Encryption and Decryption

One of the most expensive, most secret, and most mathematical organizations within our government is the NSA—the National Security Agency—the code people. They, together with ordinary businesses, spend billions of dollars annually studying how to send messages securely and how to intercept and decipher other people's messages. If you have used the internet, you have used encryption. This is big business by any standard.

The coding method we study in this course would be regarded as a joke by coding specialists. Even without a computer it can be broken quickly. Nevertheless it illustrates the basic steps in most encryption schemes. These steps are as follows:

(1) The source text is converted to a sequence of numbers.
(2) A mathematical process is applied to this sequence of numbers to scramble and redistribute the information so that any piece of the new, scrambled sequence contains information blended together from different parts of the message. (In doing this, the hope is to conceal any regularities of language that might be used by code-breakers to detect the encryption scheme and then reverse the scheme to arrive at the source text.)
(3) (optional) The number sequence may be converted back into text, which is now expected to be meaningless to anyone who does not have the key. (We do not perform this optional step in our example of coding.)

In our coding scheme, Step 1 requires a table that will always be given to you on a test. The table establishes a one-to-one correspondence between the alphabet plus a few punctuation marks plus the "space" character on one hand, and numbers on the other hand. The numbers might be ≥ 0 or they might include negatives. Here are two examples of such tables:

Example 1:

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
0	1	−1	2	−2	3	−3	4	−4	5	−5	6	−6	7	−7	8	−8	9

S	T	U	V	W	X	Y	Z	blank	$	,	.	!	?
−9	10	−10	11	−11	12	−12	13	−13	14	−14	15	−15	16

Example 2:

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

S	T	U	V	W	X	Y	Z	blank	!	?	#	,	.	:	;
18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33

Using Example 2, the message $\boxed{\text{HELLO, THERE!}}$ would be converted to the number sequence $\boxed{7,4,11,11,14,30,26,19,7,4,17,4,27}$.

In Step 2 of our coding scheme, the number sequence is rearranged into an n-by-2 matrix. This implies that we expect the number sequence to have an even number of terms, $2n$ in fact. Unfortunately, this is not always the case. In fact our "HELLO, THERE!" example has only 17 characters. In such cases we must "pad" the number sequence with an extra character to fill out the matrix. We will always use the blank for this purpose, so in our example we would add a 26 (representing the blank) to the end of the number sequence: $\boxed{7,4,11,11,14,30,26,19,7,4,17,4,27,26}$. We then rearrange the number sequence into a 7-by-2 matrix, with each row composed of an adjacent pair of numbers in the sequence, so that the first row of the matrix is the first two numbers of the sequence, the second row is the next two numbers, and so on, thus:

$$C = \begin{bmatrix} 7 & 4 \\ 11 & 11 \\ 14 & 30 \\ 26 & 19 \\ 7 & 4 \\ 17 & 4 \\ 27 & 26 \end{bmatrix}$$

Now we are ready for the "scrambling" step. We will always do this by multiplying the n-by-2 matrix on the right by a 2-by-2 integer matrix. For example, suppose we used the matrix

$$M = \begin{bmatrix} 2 & -5 \\ -3 & 7 \end{bmatrix}$$

for this purpose. We would obtain the "scrambled" matrix

$$S = CM = \begin{bmatrix} 2 & -7 \\ -11 & 22 \\ -62 & 140 \\ -5 & 3 \\ 2 & -7 \\ 22 & -57 \\ -24 & 47 \end{bmatrix}$$

Finally, we would reverse the sequence-to-matrix process to obtain the number sequence Code= $\boxed{2, -7, -11, 22, -62, 140, -5, 3, 2, -7, 22, -57, -24, 47}$. This is our coded version of the original "HELLO, THERE!" message.

To decode this numerical message, you would simply reverse the process. (Note: this is the part you most need to pay attention to because this is what you will be expected to do on the final exam.)

You would be given a numerical code sequence like Code in the box above, a 2-by-2 encoding matrix like M, and a letter-to-number table such as the one in Example 2. You would reconstruct the $n - by - 2$ "scrambled" matrix S by taking the numbers from the Code sequence two at a time to form the rows. You would then find the inverse M^{-1} of the matrix M. In this case it is

$$M^{-1} = \begin{bmatrix} -7 & -5 \\ -3 & -2 \end{bmatrix}.$$

You would multiply the scrambled matrix S by M^{-1} to get

$$C = SM^{-1} = \begin{bmatrix} 2 & -7 \\ -11 & 22 \\ -62 & 140 \\ -5 & 3 \\ 2 & -7 \\ 22 & -57 \\ -24 & 47 \end{bmatrix} \begin{bmatrix} -7 & -5 \\ -3 & -2 \end{bmatrix} = \begin{bmatrix} 7 & 4 \\ 11 & 11 \\ 14 & 30 \\ 26 & 19 \\ 7 & 4 \\ 17 & 4 \\ 27 & 26 \end{bmatrix}.$$

We would then unpack this matrix into the number sequence

$$\boxed{7,4,11,11,14,30,26,19,7,4,17,4,27,26}.$$

Finally, we would use the table to convert these numbers back to readable text: HELLO, THERE!

If, during this process, you get something that does not make sense to you, it is probably a signal that you made a mistake at some point. An example would be if your matrix C had numbers in it that did not correspond to numbers in the letter-to-number table. If this happens, check that you computed the inverse matrix correctly, including the correct value of the determinant. If that is OK, check that you computed the product correctly. If you used a calculator to do this, make sure you entered the matrix correctly.

Real encryption schemes differ from this one in several respects. In our scheme just two characters are combined; in a real scheme, many characters are combined. In our scheme, the combining is done by matrix multiplication; in a real scheme, the process would be more complex. However, like ours, it would have to be something that could be easily reversed once one knows the *decryption key*. (In our case, this is the matrix M^{-1}.)