A First Look at Encryption

Think back to the first time a teacher caught you passing a note in class: Wouldn’t it have been cool if that note looked like complete nonsense to your teacher but made sense to you and your friend?

The purpose of encryption is to allow people to communicate securely over some medium. Without secure cryptosystems (RSA, for example), we wouldn’t be able to purchase goods on the web without the fear of someone stealing your credit card information.

Suppose that Tim and Peet want to send secret messages to each other. They decide to replace letters with their order in the alphabet (that is, $A = 1$, $B = 2$, etc.). Note that $Z$ can be either 26 or 0.

They agree on the following encryption and decryption functions:

\[
\begin{align*}
\text{en}(x) &= \text{rem}(3x, 26) \\
\text{de}(m) &= \text{rem}(9m, 26)
\end{align*}
\]

Task 1

a. Tim sends Peet their next meeting place with the encrypted message $(22, 10, 11, 8, 19)$. Where are they meeting? You can use a calculator such as Wolfram Alpha.

**Solution:**

```
PLUTO
\text{rem}(9*22, 26) = 16
\text{rem}(9*10, 26) = 12
\text{rem}(9*11, 26) = 21
\text{rem}(9*8, 26) = 20
\text{rem}(9*19, 26) = 15
```

b. Now, you try encrypting a four-letter word of your choosing and then decrypt it to see if it’s the same message.

**Solution:**

Almost any four-letter words the students pick is fine, as long as their work...
looks OK and they seem to understand the process.

c. Why does Tim and Peet’s encryption scheme work for this 26 character alphabet, that is, why can any encrypted digit be recovered exactly by the decryption process?

   *Hint:* What is the relationship between 3, 9, and 26?

   **Solution:**

   First, we note that 3 has a multiplicative inverse mod 26 (specifically, 9). Therefore, rem(3x, 26) has a distinct answer for each x in [0, 25]. The map $x \rightarrow 3x \pmod{26}$ is a bijection with inverse $x \rightarrow 9x \pmod{26}$.

   Since there are 26 letters in the alphabet, we can map every letter to a unique value, allowing any encrypted digit be recovered exactly.

   *Optional:* Would this scheme work if the modulus was 23? What about 29? Assume you can change the other constants in the encryption and decryption functions.

   **Solution:**

   It would not work if the modulus was 23 because the range of the function would be reduced to [0, 22]. We’ll lose some of the letters towards the end of the alphabet.

   However, it *would* work with $m = 29$—in fact, we’ll be able to support 3 extra characters to take on the values 27, 28, and 29.

---

**RSA Encryption**

Tim and Peet realize that their encryption scheme, is nicely simple, but is altogether too insecure. They opt instead to use RSA encryption to send notes to each other in class. Recall that the RSA encryption algorithm works as follows:

1. Choose two primes $p, q$.
2. Calculate $n = pq$. We can publish $n$ publicly while keeping $p, q$ private because factorization is a hard problem.
3. Calculate $\phi(n) = (p - 1)(q - 1)$.
4. Choose some $1 < e < \phi(n)$ such that $\gcd(e, \phi(n)) = 1$. 


5. Find a multiplicative inverse \( d \) for \( e \) such that \( de \equiv 1 \pmod{\phi(n)} \). A multiplicative inverse has to exist because \( e \) is defined to be relatively prime to the modulus.

6. Publish \( n \) (the modulus) and \( e \) (the encryption exponent).

7. Keep all other numbers private to yourself, but remember \( d \): It will be your decryption exponent.

To encrypt a message \( m \), compute \( m^* = \text{rem}(m^e, n) \). To decrypt an encrypted message \( m^* \), calculate \( \text{rem}((m^*)^d, n) \).

**Task 2**

a. Now, create your group’s own personal RSA key by choosing large prime values of \( p \) and \( q \) with 10 digits—use an online list of primes as a reference. (We want a larger modulus so we can send longer messages, but not so long that calculators can’t handle it.) Write \( p, q, n \), and \( \phi(n) \) down here:

**Solution:**

TAs - make sure that \( p \) and \( q \) are large! The modulus has to be greater than any message we want to send, which we restricted on our end to be 18 digits or less.

b. Now, find a pair of multiplicative inverses \( e \) and \( d \) modulus \( \phi(n) \). You can use any online calculators/generators (we recommend Wolfram Alpha) to help you with things like prime factorization or solving congruencies.

**Solution:**

TAs - just ask how they did this part to make sure they understand what’s going on. One way of finding an \( e \) is just prime factorizing \( \phi(n) \) and making a new number that doesn’t share any of those numbers. Then, plug \( de \equiv 1 \pmod{\phi(n)} \) into Wolfram Alpha to solve for \( d \). In class, we pointed out that picking a random number in the right range and running Euclid will work quickly and accurately.

c. Now, “publish” your \( n \) and \( e \) by posting them on this EdStem post. Your TAs will send you an encrypted message via a follow-up to your comment, which you’ll have to decrypt to get checked off at the end of this section.

**Optional Checkpoint** - Call over a TA if queue is short!
Encryption and Privacy

In the wake of the deadly December 2015 terrorist attack in San Bernardino, California, the FBI asked Apple to develop software that would allow the FBI to access their devices without use of a password. When Apple refused to comply (optional reading), it sparked massive debate about encryption and a user’s right to privacy. Read this article about the FBI-Apple encryption dispute regarding FBI access to the iPhone 5C of one of the shooters. In the end, the FBI found a third party to unlock the iPhone for them, resulting in the case being dropped (if you’re interested, you can read an updated news article here).

Task 3

a. Give two reasons why Apple should have complied with the FBI request (that is, allowed the FBI to enter its devices), and two reasons why Apple shouldn’t have complied with the request (that is, refuse to develop an FBI backdoor to their devices).

Solution:

Pass condition: Students must be able to demonstrate a reasonable understanding of both sides of the multifaceted issue, and they should be able to give some justification as to why one side was right or wrong. They do not have to take a stance personally, but they are not necessarily required to remain neutral either.

Examples: Potential reasons why Apple should have complied with the FBI request:

- Public safety should come before Apple’s stance; the government should not ”roll over” simply because a private company does not wish to comply with an investigation.
- Finding motives, connections, and additional information about the San Bernardino shooters takes precedent. There is no guarantee that the government will use this new software any time in the future. In this moment, this investigation on terrorism is priority.
- Since police are allowed to search houses with a warrant, police should be able to search a phone with a warrant.

Potential reasons why Apple should not have complied:

- The user has a right to privacy over data on their phone, and forcing Apple to unlock the phone sets a dangerous precedent for future administrations.
• Developing any sort of intentional backdoor would be an easy target for any malicious hackers.

• Due to the nature of policing in America, it is far more likely that minority populations will be targeted by these laws. This is especially true when dealing with terrorism and foreign threats, where certain politicians like to revert to reductionist, xenophobic attitudes.

b. The data-encryption debate centers on one of the oldest questions about governance: Does public safety outweigh an individual’s right to privacy? Consider the rise of privacy-conscious technologies like cryptocurrencies, the Tor browser, and private-use VPNs. These technologies increase privacy for the user by encrypting and obfuscating data, but they can make it easier for malicious actors to cover their tracks. Should these technologies be regulated by the government, or is the increasingly privacy-aware public a good thing?

Solution:

*Pass condition:* Students must be able to demonstrate an understanding of both sides of the argument, and provide ample reason to support their perspective. There is no correct answer, but students should be able to recognize the potential consequences of choosing one stance and fully committing to it. Additionally, students should be able to explain how technology, in particular, makes this debate more complicated.

*Prodding questions:*

• Should the government choose public safety over the right to privacy (or vice versa) every time? In other words, should there be exceptions to each stance?

• Was this debate as complicated when there was less technology, before the internet existed? Or does the same debate manifest itself in different ways throughout history?

• Would being more aware of what personal data companies have on you impact your activity on the internet or what information you willingly give out?

c. Consider your role as a potential computer scientist or mathematician. To what extent are you responsible for protecting and encrypting user data? Is it worth collecting more data in some cases to develop a more robust, potentially insightful machine-learning program or is a user’s right to privacy more important? Where does the balance between the two lie?
Solution:

*Pass condition:* Students should be able to explain why protecting user data is an important consideration, and how they plan to be mindful of such concerns. No matter which stance they take, students should be able to justify it using their own moral guides. Students should also be able to suggest a balance between creating more accurate machine learning programs or protecting users’ privacy, and where they believe they would draw the line.

*Prodding questions:*

- Think about this situation as if you are a user. Would you prefer for there to be programs (such as targeted advertising) that result in tailored experience for you, or for your data to be entirely protected? What is your limit?
- How would you protect user data as a mathematician or computer scientist?
- What is considered sensitive user data and should not be collected for any purposes?

Lastly fill [this anonymous mid-semester Feedback form](#) for the Socially Responsible Computing (SRC) component of the course.

**Checkpoint 1 - Call a TA over!**
The Product Rule

Given finite sets $S_1, S_2, ..., S_n$, the product rule tells us that

$$|S_1 \times S_2 \times ... \times S_n| = |S_1| \cdot |S_2| \cdot \ldots \cdot |S_n|.$$ 

This rule is often useful when we are doing counting and what we are counting comes from some number of independent choices. For instance, when picking an outfit for the day, say you have 4 shirts, 3 pants, and 2 pairs of shoes to choose from. We can think of an outfit as a sequence (shirt, pant, shoe), so to find the total number of outfits, we multiply the number of choices we can make for each position, $4 \cdot 3 \cdot 2 = 24$.

Even if our choices depend on each other, we can still sometimes use this concept. For instance, let’s say that for each of our 4 shirts, one of our 3 pairs of pants looks terrible with it, so we don’t want to wear those pants if we’re wearing that shirt. Then, we still have 4 choices for our shirt, but having made that choice, we now only have 2 pants to choose from. We still have 2 shoe options. So, our total number of outfit choices is $4 \cdot 2 \cdot 2 = 16$.

**Task 4**

a. Suppose we go to the sandwich shop, and we want to order a combo special. The combo special includes a sandwich, and side, and a drink.

There are 5 different kinds of sandwiches, 6 different kinds of sides, and 8 different kinds of drinks. How many different ways could we order a combo special, and why?

**Solution:**

240. We let $A$ be the set of sandwiches, $B$ be the set of sides, and $C$ be the set of drinks. Then $A \times B \times C$ has all of the different ways we could order, and there are $5 \cdot 6 \cdot 8 = 240$ elements in this Cartesian product.

b. When we first talked about functions, these functions only took in one input. Since then, we’ve seen propositions, which are functions that can take in more than one input! In general, functions can take in one or more inputs.

Note that, if the function takes in more than one input, say $n$ inputs, we can think of it taking in one input, where each input is a tuple of length $n$.

i. Consider a function of 3 inputs, where each input value can be 0, 1, 2, or 3. If we think about this function as a function of one input, how many possible inputs does it have?
Solution:
There are 4 options for each position in the input, so $4^3$.

ii. Consider a function with the same input as described above. Its output for each input is either 0 or 1. How many unique functions are there? 
*Hint:* Two functions are identical if every input leads to the same output.

Solution:
There are $4^3$ inputs, as described above. There are 2 options for what each can map to, so $2^{4^3}$ functions.

iii. Optional: Consider functions of 2 inputs, where each input can take on either 0 or 1 and outputs either 0 or 1, but the order of the inputs does not affect that output of the function. For example, $f(x, y)$ could be $(x \land y)$. How many possible such functions are there?

Solution:
This time, we consider the possible input sequences: $(0, 0), (0, 1), (1, 0), (1, 1)$. Because $f(0, 1) = f(1, 0)$, once we choose where one maps to, we have to choose the same thing for the other, so we effectively only have a choice for 3 inputs, or $2^3$. 
Permutations and Counting Subsets

A permutation of a set $A$ is an ordered list of the elements of $A$.

**Task 5:** Explain why there are $n!$ different permutations of a set of size $n$.

**Solution:**

$n$ options for the first position, $n - 1$ for the second, etc. gives us $n \cdot (n - 1) \cdot \ldots \cdot 1$.

Let $|A| = n$. Let’s say we want a permutation of $k$ elements of $A$, where $k \leq n$. We could make such a permutation by picking one of $n$ elements for the first position, $n - 1$ elements for the second, etc, getting us $n \cdot (n - 1) \cdot \ldots \cdot (n - k + 1)$ permutations. We can also write this quantity

$$\frac{n!}{(n-k)!}.$$

How could we have gotten that same result in a different way? Well, we could have made all $n!$ permutations of $A$, and then grouped those based on the ordering of the first $k$ elements. Each ordering of $k$ elements has $(n-k)!$ possibilities for the order of the elements that follow it, hence we divide by $(n-k)!$.

Suppose we want to know the number of subsets of size $k$ of a set of size $n$. The general formula for this value, called $\binom{n}{k}$ is

$$\frac{n!}{(n-k)!k!}.$$

Why? First, we can think of the permutations of length $k$ of elements of $A$, which there are $\frac{n!}{(n-k)!}$ of. Then, we can group these with the other permutations that have the same elements but in a different order, and divide our count by the number in each group. How big is each group? For any set of $k$ elements, there are $k!$ ways to order them. So, we divide by $k!$.

Try it out with a small example set to make sure it makes sense to you—perhaps with the number of ways to pick 2 astronauts out of a crew of 9?

**Task 6:** Count the size of each set, and sort them from largest to smallest.

a. The number of permutations of all the letters in the alphabet.

b. The number of subsets of size 6 of the letters of the English alphabet (one such subset is $\{a, b, c, d, e, f\}$).
c. The number of 6 letter words made of non-repeating letters (one such 6-letter word is “abcdef”).

d. The number of (any sized) subsets of the set of the letters in the English alphabet.

e. The number of subsets of size 20 of the letters of the alphabet.

Solution:
1. \(26!\)
2. \(\binom{26}{6} = \frac{26!}{20!6!}\)
3. \(\frac{26!}{20!}\)
4. \(2^{26}\)
5. \(\binom{26}{20} = \frac{26!}{20!6!}\)

1, 3, 4, 2 = 5

Task 7: Consider a string of size \(n \geq 2\), \(S = s_1s_2...s_n\), where each \(s_i\) is a 1–9 digit. For each of the following conditions, find the number \(N\) of such strings that satisfy the condition, and prove your answer.

a. Only \(x\) types of digits are used, where \(1 \leq x \leq 9\).

Solution:
There are \(n\) digits, each of which has \(x\) options, meaning there are \(x^n\) possibilities.

b. No two consecutive digits are the same.

Solution:
The first digit has 9 options. The remaining digits have 8 options since they cannot match the preceding digits. Thus, there are \(9 \cdot 8^{n-1}\).

c. Optional: The sum of any \(k\) consecutive digits is the same, where \(1 \leq k \leq n\).

Solution:
Consider some arbitrary sequence of \(k\) digits of the number \(s_is_{i+1}...s_{i+k-1}\) with sum \(r\). Now, if we move on to the next \(k\) digits with \(s_2\) at the front, we seek
the sum of $s_{i+1}s_{i+2}...s_{i+k}$. We know that each $k$ group must have the same sum, $r$, so

$$s_{i+1} + s_{i+2} + ... + s_{i+k} = r$$

But we know from the first $k$ digits that

$$r - s_i + s_{i+k} = r$$

which implies that $s_{k+i} = s_i$. This means that, for all $k$-sequences to have the same sum, a particular $k$ sequence must be cycled through in order. This implies that counting the number of $k$ sequences possible is sufficient. Since there are $k$ digits to choose, there are $9^k$ such options.

d. Optional: The product of any set of $k$ consecutive digits is the same, where $1 \leq k \leq n$.

Solution:

The argument is the same as above. The only step that changes is

$$s_{i+k}r/s_i = r$$

which still implies that $s_{k+i} = s_i$. There are similarly $9^k$ options.

Task 8: By now, you should’ve received an encrypted message from your TAs. Use your private decryption exponent $d$ to decrypt and read the message. More information about how to interpret the plaintext message is on the EdStem post.

Solution:

Our default message was 1920011802210311, ”STARBUCK”. More creative TAs may have sent you a different message.

Checkoff - Call over a TA!