MTH6115 Cryptography Exercises 2 Solutions

Q1 (a) If $y \leq x/2$ then since $r < y$ we have $r < x/2$. On the other hand, if $y > x/2$, then we get a remainder of $r = x - y < x/2$ when we divide x by y.

We show that Euclid's algorithm requires at most $\lfloor \log_2 x + \log_2 y \rfloor$ divisions to find $\gcd(x, y)$ by induction on $x + y$.

Base case $x + y = 2$. Then $x = 1$ and $y = 1$ and we may deduce that $\gcd(x, y) = 1$ by using 0 divisions. Since $\lfloor \log_2 1 + \log_2 1 \rfloor = 0$, the statement is true when $x + y = 2$.

Inductive Hypothesis Suppose that $K \geq 2$ and that Euclid's algorithm requires at most $\lfloor \log_2 x + \log_2 y \rfloor$ divisions to find $\gcd(x, y)$ for all integers x, y with $1 \leq y \leq x$ and $x + y \leq K$.

Induction Step Suppose that $x + y = K + 1$. We need to show that Euclid's algorithm requires at most $\lfloor \log_2 x + \log_2 y \rfloor$ divisions to find $\gcd(x, y)$. In the first step of Euclid's algorithm we divide x by y to get a remainder r with $0 \leq r < y$. We know that $\gcd(x, y) = \gcd(y, r)$. Since $r < y \leq x$, we have $y + r < x + y = K + 1$, so we may apply induction to deduce that Euclid's algorithm requires at most $\lfloor \log_2 y + \log_2 r \rfloor$ divisions to find $\gcd(y, r)$. Hence Euclid's algorithm requires at most $\lfloor \log_2 y + \log_2 r \rfloor + 1$ divisions to find $\gcd(x, y)$. Since $r < x/2$ by the first part, we have

$$\lfloor \log_2 y + \log_2 r \rfloor + 1 = \lfloor \log_2 y + \log_2 r + 1 \rfloor = \lfloor \log_2 y + \log_2 2r \rfloor < \lfloor \log_2 x + \log_2 y \rfloor$$

Thus Euclid's algorithm requires at most $\lfloor \log_2 x + \log_2 y \rfloor$ divisions to find $\gcd(x, y)$.

(b) Suppose x is not prime then $x = yz$ for some $1 < y \leq z < x$. Hence $x \geq y^2$ and so $y \leq \sqrt{x}$. This implies that if x is not prime then it is divisible by some integer y with $2 \leq z \leq \sqrt{x}$. So we can find a proper factor of x or decide that x is prime by dividing x by each integer between 2 and $\sqrt{x}$.

(c) The algorithms given in (a) takes at most

$$(\log_2 10^{20} + \log_2 10^{20}) \times 0.01 = (40 \log_2 10) \times 0.01 \sim 1.33$$

seconds. The algorithms given in (b) takes at most $\sqrt{10^{20}} \times 0.01 = 10^{10} \times 0.01 = 10^8$ seconds, which is approximately 3.2 years.

Q2 (b) The most common letter in the ciphertext is N and the next most common letter is G. So we guess that N and G should be deciphered as 'e' and 't' respectively. Hence $\theta_{b,c}(13) = 4$ and $\theta_{b,c}(6) = 19$. This gives us the simultaneous equations

$$\begin{aligned} 13b + c &\equiv 4 \pmod{26} \qquad (1) \\ 6b + c &\equiv 19 \pmod{26} \qquad (2) \end{aligned}$$

Subtracting (2) from (1) gives $7b \equiv -15 \pmod{26}$. Since the multiplicative inverse of 7 in $\mathbb{Z}_{26}$ is 15, this gives $b \equiv 15(-15) \pmod{26}$. Thus $b \equiv -225 \equiv 9$

(mod 26). Substituting into (2) we obtain $6(9)+c \equiv 19 \pmod{26}$. So $c \equiv 19-6(9) \equiv -35 \equiv 17 \pmod{26}$. Thus the affine permutation for deciphering is $\theta_{9,17}$. This implies that the permutation we use for deciphering is $A \to r, B \to a, C \to j, D \to s, E \to b, F \to k, G \to t, H \to c, I \to l, J \to u, K \to d, L \to m, N \to e, O \to n, P \to w, Q \to f, R \to o, S \to x, T \to g, U \to p, V \to y, W \to h, X \to q, Y \to z, Z \to i$. This gives the plaintext as

```
thefi rstgo almig hthav ebeen handb allbu tther efere ehasg iveni
tthes econd onemi ghtha vebee nouts ideth eboxb utaga inheh asmad
ethed ecisi onaga instu satth eendo fthed aywed idntd efend themw
ellen oughi cantb ecrit icalo fther efere ebuti canbe criti calof
mydef ender snotd efend ingit wellt heywe repoo rzzzz
```

> The first goal might have been handball but the referee has given it. The second one might have been outside the box but, again, he has made the decision against us. At the end of the day, we didn't defend them well enough. I can't be critical of the referee but I can be critical of my defenders not defending it well, they were poor.

(b) We can find $\theta_{s,t}$ similarly using the facts that $\theta_{s,t}(4)=13$ and $\theta_{s,t}(19)=6$. This gives $s=3$ and $t=1$.

Q3 The trigram KLH occurs with the K in positions 1, 112, 163, 181, 361, 367. This indicates that the length of the keyword is $\gcd(111, 51, 18, 180, 6) = 3$.

We might also guess that KLH is an encryption of 'the'. This would imply that the first Caeser cipher maps 't' to K, the second Caeser cipher maps 'h' to L, and the third Caeser cipher maps 'e' to H. This would in turn imply that the keyword is RED. If we use this to decipher the message we obtain

> The boys have done brilliantly. I don't know where they found u[i]t from. They found great resolve and great inner strength to fight back for the win. We have done that in a lot of games this season and I have the utmost regard for the players for that. It seemed to be, from where I was standing, an eventful game. Sv[u]nderland scored and deserved to. It took us some time to get into our st[r]ide after the goal. It was a difficult period but we weathered that and then in the last twenty minutes of the first half we came into the gal[m]e and got a bit of rhythm about ourselves.

(There were three wrong letters and one missing letter in the ciphertext. i have indicated the correct letters with square brackets.)

Alternatively we could rewrite the cyphertext in blocks of size three and use frequency analysis for the first, second, and third letters in each block to find the keyword:

Frequencies of letters in positions congruent to one modulo three in the ciphertex:

```
A 0=0%, B 0=0%, C 3=2%, D 2=1.3%, E 11=7.3%, F 11=7.3%, G 2=1.3%,
H 0=0%, I 10=6.7%, J 4=2.7%, K 15=10%, L 5=3.3%, M 1=0.7%,
N 5=3.3%, O 0=0%, P 2=1.3%, Q 0=0%, R 15=10%, S 4=2.7%, T 2=1.3%,
U 9=6%, V 23=15.3%, W 9=6%, X 3=2%, Y 8=5.3%, Z 6=4%
```

If we compare with the expected frequencies in AAIW, we see that the best fit is a Caeser shift which takes ‘a’ to R. So we can decrypt the first letter in each block as below.

```
tLH bSB sLD vIG oRH bVL lPL aRW lCL dSQ tOQ oAZ hIU eXK eCI oYQ dMX
fVR mXK eCI oYQ dKU eEW rIV oPY eEQ dKU eEW iRQ eVV tVH nKW hXR fMJ
hXE aGN fSU tLH wMQ wIK aZH dSQ eXK aXL nEO oXR fKD mIV tLL sWH aWR
nEQ dMK aZH tLH uXP oWW rIJ aVG fSU tLH pPD yIU sJR rXK aXL tWH eQH
dXR bII rSP wLH rIL wEV sXD nHL nKD nIY eRW fYO gEP eWY nHH rPD nHV
cSU eHD nHG eWH rZH dXR iXW oSN uWV oQH tMP eXR gIW iRW oSX rWW iHH
aJW eVW hIJ oEO iXZ aWD dMI fMF uPW pIU iSG bYW wIZ eEW hIU eHW hEW
aRG tLH nMQ tLH lEV tXZ eRW yQL nYW eWR fXK eJL rWW hEO fAH cEP eMQ
tSW hIJ aPH aRG gSW aFL tSI rLB tLP aFR uXR uVV ePY eWT
```

Frequencies of letters in positions congruent to zero modulo three in the ciphertext:

```
A 0=0%, B 2=1.3%, C 0=0%, D 8=5.3%, E 1=0.7%, F 1=0.7%, G 6=4%,
H 21=14%, I 5=3.3%, J 4=2.7%, K 7=4.7%, L 11=7.3%, M 0=0%, N 2=1.7%,
O 4=2.7%, P 6=4%, Q 11=7.3%, R 11=7.3%, S 0=0%, T 0=0%, U 9=6%,
V 8=5.3%, W 22=14.7%, X 2=1.3%, Y 4=2.7%, Z 4=2.7%
```

If we compare with the expected frequencies in AAIW, we see that the best fit is a Caeser shift which takes ‘a’ to D. So we can decrypt the last letter in each block as below.

```
tLe bSy sLa vId oRe bVi lPi aRt lCi dSn tOn oAw hIr eXh eCf oYn dMu
fVo mXh eCf oYn dKr eEt rIs oPv eEn dKr eEt iRn eVs tVe nKt hXo fMg
hXb aGk fSr tLe wMn wIh aZe dSn eXh aXi nEl oXo fKa mIs tLi sWe aWo
nEn dMh aZe tLe uXm oWt rIg aVd fSr tLe pPa yIr sJo rXh aXi tWe eQe
dXo bIf rSm wLe rIi wEs sXa nHi nKa nIv eRt fYl gEm eWv nHe rPa nHs
cSr eHa nHd eWe rZe dXo iXt oSk uWs oQe tMm eXo gIt iRt oSu rWt iHe
aJt eVt hIg oEl iXw aWa dMf fMc uPt pIr iSd bYt wIw eEt hIr eHt hEt
aRd tLe nMn tLe lEs tXw eRt yQi nYt eWo fXh eJi rWt hEl fAe cEm eMn
tSt hIg aPe aRd gSt aFi tSf rLy tLm aFo uXo uVs ePv eWq
```

Frequencies of letters in positions congruent to two modulo three in the ciphertext:

```
A 2=1.3%, B 0=0%, C 3=2%, D 0=0%, E 13=8.7%, F 2=1.3%, G 1=0.7%,
H 7=4.7%, I 16=10.7%, J 3=2%, K 5=3.3%, L 11=7.3%, M 9=6%,
N 0=0%, O 1=0.7%, P 7=4.7%, Q 3=2%, R 8=5.3%, S 13=8.7%, T 0=0%,
U 0=0%, V 7=4.7%, W 12=8%, X 19=12.7%, Y 5=3.3%, Z 3=2%
```

If we compare with the expected frequencies in AAIW, we see that the best fit is a Caeser shift which takes ‘a’ to E. So we can decrypt the middle letter in each block as below.

```
the boy sha ved one bri lli ant lyi don tkn oww her eth eyf oun diu
fro mth eyf oun dgr eat res olv ean dgr eat inn ers tre ngt hto fig
htb ack for the win weh ave don eth ati nal oto fga mes thi sse aso
nan dih ave the utm ost reg ard for the pla yer sfo rth ati tse eme
dto bef rom whe rei was sta ndi nga nev ent ful gam esv nde rla nds
cor eda ndd ese rve dto itt ook uss ome tim eto get int oou rst ide
aft ert heg oal itw asa dif fic ult per iod but wew eat her edt hat
and the nin the las ttw ent ymi nut eso fth efi rst hal fwe cam ein
tot heg ale and got abi tof rhy thm abo uto urs elv esq
```