CEMC.UWATERLOO.CA | The CENTRE for EDUCATION in MATHEMATICS and COMPUTING

CEMC Math Circles - Grade 9/10

Wednesday, March 24, 2021
Surprise Party - Solution

Solutions to Activity 1:

(a) The encrypted message is VGXZE YZGXZY GZ YKBKT.

(b) The decrypted message is THE THEME IS FANTASY.

(c) One way to decode a message that is encrypted using a Caesar Cipher, when the key is unknown, is to try all possible keys until one key produces a message that makes sense. There are only 25 possible keys, so this wouldn't take too long.

A more clever way is to take advantage of letter frequencies in the English language. The most common letter in the English language is E. The most common letter in the encrypted message is P. This means that a good guess might be that the letter E has been shifted to the letter P. This would make the key equal to 11. Using a key of 11, the decrypted message is:

EUGENE WILL BRING THE CAKE AND DECORATIONS

Note: *An attempt to break a substitution cipher by using knowledge of commonly used letters or phrases in a language, as we did above, is an example of what is called frequency analysis. For frequency analysis to be as reliable as possible, we want to study as much text, encrypted using the same cipher, as we can. If we have only a short message to work with, then it is very possible that the letter* E *will not be the most frequently occurring letter in the original message (and we will be tricked). If we have a very long message, or a very large quantity of messages, chances are good that within a few tries we will have found the right match for the letter* E.

Solutions to Activity 2:

(a) The encrypted message is RHFVMAAKX APKV.

(b) The decrypted message is ALICE WILL BRING SOME BOARDGAMES.

(c) Decrypting a message that is encoded using the Vigenère Cipher, when the key is unknown, is much more difficult to do by hand, compared to the Ceaser cipher. However, it is possible! Note that there are several ways to decrypt such a message. We discuss one approach that can be taken.

We start by finding the key. Since the key is 4 letters in length, we can break the problem up into 4 different Ceaser cipher decryption problems. To do this, we sort letters of the encrypted message into 4 groups using the numbers 1 through 4 as follows:

```
W D R X W B G  C A T Y  H J Q A M  U I A J D M F  L G Z  G N W  K N Q W
1 2 3 4 1 2 3  4 1 2 3  4 1 2 3 4  1 2 3 4 1 2 3  4 1 2  3 4 1  2 3 4 1
```

You can think of the numbers 1 through 4 as place holders for the letters in the key! Next, we group together all of the letters that have a 1 directly below: WWAJUDGWW. Doing the same for the other numbers, we get the following 4 groups of letters:

WWAJUDGWW DBTQIMZK RGYAAFGN XCHMJLNQ.

The letters in a particular group have all been encrypted using the same row, or equivalently, the same Ceaser cipher. For example, the letters in the first group have been encrypted using the first letter of the key. This is because the letters in this group all have a 1 directly below them. Likewise, the second group of letters have been encrypted using the second letter of the key. And so on. So, if we decrypt one letter from a particular group, then we know how to decrypt all of the other letters in that same group.

Let's start by using the fact that the last letter of the key is G. This fact tells us that the letters in the fourth group have been decrypted using the letter G. Using row G of the table, we decrypt the letters in this group to be RWBGDFHK. And so:

Encrypted message: WDRXWBG CATY HJQAM UIAJDMF LGZ GNW KNQW
Decrypted message: ???R??? W??? B???G ???D??? F?? ?H? ??K?

We might be able to guess what some of the shorter words are now. However, lets first try to find at least one missing letter of the key. One way to find the key, or parts of it, is to take advantage of the letter frequencies in the English language, as we did in activity 1 part (c). We use the fact that the most common letter in the English letter is E. The most common letter in the first group of letters is W. It is reasonable to guess that the letter E has been encrypted with the letter W. Let's see where this guess takes us! Going down column E of the table, we see that W belongs to row S. If this guess is correct, then the first letter of the key is an S, and so the first group of letters is encrypted using the letter S. We decrypt the first group of letters to be: EEIRCLOEE. So far, we have:

Encrypted message: WDRXWBG CATY HJQAM UIAJDMF LGZ GNW KNQW
Decrypted message: E??RE?? WI?? BR??G C??DL?? FO? ?HE ??KE

The second and fourth group of letters don't contain any multiples. The letters A and G both appear twice in the third group. We could make a guess using frequencies in the English language what these letters may be, but it might take a few tries to get the correct guess. Alternatively, since we already have several letters decrypted, we can guess what some of the shorter words may be. It is reasonable to guess that the fifth words is FOR. If this guess is correct, then R has been encrypted with Z. Looking at row R we see that Z is in column I. Since Z belongs to the second group of letters above, we know that every letter in this group has been encrypted using the letter I and that the second letter of the key is I. Using row I of the table, we decrypt the second group of letters to be: VTLIAERC. We then have:

Encrypted message: WDRXWBG CATY HJQAM UIAJDMF LGZ GNW KNQW
Decrypted message: EV?RET? WIL? BRI?G CA?DLE? FOR ?HE C?KE

Similarly, we guess that the second word in the encrypted message is WILL. If this guess is correct, then the third letter of the key is N. It follows that the key us SING. We find the decryption to be:

Encrypted message: WDRXWBG CATY HJQAM UIAJDMF LGZ GNW KNQW
Decrypted message: EVERETT WILL BRING CANDLES FOR THE CAKE