1. You are given the following ciphertext:

```
GKYBEEYDYQVONNRBSNFPLBVRBCOGRSBNKIOEQCYVYEYBNYQUVILZBREBSQFEBJYQI
SKIONYBVVYNGKYVBGGIVSYPNBVDOYQRSBJIOVGURERSDUVRQBPGKBGNYYTNGIFBRS
GKYVBNBKOLCEYNGOQYSGAKINYRSGYVBJGRISNARGKBLYVRJBSNBSQVONNRBSNBVYI
WYVCEIASCPFVINYJOGIVNAKIRSQRJGYQKYVUIVBJGRSDREEYDBEEPBNBUIVYRDSBD
YSGNFYJRURJBEEPCOGRSBQYUYSNYGYBLQYNJVRCYNKYVVYEBGRISNKRFGIVONNRBS
CREERISBRVYBEYHBSQYVGIVNKRSAKIBEEYDYQEPLYSGIVYQKYVRSBGGYLFGNGIRSU
REGVBGYVYFOCERJBSFIERGRJNBNBJGOBEEPZONGBUVRYSQKYVEBAPYVNBENIRSJEO
QYQBGYHGLYNNBDYJKBRSRSAKRJKBLBVVRYQDOSVRDKGNBQWIJBGYAKILCOGRSBBEE
YDYQEPIUUYVYQGINEYYFARGKRSBSBGGYLFGGIDBRSFIERGRJBERSUEOYSJYNBRQKY
AIOEQSYWYVNEYYFARGKBVYQKYBQERTYCOGRSBCOGRSBKBNCYYSRSZBREURVNGRSAB
NKRSDGISSIARSBEYHBSQVRBWRVDRSRBNRSJYZOEPURUGYYSAKYSNKYFEYBQYQSIGD
OREGPGIGKYUIVYRDSBDYSGJKBVDYNRSQJUYQYVBEJIOVGCOGRSBNKIONYBVVYNGVY
MOYNGAREECYQRNJONNYQBGBJIOVGKYBVRSDISNYFGYSGKYZOQDYNBRQGKYNRCYVRB
SCIVSUIVLYVBLYVRJBSOSRWYVNRGPDVBQOBGYNGOQYSGBEEYDYQEPCOREGVYFOCER
JBSFBVGPJISSYJGRISNGKVIODKDOSVRDKGNDVIOFNRSJEOQRSDGKYSVBBSQGKVIOD
KGKYSBGRISBEFVBPYVCVYBTUBNGBNBABPGIFONKFVIVONNRBSRSGYVYNGNFVINYJO GIVNJKBVDY
```

(a) Assume that you learn that this is a monoalphabetic cipher of an English text.
First determine the relative frequency of the letters in this ciphertext (construct a frequency table) and compare them to the frequency of normal English.

```
plain:   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
cipher: ?2       ?2   ?1    ?3  ?4              ?3           ?5     ?3
```

What ciphertext letter do you think corresponds to "e" in the plaintext? Examining the ciphertext and considering the most popular bigrams in the English language: TH, HE, AN, IN, ER, ON, RE, ED, ND, HA, AT, EN (in this order), and the most popular trigrams: THE, AND, THA, ENT, ION, TIO, FOR, NDE, HAS, NCE, TIS, OFT, MEN (construct frequency tables—details at the end) try to determine what ciphertext letters correspond to "t" and "h" in the plaintext (you must take into account the fact that the most popular letter in the plaintext that corresponds to "e" must also be in the most popular bigrams and trigrams). Now construct the statistical frequency table of letters: top row is the plaintext letters of the English alphabet, bottom row are the corresponding ciphertext letters. *What is the plaintext of the given ciphertext*?

(b) Assume now that the monoalphabetic cipher you have determined was generated using transposition as follows: first choose a keyword, (for example "CIPHER") and write the keyword as the first row. Then write the remaining letters in the following rows. The monoalphabetic cipher to use is then read-off by the columns left to right. For example, using the keyword "CIPHER" we get:

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

```
  plain: 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 cipher:
  C A K S Y I B L T Z P D M U H F N V E G O W R J Q X Determine the 6-
```
letter *keyword* (the 6 first words of the plaintext will give you a clue). Show your work.

Answers:

For problem **1 a** I wrote code to solve my problem:

```cpp
// Decrypting the cipher


#include <iostream>
#include <vector>
#include <map>
#include <cmath>

using namespace std;

int main()
{
  string input;
  input =
"GKYBEEYDYQVONNRBSNFPLBVRBCOGRSBNKIOEQCYVYEYBNYQUVILZBREBSQFEBJYQISKIONYBVVYNGKYVBGGIVSYPN
BVDOYQRSBJIOVGURERSDUVRQBPGKBGNYYTNGIFBRSGKYVBNBKOLCEYNGOQYSGAKINYRSGYVBJGRISNARGKBLYVRJB
SNBSQVONNRBSNBVYIWYVCEIASCPFVINYJOGIVNAKIRSQRJGYQKYVUIVBJGRSDREEYDBEEPBNBUIVYRDSBDYSGNFYJRU
RJBEEPCOGRSBQYUYSNYGYBLQYNJVRCYNKYVVYEBGRISNKRFGIVONNRBSCREERISBRVYBEYHBSQYVGIVNKRSAKIBEEY
DYQEPLYSGIVYQKYVRSBGGYLFGNGIRSUREGVBGYVYFOCERJBSFIERGRJNBNBJGOBEEPZONGBUVRYSQKYVEBAPYVNBE
NIRSJEOQYQBGYHGLYNNBDYJKBRSRSAKRJKBLBVVRYQDOSVRDKGNBQWIJBGYAKILCOGRSBBEEYDYQEPIUUYVYQGINE
YYFARGKRSBSBGGYLFGGIDBRSFIERGRJBERSUEOYSJYNBRQKYAIOEQSYWYVNEYYFARGKBVYQKYBQERTYCOGRSBCOG
RSBKBNCYYSRSZBREURVNGRSABNKRSDGISSIARSBEYHBSQVRBWRVDRSRBNRSJYZOEPURUGYYSAKYSNKYFEYBQYQSI
GDOREGPGIGKYUIVYRDSBDYSGJKBVDYNRSQJUYQYVBEJIOVGCOGRSBNKIONYBVVYNGVYMOYNGAREECYQRNJONNYQB
GBJIOVGKYBVRSDISNYFGYSGKYZOQDYNBRQGKYNRCYVRBSCIVSUIVLYVBLYVRJBSOSRWYVNRGPDVBQOBGYNGOQYSG
BEEYDYQEPCOREGVYFOCERJBSFBVGPJISSYJGRISNGKVIODKDOSVRDKGNDVIOFNRSJEOQRSDGKYSVBBSQGKVIODKGK
YSBGRISBEFVBPYVCVYBTUBNGBNBABPGIFONKFVIVONNRBSRSGYVYNGNFVINYJOGIVNJKBVDY";
  map<char, double> letter_map;
  cout << "The input size is: " << input.length() << endl;

// mapping the text and adding the number of occurences
  cout << "Cipher    # of times apeared        percent appeared" << endl;
  for (char c : input)
  {
    letter_map[c]++;
```

```cpp
    }
    letter_map['X'];

    // cout the letters and the percent that it appears
    double percent;
    for (auto c : letter_map)
    {
      percent = round((c.second / input.size()) * 100.0 * 100.0) / 100.0;

      cout << c.first << "        " << c.second << "                " << percent << "%" << endl;
    }

    cout << "\nFrom examining the occurrence percentages of the letters, I conclude that: \n"
        << "1. 'Y' = 'E' because 'E' is the most common letter in the English language.\n"
        << "2. Bigrams: There are common pairs: 'YV', 'YN', 'YQ' that appear frequently, which can be something like 'er', 'en', or 'ed'.\n"
        << "3. Bigrams: When looking for 'TH' patterns, the most frequent was 'GK', so 'G' must = 'T' and 'K' must = 'H'.\n"
        << "4. Trigrams: For the trigrams: 'GKY' can be replaced with 'that'."
        << endl;

    vector<char> rel_letter = {
    // Plaintext
      'w', 'a', 'b', 'g', 'l', 'p', 't', 'x', 'o', 'c', 'h', 'm', 'q', 's', 'u', 'y', 'd', 'i', 'n', 'k', 'f', 'r', 'v', 'z', 'e', 'j'
    };
    // Cipher 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

    // creates the matching letters map
    int i =0;
    map<char, char> matching_letters;
    for (auto c : letter_map)
    {
      matching_letters[c.first] = rel_letter[i];
      i++;
    }

    cout << "\nMapping the letters:" << endl;
    string first[matching_letters.size()];
    string second[matching_letters.size()];
    for (const auto &c : matching_letters)
    {
```

3

```
    cout << c.first << "  ";
  }
  cout << endl;

  for (const auto &c : matching_letters)
  {
    cout << c.second << "  ";
  }
  cout << endl << endl;

  string decoded;
  for (char c : input) {
    decoded += matching_letters[c];
  }

  cout << decoded;
  return 0;
}
```

Here is the output:

The input size is: 1050

Cipher	# of times apeared	percent appeared
A	16	1.52%
B	102	9.71%
C	20	1.9%
D	30	2.86%
E	55	5.24%
F	22	2.1%
G	84	8%
H	3	0.29%
I	53	5.05%
J	32	3.05%
K	45	4.29%
L	13	1.24%
M	1	0.1%

N	71	6.76%
O	44	4.19%
P	17	1.62%
Q	41	3.9%
R	89	8.48%
S	82	7.81%
T	3	0.29%
U	19	1.81%
V	77	7.33%
W	5	0.48%
X	0	0%
Y	121	11.52%
Z	5	0.48%

From examining the occurrence percentages of the letters, I conclude that:

1. 'Y' = 'E' because 'E' is the most common letter in the English language.

2. Bigrams: There are common pairs: 'YV', 'YN', 'YQ' that appear frequently, which can be something like 'er', 'en', or 'ed'.

3. Bigrams: When looking for 'TH' patterns, the most frequent was 'GK', so 'G' must = 'T' and 'K' must = 'H'.

4. Trigrams: For the trigrams: 'GKY' can be replaced with 'that'.

Mapping the 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

w a b g l p t x o c h m q s u y d i n k f r v z e j

theallegedrussianspymariabutinashouldbereleasedfromjailandplacedonhousearresttheattorneysarguedinacourtfilingfrid
aythatseekstopaintherasahumblestudentwhoseinteractionswithamericansandrussiansareoverblownbyprosecutorswhoin
dictedherforactingillegallyasaforeignagentspecificallybutinadefenseteamdescribesherrelationshiptorussianbillionairealex
andertorshinwhoallegedlymentoredherinattemptstoinfiltraterepublicanpoliticsasactuallyjustafriendherlawyersalsoinclud
edatextmessagechaininwhichamarriedgunrightsadvocatewhombutinaallegedlyofferedtosleepwithinanattempttogainpolit
icalinfluencesaidhewouldneversleepwitharedheadlikebutinabutinahasbeeninjailfirstinwashingtonnowinalexandriavirginia
sincejulyfifteenwhenshepleadednotguiltytotheforeignagentchargesindcfederalcourtbutinashousearrestrequestwillbedisc
ussedatacourthearingonseptenthejudgesaidthesiberianbornformeramericanuniversitygraduatestudentallegedlybuiltrepu

blicanpartyconnectionsthroughgunrightsgroupsincludingthenraandthroughthenationalprayerbreakfastasawaytopushpror
ussianinterestsprosecutorscharge

1b)

Reverse engineering the cipher so that the receiver can send it the way that the sender sent it

What makes a keyword a keyword?

  - it has to be recognizable

  - the characters have to be unique

The answer is: BUTINA

So i remove the mappings of the characters from:

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

to:

c d e f g h j k l m o p q r s v w x y z

making the grid now:

B U T I N A

c d e f g h

j k l m o p

q r s v w x

y z

making the columns from it we get:

column 1: b c j q y

column 2: u d k r z

column 3: t e l s

column 4: i f m v

column 5: n g o w

column 6: a h p x

which now gives us the final mapping:

b c j q y u d k r z t e l s i f m v n g o w a h p x