Grade 6 Math Circles

November 14/15/16

Encryption

Alphabet Soup

What do you do when you want to send your friend a message, but don't want anyone else to be able to read it? There are many different things that you can do. For example

1. Translate the message into a different language that you and your friend know, but the people around you might not.

   While this might sound great, there might be some people around that speak your language without you knowing!

2. Create a new language that only you and your friend know.

   Learning a new language is hard! So this won't be an easy thing to do.

3. Use a secret code that only you and your friend know to transform your message into something unreadable.

This last strategy is the one that we will focus on today. Using a secret code to transform a message is called **encryption**. Transforming a message using the secret code is known as **encoding**, and turning a message that's already been transformed back into readable text is known as **decoding**.

Exercise 1: James' First Message

Below I've encoded part of a message and I only want you to be able to read it. Here's my secret code: *every letter in the alphabet has been shifted over by one.* Here's my message:

J XJMM UFMM ZPV XIBU UIF OFYU TFTTJPO JT BCPVU

Using my code, figure out what my message is.

Did you figure out the message on your first try? Notice that my code did not tell you which direction the letters were shifted, only that the letters were all shifted over by one. But even with incomplete information, you were probably able to successfully decode my message.

The code used in Example 1 in known as a **shift cipher**, because all of the letters in the alphabet

were shifted over by one. We don't have to only shift by one, but we can instead shift by any number. The important result of encoding is that the words are unrecognizable, and look something like what you would find in a can of alphabet soup.

Exercise 2: Shift Ciphers

Write a short message and encode it using a shift cipher. Trade messages with a friend and try to decode their message. Make sure you give each other your secret codes so you can successfully decode each other's messages.

Decoding Messages

Remember the goal of encryption: using a secret code to change your message into something unreadable, so only the people with the code can figure out the message. Just because someone doesn't know the secret code, doesn't mean they cannot figure it out. While shift ciphers are very easy to use, they are also very easy to decode without the secret code. Since there are 26 letters in the alphabet, there are only 25 possible codes to try (note that shifting letters over by 26 is the same as not shifting at all). With some extra techniques, you can figure out someone's secret code with even fewer tries!

Exercise 3: James' Second Message

Without knowing how far each letter has been shifted over, see if you can figure out the message below:

ZNK TKDZ YKYYOUT UL ZNK MXGJK YOD SGZN IOXIRKY CUXQYNUV OY UT ZNK IUTIKVZ UL

Hopefully you were able to decode the message, but it might have taken you quite a few tries to do so! There are some strategies that we can use to help us decode ciphers.

One of the most helpful things to know when trying to decode a message is how often different words and letters tend to appear in everyday speech. Here's a few useful lists.

Most Common Letters	Most Common Two Letter Combinations	Most Common Three Letter Combinations
E	TH	THE
T	ER	AND
A	ON	THA
O	AN	ENT
I	RE	ION
N	HE	TIO

Notice that the most common three letter word is 'THE' so if we see a lot of the same three letter words in the encoded text, it would be a reasonable guess that this word corresponds to 'THE'. Looking back at James' Second Message, we can see that the three letter word 'ZNK' appears quite frequently, so let's see if this can help us.

'ZNK' corresponding to 'THE' means that 'Z' corresponds to 'T'. To go from 'Z' to 'T', we shift backwards by 6 letters. Shifting the 'N' and 'K' backwards by 6 letters as well gives us 'H' and 'E' respectively. We can now guess that the code used here was shifting by six, and we solved this using without having to try other shifts first! This strategy for decoding messages is called **frequency analysis**.

Exercise 4: Frequency Analysis

Try to decode the following message while using frequency analysis:

N CVT NAQ N PBJ JRAG GBTRGURE GB N SNEZ NAQ N ZNEXRG

Remember to look for common letters and words.

Rail Fence Ciphers

Now that we've seen how easy it is to decode shift ciphers without knowing the code, let's move on to something that may be a little trickier to decode, and so our encoded messages will be safer.

Example 1: The Rail Fence Cipher

Consider the following text:

I like building fences

We will encode our message using the **rail fence cipher** (or simply the **rail cipher**). Here's how it works:

1. First, choose how many rails you want to use. Note that this number has to be less than the number of letters in your message. In this example, we'll choose the number of rails to be $R = 4$.
2. Moving diagonally, one letter at a time place the letters on the rails.

I
l
i
k

When you run out of rails, start going back up diagonally.

I u
l b
i e
k

Keep going up and down until you run out of letters.

I u g s
l b i n f e
i e l i e c
k d n

3. Now write down the letters one row at a time from left to right to get

Iugslbinfeielieckdn

Now this really looks like alphabet soup! This text is our encoded message.

The secret code in a rail fence cipher is the number of rails used, R. Even if we have the code, decoding the message is even trickier here. Notice that in the zig-zag pattern in example 1, the vertical positions repeat every 6 letters. That is, every 6^{th} letter will be in the top position. So the first few letters will all have 5 spaces in between them in the original word. Let's see this in action.

Example 2: Decoding the Rail Fence Cipher

Consider the blank spaces representing the original message.

_ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _

We know that the first few letters in the encoded message should have 5 spaces between them. So we have

I _ _ _ _ _ u _ _ _ _ _ g _ _ _ _ _ s

This matches our original text so far! Now that we have this outline, looking back at the original zig-zag, we can place the next letters immediately around the already placed letters.

I l _ _ _ b u _ _ _ _ _ g _ _ _ _ _ s

and then

I l _ _ _ b u i _ _ _ n g f _ _ _ e s

and

I l i _ e b u i l _ i n g f e _ c e s

and finally

I l i k e b u i l d i n g f e n c e s

which is exactly the message we started with.

Notice that the number of rails R was 4 in this example, and the number of spaces between the first set of letters was

$$2(R-1)-1$$

In our case, we had $2(R-1)-1 = 2(4-1)-1 = 2(3)-1 = 5$, which matches what we discovered above.

Exercise 5: Encoding Rail Ciphers

Encode the following message using the rail cipher and secret code $R = 4$

Ciphers are fun

Also encode the following message using the rail cipher and secret code $R = 5$

Math circles is a great workshop

Exercise 6: Decoding Rail Ciphers

Decode the following message using the rail cipher with code $R = 3$

horeeleeynlvo

And decode the following message using the rail cipher with code $R = 5$

pmirusyrturifueopeaocllvo

Notice that the method of frequency analysis does not work here since the number of each letter doesn't change after we encode our message, we're only scrambling the message. Because of this, the rail cipher can be much safer to use since it can be much harder to decode the message.

Recall that there were at most 24 ways to shift our letters to get the code in a shift cipher. In the rail cipher, we can have at most as many rails as there are letters in the message. So if our message has 1000 letters, we can choose any number between 1 and 1000 to be our secret code. For example, we can choose $R = 213$ to be the number of rails if our message has 1000 letters. It would be much harder to figure out this code than it would for any shift cipher!

Stop and Think

Not every value of R is a good value. Once the number of rails is greater than half of the number letters, the rail cipher becomes easier to decode. The closer R is to the number of letters, the easier decoding it becomes. Think about why this might be.

Exercise 7: James' Third Message

Without knowing the secret code, see if you can decode the following message using the rail cipher:

meacstir

The Importance of Encryption in Real Life

While the rail fence cipher is better than the shift cipher, it is not good enough for the modern world. A modern computer can decode rail fence ciphers in seconds, even for very long messages!

Every message that is sent over the internet is encrypted, and it is very important that these messages don't get decoded by people who aren't supposed to see the message. For example, if you buy something online, your banking information is sent to the website that you're buying from. You **do not** want strangers to be able to decode your banking information, otherwise they can spend your money! Modern problems require modern solutions. So we will consider one more method of encrypting our messages.

RSA Encryption

RSA encryption is the most commonly used encryption method in the modern world for communication. Every computer uses it for almost every single message that it sends. Before we see how it works, let's review some math concepts.

Recall

A *prime number* is a whole number, greater than or equal to 2, that is only divisible by 1 and itself.

Example 3

2, 3, 5, 7, 11, 13, and 17 are all examples of prime numbers.
6 and 9 are not a prime numbers since they are divisible by 3.

Recall

A *multiple* of a number is the product of the number with an integer.

Example 4

Some multiples of 5 are 5, 10, 15, 20, 25, and 30.
Some multiples of 3 are 3, 6, 9, 12, 15, 18, and 21.

Recall

The *lowest common multiple* of two numbers is the smallest positive multiple that is shared by both numbers.

Example 5

The lowest common multiple of 3 and 5 is 15. The lowest common multiple of 6 and 8 is 24.

Now, we are ready to see RSA encryption.

1. To begin, make sure you message is a number. For this example, we'll use $m = 2$ as our message.
2. Pick two prime numbers p and q. Let's use $p = 5$ and $q = 7$.
3. Compute $n = p \times q$. We have $n = 5 \times 7 = 35$.
4. Find the lowest common multiple, L, of $p - 1$ and $q - 1$. We have $p - 1 = 4$ and $q - 1 = 6$. Notice that 12 is the lowest common multiple of 4 and 6, so $L = 12$.
5. Choose some number e that can't be divided by any of the same prime numbers as L. e has to be smaller than L. For example, we can choose $e = 5$, but we can't choose $e = 10$ since 2 divides both 10 and $L = 12$.
6. Now we can encrypt our message. Calculate m^e where m was our message from the first step and e is the number we chose in part 5. We have $m^e = 2^5 = 32$. If we get a number larger than n, then we must perform long division of $n \div 32$ and take the remainder. This is our encoded message. That is, if our original message is 2, then 32 is the encrypted message!

Example 6: RSA Encoding

1. Let's encode another message using RSA encryption. My message is $m = 5$.
2. I pick the prime numbers $p = 11$ and $q = 5$.
3. Then $n = 11 \times 5 = 55$.
4. We want to find the lowest common multiple of $p - 1 = 10$ and $q - 1 = 4$. We find that the lowest common multiple is $L = 20$.
5. Notice that 3 can't be divided by any of the same numbers that divide $L = 20$, so let's pick $e = 3$.
6. Finally, $m^e = 5^3 = 125$. This number is bigger than $n = 55$, so we perform long division and we see that the remainder of $125 \div 55 = 15$. So we take 15 to be our encoded message!

Exercise 8: RSA Encoding

Using $p = 13$ and $q = 3$, encode the message $m = 4$ using RSA encryption.

In order to decode a message, we need the secret code. However, the secret code in RSA requires math that is beyond what we will cover today. Just like the shift and rail fence ciphers, we don't need the secret code the decode the message. But, unlike the shift and rail fence ciphers, decoding the message in RSA encryption without knowing the secret code is very hard. In fact, it is so difficult that the whole world feels safe enough to use RSA encryption in almost everything that we do over the internet.

When we pick small messages and small prime numbers like we've done here, it wouldn't take too long for a computer to decode a message. In practice, the prime numbers that are chosen are hundreds of digits long. Trying to decode the message when the prime numbers are this long could take longer than the universe has been alive for! That seems pretty safe to me.