Post's Correspondence Problem (PCP)

An instance of **PCP** consists of two finite lists of strings over some alphabet Σ. The two lists must be of equal length:

$A = w_1, w_2, \ldots, w_k$
$B = x_1, x_2, \ldots, x_k$

For each i, the pair (w_i, x_i) is said to be a *corresponding pair*. We say that this instance of PCP has a solution if there is a sequence of one or more indices:

$$i_1, i_2, \ldots, i_m \qquad m \geq 1$$

such that:

$$w_{i_1} w_{i_2} \ldots w_{i_m} = x_{i_1} x_{i_2} \ldots x_{i_m}$$

The PCP is: given (A, B), tell whether there is a solution.

Aside: To express **PCP** as a language, we let L_{PCP} be the language:

$\{\langle A, B\rangle | (A, B)$ instance of PCP with solution$\}$

Example: Consider (A, B) given by:

$A = 1, 10111, 10$
$B = 111, 10, 0$

Then $2, 1, 1, 3$ is a solution as:

$$\underbrace{10111}_{w_2}\underbrace{1}_{w_1}\underbrace{1}_{w_1}\underbrace{10}_{w_3} = \underbrace{10}_{x_2}\underbrace{111}_{x_1}\underbrace{111}_{x_1}\underbrace{0}_{x_3}$$

Note that $2, 1, 1, 3, 2, 1, 1, 3$ is another solution.

On the other hand, you can check that:

$A = 10, 011, 101$
$B = 101, 11, 011$

Does not have a solution.

Modified PCP (MPCP)

The **MPCP** has an additional requirement that the first pair in the solution must be the first pair of (A, B).

So $i_1, i_2, \ldots, i_m$, $m \geq 0$, is a solution to the (A, B) instance of **MPCP** if:

$$w_1 w_{i_1} w_{i_2} \ldots w_{i_m} = x_1 x_{i_1} x_{i_2} \ldots x_{i_m}$$

Def: We say that $i_1, i_2, \ldots, i_r$ is a **partial solution** of PCP if one of the following is the prefix of the other:

$$w_{i_1} w_{i_2} \ldots w_{i_r} \qquad x_{i_1} x_{i_2} \ldots x_{i_r}$$

Same def holds for MPCP, but w_1, x_1 must be at the beginning.

We now show:

1. If **PCP** is decidable, then so is **MPCP**.

2. If **MPCP** is decidable, then so is L_U.

3. Since L_U is *not* decidable, neither is **(M)PCP**.

$\boxed{\textbf{PCP decidable} \Longrightarrow \textbf{MPCP decidable}}$

We show that given an instance (A, B) of MPCP, we can construct an instance (A', B') of PCP such that:

$$(A, B) \text{ has solution } \iff (A', B') \text{ has solution}$$

Let (A, B) be an instance of MPCP over the alphabet Σ. Then (A', B') is an instance of PCP over the alphabet $\Sigma' = \Sigma \cup \{*, \$\}$.

If $A = w_1, w_2, w_3, \ldots, w_k$, then
$A' = *\mathbf{w}_1*, \mathbf{w}_1*, \mathbf{w}_2*, \mathbf{w}_3*, \ldots, \mathbf{w}_k*, \$$.

If $B = x_1, x_2, x_3, \ldots, x_k$, then
$B' = *\mathbf{x}_1, *\mathbf{x}_1, *\mathbf{x}_2, *\mathbf{x}_3, \ldots, *\mathbf{x}_k, *\$$.

where if $x = a_1 a_2 a_3 \ldots a_n \in \Sigma^*$, then
$\mathbf{x} = a_1 * a_2 * a_3 * \ldots * a_n$.

For example: If (A, B) is an instance if MPCP given as:

$A = 1, 10111, 10$
$B = 111, 10, 0$

Then (A', B') is an instance of PCP given as follows:

$A' = *1*, 1*, 1*0*1*1*1*, 1*0*, \$$
$B' = *1*1*1, *1*1*1, *1*0, *0, *\$$

$\boxed{\textbf{MPCP decidable} \Longrightarrow L_U \text{ decidable}}$

Given a pair (M, w) we construct an instance (A, B) of MPCP such that:

TM M accepts $w \iff (A, B)$ has a solution.

Idea: The MPCP instance (A, B) simulates, in its partial solutions, the computation of M on w.

That is, partial solutions will be of the form:

$$\#\alpha_1\#\alpha_2\#\alpha_3\#\ldots$$

where α_1 is the initial ID of M on w, and for all i, $\alpha_i \vdash \alpha_{i+1}$.

The string from the B list will always be one ID ahead of the A list; the A list will be allowed to "catch-up" only when M accepts w.

To simplify things, we assume the result of Theorem 8.12 which says that we may assume that our TM M:

1. Never prints a blank.

2. Never moves left from its initial head position.

The IDs of M will always be of the form $\alpha q \beta$, where α, β are non-blank tape symbols and q is a state.

Let M be a TM satisfying Theorem 8.12, and $w \in \Sigma^*$. We construct an instance (A, B) of MPCP as follows:

1. A: $\#$
   B: $\#q_0 w\#$

2. A: $X_1, X_2, \ldots, X_n, \#$
   B: $X_1, X_2, \ldots, X_n, \#$
   where the X_i are all the tape symbols.

3. To simulate a move of M, for all non-accepting $q \in Q$:

list A	list B	
qX	Yp	if $\delta(q, X) = (p, Y, \rightarrow)$
ZqX	pZY	if $\delta(q, X) = (p, Y, \leftarrow)$
$q\#$	$Yp\#$	if $\delta(q, B) = (p, Y, \rightarrow)$
$Zq\#$	$pZY\#$	if $\delta(q, B) = (p, Y, \leftarrow)$

4. If the ID at the end of B has an accepting state, then we need to allow A to catch up with B. So we need for all accepting states q, and all symbols X, Y:

list A	list B
XqY	q
Xq	q
qY	q

5. Finally, after using 4 and 3 above, we end up with $x\#$ and $x\#q\#$, where x is a long string. Thus we need $q\#\#$ in A and $\#$ in B to complete the catching up.

Example: consider
$M = (\{q_1, q_2, q_3\}, \{0, 1\}, \{0, 1, B\}, \delta, q_1, B, \{q_3\})$,
where

	0	1	B
q_1	$(q_2, 1, \rightarrow)$	$(q_2, 0, \leftarrow)$	$(q_2, 1, \leftarrow)$
q_2	$(q_3, 0, \leftarrow)$	$(q_1, 0, \rightarrow)$	$(q_2, 0, \rightarrow)$
q_3	—	—	—

Rule	list A	list B	Source
1	$\#$	$\#q_1 01\#$	
2	0 1 $\#$	0 1 $\#$	
3	$q_1 0$ $0q_1 1$ $1q_1 1$ $0q_1\#$ $1q_1\#$ $0q_2 0$ $1q_2 0$ $q_2 1$ $q_2\#$	$1q_2$ $q_2 00$ $q_2 10$ $q_2 01\#$ $q_2 11\#$ $q_3 00\#$ $q_3 10\#$ $0q_1$ $0q_2\#$	$\delta(q_1, 0) = (q_2, 1, \rightarrow)$ $\delta(q_1, 1) = (q_2, 0, \leftarrow)$ $\delta(q_1, 1) = (q_2, 0, \leftarrow)$ $\delta(q_1, B) = (q_2, 1, \leftarrow)$ $\delta(q_1, B) = (q_2, 1, \leftarrow)$ $\delta(q_2, 0) = (q_3, 0, \leftarrow)$ $\delta(q_2, 0) = (q_3, 0, \leftarrow)$ $\delta(q_2, 1) = (q_1, 0, \rightarrow)$ $\delta(q_2, B) = (q_2, 0, \rightarrow)$
4	$0q_3 0$ $0q_3 1$ $1q_3 0$ $1q_3 1$ $0q_3$ $1q_3$ $q_3 0$ $q_3 1$	q_3 q_3 q_3 q_3 q_3 q_3 q_3 q_3	
5	$q_3\#\#$	$\#$	

The TM M accepts the input 01 by the sequence of moves:

$$q_1 01 \vdash 1q_2 1 \vdash 10q_1 \vdash 1q_2 01 \vdash q_3 101$$

We examine the sequence of partial solutions that mimics this computation of M and eventually leads to a solution.

We must start with the first pair (MPCP):

A: #
B: $\#q_1 01\#$

The only way to extend this partial solution is with the corresponding pair $(q_1 0, 1q_2)$, so we obtain:

A: $\#q_1 0$
B: $\#q_1 01\#1q_2$

Now using copying pairs we obtain:

A: $\#q_1 01\#1$
B: $\#q_1 01\#1q_2 1\#1$

Next corresponding pair is $(q_2 1, 0q_1)$:

A: $\#q_1 01\#1q_2 1$
B: $\#q_1 01\#1q_2 1\#10q_1$

Now careful! We only copy the next two symbols to obtain:

A: $\#q_1 01\#1q_2 1\#1$
B: $\#q_1 01\#1q_2 1\#10q_1\#1$

because we need the $0q_1$ as the head now moves left, and use the next appropriate corresponding pair which is $(0q_1\#, q_2 01\#)$ and obtain:

A: $\#q_1 01\#1q_2 1\#10q_1\#$
B: $\#q_1 01\#1q_2 1\#10q_1\#1q_2 01\#$

We can now use another corresponding pair $(1q_20, q_310)$ right away to obtain:

A: $\#q_101\#1q_21\#10q_1\#1q_20$
B: $\#q_101\#1q_21\#10q_1\#1q_201\#q_310$

and note that we have an accepting state! We use two copying pairs to get:

A: $\#q_101\#1q_21\#10q_1\#1q_201\#$
B: $\#q_101\#1q_21\#10q_1\#1q_201\#q_3101\#$

and we can now start using the rules in 4. to make A catch up with B:

A: $\ldots\#q_31$
B: $\ldots\#q_3101\#q_3$

and we copy three symbols:

A: $\ldots\#q_3101\#$
B: $\ldots\#q_3101\#q_301\#$

And again catch up a little:

A: $\ldots\#q_3 101\#q_3 0$
B: $\ldots\#q_3 101\#q_3 01\#q_3$

Copy two symbols:

A: $\ldots\#q_3 101\#q_3 01\#$
B: $\ldots\#q_3 101\#q_3 01\#q_3 1\#$

and catch up:

A: $\ldots\#q_3 101\#q_3 01\#q_3 1$
B: $\ldots\#q_3 101\#q_3 01\#q_3 1\#q_3$

and copy:

A: $\ldots\#q_3 101\#q_3 01\#q_3 1\#$
B: $\ldots\#q_3 101\#q_3 01\#q_3 1\#q_3\#$

And now end it all with the corresponding pair ($q_3\#\#, \#$) given by rule 5. to get matching strings:

A: $\ldots\#q_3 101\#q_3 01\#q_3 1\#q_3\#\#$
B: $\ldots\#q_3 101\#q_3 01\#q_3 1\#q_3\#\#$

THEREFORE: we reduced L_U to the MPCP. Now, we can solve L_U by producing a carefully crafted instance of MPCP (A, B), and asking if it has a solution. If yes, then we know that M accepts w.

Since we have already shown that L_U is undecidable, MPCP must also be undecidable. Thus, PCP is undecidable.

NEXT: We can now use the fact that PCP is undecidable to show that a number of questions about CFLs are undecidable.