Math 4981 Spring 2021
Cryptology HW 2 Solutions
Due Thursday, January 28

1. Encrypt the plaintext message "NEVER ODD OR EVEN", using a Vigenère cipher with key word "potato".

   Solution: POTATO becomes 15, 14, 19, 0, 19, 14 so we have

Plaintext	N	E	V	E	R	O	D	D	O	R	E	V	E	N
Plaintext	13	4	21	4	17	14	3	3	14	17	4	21	4	13
Keystream	15	14	19	0	19	14	15	14	19	0	19	14	15	14
Ciphertext	2	18	14	4	10	2	18	17	7	17	23	9	19	1
Ciphertext	C	S	O	E	K	C	S	R	H	R	X	J	T	B

   So the ciphertext is "CSOEK CSRHR XJTB".

2. Decrypt the ciphertext "ODESL UKWGK SXMSK GEPP", which was encrypted with Vigenère cipher using the key word "octopus".

   Solution: "OCTOPUS" becomes 14, 2, 19, 14, 15, 20, 18. So we have

Ciphertext	O	D	E	S	L	U	K	W	G	K	S	X	M	S	K	G	E	P	P
Ciphertext	14	3	4	18	11	20	10	22	6	10	18	23	12	18	10	6	4	15	15
Keystream	14	2	19	14	15	20	18	14	2	19	14	15	20	18	14	2	19	14	15
Plaintext	0	1	11	4	22	0	18	8	4	17	4	8	18	0	22	4	11	1	0
Plaintext	A	B	L	E	W	A	S	I	E	R	E	I	S	A	W	E	L	B	A

   This gives "ABLEW ASIER EISAW ELBA" or "Able was I, ere I saw Elba".

3. Encrypt the plaintext message "RATS LIVE ON NO EVIL STAR", using an Autokey cipher with the key word "vital"

   Solution: VITAL is 21, 8, 19, 0, 11, so we have

Plaintext	R	A	T	S	L	I	V	E	O	N	N	O	E	V	I	L	S	T	A
Plaintext	17	0	19	18	11	8	21	4	14	13	13	14	4	21	8	11	18	19	0
Keystream	21	8	19	0	11	17	0	19	18	11	8	21	4	14	13	13	14	4	21
Ciphertext	12	8	12	18	22	25	21	23	6	24	21	9	8	9	21	24	6	23	21
Ciphertext	M	I	M	S	W	Z	V	X	G	Y	V	J	I	J	V	Y	G	X	V

   So the ciphertext is "MIMS WZVX GYVJ IJVY GXVZ".

4. Decrypt the ciphertext "UBTW SEFH TTHF", which was encrypted with an Autokey cipher using the key word "cipher".

   Solution:

"CIPHER" becomes 2, 8, 15, 7, 4, 17. So we have

Ciphertext	U	B	T	W	S	E	F	H	T	T	H	F
Ciphertext	20	1	19	22	18	4	5	7	19	19	7	5
Keystream	2	8	15	7	4	17	18	19	4	15	14	13
Plaintext	18	19	4	15	14	13	13	14	15	4	19	18
Plaintext	S	T	E	P	O	N	N	O	P	E	T	S

Thus we get "STEP ONNO PETS" or "step on no pets."

5. Compute **by hand** the indices of coincidence of each one of the following strings. Show your work.

   (a) `It is a truth universally acknowledged` (33)

   (b) `that a single man in possession of a good` (33)

   (c) `fortune must be in want of a wife` (26)

   Solution: The frequency charts for each string are

   (a)

Letter	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
Frequency	3	0	1	2	3	0	1	1	3	0	1	3	0	2	1	0	0	2	2	3	2

   length 33

   (b)

Letter	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
Frequency	4	0	0	1	2	1	2	1	3	0	0	1	1	4	5	1	0	0	5	2	0

   length 33

   (c)

Letter	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
Frequency	2	1	0	0	3	3	0	0	2	0	0	0	1	3	2	0	0	1	1	3	2

   length 26

   So the indices of coincidence are

   (a)

$$\frac{1}{33 \cdot 32}(3 \cdot 2 + 2 \cdot 1 + 3 \cdot 2 + 3 \cdot 2 + 3 \cdot 2 + 2 \cdot 1 + 2 \cdot 1 + 2 \cdot 1 + 3 \cdot 2 + 2 \cdot 1) = \frac{40}{1056} = \frac{5}{132} = .03\overline{78}.$$

   (b)

$$\frac{1}{33 \cdot 32}(4 \cdot 3 + 2 \cdot 1 + 2 \cdot 1 + 3 \cdot 2 + 4 \cdot 3 + 5 \cdot 4 + 5 \cdot 4 + 2 \cdot 1) = \frac{76}{1056} = \frac{19}{264} = .071\overline{96}.$$

   (c)

$$\frac{1}{26 \cdot 25}(2 \cdot 1 + 3 \cdot 2 + 3 \cdot 2 + 2 \cdot 1 + 3 \cdot 2 + 2 \cdot 1 + 3 \cdot 2 + 2 \cdot 1 + 2 \cdot 1) = \frac{34}{650} = \frac{17}{325} \approx .05231.$$

6. Compute the mutual index of coincidence for each pair of strings in the previous problem. Show me your work.

   Solution: To calculate the mutual indices of coincidence, we get

   (a) and (b):

$$\frac{1}{33 \cdot 33}(3 \cdot 4 + 2 \cdot 1 + 3 \cdot 2 + 1 \cdot 2 + 1 \cdot 1 + 3 \cdot 3 + 3 \cdot 1 + 2 \cdot 4 + 1 \cdot 5 + 2 \cdot 5 + 3 \cdot 2) = \frac{64}{1089} \approx .05877.$$

   (a) and (c):

$$\frac{1}{33 \cdot 26}(3 \cdot 2 + 3 \cdot 3 + 3 \cdot 2 + 2 \cdot 3 + 1 \cdot 2 + 2 \cdot 1 + 2 \cdot 1 + 3 \cdot 3 + 2 \cdot 2 + 1 \cdot 2) = \frac{48}{858} = \frac{8}{143} \approx .055944.$$

   (b) and (c):

$$\frac{1}{33 \cdot 26}(4 \cdot 2 + 2 \cdot 3 + 1 \cdot 3 + 3 \cdot 2 + 1 \cdot 1 + 4 \cdot 3 + 5 \cdot 2 + 5 \cdot 1 + 2 \cdot 3) = \frac{57}{858} = \frac{19}{286} \approx .06653.$$

7. Compute **by hand** the index of coincidence of the following string. What is unusual about this string?

   `A quick brown fox jumps over the lazy dog`

   Solution: This text has 33 letters, with four "o"s, 2 "a", "e", "r", and "u", and 1 of each other letter. Thus the index of coincidence is

$$\text{IndCo} = \frac{1}{33 \cdot 32}(4 \cdot 3 + 2 \cdot 1 + 2 \cdot 1 + 2 \cdot 1 + 2 \cdot 1) = \frac{20}{33 \cdot 32} = \frac{5}{33 \cdot 8} \approx .019.$$

   This is really unusually low—much lower than you'd expect even for completely random text. In a very long text you can't go much below .036, but this text is so short that we can have a very low index of coincidence.

   (The text is famously a *pangram*: a sentence that contains each letter at least once. It is not a perfect pangram, containing each letter exactly once; such a sentence would have an index of coincidence equal to zero. Perfect pangrams exist, but tend to be strange and extremely awkard in phrasing ("Mr Jock, TV quiz PhD, bags few lynx").)

   For the remaining problems, you may use an index of coincidence calculator like the one at `http://www.practicalcryptography.com/cryptanalysis/text-characterisation/index-coincidence/`

   You should be able to copy and paste the ciphertexts from the PDF.

8. Which of the following is likely to be a message encrypted with a simple substitution cipher? Why?

   (a) GSZES GNUBE SZGUG SNKGX CSUUE QNZOQ EOVJN VXKNG XGAHS AWSZZ BOVUE SIXCQ NQESX NGEUG AHZQA QHNSP CIPQA OIDLV JXGAK CGJCG SASUB FVQAV CIAWN VWOVP SNSXV JGPCV NODIX GJQAE VOOXC SXXCG OGOVA XGNVU BAVKX QZVQD LVJXQ EXCQO VKCQG AMVAX VWXCG OOBOX VZCSO SPPSN VAXUB DVVAX QJQAJ VSUXC SXXCV OVJCS NSJXV NOJQA MVBSZ VOOSH VSAWX QHGMV GWVSX CSXXC VBSNV ZVNVN SAWQZ ORVXJ CVOQE JCGUW NVA

   (b) DWVQP IIKOP UUYGC ZJDRU ZDSHI CXEAO AKRZC QAMSM DNQLF LUJYI IMJPJ VJZQL GCJSN XTXFL MWOLW IFUQK DBBEY HVMVF ZOJXV FYMJA RDTGT TZQKL YNHPD UPUYU XKNOI DXZXG IHIWK VXZET XFMOS GKIWU EIFDW RLXHH PZXPI VFWKL THEAS IROWC GJAYJ KKODL WXFPI ZVUIK LEXEL IOWVC YMFMR UIZUD CETFE TBPIX SWSPZ MRPKP LIYKL FGSTJ ZTPUH AEBQC QAEPQ GIAKH TDUVM KFGEU MWHAY ZGVSJ LNLJJ MLPEO YEMZU PYMEW XLG

   Solution: We compute the indices of coincidence. The index of coincidence of the first sample is $\approx .063$, and the index of coincidence of the second sample is $\approx .039$. Thus the second looks more like random text, and the first looks like it was encrypted with a simple substitution cipher.

9. Consider the following ciphertext:

   TOGMG GBYMK KCQIV DMLXK KBYIF VCUEK CUUIS VVXQS PWWEJ KOQGG PHUMT WHLSF YOVWW KNHHM RCQFQ VVHKW PSUED UGRSF CTWIJ KHVFA THKEF FWPTJ GGVIV CGDRA PGWVM OSQXG HKDVT WHUEV KCWYJ PSGSN GFWSL JSFSE OOQHW TOFSH ACIIN GFBIF GABGJ ADWSY TOPML ECQZW ASGVS FWRQS FSFVQ RHDRS NMVMK CBHRV KBLXK GZI

   (a) Use the Kasiski test (either by shifting the text over one-by-one, or by comparing repeated trigrams) to guess the keyword length. (Hint: the repeated trigrams are LXK at 17 and 232; TWH at 54 and 134; NGF at 149 and 174; and SFS at 156 and 209).

   Solution: We see that the offets that the trigrams have are $232-17 = 215; 124-54 = 80; 174-149 = 35; 209-156 = 53$. Since the first three have 5 as a common factor, we might guess the keyword has length five.

   Alternately, we can offset the text and look for coincidences. With an offset of 2 or 3 there are eight coincidences; with an offset of 4 there are 4; with an offset of 5 there are 16; with an offset of 6 there are 10; with an offset of 7 or 8 there are 9. This gives more weak evidence that the keyword has length 5.

   (b) Use the tool at `https://jaydaigle.net/substring.html` to split the text up into substrings. Test the index of coincidence of each substring, and use this to determine the key length. Does this match your answer in part (a)?

   Solution: If we use three substrings, the indices are $.044, .041, .044$. If we use four, the indices are $.039, .042, .040, .045$. If we use five, the indices are $.062, .069, .059, .072, .056$, which is high enough that this is almost certainly the correct keylength. But for completeness we see that with six substrings we get $.037, .045, .046, .044, .031, .035$, and with seven we get $.046, .029, .032, .041, .036, .032, .046$. So at this point we're pretty sure the keyword has length 5.

(c) Find the keyword and decrypt the message. You can use frequency analysis on substrings, or you can use mutual indices of coincidence, but be sure to show intermediate steps so I can see what you did.

Solution: The keyword is `codes` and the message is

> radio, envisioned by its inventor as a great humanitarian contribution, was seized upon by the generals soon after its birth...and impressed as an instrument of war....But radio turned over to the commander a copy of every enemy cryptogram it conveyed....Radio made cryptanalysis an end in itself.

10. Suppose you are decrypting the ciphertext

```
FBPL RHEF NWZS AMPM OBFK JHPT NHEQ RNMC LIXC BHPA NMDY ASQM AIYC YYZN UYEM
MCDQ XFGC CBPN XFTR RWLJ KUYB BQSG LBSY EYNM WHPA CYOR QYXU RNSY WIEF NLLL
MNZY BMFK NUXM WAEF NJZU NLDM ONSC NUCR QNSC BYAY AUEC JHOC ZOLJ BNLR RIYR
XQSG LBEF NFLU BIQL JNFP NUYB XZYY COCC BAZB NHEG CFPR QYXY MYNC WNCC BJPA
CNZR QYZN RHTM WMZD VUYI RHOP NKFG AYDR QUER QYJQ QIFJ MXPA UUCC CBPA JODC
BQSG LBTK YYWR QYXR XNSC BYAY AUEG XH
```

You have determined that the length of the keyword is four, have split the message into four substrings, and have computed the mutual index of coincidence of each substring with each other substring at each possible relative shift.

In the following table, the first two columns identify the substrings, the third identifies the amount by which the second substring has been shifted, and the fourth is the mutual index of coincidence. Thus, the entry that says $1, 2, 3, .055$ says that if you shift the second substring over by 3, and then compute the mutual index of coincidence with the first substring, you get .055.

Use this table to determine the relative shifts of the four substrings, find the keyword, and decrypt the message.

1	2	0	0.049	1	2	1	0.034	1	2	2	0.040	1	2	3	0.055
1	2	4	0.051	1	2	5	0.031	1	2	6	0.040	1	2	7	0.030
1	2	8	0.035	1	2	9	0.035	1	2	10	0.034	1	2	11	0.035
1	2	12	0.033	1	2	13	0.036	1	2	14	0.042	1	2	15	0.072
1	2	16	0.037	1	2	17	0.026	1	2	18	0.040	1	2	19	0.057
1	2	20	0.030	1	2	21	0.026	1	2	22	0.038	1	2	23	0.033
1	2	24	0.027	1	2	25	0.037								
1	3	0	0.026	1	3	1	0.036	1	3	2	0.048	1	3	3	0.033
1	3	4	0.031	1	3	5	0.033	1	3	6	0.023	1	3	7	0.030
1	3	8	0.044	1	3	9	0.049	1	3	10	0.035	1	3	11	0.039
1	3	12	0.054	1	3	13	0.048	1	3	14	0.040	1	3	15	0.032
1	3	16	0.030	1	3	17	0.033	1	3	18	0.034	1	3	19	0.036
1	3	20	0.036	1	3	21	0.036	1	3	22	0.035	1	3	23	0.050
1	3	24	0.070	1	3	25	0.042								
1	4	0	0.053	1	4	1	0.035	1	4	2	0.029	1	4	3	0.037
1	4	4	0.035	1	4	5	0.029	1	4	6	0.030	1	4	7	0.043
1	4	8	0.029	1	4	9	0.032	1	4	10	0.049	1	4	11	0.072
1	4	12	0.032	1	4	13	0.029	1	4	14	0.043	1	4	15	0.054
1	4	16	0.032	1	4	17	0.026	1	4	18	0.037	1	4	19	0.029
1	4	20	0.032	1	4	21	0.039	1	4	22	0.053	1	4	23	0.032
1	4	24	0.034	1	4	25	0.056								
2	3	0	0.027	2	3	1	0.037	2	3	2	0.044	2	3	3	0.037
2	3	4	0.032	2	3	5	0.041	2	3	6	0.035	2	3	7	0.025
2	3	8	0.045	2	3	9	0.075	2	3	10	0.040	2	3	11	0.032
2	3	12	0.029	2	3	13	0.034	2	3	14	0.030	2	3	15	0.043
2	3	16	0.045	2	3	17	0.024	2	3	18	0.035	2	3	19	0.035
2	3	20	0.048	2	3	21	0.045	2	3	22	0.043	2	3	23	0.038
2	3	24	0.041	2	3	25	0.037								
2	4	0	0.043	2	4	1	0.034	2	4	2	0.028	2	4	3	0.046
2	4	4	0.026	2	4	5	0.040	2	4	6	0.030	2	4	7	0.063
2	4	8	0.036	2	4	9	0.034	2	4	10	0.040	2	4	11	0.048
2	4	12	0.036	2	4	13	0.027	2	4	14	0.044	2	4	15	0.039
2	4	16	0.036	2	4	17	0.024	2	4	18	0.052	2	4	19	0.029
2	4	20	0.028	2	4	21	0.037	2	4	22	0.080	2	4	23	0.031
2	4	24	0.033	2	4	25	0.036								
3	4	0	0.032	3	4	1	0.044	3	4	2	0.049	3	4	3	0.041
3	4	4	0.030	3	4	5	0.032	3	4	6	0.044	3	4	7	0.033
3	4	8	0.031	3	4	9	0.039	3	4	10	0.027	3	4	11	0.032
3	4	12	0.042	3	4	13	0.073	3	4	14	0.039	3	4	15	0.029
3	4	16	0.037	3	4	17	0.043	3	4	18	0.038	3	4	19	0.029
3	4	20	0.040	3	4	21	0.034	3	4	22	0.033	3	4	23	0.040
3	4	24	0.050	3	4	25	0.041								

Solution:

The entries with large mutual index of coincidence are:

1	2	15	1	3	24	1	4	11
2	3	9	2	4	22	3	4	13

Thus we have

$$\begin{aligned} \beta_1 &= \beta_2 + 15 & \beta_1 &= \beta_3 + 24 & \beta_1 &= \beta_4 + 11 \\ \beta_2 &= \beta_3 + 9 & \beta_2 &= \beta_4 + 22 & \beta_3 &= \beta_4 + 13. \end{aligned}$$

If we set $\beta_1 = 0$, then from the first three equations we get $\beta_2 = -15 \equiv 11, \beta_3 = -24 \equiv 2$, and $\beta_4 = -11 \equiv 15$.

We check with the second triple of equations, and indeed we have $11 = 2 + 9, 11 \equiv 15 + 22 = 37$, and $2 \equiv 15 + 13 = 28$. So since 0 corresponds to `A`, 11 corresponds to `L`, 2 corresponds to `C`, and 15 corresponds to `P`, our keyword is some shift of `0-11-2-15` or `ALCP`.

At this point there are a few ways to figure out the rest, but I might just see that our text starts out with `FBPL RHEF` which corresponds to the numbers `5-1-15-11 17-7-4-5`. Subtracting our keyword gives `5-16-13-22 17-22-2-16` or `fqnw rwcq`.

This obviously isn't actual plaintext, but we can now make a chart

0	`FQNW RWCQ`	1	`epmv qvbp`	2	`dolu puao`	3	`cnkt otzn`
4	`bmjs nsym`	5	`alir mrxl`	6	`zkhq lqwk`	7	`yjgp kpvj`
8	`xifo joui`	9	`when inth`	10	`vgdm hmsg`	11	`ufcl glrf`
12	`tebk fkqe`	13	`sdaj ejpd`	14	`rczi dioc`	15	`qbyh chnb`
16	`paxg bgma`	17	`ozwf aflz`	18	`nyve zeky`	19	`mxud ydjx`
20	`lwtc xciw`	21	`kvsb wbhv`	22	`jura vagu`	23	`itqz uzft`
24	`hspy tyes`	25	`grox sxdr`				

The only one that looks reasonable is the shift of 9 which looks like `when inth`, and a keyword of `JULY`. Doing the entire decryption gives

```
when inth ecou rseo fhum anev ents itbe come snec essa ryfo rone peop leto
diss olve thep olit ical band swhi chha veco nnec tedt hemw itha noth eran
dtoa ssum eamo ngth epow erso fthe eart hthe sepa rate ande qual stat iont
owhi chth elaw sofn atur eand ofna ture sgod enti tlet hema dece ntre spec
ttot heop inio nsof mank indr equi rest hatt heys houl ddec lare thec ause
swhi chim pelt hemt othe sepa rati on
```

or

> When in the Course of human events, it becomes necessary for one people to dissolve the political bands which have connected them with another, and to assume among the powers of the earth, the separate and equal station to which the Laws of Nature and of Nature's God entitle them, a decent respect to the opinions of mankind requires that they should declare the causes which impel them to the separation.

11. Would the method we used on the Vigenère cipher work to decrypt an autokey cipher? Why or why not?

Solution: No, because the keystream does not repeat. As we saw in class this week, the autokey cipher is quite vulnerable, but since the keystream doesn't repeat you can't use substring analysis to find the keyword. However, the keystream comes from *somewhere*, and we can attack that connection.

12. The ciphertext `bxo tam det xzx pjl baj lqf asa mde ohy wpc wlb ajl` was encrypted with an autokey cipher. Decrypt it, using the knowledge that the word "the" appears.

    Solution: `itw ast heb est oft ime sit was the wor sto fti mes` or "It was the best of times, it was the worst of times".