CS741 Assignment 2 - Q3 Report

Neel Aryan Gupta 180050067
Pulkit Agrawal 180050081
Tathagat Verma 180050111

NOTE - The code has been written in C++ inside the file q3.cpp. To run this code, O2 flag is recommended for better performance.

```
g++ -O2 q3.cpp -o q3
./q3 < input.txt
```

Time complexity - $O(S * 2^{3S} + T * 2^{2N} * (N/S) + M * N * 2^{K})$

The first two terms of the complexity come from calculating the subgraph for maximum bias and have been explained in the report for Q2.

M = number of plaintext-ciphertext pairs given in the input

K = the size of the subkey (in number of bits) obtained from the number of active S-Boxes of the maximum bias subgraph

The third term involves iterating over all possible subkeys and all pairs of plaintext-ciphertext to update the score for each subkey, as discussed in the lectures. The N factor comes from xoring the N bits of the plaintext, ciphertext and the subkey.

Algorithm - The algorithm for finding the maximum bias subgraph has been explained in the report for Q2. Here, to find the subkey bits of the last round key, for each plaintext-ciphertext pair and each subkey, the ciphertext bits corresponding to subkey are xorred with the subkey and then passed through the respective inverse S-Boxes to obtain the bits involved in the bias equation. Now these bits are xorred with the plaintext bits (as per the max bias subgraph) and we increment the score if the result of xor is 0. Finally, those subkeys are shown in output whose scores **deviate** the most from M/2 where M is the number of plaintext-ciphertext pairs. Note that there may be multiple subkeys corresponding to the same maximum deviations, all such keys are reported by the program.

Input/Output - The input for plaintext-ciphertext is assumed to be of the following form:

```
Num of stages (T)
Size of plaintext (N)
Permutation (of size N, space separated)
Size of S-box (S)
S-box (of size 2^S)
Number of plaintext-ciphertext pairs (M)
plaintext1 ciphertext1
plaintext2 ciphertext2
...
```

plaintext and *ciphertext* are given as N-bit **integers.**

The subkey is reported in binary form which is split using | symbol, to denote different S-Box blocks. Dash symbol **-** is used to represent that this bit does not belong to the subkey. For example,

1011|----|0101|----

The above output corresponds to a 16-bit block and a 4-bit S-Box, where the active S-Boxes in the last round are 1st and 3rd S-Boxes respectively. The actual subkey would be *10110101*.

Sample output:

```
P5, P7, K05, K07, K13, K29, C9
Bias = 9/32

**************
0000|0101|0000
----+----+----
0000|0001|0000

0001|0000|0000
----+----+----
0010|0000|0000

0000|0000|0100
----+----+----
0000|0000|0100

0000|0000|0100
**************

Keys with Max deviation :
----|----|0011
```

Here the only active S-Box is 3rd one and *0011* is the candidate subkey.