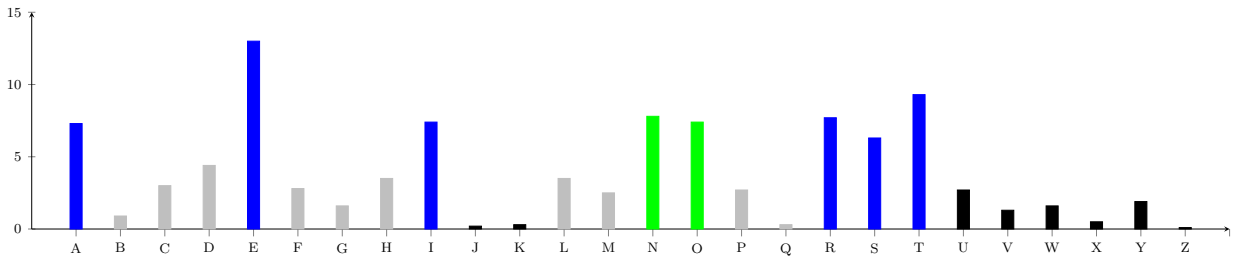
1 Useful Information

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



2 Tasks

Find the difference in keys

Given two ciphertexts Y : `M BQDEAZ ITA ZQHQD YMPQ M YUEFMWQ ZQHQD FDUQP MZKFTUZS ZQI` and Y' : `GZVMI AMJH TZNOZMYVT GDQZ AJM OJYVT CJKZ AJM OJHJMMJR`, which are two different messages encrypted by shift cipher using different keys z and z'. Find $z - z' \pmod{26}$.

Suppose cryptogram Y was encrypted with key z, and cryptogram Y' was encrypted with key z'. We need to find the difference $z - z'$. If we encrypt Y' with another key g, then we will obtain another cryptogram $E_g(E'_z(x)) = E_{g+z'}(x)$. For some value of g, the value $g + z' = z$, and this value of g will be the difference $g = z - z'$.

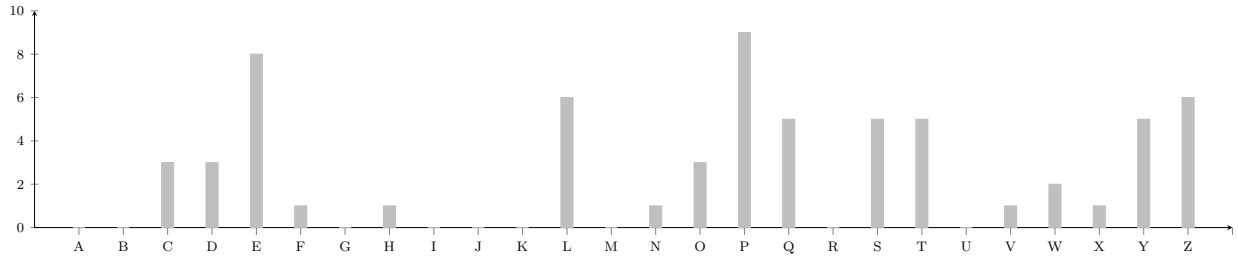
$g = 0$ $\mathrm{IC}(Y, E_g(Y')) = 0.047$
$g = 1$ $\mathrm{IC}(Y, E_g(Y')) = 0.031$
$g = 2$ $\mathrm{IC}(Y, E_g(Y')) = 0.026$
$g = 3$ $\mathrm{IC}(Y, E_g(Y')) = 0.044$
$g = 4$ $\mathrm{IC}(Y, E_g(Y')) = 0.056$
$g = 5$ $\mathrm{IC}(Y, E_g(Y')) = 0.029$
$g = 6$ $\mathrm{IC}(Y, E_g(Y')) = 0.044$
$g = 7$ $\mathrm{IC}(Y, E_g(Y')) = 0.052$
$g = 8$ $\mathrm{IC}(Y, E_g(Y')) = 0.036$
$g = 9$ $\mathrm{IC}(Y, E_g(Y')) = 0.031$
$g = 10$ $\mathrm{IC}(Y, E_g(Y')) = 0.040$
$g = 11$ $\mathrm{IC}(Y, E_g(Y')) = 0.028$
$g = 12$ $\mathrm{IC}(Y, E_g(Y')) = 0.034$
$g = 13$ $\mathrm{IC}(Y, E_g(Y')) = 0.052$
$g = 14$ $\mathrm{IC}(Y, E_g(Y')) = 0.025$
$g = 15$ $\mathrm{IC}(Y, E_g(Y')) = 0.029$
$g = 16$ $\mathrm{IC}(Y, E_g(Y')) = 0.049$
$g = 17$ $\mathrm{IC}(Y, E_g(Y')) = 0.066$
$g = 18$ $\mathrm{IC}(Y, E_g(Y')) = 0.032$
$g = 19$ $\mathrm{IC}(Y, E_g(Y')) = 0.035$
$g = 20$ $\mathrm{IC}(Y, E_g(Y')) = 0.036$
$g = 21$ $\mathrm{IC}(Y, E_g(Y')) = 0.044$
$g = 22$ $\mathrm{IC}(Y, E_g(Y')) = 0.037$
$g = 23$ $\mathrm{IC}(Y, E_g(Y')) = 0.028$
$g = 24$ $\mathrm{IC}(Y, E_g(Y')) = 0.032$
$g = 25$ $\mathrm{IC}(Y, E_g(Y')) = 0.037$

It can be seen that for $g = 17$, the value of $\mathrm{IC}(Y, E_g(Y'))$ is the closest to 0.066. Hence, $z - z' = 17$.

Decrypt the messages

1. `ESPCP TD L ETOP TY ESP LQQLTCD ZQ XPY HSTNS ELVPY LE ESP QWZZO WPLOD ZY EZ QZCEFYP`

The **IC** of the ciphertext is 0.0692 which gives us a good indication of monoalphabeticity. Let's examine the frequency distribution of the ciphertexts to get additional information. We can spot

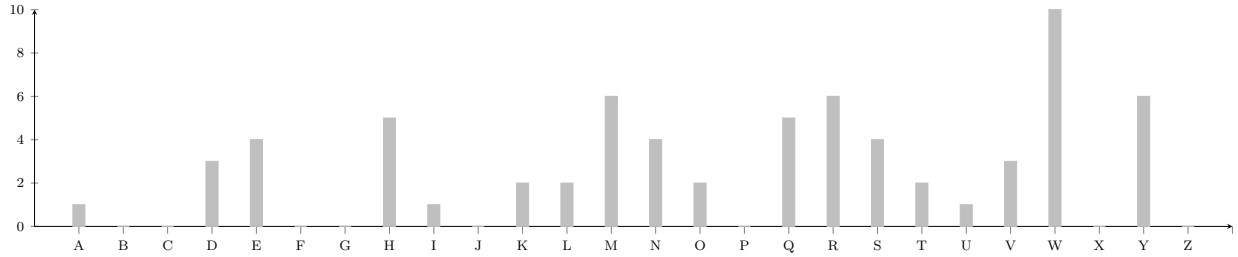


6 consecutive low frequency letters `F,G,H,I,J,K` - they might correspond to plaintext sequence `U,V,W,X,Y,Z`. Following this series of low frequency letters, we can spot 3 high frequency letters, all 3 letters apart, at positions `L,P,T` (might be plaintexts `A,E,I`), followed by 4 low frequency letters, followed by a high frequency pair `Y,Z` which might correspond to plaintexts `N,O`. The high frequency triplet of ciphertexts `C,D,E` should correspond to `R,S,T`. All this shows us that indeed, the pattern of the plain language can be seen. It is shifted, but it is not stretched. Therefore, the cipher did not use multiplication, and this ciphertext must be produced by a shift cipher.

To break the shift cipher, we need to map a ciphertext to a plaintext. One of the guesses we could do is that the ciphertext `L` standing alone in the ciphertext sequence $TDLETOP$, must be plaintext `A`. This guess is supported by the fact that ciphertext `P` is has the highest frequency among all other ciphertexts, and hence it must be plaintext `E`. If `L` $\mapsto$ `A`, then the encryption key must be 11, and hence the decryption key is 15. Let us try to decipher the message with key 15. The result is `THERE IS A TIDE IN THE AFFAIRS OF MEN WHICH TAKEN AT THE FLOOD LEADS ON TO FORTUNE`. The decrypted text is intelligible, which is a good sign that it is the plaintext that was encrypted.

2. `SV SQ VNY OWMMWR KWRTYL WD EHH MYR NWK EMWRI QW MERU MSHHSWRQ WD DEOYQ VNYLY QNWAHT`

The IC of the cryptogram is 0.0624 which is a good indication that this cryptogram was produced by a monoalphabetic cipher. Let us examine the frequency distribution of ciphertexts. Here we



cannot easily spot the pattern of the plaintext English language, and hence it is reasonable to assume that this cryptogram was produced by an affine cipher. We can spot two digraphs in the beginning of the cryptogram `SV SQ`, both starting with the same letter. These are the two digraphs

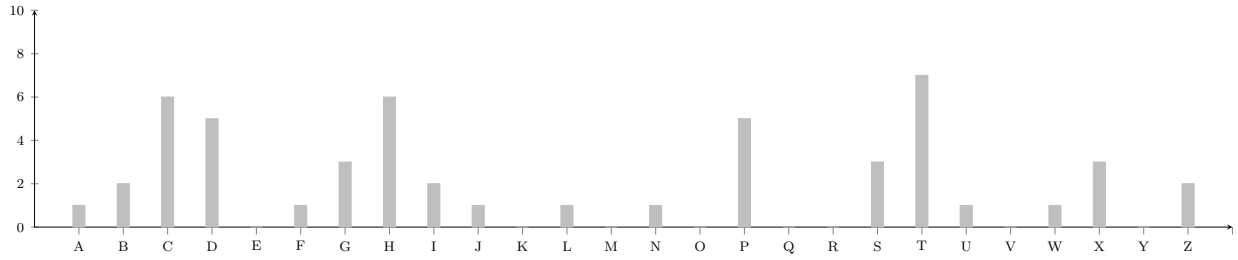
common to be placed in the beginning of the sentence. Most probably, SV SQ corresponds to IT IS. The word EHH is quite short and has a specific pattern, most probably it means ALL. If this guess is correct, then digraph WD EHH could be one of AT ALL, OF ALL. If EHH corresponds to ALL then the decryption key (a, b) should map E $\mapsto$ A and H $\mapsto$ L. We construct a system of congruences

$$\begin{aligned} 4a + b &\equiv 0 \pmod{26} \ , \\ 7a + b &\equiv 11 \pmod{26} \ . \end{aligned}$$

If we subtract the first equation from the second, we get $3a \equiv 11 \pmod{26}$. Since $3^{-1} = 9 \pmod{26}$, then multiplying both sides of the equation by 9, we get $a = 21$, and $b = 20$. Decrypting the cryptogram with key $(21, 20)$, we obtain the message IT IS THE COMMON WONDER OF ALL MEN HOW AMONG SO MANY MILLIONS OF FACES THERE SHOULD.

3. BDSTGC WXHIDGN AXZT P STPU BPC PCHLTGH FJTHIXDCH CD DCT PHZTS

The IC of the cryptogram is 0.065 which gives us a good indication of the monoalphabeticity. The frequency distribution is We clearly see 6 consecutive low frequency letters J,K,L,M,N,O,

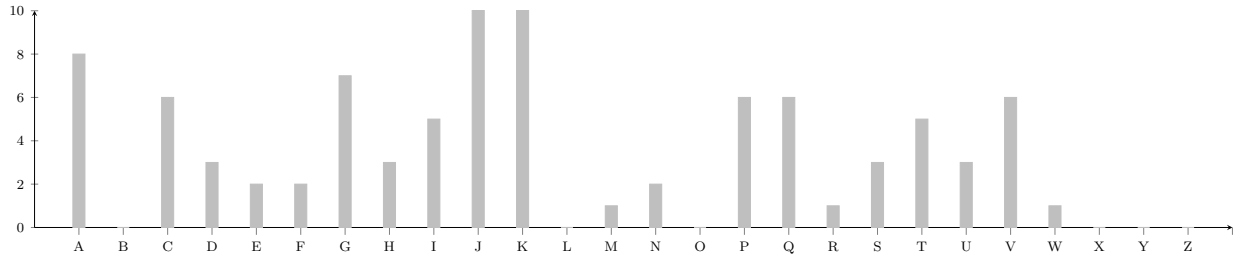


followed by 3 high frequency letters evenly separated 3 letters apart P,T,X. Then J,K,L,M,N,O must be U,V,W,X,Y,Z, and P,T,X must be A,E,I. This frequency distribution gives us a clue that multiplication was not used in this process, hence it must be the shift cipher. Since T is the most frequent letter in the cryptogram, it is reasonable to assume that T corresponds to plaintext E. If this is the case, then the decryption key would be the one satisfying equation $19 + z \equiv 4 \pmod{26}$, and hence z must be 11. Decrypting with key 11 we obtain the text MODERN HISTORY LIKE A DEAF MAN ANSWERS QUESTIONS NOONE ASKED.

4. JDI HVANGNKFKJS JDGJ EI MGS PGKF KT G EAVJDS UGQCI KC TAJ CQPPKUKITJ RQCJKPKUGJKAT PAV AQV VIPQCKTW JA CQHHAVJ KJ

The IC of the cryptogram is 0.631 which is a good indication of monoalphabeticity. The frequency distribution is We can clearly see that it must be the cryptogram obtained by means of the affine cipher. We can spot that the cryptogram begins with a trigraph JDI, which might correspond to THE, which is the most common trigraph in the English language, and the place in the beginning of the sentence is a proper position for this trigraph. If we substitute T for J, H for D, and E for I, then we get

THE HVANGNKFKTS THGT EE MGS PGKF KT G EAVTHS UGQCE KC TAT CQPPKUKETT
RQCTKPKUGTKAT PAV AQV VEPQCKTW JA CQHHAVT KT



In this text we can spot the sequence **TH**G**T** which must be THAT, hence G corresponds to A, and we have

THE HVAN**A**NKFK**T**S **THAT** **E**E MA**S** PA**K**F KT **A** EAV**TH**S U**A**QC**E** KC TA**T** CQPPKUK**E**T**T**
RQC**T**KPKU**AT**KAT PAV AQV V**E**PQCKTW JA CQHHAV**T** K**T**

Now we make a reasonable guess that the last digraph K**T** is AT, hence K corresponds to A. But this contradicts with our previous guess that G corresponds to A. Therefore, the digraph K**T** must be IT, which means that K corresponds to I.

THE HVAN**A**N**I**F**IT**S **THAT** **E**E MA**S** PA**I**F **I**T **A** EAV**TH**S U**A**QC**E** **I**C TA**T** CQPP**I**U**I**E**T**T
RQC**TI**P**I**U**AT**I**A**T PAV AQV V**E**PQC**I**TW JA CQHHAV**T** **IT**

The sequence **I**T **A** is likely to be IN A, it is reasonable to assume that T corresponds to N.

THE HVAN**A**N**I**F**IT**S **THAT** **E**E MA**S** PA**I**F **IN A** EAV**TH**S U**A**QC**E** **I**C **N**A**T** CQPP**I**U**I**E**NT**
RQC**TI**P**I**U**AT**I**AN** PAV AQV V**E**PQC**IN**W JA CQHHAV**T** **IT**

The sequence **I**C **N**A**T** could correspond to IS NOT. Hence C corresponds to S and A corresponds to O.

THE HV**O**N**A**N**I**F**IT**S **THAT** **E**E MA**S** PA**I**F **IN A** EAV**TH**S U**AQSE IS NOT** **S**QPP**I**U**I**E**NT**
RQ**ST**IP**I**U**AT**I**ON** P**O**V **O**QV V**E**PQ**S**IN**W** JA **S**QHH**O**V**T** **IT**

The sequence P**O**V **O**QV might mean FOR OUR. Hence, P corresponds to F, V corresponds to R, Q corresponds to U.

THE H**R**ON**A**N**I**F**IT**S **THAT** **E**E MA**S** **FAI**F **IN A** EA**R**TH**S** U**AQSE IS NOT SUFFI**U**IENT**
RU**ST**IFI**U**A**T**I**ON FOR OUR REFUSIN**W** JA **S**U**H**H**O**R**T** **IT**

The sequence **SUFFI**U**IENT** is SUFFICIENT, and hence U corresponds to C. The sequence **REFUSIN**W might be REFUSING, and hence W corresponds to G. The sequence **SU**HH**ORT** is SUPPORT, and hence H corresponds to P.

THE PRON**A**N**I**F**IT**S **THAT** **E**E MA**S** **FAI**F **IN A** EA**R**TH**S** **CAQSE IS NOT SUFFICIENT**
RU**ST**IFI**CAT**I**ON FOR OUR REFUSING** JA **SUPPORT IT**

The word R**USTIFICATION** suggests JUSTIFICATION, and hence R corresponds to J. JA also must be TO, and **FAI**F must be FAIL.

THE PRON**A**N**I**F**IT**S **THAT** **E**E MA**S** **FAIL IN A** E**O**R**TH**S **CAQSE IS NOT SUFFICIENT**
JUSTIFICATION FOR OUR REFUSING TO SUPPORT IT

EE MAS corresponds to one of HE MAY or WE MAY. Since we have already mapped H, it is more likely to be WE MAY.

THE PRONANIFITY THAT WE MAY FAIL IN A EORTHY CAQSE IS NOT SUFFICIENT JUSTIFICATION FOR OUR REFUSING TO SUPPORT IT

EORTHY CAQSE is a good candidate for WORTHY CAUSE, and PRONANIFITY is likely to be PROBABILITY. Hence, the plaintext is

THE PROBABILITY THAT WE MAY FAIL IN A WORTHY CAUSE IS NOT SUFFICIENT JUSTIFICATION FOR OUR REFUSING TO SUPPORT IT

Alternatively, we could consider that the decryption key (a, b) maps JDI to THE. We could construct a system of congruences

$$\begin{aligned} 9a + b &\equiv 19 \pmod{26} \ , \\ 3a + b &\equiv 7 \pmod{26} \ , \\ 8a + b &\equiv 4 \pmod{26} \ . \end{aligned}$$

If we subtract the third equation from the first one, we get $a \equiv 15 \pmod{26}$. Plugging this value into equation 2, we get $19+b \equiv 7 \pmod{26}$ and from there we obtain the value of $b = 14$. Decrypting the message using key $(15, 14)$ we get the plaintext THE PROBABILITY THAT WE MAY FAIL IN A WORTHY CAUSE IS NOT SUFFICIENT JUSTIFICATION FOR OUR REFUSING TO SUPPORT IT