CSCI361 Assignment 1

Asjad Athick
4970512
mama158

Part 1

Ctext-1

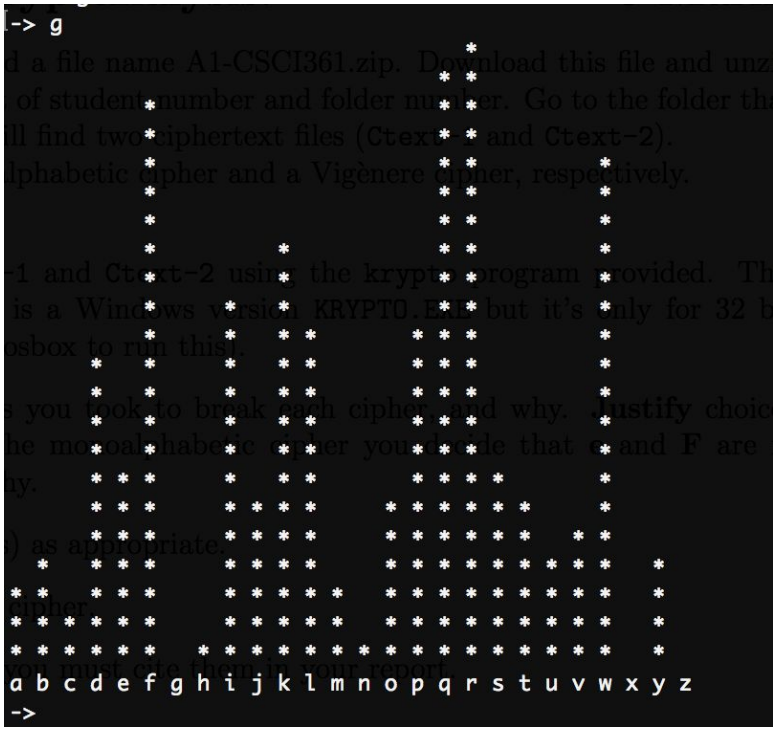
Cipher text

```
fk vdftd rwtd fkefufeswi jlurjrkq fp krsqowifzre ay pqwqfpqftwi
iwvp
jlirtsirp fk w cwp blo rxwjmir vfii molestr vloh eofur w mfpqlk
asq vdror jlqflk fp fkrbbrtqfur lk qdr pmlq wke drwere fk wii
efortqflkp wq qdr pwjr qfjr rkrocy vfii ar morprkq asq fkrbbrtqfur
lkr jfcdq pwy qdwq qdr psj lb wii qdr nswkqfqfrp lb drwq ilpq fk
qdr
tlsopr lb wii qdr wtqfufqfrp qdwq dwur qwhrk miwtr fk qdr skfuropr
fk lso dljrp wke lsq lb qdrj ay ewy wke ay kfcdq rnswiiy wq wii
qfjrp
wke fk wii miwtrp vdrktr vr hklv klq fq vwp fk lie ewyp vfqd lso
```

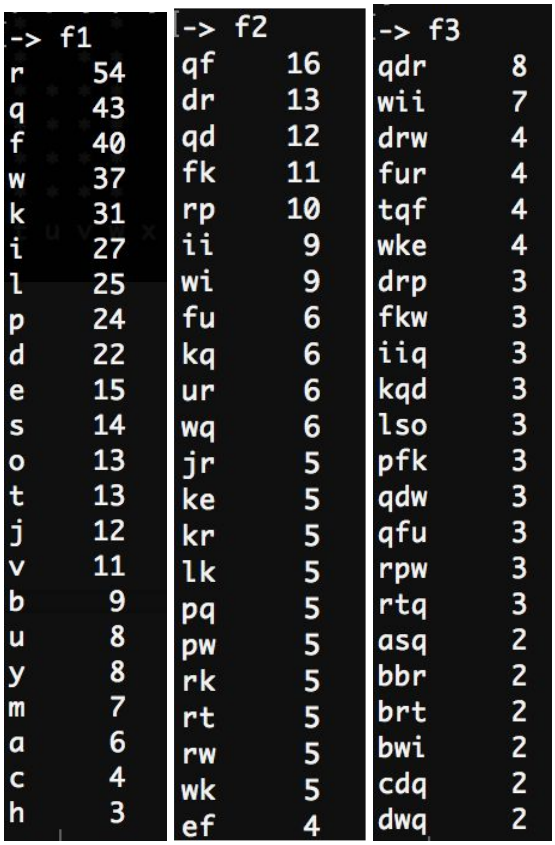
Index of Coincidence (calculated using Krypto)

IC = 0.065
Average = 0.065

Frequency Distribution Graph



Steps to break the cipher:



-> f1	
r	54
q	43
f	40
w	37
k	31
i	27
l	25
p	24
d	22
e	15
s	14
o	13
t	13
j	12
v	11
b	9
u	8
y	8
m	7
a	6
c	4
h	3

-> f2	
qf	16
dr	13
qd	12
fk	11
rp	10
ii	9
wi	9
fu	6
kq	6
ur	6
wq	6
jr	5
ke	5
kr	5
lk	5
pq	5
pw	5
rk	5
rt	5
rw	5
wk	5
ef	4

-> f3	
qdr	8
wii	7
drw	4
fur	4
tqf	4
wke	4
drp	3
fkw	3
iiq	3
kqd	3
lso	3
pfk	3
qdw	3
qfu	3
rpw	3
rtq	3
asq	2
bbr	2
brt	2
bwi	2
cdq	2
dwq	2

1. Apply substitutions for highest frequency words according to statistics of English

(A simple C++ program to substitute characters was written, source code in appendix A. Krypto substitution functionality had a glitch).

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

Substitution Logs

fk vhfqh rwqh fkefufeswi jlurjrkt fp krstowifzre ay ptwtfptfqwi iwvp jlirqsirp fk w cwp blo rxwjmir vfii molesqr vlod eofur w mfptlk ast vhror jltflk fp fkrbbrqtfur lk thr pmlt wke hrwere fk wii eforqtflkp wt thr pwjr tfjr rkrocy vfii ar morprkt ast fkrbbrqtfur lkr jfcht pwy thwt thr psj lb wii thr nswktftfrp lb hrwt ilpt fk thr qlsopr lb wii thr wqtfuftfrp thwt hwur twdrk miwqr fk thr skfuropr fk lso hljrp wke lst lb thrj ay ewy wke ay kfcht rnswiiy wt wii tfjrp wke fk wii miwqrp vhrkqr vr dklv klt ft vwp fk lie ewyp vfth lso
Replace: r
With: e

fk vdftd rwtd fkefufeswi jlurjrkq fp krsqowifzre ay pqwqfpqftwi iwvp jlirtsirp fk w cwp blo rxwjmir vfii molestr vloh eofur w mfpqlk asq vdror jlqflk fp fkrbbrtqfur lk qdr pmlq wke drwere fk wii efortqflkp wq qdr pwjr qfjr rkrocy vfii ar morprkq asq fkrbbrtqfur lkr jfcdq pwy qdwq qdr psj lb wii qdr nswkqfqfrp lb drwq ilpq fk qdr tlsopr lb wii qdr wtqfufqfrp qdwq dwur qwhrk miwtr fk qdr skfuropr fk lso dljrp wke lsq lb qdrj ay ewy wke ay kfcdq rnswiiy wq wii qfjrp wke fk wii miwtrp vdrktr vr hklv klq fq vwp fk lie ewyp vfqd lso
Replace: q
With: t

…
ik vzicz ewcz ikriuirswf jluejekt ip kestowfiher ay ptwtipticwf fwvp jlfecsfep ik w qwp blo exwjmfe viff molrsce vlod roiue w miptlk ast vzeoe jltilk ip ikebbectiue lk tze pmlt wkr zewrer ik wff rioectilkp wt tze pwje tije ekeoqy viff ae moepekt ast ikebbectiue lke jiqzt pwy tzwt tze psj lb wff tze nswktitiep lb zewt flpt ik tze clsope lb wff tze wctiuitiep tzwt zwue twdek mfwce ik tze skiueope ik lso zljep wkr lst lb tzej ay rwy wkr ay kiqzt enswffy wt wff tijep wkr ik wff mfwcep vzekce ve dklv klt it vwp ik lfr rwyp vitz lso
Replace: k
With: n
…

in vhich each inriuirsal jfuejent ip nestoalizer wy ptatiptical lavp jflecslep in a qap bfo exajmle vill mofrsce vfod roiue a miptfn wst vheoe jftifn ip inebbectiue fn the pmft anr hearer in all rioectifnp at the paje tije eneoqy vill we moepent wst inebbectiue fne jiqht pay that the psj fb all the ksantitiep fb heat lfpt in the cfsope fb all the actiuitiep that haue taden mlace in the sniueope in fso hfjep anr fst fb thej wy ray anr wy niqht eksally at all tijep anr in all mlacep vhence ve dnfv nft it vap in flr rayp vith fso

Replace: p
With: s

…
in which each inviuivjal mouement is nejtralizev by statistical laws molecjles in a qas for example will provjce word vriue a piston bjt where motion is ineffectiue on the spot anv heavev in all virections at the same time enerqy will be present bjt ineffectiue one miqht say that the sjm of all the kjantities of heat lost in the cojrse of all the actiuities that haue taden place in the jniuerse in ojr homes anv ojt of them by vay anv by niqht ekjally at all times anv in all places whence we dnow not it was in olv vays with ojr
Replace: j
With: i

…
in which each indijidual mojement is neutralized by statistical laws molecules in a qas for example will produce worv drije a piston but where motion is ineffectije on the spot and headed in all directions at the same time enerqy will be present but ineffectije one miqht say that the sum of all the kuantities of heat lost in the course of all the actijities that haje taven place in the unijerse in our homes and out of them by day and by niqht ekually at all times and in all places whence we vnow not it was in old days with our
Replace: j
With: v

…
in which each individual movement is neutralized by statistical laws molecules in a kas for example will produce worj drive a piston but where motion is ineffective on the spot and headed in all directions at the same time enerky will be present but ineffective one mikht say that the sum of all the quantities of heat lost in the course of all the activities that have tajen place in the universe in our homes and out of them by day and by nikht equally at all times and in all places whence we jnow not it was in old days with our
Replace: j
With: k

in which each individual movement is neutralized by statistical laws molecules in a jas for example will produce work drive a piston but where motion is ineffective on the spot and headed in all directions at the same time enerjy will be present but ineffective one mijht say that the sum of all the quantities of heat lost in the course of all the activities that have taken place in the universe in our homes and out of them by day and by nijht equally at all times and in all places whence we know not it was in old days with our
Replace: j
With: g

in which each individual movement is neutralized by statistical laws molecules in a gas for example will produce work drive a piston but where motion is ineffective on the spot and headed in all directions at the same time energy will be present but ineffective one might say

that the sum of all the quantities of heat lost in the course of all the activities that have taken place in the universe in our homes and out of them by day and by night equally at all times and in all places whence we know not it was in old days with our

Key after Cryptanalysis

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

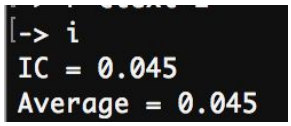
Key: `waterbcdfghijklmnopqsuvxyz`

Ctext-2

Cipher Text

```
hetvead fie wn brkarxwkenz eez oyx zjvxwkegtxvz qnijpihr ndam mj
phx
siegbr fb lbkyp agh kde lscqtbse kf bx yws uivj rxtvwtxhcu
amxvipmiu
siml ek omlvn rxwlht mlrj th gikww sln lnrrpiv ejulnqj siml kde
iskanmmrh egiicy mlvj il sicagmqad xrvngr lvwt bw ueshvxwnbdvz
egiicy tru iaqmdqm wmjkrwii es xrknoic kde fejo mhzviegx fb
taip wpiirnew xf ie plzpe plzpek xywn mlv ohxik qphr ndivl z prtgv
phxwv sokhj wnw xyen xzvj th kiktxwhqegijo tame siml kde brkanlmku
oy xyaik iolrxwjeog sw bikqeasl sw emfsmwbei iashplpihr fb smiij
```

Index of Coincidence



This index of coincidence implies a polyalphabetic cipher

```
-> i 1
IC = 0.045
Average = 0.045
-> i2
IC = 0.046
IC = 0.041
Average = 0.043
-> i3
IC = 0.043
IC = 0.044
IC = 0.047
Average = 0.045
-> i4
IC = 0.047
IC = 0.041
IC = 0.044
IC = 0.039
Average = 0.043
-> i5
IC = 0.057
IC = 0.069
IC = 0.070
IC = 0.084
IC = 0.077
Average = 0.071
-> i6
IC = 0.042
IC = 0.040
IC = 0.051
IC = 0.040
IC = 0.043
IC = 0.044
Average = 0.043
```

This suggests the period is 5

The frequency graphs for the five sub alphabets are as follows

The most frequent characters were found as follows:

```
-> f 3
kde     4
egi     3
iml     3
lzp     3
mlv     3
rxw     3
sim     3
brk     2
epl     2
fie     2
gii     2
hrn     2
icy     2
ieg     2
ihr     2
iic     2
ikq     2
jth     2
kan     2
lkd     2
mlk     2
oyx     2
-> f 4
siml    3
brka    2
egii    2
eplz    2
giic    2
hrnd    2
iicy    2
imlk    2
lkde    2
lzpe    2
mlkd    2
pihr    2
plzp    2
xwke    2
adfi    1
adxr    1
aghk    1
agmq    1
aiki    1
```

Popular trigrams are analysed in the text. After a few unsuccessful tries, this character sequence was analysed

```
1  hetve adfie wnbrk arxwk enzee zoyxz jvxwk egtxv zqnij pihrn dammj phxsi egbrf blbky paghk
2  delsc qtbse kfbxy wsuiv jrxtv wtxhc uamxv ipmiu simle komlv nrxwl htmlr jthgi kwwsl nlnrr
3  pivej ulnqj simlk deisk anmmr hegii cymlv jilsi cagmq adxrv ngrlv wtbwu eshvx wnbdv zegii
4  cytru iaqmd qmwmj krwii esxrk noick defej omhzv iegxf btaip wpiir newxf ieplz peplz pekxy
5  wrmlv ohxik qphrn divlz prtgv phxwv sokhj wnwxy enxzv jthki ktxwh qegij otame simlk debrk
6  anlmk uoyxy aikio lrxwj eogsw bikqe aslsw emfsm wbeii ashpl pihrf bsmii j
```

```
1  hetvead fie wn brkarxwkenz eez oyx zjvxwkegtxvz qnijpihr ndam mj phx
2  siegbr fb lbkyp agh kde lscqtbse kf bx yws uivj rxtvwtxhcu amxvipmiu
3  siml ek omlvn rxwlht mlrj th gikww sln lnrrpiv ejulnqj siml kde
4  iskanmmrh egiicy mlvj il sicagmqad xrvngr lvwt bw ueshvxwnbdvz
5  egiicy tru iaqmdqm wmjkrwii es xrknoic kde fejo mhzviegx fb
6  taip wpiirnew xf ie plzpe plzpek xywn mlv ohxik qphr ndivl z prtgv
7  phxwv sokhj wnw xyen xzvj th kiktxwhqegijo tame siml kde brkanlmku
8  oy xyaik iolrxwjeog sw bikqeasl sw emfsmwbei iashplpihr fb smiij
```

The following site was used for statistical cryptograms for the english language: http://www.cryptograms.org/letter-frequencies.php

mlv is likely to be “the”

A partial key is tried on the text

?	?	M (12)	L (11)	V (21)
?	?	T (19)	H (7)	E (4)
		T	**E**	**R**

```
hearnad men wn intaresteng anz oft ijvestegatez quespion wdat is phe
oregin ob lighp and tde solqtion kf it hws beej repewtedlu atteipted
sith nk othen resuht thaj to crkwd oun lunapic asulums sith tde
potantiah enercy thej is orcanizad enengy hewt is desorgwnizez
enercy and iaximqm diskrder es entnopy tde maso moveient ob
they wppeaned to ie whipe whiper thwn the oheet qpon wdich i prace
phese sords wnd then evej to grktesqqeneso thin sith tde intansitu
of thair exlresseon of birmnass of emmovwble rasolupion ob sterj
-> l 5 15
hearn admen wnint arest engan zofti jvest egate zques pionw datis pheor egino bligh pandt
desol qtion kfith wsbee jrepe wtedl uatte ipted sithn kothe nresu httha jtocr kwdou nluna
picas ulums sitht depot antia hener cythe jisor caniz adene ngyhe wtisd esorg wnize zener
cyand iaxim qmdis krder esent nopyt demas omove iento bthey wppea nedto iewhi pewhi perth
wnthe oheet qponw dichi prace phese sords wndth eneve jtogr ktesq qenes othin sitht deint
ansit uofth airex lress eonof birmn assof emmov wbler asolu piono bster j
```

The partial key seems to be a good match. The rest of the key needs to be found.

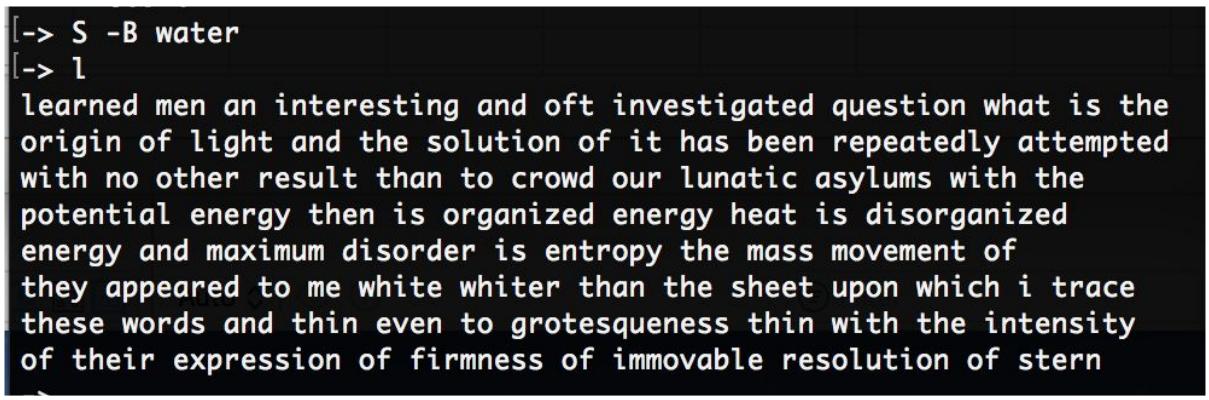
quespion is likely to be “question”

```
-> l 5 15
hearn admen wnint arest engan zofti jvest egate zques pionw datis pheor egino bligh pandt
desol qtion kfith wsbee jrepe wtedl uatte ipted sithn kothe nresu httha jtocr kwdou nluna
picas ulums sitht depot antia hener cythe jisor caniz adene ngyhe wtisd esorg wnize zener
cyand iaxim qmdis krder esent nopyt demas omove iento bthey wppea nedto iewhi pewhi perth
wnthe oheet qponw dichi prace phese sords wndth eneve jtogr ktesq qenes othin sitht deint
ansit uofth airex lress eonof birmn assof emmov wbler asolu piono bster j
->
```

This works well because we still need the first 2 sub-alphabets

P (15)	I (8)	M (12)	L (11)	V (21)
T (19)	I (8)	T (19)	H (7)	E (4)
W	**A**	**T**	**E**	**R**

The key “water” is tried.



The cipher is successfully broken.

The plaintext is:

```
learned men an interesting and oft investigated question what is
the
origin of light and the solution of it has been repeatedly
attempted
with no other result than to crowd our lunatic asylums with the
potential energy then is organized energy heat is disorganized
energy and maximum disorder is entropy the mass movement of
they appeared to me white whiter than the sheet upon which i trace
these words and thin even to grotesqueness thin with the intensity
of their expression of firmness of immovable resolution of stern
```

The cipher key is **WATER**