BASIC CRYPTOGRAPHY

MATH CIRCLE (HS1) 4/26/2015

Warmups

Try to decode the following messages.

1) TSEISAEEHTYLBABORPSISIHT.

2) TIMSAESOTOADIHRNISMLROHFUT
HSESGINTOHRETEADSIIATTEORH.

3) JU JT BT FBTZ BT BCD UP CSFBL UIJT DPEF. USZ ZPVS MVDL XJUI UIF GJGUI!

4) DYUGETATIOESHHRETOOAREHTHSNITEADS.

5) VA GUVF PBAGRKG, GURER'F AB QVFERFCRPG,
FB, JURA V OHFG ZL EULZR, LBH OERNX LBHE ARPXF.
JR TBG SVIR ZVAHGRF SBE HF GB QVFPBAARPG,
SEBZ NYY VAGRYYRPG PBYYRPG GUR EULGUZ RSSRPG.
FB YBFR NA VAUVOVGVBA, SBYYBJ LBHE VAGHVGVBA,
SERR LBHE VAARE FBHY NAQ OERNX NJNL SEBZ GENQVGVBA.

This handout is based on previous math circle handouts by Shannon Rogers and Professor Christian Haesemeyer.

Caesar Cipher

The method to code the messages in 3) and 5) was similar. In 3), every letter was shifted 1 forward (A$\rightarrow$B,B$\rightarrow$C, etc.) called a *Caesar Shift of length 1*. In 5), every letter is shifted forward by 13. (A$\rightarrow$N,B$\rightarrow$O,etc.) called a *Caesar Shift of length 13*.

1) Find a partner and exchange coded messages using a Caesar Cipher. What information do you need to share between each other to (simply) decode each others' messages?

2) Encode the message "Hello World" with Caesar Shifts of length $1, 2, 3, 4$. Hint: Try to organize your work!

3) Decode the message "WHVWLQJRQHWZRWKUHH". Hint: It is a Caesar Shift of length less than 6.

4) Decode the message "RFPCCESWQUYJIGLRMYZYP". Hint: It is a Caesar Shift of length more than 20.

5) Explain why Caesar Ciphers are not very secure. Use some math in your answer!

Monoalphabetic Cipher

A Caesar Cipher is just one of many *monoalphabetic ciphers*, where we code a message using some permutation/rearrangement of the alphabet. Caesar ciphers are some examples of monoalphabetic ciphers. Another example is the *reverse monoalphabetic cipher*, where A$\rightarrow$Z,B$\rightarrow$Y,etc.

1) a) Encode the joke "Why is a math book always unhappy?" using the reverse monoalphabetic cipher.

b) Decode the answer to the joke (using the same cipher) "YVXZFHVRGZODZBHSZHOLGHLUKILYOVNH".

2) Find a partner and exchange coded messages using your own monoalphabetic cipher. What information do you need to share between each other to (simply) decode each others' messages?

3) How many different Monoalphabetic Ciphers are there?

While the answer to 3) may make it seem like monoalphabetic ciphers are somewhat secure, one way to help break them is using letter frequencies. The idea is that (especially in longer messages) the most frequent letters in the coded message will correspond to the most frequent letters in usual English text.

The relative frequencies of letters in English are:

e 12.702%, t 9.056%, a 8.167%, o 7.507%, i 6.966%, n 6.749%, s 6.327%, h 6.094%, r 5.987%, d 4.253%, l 4.025%, c 2.782%, u 2.758%, m 2.406%, w 2.360%, f 2.228%, g 2.015%, y 1.974%, p 1.929%, b 1.492%, v 0.978%, k 0.772%, j 0.153%, x 0.150%, q 0.095%, z 0.074%.

4) Use relative frequencies to help decode the following message:

V FXZY KDY EWMN. EWMN VZ YGW FVKA PVOOWN. EWMN VZ YGW OVYYOW AWMYG YGMY HNVKUZ YDYMO DHOVYWNMYVDK. V SVOO EMBW FQ EWMN. V SVOO CWNFVY VY YD CMZZ DIWN FW MKA YGNDXUG FW. MKA SGWK VY GMZ UDKW CMZY V SVOO YXNK YGW VKKWN WQW YD ZWW VYZ CMYG. SGWNW YGW EWMN GMZ UDKW YGWNW SVOO HW KDYGVKU. DKOQ V SVOO NWFMVK.

Hint: I've also preserved words, so use that to your advantage!

Vigenere Cipher

This way of coding that is resistant to frequency analysis. Here a message is encoded using multiple Caesar shifts with the help of a key word, associating each letter to a Caesar shift of a different length: A$\rightarrow 0$,B$\rightarrow 1$,etc.

For example, suppose we want to encode the message "circle" using the key word "math". C is shifted by 12 (m), resulting in O, I is shifted by 0 (a) resulting in I, R is shifted by 19 (t) resulting in K, C is now shifted by 7 (h) resulting in J, L is shifted by 12 (m again) resulting in X, and finally E is shifted by 0 (a again) resulting in E. In summary, "CIRCLE" is coded as "OIKJXE".

1) a) Encode the joke "Why did the chicken cross the mobius strip?" using a Vigenere cipher and key word "ABCD".

b) Decode the answer to the joke "TPIHTUQWHFUDMFULDF".

2) Find a partner and exchange coded messages using your own Vigenere cipher. What information do you need to share between each other to (simply) decode each others' messages?

Recall we've discussed earlier how Caesar Cipher's are relatively unsecure. Thus, the main security comes from the unknown length of a key word (or a really long keyword). There are, however, methods to help guess the key length.

3) (Kasiski Examination)

a) Encode the message "Crypto is short for cryptography" using the key word "ABCD".

b) What do you notice about how "CRYPTO" was encoded (both times) in this message? What is the relationship between the length of the key word (length is 4) and how far apart the two instances of "CRYPTO" appear in the message (they are 16 apart)?

c) Using b) as an example, come up with a general way to (possibly) narrow down key word lengths.

4) Decode the following message, knowing that the possible key words are fruit:

HMDIW MMZTV SXZOQ XSMDF FBEXT TXMEA RECMI IWTMR OUMCM MIIWT OSXMN ENSFS PIDXS ETSNI ZXPTX MRCQW DBTIC MUWYW ARPBA FWIYI HMDIW MMZTV SKCWG ROEUP WBTMY OEIGM DFPBT IDIYI LOMMY

5) Discuss ways to make the Vigenere Cipher relatively secure. What are the positives and negatives of these approaches?