Post Correspondence Problem (PCP)

The PCP was first introduced by Emil Post in 1946. Later the problem was found to have many applications in the theory of formal languages.

Definition :-

An instance of PCP consists of two lists $A = w_1 --- w_k$ and $B = x_1 --- x_k$ of strings over some alphabet Σ. This instance of PCP has a solution if there is any sequence of integers $i_1, i_2, i_3 --- i_m$, with $m \geq 1$ such that

$$w_{i_1} w_{i_2} w_{i_3} --- w_{i_m} = x_{i_1} x_{i_2} ---- x_{i_m}$$

The sequence $i_1 i_2 i_3 ---- i_m$ is a solution to this instance of PCP

Example :-

①

	List A	List B
i	w_i	x_i
1	110	110110
2	0011	00
3	0110	110

Solution ① (2,3,1)
② (2,1,1,3,2,1,1,3)

②

i	w_i	x_i
1	011	101
2	11	011
3	1101	110

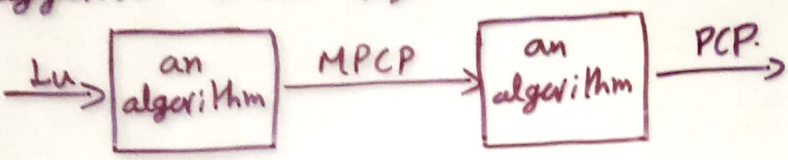
No solution.

③

i	w_i	x_i
1	b	b^3
2	bab^3	ba
3	ba	a

Solution : (2, 1, 1, 3)

We shall prove PCP undecidable by reducing L_u to PCP. To facilitate the proof we introduce "modified PCP". We reduce L_u to the modified PCP. The chain of reductions is suggested in the figure.



Since the original L_u is known to be undecidable, we conclude that PCP is undecidable.

__The "Modified" PCP__:

In the modified PCP, there is the additional requirement on a solution that the first pair on A & B lists must be the first pair in the solution.

An instance of MPCP is two lists $A = w_1, w_2 \ldots w_m$

$B = x_1, x_2 \ldots x_m$.

Solution is a list of 0 or more integers $i_1, i_2 \ldots i_m$

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

Notice that the pair (w_1, x_1) is forced to be at the beginning of two strings even though the index 1 is not mentioned at the front of the list that is the solution.

An important step in showing PCP is undecidable is reducing MPCP to PCP. Later we show MPCP is undecidable by reducing L_u to MPCP.

Lemma:- If PCP were decidable, then MPCP would be decidable ③

(or)

Reducing MPCP in to PCP.

Proof:- For every instance of MPCP, with alphabet Σ, we can construct an instance of PCP as follows.

→ In the strings of A list, the *'s follow the symbols of Σ

In the strings of B list, the *'s precede the symbols of Σ

We are given an instance of MPCP with lists $A = \omega_1\, w_2 --- \omega_k$

$B = x_1\, x_2 --- x_k$

We construct a PCP instance $C = y_0\, y_1 --- y_{k+1}$

$D = z_0\, z_1 --- z_{k+1}$ as follows.

1. For $i = 1, 2, ---k$. let $Y_i \rightarrow \omega_i$ with * after each symbol of ω_i, & let $z_i \rightarrow x_i$ with a * before each symbol of x_i

2. $y_0 = *y_1$ & $z_0 = z_1$ That is, the 0th pair looks like pair 1, except that there is an extra * at the begining of the string from the first list. Note that the 0th pair will be the only pair in the PCP instance where both strings begin with the same symbol, so any solution to this PCP instance will have to begin with index 0

3. $y_{k+1} = \$$ and $z_{k+1} = *\$$.