4.1 Playfair Cipher

You create a grid similar to that described in Nihilist cipher. Consider picking a keyword (without repetition) such as "PROBLEMS":

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

The only difference here is that we don't have the numerical labels for the rows and columns.

In this cipher, we will encipher letters pairs at a time. Consider the following plaintext:

```
SHE WENT TO THE STORE
```

When we pair up the letters they get grouped as follows:

```
SH EW EN TT OT HE ST OR E
```

But, we are not allowed to encipher any double letters. So, in this case, we will insert an Q into the plaintext. (If Q is a double letter, then insert another infrequent letter, say X.)

```
SH EW EN TQ TO TH ES TO RE
```

To encipher pairs of letters, adhere to the following rules:

1) If the two letters are on the same row of the chart, like "ES", then replace each letter by the letter to the right. (If necessary, wrap around to the left end of the row. So "ES" encrypts to "MA".

2) If the two letters are on the same column of the chart, like, "TH", then replace each letter by the letter below it. (If necessary, wrap around to the top end of the column.) So "TH" encrypts to "YT".

3) If two letters are on a different row and column, like, "SH", then replace each letter by another letter on its same row, but in the column of the other letter. So "SH" encrypts to "AG".

Using these rules, here is the encryption of the plaintext above:

```
Plaintext : SH EW EN TQ TO TH ES TO RE
Ciphertext: AG MV MK UT QB YT MA QB PM
```

For decryption, if two ciphertext letters are on the same row or column, replace them with the two letters to the left or above, respectively. Otherwise, for each letter choose the letter on the same row and the other letter's column for decryption. (So this is the original operation, it is the reverse of itself.)

To cryptanalyze Playfair, we first might want to try to determine if a ciphertext is using Playfair. Here are some clues that it is:

1) There must be an even number of characters in the cipher text.

2) The rare consonants (j,k,q,x,z) will appear more frequently in the plaintext.

3) When divided into digraphs, no repeated letters will appear.

4) The frequency distribution of digraphs will approximate that of plaintext.

Here are some other unique characteristics of the Playfair cipher:

1) No single letter ever encrypts to itself.

2) Two reversed digraphs in the plaintext will always be represented by reverse digraphs in the ciphertext.

3) Every single letter from the plaintext can be enciphered by one of only five other letters – the one directly below it in the Playfair square or the other four in its row.

To perform a known-plaintext attack on the Playfair cipher, you try different positions of the known-plaintext to match with the ciphertext, and cross-check results with the rules above.

For example, if you tried to match the following:

```
Plaintext : asample
Ciphertext: pkkmkme
```

You can rule this out because this matches m to m and e to e, which is impossible in Playfair.

When you get a matching piece of plain and cipher text, you can start putting together possible placements of the key.

For example, consider the following matching:

```
Plaintext : asample
Ciphertext: ewdwqnb
```

From this you can make the following deductions:

a, s, e, and w are all on the same row, column or "box."
a, m, d and w are also on the same, row, column or "box."

It is quite likely that d and e are in the same row as a, and that w is in the same column as a.

In this manner you can attempt to start filling out the Playfair grid. Knowing that the keyword is in the beginning and the rest of the letters will roughly appear in alphabetical order can also help. From here, trial and error can eventually yield a solution.

Here is one more example:

Using the CAP tool, you find that a possible match for the playfair cipher is the following:

PL -> QK
AY -> FV
FA -> GB
IR -> LE

Looking at the first set of letters, it is likely that they are in a box:

```
K  L
P  Q
```

The reason for this is that it would be bizarre for these letters to be on the same row, because they are spaced out quite a bit (M, N, and O are missing) – though that could be plausible for all three of those to be in the keyword, or at least two of them. If the arrangement above is correct, then the two columns and rows are likely consecutive.

It is also likely that the second set of letters forms a box as well, with V and Y on the last row:

```
A  F
V  Y
```

It is unlikely that they are all in the same column because V and Y are too close together, and nearly impossible unless they are all in the keyword that the letters are all on the same row.

The third set of letters strongly suggests that the letters are all on the same row, something like this:

```
A B   F G
```

The guess is that the middle letter in the row is missing and the other letters are in the keyword. Now, this gives us the following:

```
A B   F G
    K L
    P Q
V     Y
```

The placement of K, L, P and Q is a guess. It could be shifted to the left or the right in the square. It is quite safe to guess that the last letters of the alphabet fill out the bottom row since there's no extra room:

```
A B   F G
    K L
    P Q
V W X Y Z
```

Since I goes to L, it's likely that I is on the same row as L and that R is in the same column as L. This means there is only one place for R, It also strongly insinuates one place for I, since I/J go together:

```
      R
A B   F G
  I K L
    P Q
V W X Y Z
```

We also see that M, N, O fill out perfectly in between L and P, so none of these letters are in the key word:

```
      R
A B   F G
  I K L M
N O P Q
V W X Y Z
```

Two of S, T and U are in the keyword. Since S and T are very frequent, let's guess that U is NOT in the keyword, and H also falls into place:

```
      R
A B   F G
H I K L M
N O P Q U
V W X Y Z
```

Finally, we have that S, and T are in the keyword along with two of the following letters: C, D, E. A bit of trial and error (based on the cipher text) will lead to the following arrangement:

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

which uses the keyword "secret."

4.2 Hill Cipher

The Hill cipher uses matrix multiplication, mod 26. In particular, the encryption key is an nxn matrix with an inverse mod 26, where n is the block size. For our purposes, we will illustrate the cipher with n=2. Consider the following key:

$$\begin{pmatrix} 3 & 1 \\ 6 & 5 \end{pmatrix}$$

To encrypt a plaintext, group the plaintext in pairs: "MA" and "TH", for example. Convert each letter to its numerical equivalent, mod 26, and write it in a nx1 matrix as follows:

$\begin{pmatrix} 12 \\ 0 \end{pmatrix}$ stands for "MA"

Now, multiply the encryption key by the plaintext and reduce mod 26 to get the ciphertext:

$\begin{pmatrix} 3 & 1 \\ 6 & 5 \end{pmatrix}\begin{pmatrix} 12 \\ 0 \end{pmatrix} = \begin{pmatrix} 36 \\ 72 \end{pmatrix} = \begin{pmatrix} 10 \\ 20 \end{pmatrix} \bmod 26$, which corresponds to the ciphertext KU.

Here is the encryption of "TH":

$\begin{pmatrix} 3 & 1 \\ 6 & 5 \end{pmatrix}\begin{pmatrix} 19 \\ 7 \end{pmatrix} = \begin{pmatrix} 64 \\ 149 \end{pmatrix} \equiv \begin{pmatrix} 12 \\ 19 \end{pmatrix} \bmod 26$, which corresponds to the ciphertext MT.

To decrypt, you need the inverse matrix. Although it's beyond the scope of this class, here's the derivation of the inverse matrix:

$$\begin{pmatrix} 3 & 1 \\ 6 & 5 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \bmod 26$$

This yields the equations

$$\begin{array}{ll}
3a + c \equiv 1 \bmod 26 & 3b + d \equiv 0 \bmod 26 \\
\quad 6a + 5c \equiv 0 \bmod 26 & 6b + 5d \equiv 1 \bmod 26 \\
-(6a + 2c \equiv 2 \bmod 26) & -(6b + 2d) \equiv 0 \bmod 26 \\
\text{-------------------------} & \text{--------------------------} \\
\qquad 3c \equiv 24 \bmod 26 & \qquad 3d \equiv 1 \bmod 26 \\
\qquad\; c \equiv 8 \bmod 26 & \quad 9(3d) \equiv 9 \bmod 26 \\
 & \qquad\; d \equiv 9 \bmod 26 \\
3a + c \equiv 1 \bmod 26 & 3b + d \equiv 0 \bmod 26 \\
3a + 8 \equiv 1 \bmod 26 & 3b + 9 \equiv 0 \bmod 26 \\
3a \equiv 19 \bmod 26 & 3b \equiv 17 \bmod 26 \\
9(3a) \equiv 9(19) \bmod 26 & 9(3b) \equiv 9(17) \bmod 26 \\
a \equiv 171 \bmod 26 & b \equiv 153 \bmod 26 \\
\quad \equiv 15 \bmod 26 & \quad \equiv 23 \bmod 26
\end{array}$$

Thus, the desired inverse matrix is $\begin{pmatrix} 15 & 23 \\ 8 & 9 \end{pmatrix}$.

Now, we can corroborate that this is the case by decrypting the example above.

$$\begin{pmatrix} 15 & 23 \\ 8 & 9 \end{pmatrix}\begin{pmatrix} 10 \\ 20 \end{pmatrix} \equiv \begin{pmatrix} 610 \\ 260 \end{pmatrix} \equiv \begin{pmatrix} 12 \\ 0 \end{pmatrix} \bmod 26$$
$$\begin{pmatrix} 15 & 23 \\ 8 & 9 \end{pmatrix}\begin{pmatrix} 12 \\ 19 \end{pmatrix} \equiv \begin{pmatrix} 617 \\ 267 \end{pmatrix} \equiv \begin{pmatrix} 19 \\ 7 \end{pmatrix} \bmod 26$$

We can also verify this by multiplying both matrices in question together:

$$\begin{pmatrix} 3 & 1 \\ 6 & 5 \end{pmatrix}\begin{pmatrix} 15 & 23 \\ 8 & 9 \end{pmatrix} \equiv \begin{pmatrix} 53 & 78 \\ 130 & 183 \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \bmod 26, \text{ as desired.}$$

Using a known plaintext attack, we can break the Hill cipher with n blocks (of n letters) of matching plain-ciphertext pairs.

Consider the following:

Let's say we know that "MA" encrypts to "KU" and that "TH" encrypts to "MT".

Let the unknown key be $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Then we can set up the following equations:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 12 \\ 0 \end{pmatrix} \equiv \begin{pmatrix} 10 \\ 20 \end{pmatrix} \bmod 26$$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 19 \\ 7 \end{pmatrix} \equiv \begin{pmatrix} 12 \\ 19 \end{pmatrix} \bmod 26$$

12a ≡ 10 (mod 26)
12c ≡ 20 (mod 26)

19a + 7b ≡ 12 (mod 26)
19c + 7d ≡ 19 (mod 26)

6a ≡ 5 (mod 13), based on a derivation shown in the mod notes.
6c ≡ 10 (mod 13)

11(6a) ≡ 11(5) (mod 13), 11 is the desired inverse
11(6c) ≡ 11(10) (mod 13)

a ≡ 55 (mod 13)
≡ 3 (mod 13)
c ≡ 110 (mod 13)
≡ 6 (mod 13)

So, a could be either 3 or 16 (mod 26).
and c could be either 6 or 19 (mod 26).

19a + 7b ≡ 12 (mod 26)
19(3) + 7b ≡ 12 (mod 26)
7b ≡ -45 (mod 26)
7b ≡ 7 (mod 26), so b = 1 (mod 26)

19c + 7d = 19 (mod 26)
19(6) + 7d = 19 (mod 26)
7d = -95 (mod 26)
7d = 9 (mod 26)
15(7d) = 15(9) (mod 26)
d = 135 (mod 26)
d = 5 (mod 26)

Plugging in a = 16 (mod 26) yields the solution b = 14 (mod 26).
Similarly, plugging in c = 19 (mod 26) yields the solution d = 18 (mod 26).

Thus, using the given information, we've narrowed down the possible keys to these four:

$$\begin{pmatrix} 3 & 1 \\ 6 & 5 \end{pmatrix}, \begin{pmatrix} 3 & 1 \\ 19 & 18 \end{pmatrix}, \begin{pmatrix} 16 & 14 \\ 6 & 5 \end{pmatrix}, \begin{pmatrix} 16 & 14 \\ 19 & 18 \end{pmatrix}$$

Trying each one out on the rest of the ciphertext will lead to the correct one. (Technically, you'd have to first invert it, and then apply that inverted matrix to the rest of the ciphertext.)