1 Theory

Indices of letters:

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

Measure of Roughness (**MR**) is a measure how much a distribution differs from a uniform distribution.

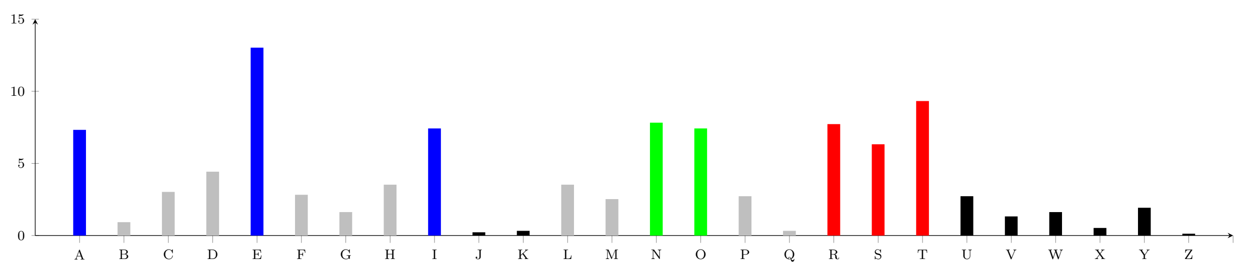
$$\mathbf{MR} = \sum_i \left(p_i - \frac{1}{26}\right)^2 = \sum_i p_i^2 - 2\frac{1}{26}\underbrace{\sum_i p_i}_{=1} + \underbrace{\sum_i \left(\frac{1}{26}\right)^2}_{=26\cdot\frac{1}{26^2}} = \sum_i p_i^2 - \frac{1}{26} \approx \sum_i p_i^2 - 0.038 \ .$$

Index of coincidence **IC** is an approximation to $\sum_i p_i^2$. For a ciphertexts $Y = y_1 y_2 \dots y_N$ and $Y' = y'_1 y'_2 \dots y'_M$ with characteristic frequency of letter i denoted as f_i, the index of coincidence (**IC**) is

$$\mathbf{IC}(\mathrm{Y}) = \frac{\sum_i f_i(f_i - 1)}{N(N-1)} \ , \qquad \mathbf{IC}(\mathrm{Y,Y'}) = \sum_i \frac{f_i f'_i}{NM} \ .$$

I.C. approximates the probability that any two letters randomly sampled from a distribution (or from two different distributions) will be the same. Since IC approximates $\sum_i p_i^2$, it has the same range of variation 0.038 to 0.066. The lower bound corresponds to a uniform distribution, and the upper bound correponds to monoalphabeticity. The number 0.066 is obtained by summing up the squared characteristic frequencies of English letters. On average, in a 1000 letter long sample of English text, the letters are distributed as follows:

A	73	*B*	9	*C*	30	*D*	44	*E*	130	*F*	28
G	16	*H*	35	*I*	74	*J*	2	*K*	3	*L*	35
M	25	*N*	78	*O*	74	*P*	27	*Q*	3	*R*	77
S	63	*T*	93	*U*	27	*V*	13	*W*	16	*X*	5
Y	19	*Z*	1								



The same picture would result from the examination of any reasonably long plain language text. Relative frequencies may vary slightly, but the basic facts remain the same:

- Evenly spaced vowels `A E I` with high frequency are evenly spaced 4 letters apart.

- Letter `E` is the most frequent of all the letters
- Consecutive part `N,O` have high frequency
- Consecutive triplet `R,S,T` has high frequency
- The pair `J,K` has low frequency
- The string `U,V,W,X,Y,Z` has low frequency.

2 Tasks

1. An additive cipher maps plaintext G to ciphertext X. What is the encryption key? Which decryption key will allow to reconstruct the plaintext?

   Solution. If $G \mapsto X$ by an additive cipher, it means that $E_z(6) = 23$ or $23 = 6 + z \pmod{26}$. In turn, $z = 23 - 6 = 17 \pmod{26}$. The encryption key is 17. The decryption key is $26 - 17 = 9$. Given X, we can get G as $23 + 9 \equiv 6 \pmod{26}$.

2. We know that a ciphertext was produced by a shift cipher, and that the encryption key was 17. What is the decryption key?

   Solution. For an encryption key e the corresponding decryption key is $26 - e$. Hence, the decryption key is $26 - 17 = 9$.

3. We know that the plaintext word `THE` is encrypted by an affine cipher into trigam `NHM`. What is the encryption key? What is the decryption key?

   Solution. To obtain the encryption key (a, b), we construct a system of congruences:

   $$19a + b = 13 \ ,$$
   $$7a + b = 7 \ ,$$
   $$4a + b = 12 \ .$$

   If we subtract the third equation from the second, we get $3a = 21$, and hence $a = 7$. To get the value of b, we put value of a in the first equation to get $3 + b = 13$, and hence $b = 10$.

   To get the decryption key, we construct another system of congruences

   $$13a + b = 19 \ ,$$
   $$7a + b = 7 \ ,$$
   $$12a + b = 4 \ .$$

   Subtracting the third equation from the first one, we get $a = 15$, and plugging this value into the second equation, we have $1 + b = 7$, and hence $b = 6$. So the encryption key is $(7, 10)$, and the decryption key is $(15, 6)$.

   For any encryption key (a, b) there exists a corresponding decryption key $(a^{-1}, -a^{-1}b)$. Since $7^{-1} \equiv 15 \pmod{26}$ and $-15 \cdot 10 \equiv 6 \pmod{26}$, then encryption key $(7, 10)$ has corresponding decryption key $(15, 6)$.

4. A ciphertext obtained by an affine cipher with key $(3, 17)$. Which key will you use to decrypt it?

   Solution. It can be seen that the encryption key $(3, 17)$ maps input 2 to 23, and input 3 to 0 as shown below.

$$f(2) = 2 \cdot 3 + 17 = 23 \pmod{26} \Longleftrightarrow 2 \mapsto 23 \ ,$$
$$f(3) = 3 \cdot 3 + 17 \equiv 0 \pmod{26} \Longleftrightarrow 3 \mapsto 0 \ .$$

   To reconstruct the decryption key, we construct a system of congruences

$$23a + b = 2 \ ,$$
$$0a + b = 3 \ .$$

   From this equation, we immediately get the value of $b = 3$. Plugging it into the first equation, we have $23a = 25$. Since $23^{-1} = 17$, multiplying both sides of the equation by 17, we get $a = 9$. So the decryption key is $(9, 3)$.

5. What is the I.C. of the ciphertext `EPYEPOPDZSZUFPO`?

   Solution.

$$\begin{aligned} \text{IC} &= \frac{f_E \cdot (f_E - 1) + f_P \cdot (f_P - 1) + f_O \cdot (f_O - 1) + f_Z \cdot (f_Z - 1)}{15 \cdot 14} \\ &= \frac{2 \cdot 1 + 4 \cdot 3 + 2 \cdot 1 + 2 \cdot 1}{15 \cdot 14} = \frac{18}{210} = 0.086 \ . \end{aligned}$$

6. Encrypt the word `MORNING` using a shift cipher with key 11.

   Solution.

$$\begin{aligned} &\text{M} \Rightarrow 12 + 11 \equiv 23 \pmod{26} \Rightarrow \text{X} \\ &\text{O} = 14 + 11 \equiv 25 \pmod{26} \Rightarrow \text{Z} \\ &\text{R} = 17 + 11 \equiv 2 \pmod{26} \Rightarrow \text{C} \\ &\text{N} = 13 + 11 \equiv 24 \pmod{26} \Rightarrow \text{Y} \\ &\text{I} = 8 + 11 \equiv 19 \pmod{26} \Rightarrow \text{T} \\ &\text{N} = 13 + 11 \equiv 24 \pmod{26} \Rightarrow \text{Y} \\ &\text{G} = 6 + 11 \equiv 17 \pmod{26} \Rightarrow \text{R} \end{aligned}$$

   Hence, `MORNING` corresponds to the ciphertext `XZCYTYR`.

7. Encrypt the word `SYMBOL` using an affine cipher with key $(3, 2)$.

Solution.

$$
\begin{aligned}
\text{S} &\Rightarrow 18 \cdot 3 + 2 \equiv 4 \pmod{26} \Rightarrow \text{E} \\
\text{Y} &\Rightarrow 24 \cdot 3 + 2 \equiv 22 \pmod{26} \Rightarrow \text{W} \\
\text{M} &\Rightarrow 12 \cdot 3 + 2 \equiv 12 \pmod{26} \Rightarrow \text{M} \\
\text{B} &\Rightarrow 1 \cdot 3 + 2 \equiv 5 \pmod{26} \Rightarrow \text{F} \\
\text{O} &\Rightarrow 14 \cdot 3 + 2 \equiv 18 \pmod{26} \Rightarrow \text{S} \\
\text{L} &\Rightarrow 11 \cdot 3 + 2 \equiv 9 \pmod{26} \Rightarrow \text{J}
\end{aligned}
$$

Hence, `SYMBOL` corresponds to `EWMFSJ`.

8. Encrypt the word `PARADOX` using a Vigenère cipher with key `YESTERDAY`.

Solution. Let us construct the Vigenère table for key `YESTERDAY`. The first letter of the

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X

plaintext is encrypted using the first row of the table, hence $P \mapsto N$, the second letter of the plaintext is encrypted using the second alphabet, hence $A \mapsto E$, etc. The plaintext `PARADOX` corresponds to the ciphertext `NEJTHFO`.