Assignment #2:
Classical Ciphers

Sep. 16, 2014

Deadline: Oct. 6, 5:00 PM.
Total: 100

All written assignments should be created using a word processor (e.g., Word or Latex). Handwritten work will not be accepted.

All students should submit a *pdf file* of their answers in the Scholar's *Assignments* function. Be sure to name the file as follows: `HW2_LastName_FirstInitial_PID.pdf`. For example, if a student's name is John Doe and his PID is "doe3," then he would name the file as `HW2_Doe_J_doe3.pdf`.

Also, submit the source code (including any header files that you may have used) to Scholar. Name the file as: `HW2code_lastName_firstInitial_PID.fileExtension` (e.g., `HW2code_Smith_J_smj.cpp`).

Zip up all files into ***one zip file*** and submit it using Scholar's Assignment function.

1. The objective of this problem is to learn about monoalphabetic substitution ciphers and to use letter frequencies to cryptanalyze a given ciphertext.

Write a program (in any programming language) to analyze the letter frequencies in a block of text. Ignore case (upper and lower count the same). It should count the number of times each letter appears in the text, the number of times each pair of letters appears, and the number of times each sequence of 3 letters appears. Your program should sort the counts of single letters, pairs of letters (bigrams), and triples of letters (trigrams), then print the non-zero values in decreasing order. For single letters, print all non-zero values; for bigrams and trigrams, print the first 40 non-zero values.

Turn in the following items by the deadline.

- (a) (50 pts) As instructed above, submit a soft copy of your program via Scholar. *Write your own program! Do not copy other peoples' programs.* Violations will be reported to the VT Honor System.
- (b) (25 pts) Output of your program run against at least 2 different texts of *sufficient length*. Make sure that the texts are sufficiently long enough for frequency analysis.
- (c) (5 pts) Compare the sets of frequencies you produced for part (b). Are they similar or different? Explain why they are similar or different.
- (d) (20 pts) Using your frequency analysis results, decode the ciphertext given below.

For part (b), you can run your program on texts found on the Internet. For example, you can find the full text for Hamlet online (http://www.bibliomania.com/0/6/3/1057/frameset.html).

Ciphertext for (d):

bt jpx rmlx pcuv amlx icvjp ibtwxvr ci m lmt'r pmtn, mtn yvcjx cdxv mwmbtrj jpx amtngxrjbah uqct jpx qgmrjxv ci jpx ymgg ci jpx hbtw'r qmgmax; mtn jpx hbtw rmy jpx qmvj ci jpx pmtn jpmj yvcjx. jpxt jpx hbtw'r acutjxtmtax ymr apmtwxn, mtn pbr jpcuwpjr jvcufgxn pbl, rc jpmj jpx scbtjr ci pbr gcbtr yxvx gccrxn, mtn pbr htxxr rlcjx ctx mwmbtrj mtcjpxv. jpx hbtw avbxn mgcun jc fvbtw bt jpx mrjvcgcwxvr, jpx apmgnxmtr, mtn jpx rccjprmexvr. mtn jpx hbtw rqmhx, mtn rmbn jc jpx ybrx lxt ci fmfegct, ypcrcxdxv rpmgg vxmn jpbr yvbjbtw, mtn rpcy lx jpx btjxvqvxjmjbct jpxvxci, rpmgg fx agcjpxn ybjp ramvgxj, mtn pmdx m apmbt ci wcgn mfcuj pbr txah, mtn rpmgg fx jpx jpbvn vugxv bt jpx hbtwncl. jpxt amlx bt mgg jpx hbtw'r ybrx lxt; fuj jpxe acugn tcj vxmn jpx yvbjbtw, tcv lmhx htcyt jc jpx hbtw jpx btjxvqvxjmjbct jpxvxci. jpxt ymr hbtw fxgrpmoomv wvxmjge jvcufgxn, mtn pbr acutjxtmtax ymr apmtwxn bt pbl, mtn pbr gcvnr yxvx mrjctbrpxn. tcy jpx kuxxt, fe vxmrct ci jpx ycvnr ci jpx hbtw mtn pbr gcvnr, amlx btjc jpx fmtkuxj pcurx; mtn jpx kuxxt rqmhx mtn rmbn, c hbtw, gbdx icvxdxv; gxj tcj jpe jpcuwpjr jvcufgx jpxx, tcv gxj jpe acutjxtmtax fx apmtwxn; jpxvx br m lmt bt jpe hbtwncl, bt ypcl br jpx rqbvbj ci jpx pcge wcnr; mtn bt jpx nmer ci jpe ybrncl ci jpx wcnr, ymr icutn bt pbl; ypcl jpx hbtw txfuapmntxoomv jpe imjpxv, jpx hbtw, b rme, jpe imjpxv, lmnx lmrjxv ci jpx lmwbabmtr, mrjvcgcwxvr, apmgnxmtr, mtn rccjprmexvr; icvmrluap mr mt xzaxggxtj rqbvbj, mtn htcygxnwx, mtn utnxvrjmtnbtw, btjxvqvxjbtw ci nvxmlr, mtn rpcybtw ci pmvn rxtjxtaxr, mtn nbrrcgdbtw ci ncufjr, yxvx icutn bt jpx rmlx nmtbxg, ypcl jpx hbtw tmlxn fxgjxrpmoomv; tcy gxj nmtbxg fx amggxn, mtn px ybgg rpcy jpx btjxvqvxjmjbct.