# Assignment 1
# Monoalphabetic Substitution Ciphers Cryptanalysis

Kunal Dhyani (2025MCS2112)

## 1 Introduction

This report describes the methodology used to crack multiple monoalphabetic substitution ciphers containing lowercase letters, digits, and symbols. The plaintexts are English paragraphs with punctuation preserved (comma, semicolon, dot).

The objective was to decipher or recover the plaintext and the substitution key using frequency analysis and heuristic search.

## 2 Cipher Model

Each cipher follows these rules:(assumed)

- Plaintext contains only lowercase letters and spaces.
- Ciphertext contains lowercase letters, digits, and symbols.
- Punctuation characters `,` `;` `.` `!` remain unchanged.
- One-to-one monoalphabetic substitution is used.

## 3 Approach

Cryptanalysis algorithm consists of the following stages:

### 3.1 1. Frequency Analysis

For initial guessed key generation, based on general frequency distribution pattern for english language. The frequency of each cipher character is computed and mapped to the expected English letter frequency ordering:

This is used for an initial guess for the key:

```
"etaonihsrdlumcwfygpbvkxjqz"
```

### 3.2 3. Hill Climbing with Random Restarts

A hill-climbing search is performed by mutating the key:

- Random swaps of key letters
- Partial shuffling
- Rotations and reversals

Multiple restarts are used to avoid local optima.

### 3.3 2. Scoring Function

Each decrypted candidate is scored using:

- Dictionary word matching (used dictionary containing approx. 4 Lakh words)
- Penalty for vowel-less words and unmapped characters
- Bonus for most common English words

The final score is:

$$\text{Score} = \text{dictionary matches} - \text{penalties} - \text{vowel less words}$$

## 4 Ciphertext 1

```
qots4o 7#8417o 17 z 4syz831r x4zyz t1wy x14or3ox q5 41rnz4x w18uwz3o4 z8x 24133o8 q5
w18uwz3o4 z8x u1y u41@z8. 3no t1473 1873zwwyo83 18 3no qots4o 341ws$5, 13 tswws27 vo77o,
o3nz8 nz2uo z8x row18o, v#w1o xowp5 z7 3no5 yoo3 s8 z o#4z1w 34z18 z8x x17oyqz4u 18
91o88z 3s 7po8x 3no 81$n3 3s$o3no4. 187p14ox q5 po47s8zw o0po41o8ro7, w18uwz3o4
rswwzqs4z3ox 213n u1y u41@z8 3s xo9owsp 3no 7r4oo8pwz5, 2ns p4o91s#7w5 zppoz4ox 18 n17
t1wy7 7wzruo4 z8x xz@ox z8x rs8t#7ox. rz7318$ 2z7 o03o8719o, 13 3ssu 818o ys83n7 ts4
nz2uo z8x xowp5 3s qo rz73, 213n 3no pz14 zw7s rs8341q#318$ #8r4ox13ox 4o2413o7.
p418r1pzw pns3s$4zpn5 3ssu pwzro o8314ow5 18 91o88z. $ssx t1wy 3s 2z3rn.
```

### 4.1 Deciphered Key

Deciphered Key: zqrxot$n1vuwy8spx473#9205@

### 4.2 Deciphered Plaintext

> before sunrise is a romantic drama film directed by richard linklater and written by linklater and kim krizan. the first installment in the before trilogy, it follows jesse, ethan hawke and celine, julie delpy as they meet on a eurail train and disembark in vienna to spend the night together. inspired by personal experiences, linklater collaborated with kim krizan to develop the screenplay, who previously appeared in his films slacker and dazed and confused. casting was extensive, it took nine months

for hawke and delpy to be cast, with the pair also contributing uncredited rewrites. principal photography took place entirely in vienna. good film to watch.

# 5 Ciphertext 2

```
q#qp1#x #z p1 p5t@#xp1 t4#x @v5p1xt p1o o#zpzqt@ $#05 o#@txqto, w@#qqt1, 4@vorxto, p1o
xv to#qto 68 9p5tz xp5t@v1. #1xv@4v@pq#1n 6vq3 3#zqv@#xp0 p1o $#xq#v1p0#sto pz4txqz, #q
#z 6pzto v1 pxxvr1qz v$ q3t z#12#1n v$ q3t @5z q#qp1#x, p1o zqp@z 0tv1p@ov o#xp4@#v p1o
2pqt w#1z0tq pz 5t56t@z v$ o#$$t@t1q zvx#p0 x0pzztz w3v $p00 #1 0vut p6vp@o q3t z3#4
or@#1n #qz #00 $pqto 5p#ot1 uv8pnt. xp5t@v1 #1z4#@pq#v1 $v@ q3t $#05 xp5t $@v5 3#z
$pzx#1pq#v1 w#q3 z3#4w@tx2z, 3t $t0q p 0vut zqv@8 #1qt@z4t@zto w#q3 q3t 3r5p1 0vzz wvr0o
6t tzzt1q#p0 qv xv1ut8 q3t t5vq#v1p0 #54pxq v$ q3t o#zpzqt@. 4@vorxq#v1 6tnp1 w3t1 xp5t@v1
z3vq $vvqpnt v$ q3t pxqrp0 q#qp1#x w@tx2. p1vq3t@ nvvo $#05.
```

## 5.1 Deciphered Key

Deciphered Key: p6xot$n3#92051v4x@zqruwx8s

## 5.2 Deciphered Plaintext

titanic is an american epic romance and disaster film directed written produced and co edited by james cameron incorporating both historical and fictionalized aspects it is based on accounts of the sinking of the rms titanic and stars leonardo dicaprio and kate winslet as members of different social classes who fall in love aboard the ship during its ill fated maiden voyage cameron inspiration for the film came from his fascination with shipwrecks he felt a love story interspersed with the human loss would be essential to convey the emotional impact of the disaster production began when cameron shot footage of the actual titanic wreck another good film

# 6 Ciphertext 3

```
vy04p0t x08 ox8n0ot0n0, 5#tt#x @4734 5u y8o p#4 40q# px#q1y04v, 30o 04 84v804 3x8t#x
90q7$o 97x y8o q7v#x4 y84v$ot048 z8t#x0t$x#. q$4oy8 px#q1y04v 30o 0 p874##x 79 y84v8
04v $xv$ o7180z 981t874. y# 30o 74# 79 ty# 98xot 0$ty7xo t7 3x8t# 057$t 10ot#
y8#x0x1y8#o 04v ty# pz82yto 79 37q#4 04v z057$x#xo px#n0z#4t 84 ty# o718#tu. y# 8o 74#
79 ty# q7ot 1#z#5x0t#v 3x8t#xo 79 ty# 84v804 o$5174t84#4t,04v 8o x#20xv#v 0o 74# 79 ty#
97x#q7ot y84v8 3x8t#xo 79 ty# #0xzu t3#4t8#ty 1#4t$xu. y8o 37x@o 841z$v# 27v004,
@0xq05y77q8, 20504, q04o0x7n0x, 8v20y.
```

## 6.1 Deciphered Key

Deciphered Key: 051v#92y8x@zq47pxxot$n3xux

### 6.2 Deciphered Plaintext

dhanpat rai srivastava better known by his pen name premchand was an indian writer famous for his modern hindustani literature munshi premchand was a pioneer of hindi and urdu social fiction he was one of the first authors to write about caste hierarchies and the plights of women and labourers prevalent in the society he is one of the most celebrated writers of the indian subcontinentand is regarded as one of the foremost hindi writers of the early twentieth century his works include godaan karmabhoomi gaban mansarovar idgah

## 7 Implementation Details

The implementation was written in **C++** and includes:

- Dictionary-based validation
- Preferred most frequently words used in English langauage
- Randomized hill climbing with restarts

The algorithm successfully recovers readable plaintext for long ciphertexts without human intervention.

## 8 Resources used

1. Dictionary: words_alpha.txt
    - Source: Github Link