CS243: Discrete Structures

More on Cryptography and Mathematical Induction

Işıl Dillig

Announcements

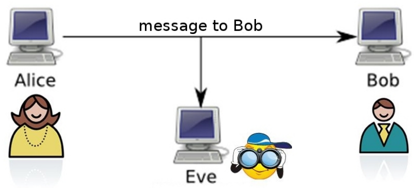
- Class canceled next Thursday – I am out of town
- Homework 4 due Oct 22 instead of next Thursday (Oct 18)

Agenda for Today

- Talk about RSA algorithm for public-key cryptography
- Start discussion of mathematical induction
- Will spend 2 lectures on mathematical induction

Cryptography

- Cryptography is the study of techniques for secure transmission of information in the presence of adversaries

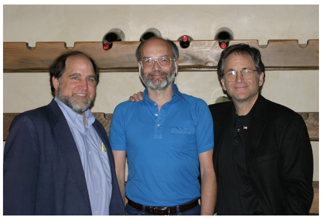


- How can Alice send secrete messages to Bob without Eve being able to read them?

Private vs. Public Crypto Systems

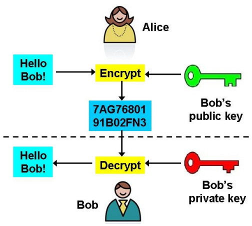
- Two different kinds of cryptography systems:
  1. Private (secret) key cryptography
  2. Public key cryptography
- In private key cryptography, sender and receiver agree on secret key that both use to encrypt/decrypt the message
- In public key crytography, a public key is used to encrypt the message, and private key is used to decrypt the message
- Modern systems use puiblic key crypto, best known public key encryption algorithm is RSA algorithm

RSA History



- Named after its inventors Rivest, Shamir, and Adlemann, all researchers at MIT (1978)
- Actually, similar system invented earlier by British researcher Clifford Cocks, but classified – unknown until 90's

RSA Overview



- Bob has two keys: public and private
- Everyone knows Bob's public key, but only he knows his private key
- Alice encrypts message using Bob's public key
- Bob decrypts message using private key
- Public key can encrypt, but not decrypt
- Therefore, noone can read message accept Bob

High Level Math Behind RSA

- In the RSA system, private key consists of two very large prime numbers p, q
- Public key consists of a number n, which is the product of p, q and another number e
- e is a number relatively prime with $(p-1)(q-1)$ ($\phi(n)$, Euler's totient function, which gives number of integers $\leq n$ and relatively prime with n)
- Encrypt messages using n, e, but to decrypt, must know p, q
- In theory, can extract p, q from n using prime factorization, but this is intractable for very large numbers
- **Security of RSA relies on inherent computational difficulty of prime factorization**

Encryption in RSA

- To send message to Bob, Alice first represents message as a sequence of numbers
- Call this number representing message M
- Alice then uses Bob's public key n, e to perform encryption as:

$$C = M^e \pmod{n}$$

- C is called the ciphertext

Encryption Example

- Encrypt message "STOP" using RSA with $n = 2537,\ e = 13$
- First convert each letter to a number in $[0, 25]$: $S = 18, T = 19, O = 14, P = 15$
- Group sequence into blocks of 4 digits:

$$M = 1819\ 1415$$

- Now encrypt each block as $C = M^{13} \pmod{2537}$
- For first block, $1819^{13} \pmod{2537} = 2081$; for second block $1415^{13} \pmod{2537} = 2182$
- Ciphertext: 2081 2182

RSA Decryption

- How do we decrypt cipher text using private keys p, q?
- Decryption key d is the inverse of e modulo $(p-1)(q-1)$:

$$d \cdot e \equiv 1 \pmod{(p-1)(q-1)}$$

- As we saw earlier, inverse of e mod $(p-1)(q-1)$ can be computed reasonably efficiently if we know $(p-1)(q-1)$
- However, since adversaries do not know p, q, they cannot compute d with reasonable computational effort!

RSA Decryption, cont.

- Using the Chinese remainder theorem and another theorem called Fermat's Little Theorem, it can be shown that:

$$(M^e)^d \equiv M \pmod{n}$$

- Since the ciphertext C is just M^e, $C^d \pmod{n}$ allows decrypting the message
- Since Bob can compute d using p, q, Bob can easily decrypt message, but no one else can!

Decryption Example

- Decrypt the cipher text 0981 0461 for the RSA cipher with $p = 43,\ q = 59$, and $e = 13$.
- First we need to compute d, the inverse of e modulo $(p-1)(q-1)$
- Here, $(p-1)(q-1) = 2436$; thus solve:

$$13x \equiv 1 \pmod{2436}$$

- To solve this, first compute s, t such that:

$$13s + 2436t = 1$$

- Apply extended Euclidian algorithm: $s = 937,\ t = -5$

Example, cont.

Decrypt 0981 0461 using $p = 43,\ q = 59, n = 2537$, and $e = 13$.

- To solve $13x \equiv 1 \pmod{2436}$, computed $s = 937,\ t = -5$
- Recall: Solution to this sytem is given by:

$$x = \frac{sb}{d} + \frac{m}{d}u \ \text{ where } u \in \mathbb{Z}$$

- Here, $s = 937, b, d = 1, m = 2436$, thus solution: $\overline{e} = 937$
- $0981^{937} \pmod{2537} = 0704$; $0461^{937} \pmod{2537} = 1115$
- Thus, decrypted message is 0704 1115, or in English, "HELP"

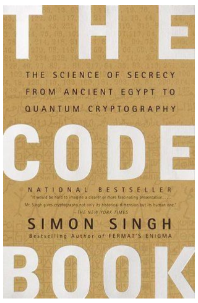
Security of RSA

- The encryption function used in RSA is a trapdoor function
- Trapdoor function is easy to compute in one direction, but very difficult in reverse direction without additional knowledge
- Encryption direction is easy because just requires exponentiation and mod
- Decryption without private key is very hard because requires prime factorization
- Therefore, security of RSA depends on difficulty of prime factorization

Security of RSA, cont.

- However, as computers get more powerful and factorization algorithms better, possible to factor larger and larger integers
- Therefore, over time, necessary to use larger and larger prime numbers to ensure secure communication
- For quantum computing, there are very efficient algorithms for computing prime factors (Shor's algorithm)
- If we could build quantum computers with sufficient "qubits", RSA would no longer be secure!
- However, today, RSA is considered secure if you use sufficiently large prime numbers (> 200 digits)

Book Recommendation



If you are interested in (history of) cryptography, read "The Code Book" by Simon Singh!

Introduction to Mathematical Induction

- Many mathematical theorems assert that a property holds for all natural numbers, odd positive integers, etc.
- Mathematical induction: very important proof technique for proving such universally quantified statements
- Induction will come up over and over again in other classes:
  - algorithms, programming languages, automata theory, ...

Analogy

- Suppose we have an infinite ladder, and we know two things:
  1. We can reach the first rung of the ladder
  2. If we reach a particular rung, then we can also reach the next rung
- From these two facts, can we conclude we can reach every step of the infinite ladder?
- Answer is yes, and mathematical induction allows us to make arguments like this

Mathematical Induction

- Used to prove statements of the form $\forall x \in \mathbb{Z}^+.\ P(x)$
- An inductive proof has two steps:
  1. Base case: Prove that $P(1)$ is true
  2. Inductive step: Prove $\forall n \in \mathbb{Z}^+.\ P(n) \rightarrow P(n+1)$
- Induction says if you can prove (1) and (2), you can conclude:

$$\forall x \in \mathbb{Z}^+.\ P(x)$$

Inductive Hypothesis

- In the inductive step, need to show:

$$\forall n \in \mathbb{Z}^+.\ P(n) \rightarrow P(n+1)$$

- To prove this, we assume $P(n)$ holds, and based on this assumption, prove $P(n+1)$
- The assumption that $P(n)$ holds is called the inductive hypothesis

Example 1

- Prove the following statement by induction:

$$\forall n \in \mathbb{Z}^+.\ \sum_{i=1}^{n} i = \frac{(n)(n+1)}{2}$$

- Base case: $n = 1$. In this case, $\sum_{i=1}^{1} i = 1$ and $\frac{(1)(1+1)}{2} = 1$; thus, the base case holds.
- Inductive step: By the inductive hypothesis, we assume $P(k)$:

$$\sum_{i=1}^{k} i = \frac{(k)(k+1)}{2}$$

- Now, we want to show $P(k+1)$:

$$\sum_{i=1}^{k+1} i = \frac{(k+1)(k+2)}{2}$$

Example 1, cont.

- First, observe:

$$\sum_{i=1}^{k+1} i = \sum_{i=1}^{k} i + (k+1)$$

- By the inductive hypothesis, $\sum_{i=1}^{k} i = \frac{(k)(k+1)}{2}$; thus:

$$\sum_{i=1}^{k+1} i = \frac{(k)(k+1)}{2} + (k+1)$$

- Rewrite left hand side as:

$$\sum_{i=1}^{k+1} i = \frac{k^2+3k+2}{2} = \frac{(k+1)(k+2)}{2}$$

- Since we proved both base case and inductive step, property holds.

Example 2

- Prove the following statement for all non-negative integers n:

$$\sum_{i=0}^{n} 2^i = 2^{n+1} - 1$$

- Since need to show for all $n \geq 0$, base case is $P(0)$, not $P(1)$!
- Base case ($n = 0$): $2^0 = 1 = 2^1 - 1$
- Inductive step:

$$\sum_{i=0}^{k+1} 2^i = \sum_{i=0}^{k} 2^i + 2^{k+1}$$

Example 2, cont.

$$\sum_{i=0}^{k+1} 2^i = \sum_{i=0}^{k} 2^i + 2^{k+1}$$

- By the inductive hypothesis, we have:

$$\sum_{i=0}^{k} 2^i = 2^{k+1} - 1$$

- Therefore:

$$\sum_{i=0}^{k+1} 2^i = 2^{k+1} - 1 + 2^{k+1}$$

- Rewrite as:

$$\sum_{i=0}^{k+1} 2^i = 2 \cdot 2^{k+1} - 1 + = 2^{k+2} - 1$$

Example 3

- Prove that $n < 2^n$ for all positive integers n
- Base case (n=0): $0 < 2^0$
- Inductive step: Need to show $k+1 < 2^{k+1}$ assuming $k < 2^k$
- From the inductive hypothesis, we know $k+1 < 2^k + 1$
- Since $1 \leq 2^k$ for $k \in \mathbb{Z}^+$, this implies $k+1 < 2^k + 2^k$
- Thus, $k+1 < 2 \cdot 2^k = 2^{k+1}$

$\square$

Example 4

- Prove that $2^n < n!$ for all integers $n \geq 4$
- Base case (n=4): $2^4 = 16 < 24 = 4!$
- Inductive step: By the inductive hypothesis, we know $2^k < k!$
- Multiply both sides by 2: $2 \cdot 2^k < 2 \cdot k!$
- Since $k \geq 4$, $2 < k+1$, therefore:

$$2^{k+1} < (k+1)k! = (k+1)!$$

Example 5

- Prove that $3 \mid (n^3 - n)$ for all positive integers n.
- Base case (n=1): $3 | (0 = 1^3 - 1)$
- Inductive step: Show $3 | (k+1)^3 - (k+1)$ assuming $3 | k^3 - k$
- First, rewrite $(k+1)^3 - (k+1)$ as

$$k^3 + 3k^2 + 3k + 1 - k - 1 = (k^3 - k) + 3(k^2 + k)$$

- By the inductive hypothesis, $3 | (k^2 + k)$; and also $3 | 3(k^2 + k)$
- By Thm from earlier, it follows that $3 | (k^3 - k + 3(k^2 + k))$

Example 6

- We can also use induction to prove results about sets.
- Use induction to prove generalized DeMorgan's law for sets:

$$\overline{\bigcap_{j=1}^{n} A_j} = \bigcup_{j=1}^{n} \overline{A_j} \quad \text{for } n \geq 2$$

- We'll prove this by induction on the number of sets
- Base case (n=2):

$$\overline{A_1 \cap A_2} = \overline{A_1} \cup \overline{A_2}$$

- We already proved this earlier!

Example 6, cont.

- Inductive step:

$$\overline{\bigcap_{j=1}^{k+1} A_j} = \overline{(\bigcap_{j=1}^{k} A_j) \cap A_{k+1}}$$

- Now, using De Morgan's law, rewrite left hand side:

$$\overline{\bigcap_{j=1}^{k+1} A_j} = \overline{(\bigcap_{j=1}^{k} A_j)} \cup \overline{A_{k+1}}$$

- From inductive hypothesis, we have $\overline{(\bigcap_{j=1}^{k} A_j)} = \cup_{j=1}^{k} \overline{A_j}$, thus:

$$\overline{\bigcap_{j=1}^{k+1} A_j} = \bigcup_{j=1}^{k} \overline{A_j} \cup \overline{A_{k+1}} = \bigcup_{j=1}^{k+1} \overline{A_j}$$

Correctness of Induction

- Why is induction a valid proof technique?
- Suppose we can prove the base case and inductive step, but $\forall n.P(n)$ does not hold for positive integers.
- There must be a least element k for which $P(k)$ doesn't hold.
- Two possibilities: Either (i) $k = 1$ or (ii) $k \geq 2$
- (i) k cannot be 1 because we proved $P(1)$ in base case
- (ii) Since k is the least element, we know $P(k-1)$ holds
- But, in the inductive step we proved $P(k-1) \rightarrow P(k)$; thus, $P(k)$ must also hold!

Strong Induction

- Slight variation on the inductive proof technique is strong induction
- Regular and strong induction only differ in the inductive step
- Regular induction: assume $P(k)$ holds and prove $P(k+1)$
- Strong induction: assume $P(1), P(2), .., P(k)$; prove $P(k+1)$
- Regular induction and strong induction are equivalent, but strong induction can sometimes make proofs easier

Motivation for Strong Induction

- Prove that if n is an integer greater than 1, then it is either a prime or can be written as the product of primes.
- Let's first try to prove the property using regular induction.
- Base case (n=2): Since 2 is a prime number, $P(2)$ holds.
- Inductive step: Assume k is either a prime or the product of primes.
- But this doesn't really help us prove the property about $k+1$!
- Claim is proven much easier using strong induction!

Proof Using Strong Induction

Prove that if n is an integer greater than 1, then it is either a prime or can be written as the product of primes.

- Base case: same as before.
- Inductive step: Assume each of $2, 3, \ldots, k$ is either prime or product of primes.
- Now, we want to prove the same thing about $k+1$
- Two cases: k is either (i) prime or (ii) composite
- If it is prime, property holds.

Proof, cont.

- If composite, $k+1$ can be written as pq where $2 \geq p, q \geq k$
- By the IH, p, q are either primes or product of primes.
- Thus, $k+1$ can also be written as product of primes
- Observe: Much easier to prove this property using strong induction!

A Word about Base Cases

- In all examples so far, we had only one base case
  - i.e., only proved the base case for one integer
- In some inductive proofs, there may be multiple base cases
  - i.e., prove base case for the first k numbers
- In the latter case, inductive step only needs to consider numbers greater than k

Example

- Prove that every integer $n \geq 12$ can be written as $n = 4a + 5b$ for some non-negative integers a, b.
- Proof by strong induction on n and consider 4 base cases
- Base case 1 (n=12): $12 = 3 \cdot 4 + 0 \cdot 5$
- Base case 2 (n=13): $13 = 2 \cdot 4 + 1 \cdot 5$
- Base case 3 (n=14): $14 = 1 \cdot 4 + 2 \cdot 5$
- Base case 4 (n=15): $15 = 0 \cdot 4 + 3 \cdot 5$

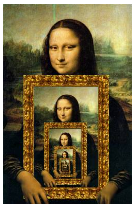
Example, cont.

Prove that every integer $n \geq 12$ can be written as $n = 4a + 5b$ for some non-negative integers a, b.

- Inductive hypothesis: Suppose every $12 \leq i \leq k$ can be written as $i = 4a + 5b$.
- Inductive step: We want to show $k + 1$ can also be written this way for $k + 1 \geq 16$
- Observe: $k + 1 = (k - 3) + 4$
- By the inductive hypothesis, $k - 3 = 4a + 5b$ for some a, b because $k - 3 \geq 12$
- But then, $k + 1$ can be written as $4(a + 1) + 5b$

Recursive Definitions

- In some cases, it is easier to define objects/functions in terms of themselves rather than directly
- Such definitions are called recursive definitions
- Picture below is "defined" recursively because each picture containts a smaller version of itself



Recursive Definitions in Math

- Recursive definitions come up a lot in discrete math
- For example, consider the following sequence:

$$1, 3, 9, 27, 81, \ldots$$

- This sequence can be defined recursively as follows:

$$\begin{aligned} a_0 &= 1 \\ a_n &= 3 \cdot a_{n-1} \end{aligned}$$

- First part called base case; second part called recursive step
- Very similar to induction; in fact, recursive definitions sometimes also called inductive definitions

Recursively Defined Functions

- Just like sequences, functions can also be defined recursively
- Example:

$$\begin{aligned} f(0) &= 3 \\ f(n+1) &= 2f(n) + 3 \quad (n \geq 1) \end{aligned}$$

- What is $f(1)$?
- What is $f(2)$?
- What is $f(3)$?

Recursive Definition Examples

- Consider $f(n) = 2n + 1$ where n is non-negative integer
- What's a recursive definition for f?
- Consider the sequence $1, 4, 9, 16, \ldots$
- What is a recursive definition for this sequence?
- Recursive definition of function defined as $f(n) = \sum_{i=1}^{n} i$?

Recursive Definitions of Important Functions

- Some important functions/sequences defined recursively
- Factorial function:

$$\begin{aligned} f(1) &= 1 \\ f(n) &= n \cdot f(n-1) \quad (n \geq 2) \end{aligned}$$

- Fibonacci numbers: $1, 2, 3, 5, 8, 13, 21, \ldots$

$$\begin{aligned} a_1 &= 1 \\ a_2 &= 2 \\ a_n &= a_{n-1} + a_{n-2} \quad (n \geq 3) \end{aligned}$$

- Just like there can be multiple bases cases in inductive proofs, there can be multiple base cases in recursive definitions

Inductive Proofs for Recursively Defined Structures

- Recursive definitions and inductive proofs are very similar
- Natural to use induction to prove properties about recursively defined structures (sequences, functions etc.)

Example

- Let f_n denote the n'th element of the Fibonacci sequence
- Prove: For $n \geq 3$, $f_n > \alpha^{n-2}$ where $\alpha = \frac{1+\sqrt{5}}{2}$
- Proof is by strong induction on n with two base cases
- Intuition 1: Definition of f_n has two base cases
- Intuition 2: Recursive step uses f_{n-1}, f_{n-2} $\Rightarrow$ strong induction
- Base case 1 (n=3): $f_3 = 2$, and $\alpha < 2$, thus $f_3 > \alpha$
- Base case 2 (n=4): $f_4 = 3$ and $\alpha^2 = \frac{(3+\sqrt{5})}{2} < 3$

Example, cont.

Prove: For $n \geq 3$, $f_n > \alpha^{n-2}$ where $\alpha = \frac{1+\sqrt{5}}{2}$

- Inductive step: Assuming property holds for f_i where $3 \leq i \leq k$, need to show $f_{k+1} > \alpha^{k-1}$
- First, rewrite α^{k-1} as $\alpha^2\alpha^{k-3}$
- α^2 is equal to $1 + \alpha$ because:

$$\alpha^2 = \left(\frac{1+\sqrt{5}}{2}\right)^2 = \frac{\sqrt{5}+3}{2} = \alpha + 1$$

- Thus, $\alpha^{k-1} = (\alpha+1)(\alpha^{k-3}) = \alpha^{k-2}\alpha^{k-3}$

Example, cont.

- $\alpha^{k-1} = \alpha^{k-2}\alpha^{k-3}$
- By recursive definition, we know $f_{k+1} = f_k + f_{k-1}$
- Furthermore, by inductive hypothesis:

$$f_k > \alpha^{k-2} \qquad f_{k-1} > \alpha^{k-3}$$

- Therefore, $f_{k+1} > \alpha^{k-2}\alpha^{k-3} = \alpha^{k-1}$

$\square$