves.
To win the game we need to reach a situation where the same word appears on the top halves as well as on the bottom half of the dominoes.

If we start with a set of three dominoes

$\left[\frac{a}{aa}\right], \left[\frac{bb}{b}\right], \left[\frac{a}{bb}\right]$

one could have the following game:

$\left[\frac{a}{aa}\right] \left[\frac{a}{bb}\right] \left[\frac{bb}{b}\right] \left[\frac{bb}{b}\right]$

Note that:

- one of the dominoes is used twice;
- the word *aabbbb* appear both on the top half and the lower half.

A Post correspondence system has a solution if and only if it is possible to win the game defined by the system.

Theorem

There is no algorithm that can determine whether a Post correspondence system has a solution.

Proof:
We know that there exists a semi-Thue process Π on $\{a, b\}$ whose word problem is unsolvable. For every production $x \to y$ we can assume that $x, y \neq 0$.
We add to the productions of Π the productions $a \to a$ and $b \to b$ which have no effect on whether $u \overset{*}{\underset{\Pi}{\Rightarrow}} v$. However, it guarantees that whenever $u \overset{*}{\underset{\Pi}{\Rightarrow}} v$, there is a derivation

$$u = u_1 \underset{\Pi}{\Rightarrow} u_2 \underset{\Pi}{\Rightarrow} \cdots \underset{\Pi}{\Rightarrow} u_m = v$$

such that m is an odd number.

Proof cont'd:

Indeed, because of the added productions we have

$$u_i \underset{\Pi}{\Rightarrow} u_i$$

for every i so any step of the derivation can be repeated in order to change the length of the derivation.

Let $u, v \in \{a, b\}^*$. We construct a Post correspondence problem $\Pi_{u,v}$ which depends on Π, u, and v such that $\Pi_{u,v}$ has a solution if and only if $u \overset{*}{\underset{\Pi}{\Rightarrow}} v$.

Proving the existence of $\Pi_{u,v}$ with the above mentioned property implies the theorem. Indeed, if there were an algorithm for determining whether a PCP has a solution, this algorithm applied to $\Pi_{u,v}$ would determine whether $u \overset{*}{\underset{\Pi}{\Rightarrow}} v$. Since Π has an unsolvable word problem, this is impossible.

Proof cont'd:

Let the n productions of Π be $x_i \to y_i$ for $1 \leqslant i \leqslant n$, including $a \to a$ and $b \to b$.

The alphabet of the Post correspondence problem $\Pi_{u,v}$ consists of eight symbols:

$$a\ b\ \tilde{a}\ \tilde{b}\ [\ ]\ *\ \tilde{*}.$$

If $w \in \{a, b\}^*$ we denote by $\tilde{w}$ the word on $\{\tilde{a}, \tilde{b}\}$ obtained by replacing a, b with $\tilde{a}, \tilde{b}$, respectively.

Proof cont'd:

$\Pi_{u,v}$ consists of $2n + 4$ dominoes shown below:

- 4 dominoes:

$[u*$	$*$	$\tilde{*}$	$]$
$[$	$\tilde{*}$	$*$	$\tilde{*}v]$

- 2n dominoes corresponding to the n productions $x_i \to y_i$ of Π for $1 \leqslant i \leqslant n$:

$\tilde{y}_i$	y_i
x_i	$\tilde{x}_i$

The second group includes the four dominoes:

$$\left[\frac{a}{\tilde{a}}\right], \left[\frac{\tilde{a}}{a}\right], \left[\frac{b}{\tilde{b}}\right], \left[\frac{\tilde{b}}{b}\right]$$

Note that it is OK to use dominoes of the form

$\left[\dfrac{\tilde{p}}{p}\right]$ or $\left[\dfrac{\mathrm{p}}{\tilde{p}}\right]$

because any such dominoes can be assembled from dominoes in the list

$\left[\dfrac{a}{\tilde{a}}\right], \left[\dfrac{\tilde{a}}{a}\right], \left[\dfrac{b}{\tilde{b}}\right], \left[\dfrac{\tilde{b}}{b}\right]$

We prove first that if $u \overset{*}{\underset{\Pi}{\Rightarrow}} v$, then there is a solution to $\Pi_{u,v}$.

Since $u \overset{*}{\underset{\Pi}{\Rightarrow}} v$, there is a derivation

$$u = u_1 \underset{\Pi}{\Rightarrow} u_2 \underset{\Pi}{\Rightarrow} \cdots \underset{\Pi}{\Rightarrow} u_m = v,$$

where m is an odd number. Thus, for each i, $1 \leqslant i < m$ we have

$$u_i = p_i x_{j_i} q_i, u_{i+1} = p_i y_{j_i} q_i,$$

where we used the j_i^{th} production of Π, $x_{j_i} \rightarrow y_{j_i}$.

We claim that the word

$$[u_1 * \tilde{u}_2 \tilde{*} u_3 * \cdots * \tilde{u}_{m-1} \tilde{*} u_m]$$

is a solution of $\Pi_{u,v}$.
The beginning of the sequence of dominoes is:

$$\begin{bmatrix} [u_1 * \\ [\end{bmatrix} \begin{bmatrix} \tilde{p}_1 \\ p_1 \end{bmatrix} \begin{bmatrix} \tilde{y}_{j_1} \\ x_{j_1} \end{bmatrix} \begin{bmatrix} \tilde{q}_1 \\ q_1 \end{bmatrix} \begin{bmatrix} \tilde{*} \\ * \end{bmatrix} \ldots$$

At this stage the word on top is $[u_1 * \tilde{u}_2 \tilde{*}$ while at the bottom we have $[u_1 *$.

The play continues as follows:

$[u*$	$\tilde{p}_1$	$\tilde{y}_{j_1}$	$\tilde{q}_1$	$\tilde{*}$	p_2	y_{j_2}	q_2	$*$
$[$	p_1	x_{j_1}	q_1	$*$	$\tilde{p}_2$	$\tilde{x}_{j_2}$	$\tilde{q}_2$	$\tilde{*}$

The word at the top is $[u_1 * \tilde{u}_2 \tilde{*} u_3 *$ and the word at the bottom is $[u_1 * \tilde{u}_2 \tilde{*}$.

Since m is an odd number we can win as follows:

$[u*$	$\tilde{p}_1$		$\tilde{*}$	p_{m-1}	$y_{j_{m-1}}$	q_{m-1}	$]$
$[$	p_1	$\ldots$	$*$	$\tilde{p}_{m-1}$	$\tilde{x}_{j_{m-1}}$	$\tilde{q}_{m-1}$	$\tilde{*}v$

because at this point the top and the bottom words are equal to $[u_1 * \tilde{u}_2 \tilde{*} u_3 * \cdots * \tilde{u}_{m-1} \tilde{*} u_m]$.

Conversely, suppose that $\Pi_{u,v}$ has a solution w. Examining the set of tiles the only possible way to win is to play

$[u*$		$]$
$[$	and	$\tilde{*}v]$

first and last, respectively, because none of the other dominoes in Π have tops and bottoms that begin or end with the same symbols. Thus, w must begin with [and end with].

Since we know that the game looks like

$[u*$		$]$
$[$	$\ldots$	$\tilde{*}v]$

a solution looks line

$$[u * \cdots \tilde{*} v]$$

Continuing from the left the play must go

$[u*$	$\tilde{y}_{i_1}$	$\tilde{y}_{i_2}$		$\tilde{y}_{i_k}$	$\tilde{*}$
$[$	x_{i_1}	x_{i_2}	$\ldots$	x_{i_k}	$*$

where $x_{i_1} x_{i_2} \cdots x_{i_k} = u$. This is necessary in order for the bottom to catch up with $u*$ which is already at the top.

Writing $u = u_1$ and $u_2 = y_{i_1} y_{i_2} \cdots y_{i_k}$ we see that $u_1 \overset{*}{\underset{\Pi}{\Rightarrow}} u_2$ and that the solution has he form

$$[u_1 * \tilde{u}_2 \tilde{*} \cdots \tilde{*} v]$$

The play continues as

$[u*$	$\tilde{y}_{i_1}$	$\tilde{y}_{i_2}$		$\tilde{y}_{i_k}$	$\tilde{*}$	y_{j_1}	y_{j_2}		y_{j_ℓ}
$[$	x_{i_1}	x_{i_2}	$\ldots$	x_{i_k}	$*$	$\tilde{x}_{j_1}$	$\tilde{x}_{j_2}$	$\ldots$	$\tilde{x}_{j_\ell}$

$*$
$\tilde{*}$

where $u_2 = x_{j_1} x_{j_2} \cdots x_{j_k}$. Again, writing

$$u_3 = y_{j_1} y_{j_2} \cdots y_{j_\ell}$$

we have $u_2 \overset{*}{\underset{\Pi}{\Rightarrow}} u_3$ and the solution has the form

$$[u_1 * \tilde{u}_2 \tilde{*} u_3 * \cdots \tilde{*} v].$$

Continuing, it is clear that the solution can be written as

$$[u_1 * \tilde{u}_2 \tilde{*} u_3 * \cdots \tilde{u}_{m-1} \tilde{*} u_m]$$

where

$$u = u_1 \overset{*}{\underset{\Pi}{\Rightarrow}} u_2 \overset{*}{\underset{\Pi}{\Rightarrow}} u_3 \overset{*}{\underset{\Pi}{\Rightarrow}} \cdots \overset{*}{\underset{\Pi}{\Rightarrow}} u_{m-1} \overset{*}{\underset{\Pi}{\Rightarrow}} u_m = v,$$

so that $u \overset{*}{\underset{\Pi}{\Rightarrow}} v$.