

## Course Information

Welcome to 6.042! In this course, we'll teach you some mathematics that we think you'll find useful in your study of computer science. This handout contains basic information about the class, but all this and more is also available on the course website:

<https://courses.csail.mit.edu/6.042/fall14/>

**Prerequisites.** The only prerequisite is 18.01. If you have already taken 18.310 or 6.046, then you should not take 6.042.

**Lecture.** There are 90-minute lectures on Tuesday and Thursday in 26-100 at 2:30 PM.

**Recitation.** There are 1-hour mandatory recitations on Wednesday and Friday focused on solving problems in small groups. Please enter your preferences for recitation section on the course website today (Thursday) by 7:30PM. We will try to have section assignments posted tonight by 10PM.

**Office hours.** Everyone on the course staff has office hours every week. Times and locations will be posted on the course website.

★ **Reading.** The text is *Mathematics for Computer Science*. A draft copy is available through the course website. Reading will be assigned each week with the problem sets.

**Homework.** There is a problem set each week, for a total of 11. Problem sets are generally released on Tuesday, due the following Monday evening at 7:30PM in the locked boxes at the elevator lobby in 32-G5, and is returned in recitation on Friday. Be neat! Graders may deduct for sloppiness. Late homework is generally not accepted.

Each problem will be submitted separately, so don't staple your solutions together. Make sure that everything you submit has your name, your recitation number, your recitation instructor's name, and a collaboration statement on it.

**Collaboration.** You are welcome to work with other students on homework, but your writeup must be entirely your own. Please do not refer to course materials from previous terms. On the top of your homework, list:

- all collaborators, other than course staff
- all written sources that you consulted, other than the text and course handouts

**Exams.** There are a 2-hour midterm in class on 10/21 and a 3-hour final during finals week.

## Mathematics for computer science

\* See course information handout

A proof is a method of ascertaining the truth.

A mathematical proof is a verification of a proposition by a chain of logical deduction from a base set of axioms.

Def: A proposition is a statement that is either True or False

Ex:  $2+3=5$ , natural #'s

Ex:  $\forall n \in \mathbb{N} \ n^2+n+41$  is a prime number  
 $\uparrow$  For all     $\exists$  exists

Predicate - Proposition whose truth depends on the value of variables

n	$n^2+n+41$	Prime?
1	41	✓
2	43	✓
3	47	
39	1601	✓

does not work after 40

Ex:  $a^4+b^4+c^4=d^4$  has no positive integer solution.

aka False. It actually works with huge #'s

$\exists a, b, c, d \in \mathbb{N}^+$   $a^4+b^4+c^4=d^4$   
 $\uparrow$  there exists       $\subset$  Positive natural #'s 1, 2, 3 ...

Ex: The regions in any map can be colored in 4 colors so that adjacent regions have different colors. Proven true by computer. 4-color theorem

$\forall n \in \mathbb{Z}, n \geq 2 \Rightarrow n^2 \geq 4$   
 integers                  implies  
 $\{0, 1, -1, 2, \dots\}$

Def: An implication  $p \Rightarrow q$  is true if  $p$  is F or  $q$  is T

Truth Table

p	q	$p \Rightarrow q$
T	T	T
F	F	T
F	T	F

"False  $\Rightarrow$  anything" is True  
pigs fly  $\Rightarrow$  I am king

Def: An axiom is a proposition that we assume to be True

Ex: If  $a=b$  and  $b=c$  then  $a=c$

Eucleidian Geometry: Given a line  $l$  and a point  $p$  not on  $l$ ,  
there is exactly 1 line through  $p$  parallel to  $l$

Axioms should be consistent and complete

no proposition can be  
proved true AND False

Every proposition is  
either True or False

## Boston Globe Article

This is an article from the Boston Globe on February 6, 1995.

# The great unsolved mysteries

Some of math's biggest remaining puzzles have fallen this century. But three major conjectures remain unsolved, not counting Fermat's last theorem, which is expected to be proved soon by Princeton professor Andrew Wiles. The remaining challenges are:

■ Goldbach's conjecture, after a proposition penned by Prussian mathematician Christian Goldbach in 1742, holds that every even number can be expressed by the sum of two prime numbers (thus, 20 can be expressed by 9 and 11). It has never been fully proven.

■ The Riemann zeta hypoth-

esis, after an 1859 paper written by Bernhard Riemann, suggests that zeros in an infinite series of numbers known as the zeta function form along a straight line in a complex plane. The hypothesis has been proven to 1.5 billion zeros — not far enough to prove it completely.

■ Poincare's conjecture, advanced in 1904 by French mathematician Jules-Henri Poincare, asserts that objects can be identified by key surface characteristics in the area of geometry known as topology. The conjecture has yet to be proven for the three-dimensional realm.

ANTHONY FLINT

# 6.042 Lecture #1

## Introduction and Proofs

Proof = method for ascertaining the truth

## Experimentation + observation

## Sampling and counterexamples

Judge, Jury

etc.

A mathematical proof is a verification of a proposition by a chain of logical deductions from a set of axioms. ↑

1

Statement that is True or False  
Ex.  $2+3=5$

$$\text{Ex. } 2+3=5$$

$\forall n \in \mathbb{N}, n^2+n+41$  is prime

↓  
"forall" ↓  
natural  
#s

### Predicate:

Proposition whose truth depends on value of a variable.

check:

<u>n</u>	<u><math>n^2+n+41</math></u>	<u>Prime?</u>
0	41	✓
1	43	✓
2	47	✓
3	53	✓
⋮		
40	1681	✗

FALSE!

Ex.  $a^4 + b^4 + c^4 = d^4$  has no positive integer solutions

$$\exists \underset{\text{positive natural}}{a, b, c, d} \in \mathbb{N}^+, \underbrace{a^4 + b^4 + c^4 = d^4}_{\text{Predicate}}$$

"There exist"

False:

$$\begin{array}{r} a = 95,800 \\ b = 217519 \\ c = 414560 \\ d = 412248 \end{array}$$

12

313  $(x^3 + y^3) = z^3$  has no positive integer solution.

False: solution has 1000+ digits

## Why

→ Factoring → cryptosystem → RSA

$\checkmark$  4 color Theorem

Proposition

Ex: Every even integer but 2, is the sum of two primes

No one knows!

Goldbach's conjecture

$$\forall n \in \mathbb{Z}, n \geq 2 \Rightarrow n^2 \geq 4$$

↑  
integers

Implies

↳ An implication  $p \Rightarrow q$  is true if  $p$  is False or  $q$  is true

Truth Table

p	q	$p \Rightarrow q$
T	T	T
T	F	F
F	T	T
F	F	T

$$\forall n \in \mathbb{Z}, n \geq 2 \Leftrightarrow n^2 \geq 4$$

IF and only IF

False.  $n = -3$

Prove both ways!

$\Leftrightarrow$  =  $\Rightarrow$  and  $\Leftarrow$   
Truth Table

p	q	$p \Rightarrow q$	$q \Rightarrow p$	$p \Leftrightarrow q$
T	T	T	T	T
T	F	F	T	F
F	T	T	F	F
F	F	T	T	T

Axiom - proposition that is assumed to be true.

Ex: If  $a=b$  and  $b=c$  then  $a=c$

no proof for this, but we assume true

Axioms should be

1. Consistent - no proposition can be proved to be True and False

2. Complete - can be used to prove every proposition is either true or False

## Objectives:

\* See problems sheet

- 1) Review
- 2) Truth table
- 3) Inference Rules
- 4) Contrapositives
- 5) Logical connectives/quantifiers

## 1) Review

Propositions - statements T or F

Axioms - Propositions accepted as Truth

Consistency -

Completeness -

Implications -  $p \Rightarrow q$ ,  $p \Leftrightarrow$   
 $\uparrow$        $\uparrow$   
 P implies q    "if and only if"

Ex:  $\forall n \in \mathbb{Z}, n \geq 2 \Leftrightarrow n^2 \geq 4$   
 $\uparrow$        $\uparrow$   
 For all in integers

False,  $n = -3$ 

V=OR

 $(p \Rightarrow q) \wedge (q \Rightarrow p)$ 

## 2) Truth Table

p	q	$p \Rightarrow q$	$q \Rightarrow p$	$p \Leftrightarrow q$
T	T	T	T	T
T	F	F	T	F
F	T	T	F	F
F	F	T	T	T

## 3) Inference Rules

Ex Modus Ponens

IF p is true and  $p \Rightarrow q$  is true then  $q$  is true.can check with first row  
truth table above

## 4) Contrapositives

Given  $p \Rightarrow q$ , the contrapositive is  $\neg q \Rightarrow \neg p$

or  $\neg q \Rightarrow \neg \bar{q}$

## Problems for Recitation 1

### 1 Team Problem: Contrapositive

Prove by truth table that an implication is equivalent to its contrapositive.

p	q	$p \Rightarrow q$	$q \Rightarrow p$	$\bar{q} \Rightarrow \bar{p}$	$p \Leftrightarrow F$
T	T	T	T	T	
T	F	F	T	F	
F	T	T	F	T	
F	F	T	T	T	✓

$$\therefore p \Rightarrow q \equiv \bar{q} \Rightarrow \bar{p}$$

$\uparrow$   
equivalent

Proof by Contradiction

To prove  $P$  is True, we assume  $P$  is F and then derive a falsehood.

$$\Rightarrow "P \text{ is False}" \text{ is False}$$

$$\Rightarrow P \text{ is True}$$

Example:

$$\text{Thm: } \frac{\sqrt{2} \text{ is irrational}}{P}$$

PF: (by cont)

Assume For the purposes of contradiction that  $\sqrt{2}$  is rational

$$\Rightarrow \exists a, b \in \mathbb{N}^+ \text{ s.t. } \sqrt{2} = \frac{a}{b} \text{ & } a, b \text{ have no common divisors}$$

$$\Rightarrow 2 = \frac{a^2}{b^2}$$

$$\Rightarrow 2b^2 = a^2$$

$\Rightarrow a$  is even ( $2|a$ )

$$\Rightarrow (4|a^2) \Rightarrow 4|2b^2$$

$$\Rightarrow 2|b^2$$

$\Rightarrow 2$  is even

$$\Rightarrow 2|b \quad \times \quad \leftarrow \text{contradiction sign}$$

$\Rightarrow \sqrt{2}$  is irrational

□

Induction Axiom

Let  $P(n)$  be a predicate. IF  $P(0)$  is T & " $\forall n \in \mathbb{N} \ P(n) \rightarrow P(n+1)$ " is T, then  $\forall n \ P(n)$  is True.

I.e. IF  $P(0)$  is true  $P(0) \rightarrow P(1)$  is true and  $P(1) \rightarrow P(2)$  is true... then  $P(0), P(1), \dots$  is True.

Thm:  $\forall n \geq 0 \quad 1+2+3+\dots+n = \frac{n(n+1)}{2}$

if  $n=1 \quad 1+2+3+\dots+n = 1 \quad (1 \text{ term})$

if  $n=0 \quad 1+2+\dots+n = 0 \quad (0 \text{ terms})$

↑ be careful with ...

PF: (by induction)

Let  $P(n)$  be the proposition that

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

| Inductive Hypothesis

Base case  $P(0)$  is True since

$$\sum_{i=1}^0 i = 0 = \frac{0(0+1)}{2} = 0 \checkmark$$

Inductive step

For  $n \geq 0$ , show  $P(n) \Rightarrow P(n+1)$

Assume

$P(n)$  is T for the purpose of induction

Need to show  $\sum_{i=1}^{n+1} i = \frac{(n+1)(n+2)}{2}$

$$= \frac{n(n+1)}{2} + n+1$$

$$= \frac{n^2+n+2n+2}{2} = \frac{n^2+n+2n+2}{2}$$

$\Rightarrow P(n+1)$

So.  $\forall n \geq 0$  " $P(n) \Rightarrow P(n+1)$ " is T. By induction

$\Rightarrow \forall n \geq 0, P(n)$  is T

## ~ Prooving a falsehood by induction

Theorem (Not!): All horses are the same color.

Pf: (by induction)

Inductive Hypothesis: In any set of  $n \geq 1$  horses, all the horses in the set have the same color.

Base case:  $P(1)$  T since only one horse ✓

Inductive step: Assume  $P(n)$  is True to verify IH.

Consider any set of  $n+1$  horses,  $H_1, H_2, H_3, \dots, H_{n+1}$

$H_1, H_2, \dots, H_n$ , are all the same color by  $P(n)$

$H_2, H_3, \dots, H_{n+1}$ , are all the same color by  $P(n)$

$\Rightarrow$  since  $\text{color}(H_1) = \text{color}(H_2 \dots H_n) = \text{color}(H_{n+1})$

$\Rightarrow$  All  $n+1$  are the same color

$\Rightarrow P(n+1)$  is True

□

What went wrong?

When  $n=1$

$\text{color}(H_2 \dots H_n)$  is the empty set

so because  $P(2)$  is False,  
 $P(2) \Rightarrow P(3)$        $F \Rightarrow F$

$P(3) \Rightarrow P(4) \dots$

$P(1)$

was missing  $P(1) \Rightarrow P(2)$ !

# 6.042 Lecture #2

OCW

## Induction

Last class: Direct proofs

Today: Indirect proofs.

### Proof by contradiction

To prove proposition  $P$  is True. Assume  $P$  is False.

then use hypothesis to derive a falsehood/contradiction

IF  $\neg P \Rightarrow F$  is True

$$\begin{array}{c} \xrightarrow{\text{not}} \\ F \\ \text{so } P \text{ is True} \end{array}$$

Ex. Thm:  $\sqrt{2}$  is irrational

PF. (by contr)

Assume, For purpose of contradiction, that  $\sqrt{2}$  is rational.

$\Rightarrow \sqrt{2} = \frac{a}{b}$  (Fraction in lowest terms)

$$2 = \frac{a^2}{b^2}$$

$2b^2 = a^2 \Rightarrow a^2$  is even  $\Rightarrow a$  is even (2|a)  
multiple of 4  $\qquad\qquad\qquad$   $\begin{matrix} \nearrow \\ 2 \text{ divides } a^2 \end{matrix}$

$$\Rightarrow 4 | a^2$$

$$(4 | 2b^2)$$

$2 | b^2 \Rightarrow b$  is even

$\Rightarrow \frac{a}{b}$  is not in lowest terms

$\Rightarrow$  contradiction



Proof is done

$\Rightarrow \sqrt{2}$  is irrational.  $\square$

### \* Induction Axiom

Let  $P(n)$  be a predicate. IF  $P(0)$  is true

and  $\forall n \in \mathbb{N} \quad (P(n) \Rightarrow P(n+1))$  is true

then  $\forall n \in \mathbb{N} \quad P(n)$  is true

Aka.

If  $P(0)$ ,  $P(0) \Rightarrow P(1)$ ,  $P(1) \Rightarrow P(2)$  is True  $\Rightarrow P(0), P(1), P(2) \dots$  are true

Ex.

Thm.  $\forall n \geq 0 \quad 1+2+3+\dots+n = \frac{n(n+1)}{2}$

$$\sum_{i=1}^n i = \sum_{\substack{i \leq i \leq n}} i = \sum_{1 \leq i \leq n} i$$

IF  $n=1 \quad 1+2+3+\dots+n = 1$

IF  $n \leq 0 \quad 1+2+3+\dots+n = 0$

PF. ① By induction

① Let  $P(n)$  be the proposition that  $\sum_{i=1}^n i = \frac{n(n+1)}{2}$

② Check base case  $P(0)$

Base case

$$P(0) \text{ is true. } \sum_{i=1}^0 i = 0 = \frac{0(0+1)}{2} = 0 \quad \checkmark$$

③ Inductive Step

For  $n \geq 0$ , show  $P(n) \Rightarrow P(n+1)$  is true

Assume  $P(n)$  is true For purposes of induction

(i.e., assume  $1+2+\dots+n = \frac{n(n+1)}{2}$ )

need to show  $1+2+\dots+(n+1) = \frac{(n+1)(n+2)}{2}$

$$\underbrace{1+2+\dots+n+(n+1)}$$

$$\frac{n(n+1)}{2} + (n+1) = \frac{n^2+n+2n+2}{2} = \frac{(n+1)(n+2)}{2} \quad \checkmark$$

<sup>so</sup>  $P(n) \Rightarrow P(n+1) \quad \square$

Ex

Thm.  $\forall n \in \mathbb{N} \quad 3 \mid (n^3 - n)$

Ex.  $n=5 \quad 3 \mid (125-5)$

PF. by induction

let  $P(n) \equiv 3 \mid (n^3 - n)$

Base case  $n=0 \quad 3 \mid (0-0) \quad \checkmark$

Inductive Step <sup>show</sup>  
For  $n \geq 0 \quad P(n) \Rightarrow P(n+1)$  is True

Assume  $P(n)$  is true

Assume  $3 \mid (n^3 - n)$

Examine

$$\begin{aligned} 3 \mid & \underbrace{(n+1)^3 - (n+1)}_{\rightarrow} \\ & n^3 + 3n^2 + 3n + 1 - (n+1) \\ & = n^3 + 3n^2 + 2n \\ & = \underbrace{n^3 - n}_{\substack{\downarrow \\ 3|(n^3-n)}} + 3n^2 + 3n \\ & \quad \text{by } P(n) \\ & 3 \mid 3n^2 \\ & 3 \mid 3n \\ \Rightarrow & 3 \mid (n+1)^3 - (n+1) \quad \square \end{aligned}$$

Induction steps

- ① Write "Proof by induction". Identify  $P(n)$
- ② Base Case
- ③ Inductive Step

start induction at any value

Ex Base case  $P(b)$  is true

Ind step  $\forall n \geq b \quad P(n) \Rightarrow P(n+1)$

(Conclusion)  $\forall n \geq b \quad P(n)$

Ex Thm (NOT)  
All horses are the same color

Proof by Induction

$P(n)$  In any set of  $n \geq 1$  horses, the horses are all the same color

Base Case

$P(1)$  True since just 1 horse  $\rightarrow$  same color

Inductive Step

Assume  $P(n)$  is true to show  $P(n+1)$  is True

Consider any set of  $n+1$  horses  $h_1, h_2, \dots, h_{n+1}$

then  $h_1, \dots, h_n$  are the same color

also  $h_2, h_3, \dots, h_{n+1}$  are the same color

Since  $\text{color}(H_1) = \text{color}(h_2, \dots, h_n) = \text{color}(H_{n+1})$

$\Rightarrow P(n+1) \quad \square$

What went wrong?

$n=1$  on inductive step should prove  $P(1) \Rightarrow P(2)$

$H_1, H_2, \dots, H_n$   
Empty Set!  
 $n-1$  horses  
but  $n=1$

Moral: ... notation is deceiving  
Establish inductive step For all  $n \geq$  base case.

$P(1) \rightarrow P(2) \Rightarrow P(3) = P(4)$

did not prove  
 $P(1) \Rightarrow P(2)$   
no proof!

1.

$\forall n, \exists$  way to tile a  $2^n \times 2^n$  region with a center square missing.

For Bill

PrF: By Induction

$P(n) =$

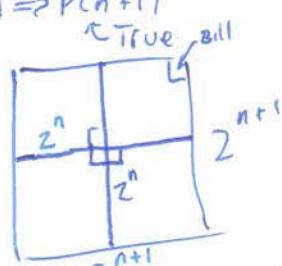
Base case

$P(0) \sim$  no tiles, one left ✓

Inductive step

For  $n \geq 0$  assume  $P(n)$  true to show  $P(n) \Rightarrow P(n+1)$

Consider a  $2^{n+1} \times 2^{n+1}$  courtyard



Hard to prove but IF we change predicate to  
... region with any square missing  
 $\Rightarrow$  Stronger  $P(n)$

# 6.042 Chapter 1 Notes

Theorem - Important proposition

Corollary - Proposition that follows from theorem

Logical Deductions

P, P implies Q ← antecedents  
Q ← conclusion. True if antecedents true

Implications - If P, then Q aka P implies Q

Method #1      ① Assume 'P'  
                  ② Show that Q logically flows

Method #2

Prove the contrapositive

P implies Q = NOT(Q) implies NOT(P)

① State contrapositive  
② Method #1

Proving If and only if  $\iff$

Method #1

Prove each statement implies the other

P iff Q = P implies Q and Q implies P

Method #2

Chain of IFFs

① Construct a chain of iff implication  
②  $P \rightarrow Q \rightarrow R \rightarrow S \rightarrow T \rightarrow Q$

Proof by contradiction

Ex  $\sqrt{2}$  is irrational

① Assume opposite  $\sqrt{2} = \frac{n}{d}$  ← integers  
 $\rightarrow \frac{n^2}{d^2} = 2$  ← Later proof... here!  
smaller positive integer denominator  
so false.  $\square$

$$\begin{aligned} 2 &= \frac{n^2}{d^2} \\ 2d^2 &= n^2 \\ d^2 &\text{ is a factor of } n^2 \rightarrow n \\ n &= 2k \\ n^2 &= (2k)^2 = 4k^2 \\ 2d^2 &= 4k^2 \\ d^2 &= 2k^2 \end{aligned}$$

# 6.042 Chapter 2 Notes

## The well ordering Principle

"Every nonempty set of nonnegative integers has a smallest element!"

Template:

Prove that  $P(n)$  is true for all  $n \in \mathbb{N}$

• Define set,  $C$  of counter-examples

$$C := \{n \in \mathbb{N} \mid P(n) \text{ is false}\}$$

For which

- Assume that  $C$  is nonempty
- By W.o.P  $\rightarrow$  smallest element,  $n$ , in  $C$
- Reach contradiction
- Conclude that  $C$  must be empty  $\rightarrow$  no counterexample exists  $\square$

# 6.042 Notes

## Chapter 1: Propositions

Truth Table

P	Q	P implies Q aka $P \rightarrow Q$
T	T	T
T	F	F
F	T	T
F	F	T

IF P is False or Q is True, P implies Q

P	Q	P:FF Q aka $P \leftrightarrow Q$
T	T	T
T	F	F
F	T	F
F	F	T

Either both True or False

$$P \rightarrow Q = \frac{\neg Q \rightarrow \neg P}{\text{NOT}(Q) \rightarrow \text{NOT}(P)}$$

contrapositive

$$\begin{array}{c} P \rightarrow Q \\ \hline \text{Implication} \end{array} \neq \begin{array}{c} Q \rightarrow P \\ \hline \text{converse} \end{array}$$

Implication and converse = :FF

Propositions:

True or False

Satisfiability - If a variable makes proposition True

Principle of InductionLet  $P(n)$  be predicate

goal of induction

$\forall n P(n)$

·  $P(0)$  is true ← Base case·  $P(n) \rightarrow P(n+1)$  ← Inductive Step

then

·  $P(n) = \text{true } \forall n \geq 0$ 

Template

- 1) State Proof by induction
- 2) Define predicate  $P(n)$
- 3) Base case  $P(0)$
- 4) Induction Hypothesis (IH)  $P(n)$
- 5)  $P(n) \Rightarrow P(n+1)$
- 6) By induction  $P(n) = \text{True } \forall n \geq 0$

Thm  $\forall n \in \mathbb{N} \quad 3 \mid n^3 - n$ Proof

By induction

use induction hypothesis to prove  $p(n) \rightarrow p(n+1)$ 

$P(n) = "3 \mid n^3 - n"$

$P(0): 0^3 - 0 = 0 \checkmark$

$I.H: P(n) \quad 3 \mid n^3 - n$

$\xrightarrow{P(n)} \xrightarrow{P(n+1)}$   
 Assume True  $3 \mid (n+1)^3 - (n+1)$

$n^3 + 3n^2 + 3n - 1 - (n+1)$

$n^3 + 3n^2 + 3n - n$

$$\begin{array}{c} (n^3 - n) + 3n^2 + 3n \\ \hline P(n) \end{array}$$

$3 \mid (n+1)^3 - (n+1)$

 $\square$

## Problems for Recitation 2

### 1 Problem: A Geometric Sum

Perhaps you encountered this classic formula in school:

$$1 + r + r^2 + r^3 + \dots + r^n = \frac{1 - r^{n+1}}{1 - r}$$

Use induction to prove that this formula is correct for all real values  $r \neq 1$ .

*Prepare a complete, careful solution. You'll be passing your proof to another group for "constructive criticism"!*

### 3 Problem: False Proof

Find the error in this false proof.

**Theorem 1.** *I can lift all the sand on the beach.*

*Proof.* The proof is by induction.  $P(n)$  : I can lift  $n$  grains of sand.

*Base case:*  $n = 0$ . Trivial!

*Induction step:* Assume I can lift  $n$  grains of sand. Then surely I can lift  $n + 1$  grains of sand. Hence  $P(n) \rightarrow P(n + 1)$ .

So  $\forall n P(n)$  is true.

Let  $M$  be the number of grains of sand on the beach, then  $P(M) \rightarrow$  I can lift all the grains on the beach.  $\square$

# 6.042 Lecture #3

OCW

## Strong Induction

Good proofs:

- correct
- brief
- in order
- complete
- "elegant"
- well organized

Fermat's Last Thm

$$\forall n > 2 \ \exists x, y, z \in \mathbb{N}^+$$

$$x^n + y^n = z^n$$

Problem:

Find sequence of moves to go from

A	B	C
D	E	F
H	G	

to

A	B	C
D	E	F
6	H	

Legal move: slide letter into adjacent blank square

Thm: There is no sequence of legal moves to invert H and 6 and return other letters to original position

Invariant:

Property that holds in initial state and preserved by every legal move.

shows that you can't reach goal state.

But not in goal state.

Row move:	A B C	A B C	Natural order
	D C	D G	
	E F H	E F H	

1	2	3
4	5	6
7	8	9

Lemma 1: A row move does not change the order of the items

Proof: In a row move, we move an item from cell  $i$  into an adjacent cell  $i+1$  or  $i-1$ . Nothing else moves. Hence, order of items is preserved.  $\square$

Column moves:

A	B	C
D	F	
H	E	G

→

A	B	C
D	F	G
H	E	

Moves 3 position

1	2	3
4	5	6
7	8	9

A	B	C
D		G
H	E	F

A		I
D	B	G
H	E	F

Lemma 2: Column move changes the relative order of precisely 2 pairs of items.

Proof: In a column move, we move an item in cell  $i$  to a blank spot in cell  $i-3$  or  $i+3$ . When an item moves 3 positions, it changes relative order with 2 other items  $i-1, i-2$  or  $i+1, i+2$ .

Def: A pair of letters  $L_1, L_2$  is an inversion (aka inverted pair) if  $L_1$  precedes  $L_2$  in alphabet but  $L_1$  is after  $L_2$  in the puzzle

A	B	C
F	O	G
E	H	

(D, F), (E, F), (E, G)  
3 inversions

Lemma 3: During a move the # of inversions can only increase by 2, decrease by 2, or stay the same.

Proof: Row move: no changes by Lemma 1

Col move: 2 pairs change order by Lemma 2

Case A: Both pairs were in order  $\Rightarrow$  # inversions 0

Case B: 1 inverted  $\Rightarrow$  # inversions 1

Case C: one pair inverted  $\Rightarrow$  # inversions stays the same  $\square$

Corollary: During a move, the parity (even/odd) of # of inversions does not change

Proof: Adding or subtracting 2 does not change the parity

Invariant:

Lemma 4: In every state reachable from start state, the parity of the # of inversions is odd.

Proof: by induction

$P(n)$  After any sequence of  $n$  moves

Base case  $n=0$ . # inversions = 1. odd

Inductive step: For  $n \geq 0$ , show  $P(n) \rightarrow P(n+1)$

Consider any sequence of  $n+1$  moves  $M_1 \dots M_{n+1}$

By  $P(n)$ , we know that parity after moves  $M_1 \dots M_n$  is odd.

By Corollary 1, parity of # inversions does not change during  $M_{n+1} \Rightarrow$  parity after  $M_1 M_2 \dots M_{n+1}$  is odd.  
 $\Rightarrow P(n+1)$

so  $P(n) \rightarrow P(n+1)$

□

Proof of Theorem: The parity of # inversions in desired state is even (0).

By Lemma 4, desired state cannot be reached from start state

□

## Strong Induction Axiom

Let  $P(n)$  be any predicate. If  $P(0)$  is true

&  $\forall n \quad (P(0) \wedge P(1) \wedge \dots \wedge P(n)) \Rightarrow P(n+1)$  is true,

then  $\forall n \quad P(n)$  is true;

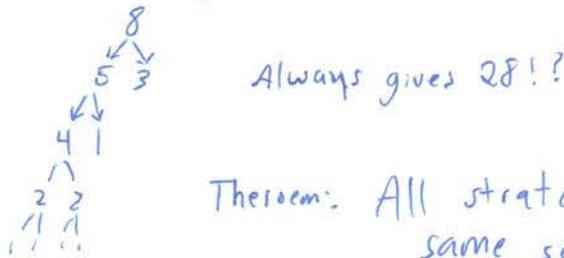
can assume all of these

are true as part of the

inductive step!  $\Rightarrow$  can make some proofs easier

Example.

## Unstacking Game



Theorem: All strategies for the  $n$ -block game produce the same score  $S(n)$

$$\text{Ex. } S(8) = 28$$

Proof:

By strong induction

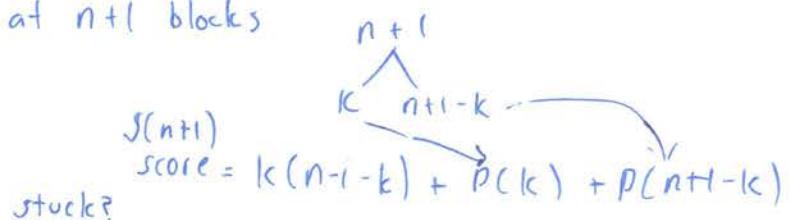
① Predicate:  $P(n)$

② Base case  $n=1$   
 $S(1) = 0 \checkmark$

③ Inductive step.

Assume  $P(1), P(2), \dots, P(n)$  to prove  $P(n+1)$

Look at  $n+1$  blocks



Make induction hypothesis stronger

Thm.  $\dots \rightarrow \dots \rightarrow \dots$

$$\textcircled{1} P(n) \rightarrow \dots \rightarrow S(n) = \frac{n(n-1)}{2}$$

\textcircled{2} Base case  $\checkmark$

\textcircled{3} Inductive step

$$\begin{aligned} \text{score} &= k(n+1-k) + \frac{k(k+1)}{2} + \frac{(n+1-k)(n-k)}{2} \\ &= \cancel{2kn + 2k - 2k^2 + k^2 - k} + \frac{(n+1)n - kn - k - k^2}{2} \\ &= S(n+1) \end{aligned}$$

# 6.042 Notes

## Chapter 3: Induction

### Well Ordering Principle

- Every non empty set of nonnegative integers has a smallest element

### Ordinary Induction

Let  $P(0)$  be a predicate.

If  $P(0)$  is true and

$P(n) \rightarrow P(n+1)$  for all nonnegative int  $n$

then

$P(m)$  is true for all nonneg int  $m$

aka

$\frac{P(0), \forall n \in N. P(n) \rightarrow P(n+1)}{\forall m \in N. P(m)}$

### Example

Thm: For all  $n \in N$

$$1 + 2 + 3 + \dots + n = \frac{n(n+1)}{2}$$

① Show that  $P(0)$  is true

$$0(0+1)/2 = 0 \checkmark$$

② Show that  $P(n) \rightarrow P(n+1)$

• Assume  $P(n)$  to prove  $P(n+1)$  Prove this

$$\frac{1+2+\dots+n}{F(n+1)} + n+1 = \frac{(n+1)(n+2)}{2} = \frac{n(n+1)}{2} + n+1$$

### Template:

① Write "Proof using induction"

② Define Appropriate predicate  $P(n)$

③ Prove  $P(0)$  True

④ Prove  $P(n) \rightarrow P(n+1)$

⑤ Invoke Induction

□

## Invariants

- ↳ Property that is preserved after series of steps
- useful for state machines. (Defined start state)

- $P(t) = \text{invariant after step } t$
- Show  $P(0)$  is true (invariant holds for start state)
- Show that

$$\forall t \in \mathbb{N}. P(t) \rightarrow P(t+1)$$

aka invariant holds for current and next state

- See puzzle example in notes.

## Strong Induction

- state : if you are using strong induction.
- $P(n+1)$  depends on  $P(a)$   $a < n$
- all the previous ones
- can assume all previous  $a$  are true during inductive step

Principle:

IF

$P(0)$  is true and for all  $n \in \mathbb{N}$ ,  $P(0), P(1), \dots, P(n)$  together imply  $P(n+1)$   
then  $P(n)$  is true for all  $n \in \mathbb{N}$

Rule:

$$\begin{aligned} P(0), \forall n \in \mathbb{N}. (P(0) \wedge P(1) \wedge \dots \wedge P(m)) \rightarrow P(n+1) \\ \forall m \in \mathbb{N}. P(m) \end{aligned}$$

- See stacking game example

## Structural Induction

wtf are they trying  
to tell me here?

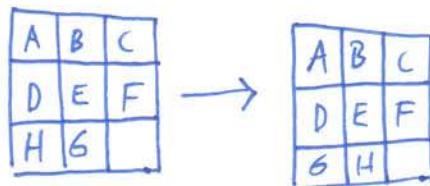
## Top 10 Proof Techniques NOT Allowed in 6.042

10. **Proof by throwing in the kitchen sink:** The author writes down every theorem or result known to mankind and then adds a few more just for good measure. When questioned later, the author correctly observes that the proof contains all the key facts needed to actually prove the result. Very popular strategy on 6.042 exams. Known to result in extra credit with sufficient whining.
9. **Proof by example:** The author gives only the case  $n = 2$  and suggests that it contains most of the ideas of the general proof.
8. **Proof by vigorous handwaving:** A faculty favorite. Works well in any classroom or seminar setting.
7. **Proof by cumbersome notation:** Best done with access to at least four alphabets and special symbols. Helps to speak several foreign languages.
6. **Proof by exhaustion:** An issue or two of a journal devoted to your proof is useful. Works well in combination with proof by throwing in the kitchen sink and proof by cumbersome notation.
5. **Proof by omission:**
  - “The reader may easily supply the details.”
  - “The other 253 cases are analogous.”
  - “...”
4. **Proof by picture:** A more convincing form of proof by example. Combines well with proof by omission.
3. **Proof by vehement assertion:** It is useful to have some kind of authority in relation to the audience.
2. **Proof by appeal to intuition:** Cloud-shaped drawings frequently help here. Can be seen on 6.042 exams when there was not time to include a complete proof by throwing in the kitchen sink.
1. **Proof by reference to eminent authority:**
  - “I saw Fermat in the elevator and he said he had a proof . . .”

Good Proofs are:

- correct
- elegant
- complete
- well organized
- clear
- in order
- brief

Problem: Find a sequence of moves to go

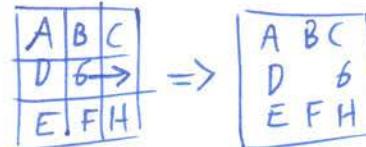


Legal move: Slide letter into adjacent blank square, up, down, left, write

Thm: The puzzle is not solvable

Row Move:

Ex



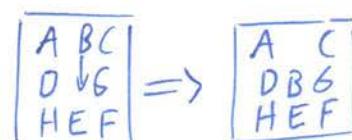
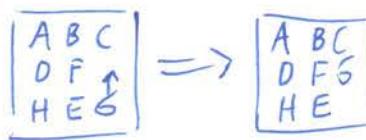
Row-major-order

1	2	3
4	5	6
7	8	9

Fact: A row move does not change the order of the letters

Column Move:

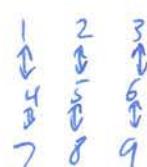
Ex



Lemma 1: A column move changes the relative order of precisely 2 pairs of items.

Proof:

In a col move, we move a letter from cell  $i$  to a blank spot in cell  $i-3$  or  $i+3$



When a letter moves 3 positions, it changes relative order with 2 letters in between  $\square$

Def:  $L_1 \text{ FL}_2$  are inverted if  $L_1$  precedes  $L_2$  in the alphabet but  $L_1$  appears after  $L_2$  in the puzzle

A	B	C
F	D	G
E	H	

(F,D), (F,E), (G,E)

Lemma 2: During a legal move, the # of inverted pairs can only increase by 2, decrease by 2, or stay the same.

Proof:

Row move: no change (Fact 1)

Column move: 3 cases By Lemma 1, 2 pairs change orders

- 1) Both pairs were in order: # inverted pairs + = 2
- 2) " not " " - = 2
- 3) One of each : no change

□

Lemma 3: (Invariant). During a move, the parity<sup>even/odd</sup> of the # of inverted pairs does not change

Proof:

Adding or subtracting 2 from a # does not change parity

□

Lemma 4: In every configuration reachable from  $\begin{array}{|c|c|c|} \hline A & B & C \\ \hline D & E & F \\ \hline H & G & \\ \hline \end{array}$ , the parity of the # of inverted pairs is odd.

A	B	C
D	E	F
H	G	

Proof: (by induction)

I-H:  $P(n)$ : After any set of  $n$  moves From  $\begin{array}{|c|c|c|} \hline A & B & C \\ \hline D & E & F \\ \hline H & G & \\ \hline \end{array}$ , the parity of the # of inverted pairs is odd.

Base case:  $n=0$  # inverted pairs = 1, odd

Inductive step: For  $n \geq 0$ , show  $P(n) \Rightarrow P(n+1)$ . Assume  $P(n)$  true for purpose of induction

Consider any sequence of  $n+1$  moves:  $M_1, M_2, \dots, M_{n+1}$ .

odd parity by  $P(n)$

By Lemma 3, parity does not change in  $M_{n+1} \Rightarrow$  parity odd.  $\Rightarrow P(n+1)$

□

## Proof of Theorem :

The parity in desired state is even (0), so by Lemma 4 it's not reachable.  $\square$

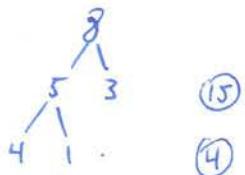
## Strong Induction Axiom

Let  $P(n)$  be a predicate.

IF  $P(0)$  is T  $\wedge \forall n \geq 0 \quad (P(0) \wedge P(1) \wedge P(2) \wedge \dots \wedge P(n)) \Rightarrow P(n+1)$   
then  $P(n)$  is T for all  $n \geq 0$

↖ Can assume more!  
(Just  $P(n)$  for regular induction)

## Unstacking Game



$$= 20$$

Thm : All strategies for the  $n$ -block game produces the same score  $S(n)$ .

$$\text{Ex } S(8) = 28$$

## Proof

By strong induction

I.H  $P(n)$

Base Case:  $n=1 \quad S(1)=0$  always ✓

Inductive step: Assume  $P(1), P(2), \dots, P(n)$  to prove  $P(n+1)$

Look at  $n+1$  blocks.  $\sum_{k=1}^{n+1} S(k) \quad 1 \leq k \leq n$

$$S(n+1) = \sum_{k=1}^{n+1} S(k) = S(n+1-k) + S(k) + S(n+1-k)$$

↖ Too hard to prove that it doesn't depend on  $k$

~ Change Inductive hypothesis!

I-H: ... the same score  $s(n) = \frac{(n-1)n}{2}$   
(can now plug in to  $s(n+1)$ !)

# 6.042 Lecture #4

OCW

## Number Theory

Number Theory - The study of integers

$m|a$  ( $m$  divides  $a$ )

IFF  $\exists_k \in \text{integer}_K \quad a = km$

Ex.  
Suppose

$a$ -gallon jug ( $a=3$ ),  $b$ -gallon jug ( $b=5$ ), get exactly 4 gallons

Thm.

$m|a$  and  $m|b$ , then  $m|\text{any result}$

State machine:

States: pairs  $(x, y)$ , where

$x = \# \text{ gallons in } a\text{-jug}$

$y = \# \text{ " } b\text{-jug}$

Start state:  $(0, 0)$  & both empty

Transitions

• emptying

$$(x, y) \rightarrow (0, y)$$

$$(x, y) \rightarrow (x, 0)$$

• Filling

$$(x, y) \rightarrow (a, y)$$

$$(x, y) \rightarrow (x, b)$$

• pouring one into other

$$(x, y) \rightarrow (0, x+y), \quad x+y \leq b$$

$$(x, y) \rightarrow (x-(b-y), b) = (x+y-b, b), \quad x+y \geq b$$

$$(x, y) \rightarrow (x+y, 0), \quad x+y \leq a$$

$$(x, y) \rightarrow (a, y-(a-x)) = (a, x+y-a), \quad x+y \geq a$$

Example

$$a=3 \quad b=5 \quad (0, 0) \rightarrow (0, 5) \rightarrow (3, 2) \rightarrow (0, 2) \rightarrow (2, 0) \rightarrow (2, 5) \rightarrow (3, 4)$$

Proof: By induction

Assume  $m|a$ ,  $m|b$

Invariant:  $P(n) = \text{IF } (x, y) \text{ is the state after } n \text{ transitions, then } m|x, m|y$

Base case:  $(0, 0) \quad m|0 \Rightarrow P(0) \checkmark$

Inductive step

Assume  $P(n)$

Suppose that  $(x, y)$  state after  $n$  transitions

$P(n) \Rightarrow m|x$  and  $m|y$

After another transition, each jugs is filled with  $0, a, b, x, y, x+y-a, x+y-b, x+y$  gallons.

$m|0, m|a, m|b, m|x, m|y$

linear combination of  $a, x, a, b$  also divisible

$\Rightarrow m$  divides any of the above

$P(n+1)$

□

Def

$\gcd(a, b)$  = greatest common divisor of  $a$  and  $b$

$$\gcd(3, 5) = 1$$

Def:  $a$  and  $b$  are relatively prime if  $\gcd(a, b) = 1$

Corollary:  $\gcd(a, b) |$  any result

Thm: Any Linear combination  $L = s'a + t'b$  of  $a$  and  $b$  with  $0 \leq L \leq b$  can be reached

$$\begin{array}{rcl} \text{Ex: } & s=2, t=3 & \uparrow \text{int} \\ & (2)(3) + (2)(5) & \nearrow \nwarrow \\ & 6+10 & \\ & \underline{-3-5} & \\ & 3 & \end{array}$$

PF: Notice  $L = sa + tb = (s+mb)a + (t-ma)b$  ↴

so  $\exists s', t' \quad L = s'a + t'b'$  with  $s' > 0$

Assume  $0 < L < b$

makes  $s'$  positive

Algorithm:

- helpful because ↴

To obtain  $L$  gallons, repeat  $s'$  times

- Fill a jug

- Pour into b-jug

When b-jug满, empty it out and continue pouring until a-jug empty.

Example.  $s=3$

First loop:  $(0,0) \rightarrow (3,0) \rightarrow (0,3)$

Second loop:  $(0,3) \rightarrow (3,3) \rightarrow (1,5) \rightarrow (1,0) \rightarrow (0,1)$

Third loop:  $(0,1) \rightarrow (3,1) \rightarrow (0,4)$

Filled the  $a$ -jug  $s$  times

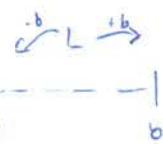
Suppose that  $b$ -jug is emptied  $v$  times

Let  $r$  be the remainder in  $b$ -jug

$$0 \leq r \leq b \quad 0 \leq L < b$$

$$r = s \cdot a - v \cdot b \quad L = s \cdot a + t \cdot b$$

$$\begin{aligned} r &= \underbrace{s'a + t'b}_{L} - t'b - v \cdot b \\ &= L - (t' + v)b \end{aligned}$$



$$t' + v \neq 0 \Rightarrow [r < 0 \vee r > b] \text{ - can't happen}$$

$$t' + v = 0 \Rightarrow v = -t' \Rightarrow r = L$$

□

There exists a unique  $\frac{q}{\uparrow}$  and  $\frac{r}{\uparrow}$  such that  $b = qa + r$   
quotient remainder  $\text{rem}(b,a)$  with  $0 \leq r \leq a$

Lemma Euclid's Algorithm

$$\gcd(a,b) = \gcd(\text{rem}(b,a), a)$$

Ex

$$\begin{aligned} \gcd(105, 224) &= \gcd(\text{rem}(224, 105), 105) & 224 = 2 \cdot 105 + 40 \\ &\stackrel{\text{use again}}{=} \gcd(\text{rem}(105, 14), 14) & = \gcd(14, 105) \\ &\stackrel{\text{repeat}}{=} \gcd(\text{rem}(14, 7), 7) & = \gcd(7, 14) \\ &= \gcd(7, 7) & 105 = 7 \cdot 14 + 0 \\ &= 7 \end{aligned}$$

Proof

$$[m|a \wedge m|b] \Rightarrow [m|b - qa = \text{rem}(b,a) \wedge m|a]$$

If  $\text{rem}(b,a) \neq 0$  then  $[m|\text{rem}(b,a) = b - qa \text{ and } m|a]$

$$\Rightarrow [m \mid m \wedge m \mid b]$$

IF  $\text{rem}(b, a) = 0 = b - qa$

$$m \mid a \Rightarrow m \mid b \\ \Leftrightarrow b = qa$$

$$m \mid a \Rightarrow m \mid b \quad \text{Reverse Argument}$$

Theorem:

$\gcd(a, b)$  is linear combination of  $a$  and  $b$

Proof: By induction

Invariant:  $P(n)$  = IF Euclid's Algorithm reaches  $\gcd(x, y)$  after  $n$  steps, then both  $x$  and  $y$  are linear combinations of  $a, b$   
 $\gcd(a, b) = \gcd(x, y)$

Base case:

$P(0)$  true

Inductive step:

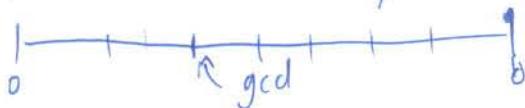
Assume  $P(n)$  Prove  $P(n+1)$

Notice that  $\exists q \quad \text{rem}(y, x) = y - qx \rightarrow$  linear combination of  $a$  and  $b$   
 $= P(n+1) \checkmark$

Last step  $\gcd(0, y) = y$

Theorem:

$\gcd(a, b)$  is the smallest positive linear combination of  $a$  and  $b$ .



combine last three theorems to prove this on hw.

11/12/14

## 6.042 Recitation

Strong Induction

$$P(0) = T$$

$$P(0) \wedge P(1) \wedge P(2) \wedge \dots \wedge P(n) \quad \checkmark \text{ Assume all of the ones before.}$$

$$\Rightarrow P(n+1)$$

$$P(n) = T \quad \forall n > 0$$

Example:

Show that you can make any amount of money  $\geq 8$  from 3¢ & 5¢

$$P(8) = 3+5 \quad \checkmark$$

Assume  $P(8) \vee \dots \vee P(n)$  For induction

$$P(n+1)? \quad \downarrow \\ n+1 = (n-2) + 3 = n+1$$

\* Need to prove  $n-2 \geq 8$  with multiple base cases  
- so two results back

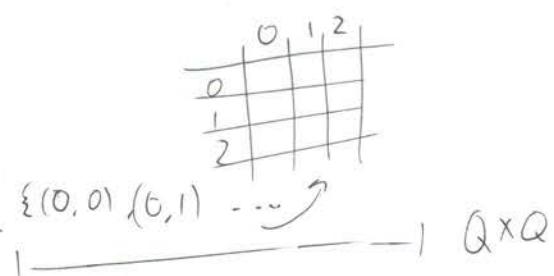
$$P(9) = 3+3+3 \quad \checkmark$$

State Machine

$$Q = \{\text{states}\}$$

$$\delta \subseteq Q \times Q \xrightarrow{\text{set notation}}$$

$$\delta = \{(0,1), (1,2)\}$$



## Problems for Recitation 3

### 1 Problem: Breaking a chocolate bar

We are given a chocolate bar with  $m \times n$  squares of chocolate, and our task is to divide it into  $mn$  individual squares. We are only allowed to split one piece of chocolate at a time using a vertical or a horizontal break.

For example, suppose that the chocolate bar is  $2 \times 2$ . The first split makes two pieces, both  $2 \times 1$ . Each of these pieces requires one more split to form single squares. This gives a total of three splits.

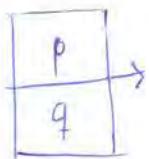
Prove that the number of times you split the bar does not depend on the sequence of splits you make.

$$k = mn$$

$P(k)$  = For a block of size  $k$ , we ~~can~~ make  $k-1$  cuts

$$P(1) : \square^0 | | - 1 = 0 \checkmark$$

$$P(k+1)$$



$$k+1 \geq 2$$

$$k+1 = p+q \quad \leftarrow \text{size}$$

$$p-1 + q-1 + 1 \quad \leftarrow \# \text{ of cuts}$$

$$p+q-1$$

$$k+1-1 = k$$

## The Euclidean Algorithm and Applications

Problem: 5 gallon jug, 3 gallon jug. Set exactly 4.

$$(0,0) \rightarrow (5,0) \rightarrow (2,3) \rightarrow (2,0) \rightarrow (0,2) \rightarrow (5,2) \rightarrow (4,3) \rightarrow (4,0) \checkmark$$

Will this work with  
6 gallon, 3 gallon jugs?

Strategy:

- Define invariant  $J$
- Show  $J$  holds at start
- Show  $J$  always holds (induction)
- Show  $J$  does not hold at end

$J$ : At state  $(x,y)$ , both  $x$  and  $y$  are  $\nearrow$  divisible by  $\text{GCD}(a,b)$   
~~multiples of 3~~

def  $a|b$  "a divides b" ; if  $ak = b$  for some integer  $k$

convention: any integer divides zero

Facts about divisibility:

- (1) IF  $a|b$  then  $a|bc$   $\forall c$
- (2) If  $a|b$  and  $b|c \Rightarrow a|c$
- (3) If  $a|b$  and  $a|c \Rightarrow a|sb+tc \ \forall s,t$
- (4) If  $a|bc$  and  $\text{GCD}(a,b)=1 \Rightarrow a|c$

def  $\text{GCD}(a,b)$  (Greatest Common Divisor)

$\Rightarrow$  the largest possible  $k$ ,  $k|a$  and  $k|b$

$$\text{GCD}(3,6)=3$$

$$\text{GCD}(3,5)=1$$

Thm: IF we have jugs of size  $(a,b)$ , then the amount of water in each is always a multiple of the  $\text{GCD}(a,b)$

Informal:  $J$  always holds



## → Proof (By induction)

$P(n) \triangleq J$  holds after  $n$  steps

Base case:

Initially  $(0,0)$  then  $k|0$  ✓

Inductive Step:  $P(n) \Rightarrow P(n+1)$

Proof by Case Analysis

Say we are at  $(x,y)$  at step  $n$

- $(0,y)$  or  $(x,0)$  Pour out
- $(a,y)$  or  $(x,b)$  Filling up
- $(0,x+y)$  if  $x+y \leq b$  pouring from Jug 1 → Jug 2  
 $(x+y-b, b)$  if  $x+y > b$
- $\dots$  pouring from Jug 2 → Jug 1

$(x,y) \rightarrow (0,y)$  since  $P(n)$ ,  $k|y$  and  $k|0$  ✓

$(x,y) \rightarrow (a,y)$  then  $k|y$  (by  $P(n)$ ),  $k|a$  b/c  $k = \text{GCD}(a,b)$  ✓

$(x,y) \rightarrow (0,x+y)$   $k|x$ ,  $k|y \stackrel{(3)}{\Rightarrow} k|(x+y)$  ✓

$(x,y) \rightarrow (x+y-b, b)$   $k|b$ ,  $k|(x+y) \stackrel{(3)}{\Rightarrow} k|(x+y-b)$  ✓

By induction,  $J$  always holds □

But at  $(4,0)$   $3 \nmid 4$  hence we can never reach it

□

Difference between Finding GCD and Factoring

def: A prime  $p$  is a positive integer whose only divisors are  $1, p$

→ Lemma IF  $p$  is prime and  $p|a \cdot b$ , either  $p|a$  or  $p|b$  or both

Proof: b/c  $p$  is prime, the  $\text{GCD}(a,p)$  is either  $p$  or  $1$

If it's  $p$ , then done  $\Rightarrow p|a$

Else if it's  $1$ , use Fact (4) to conclude  $p|b$

□

Lemma: Let  $p$  be a prime. IF  $p \mid a_1, a_2, \dots, a_n$  then there is some  $i$  where  $p \mid a_i$ .

Theorem:

Every positive integer  $n \geq 2$  can be written uniquely as a product of primes.

$$\text{Ex: } a = p_1 p_2 \dots p_i \quad (p_1 \leq p_2 \leq \dots)$$

$$15 = 2^0 3^1 5^0 \dots \quad \begin{matrix} \text{powers} \\ 0 \\ 1 \\ 0 \end{matrix} \quad \begin{matrix} 2 \\ 3 \\ 5 \end{matrix}$$

$$126 = 2^1 3^2 5^0 7^1 \dots \quad \begin{matrix} \text{powers} \\ 1 \\ 2 \\ 0 \\ 1 \end{matrix} \quad \begin{matrix} 2 \\ 3 \\ 5 \\ 7 \end{matrix}$$

$$\text{GCD} = 2^{\min(0,1)} 3^{\min(1,2)} 5^{\min(0,0)} 7^{\min(0,1)}$$

Lemma:

$$\text{IF } a = p_1^{d_1} p_2^{d_2} \dots \quad \left. \begin{matrix} p_1, p_2, \dots \text{ are all the primes} \\ 2, 3, 5, \dots \end{matrix} \right\}$$

$$b = p_1^{B_1} p_2^{B_2} \dots$$

$$\text{GCD}(a, b) = p_1^{\min(d_1, B_1)} p_2^{\min(d_2, B_2)} \dots$$

Bad way to calculate

$$\text{GCD}(a, b) \stackrel{?}{=} \text{GCD}(\cancel{a+b}, b) \stackrel{(a=qb, b)}{=} (\text{rem}(a, b), b)$$

Division Alg:  $\forall a, b \quad b > 0$ , there is a unique pair  $(q, r)$

$$a = q \underset{\substack{\text{quotient} \\ \vdots}}{b} + \underset{\substack{\text{remainder} \\ r}}{r} \quad 0 \leq r \leq b \quad (\text{shorthand } r = \text{rem}(a, b))$$

Lemma:  $\text{GCD}(a, b) = \text{GCD}(\text{rem}(a, b), b)$

Proof  $\stackrel{\text{one direction}}{\Rightarrow}$  suppose  $k \mid a$  and  $k \mid b$  then  $r = a - qb \Rightarrow k \mid r \checkmark$

$\stackrel{\text{one direction}}{\Leftarrow}$  suppose  $k \mid r$  and  $k \mid b$

$$a = ab + r \Rightarrow k \mid a \checkmark$$

$k \mid a$  and  $k \mid b \Leftrightarrow k \mid r \wedge k \mid b \quad \square$

## 6.042 Recitation #4

Divisibility

$$D(1) \quad a|b \quad b|c \Rightarrow a|c$$

$$D(2) \quad a|b \quad a|c \Rightarrow a|sb + tc \quad \forall s, t$$

$$D(3) \quad \forall c \neq 0 \quad a|b \quad a|cb$$

GCD

$$G(1) \quad \gcd(ka, kb) = k \gcd(a, b)$$

$$G(2) \quad \gcd(a, b) = 1 \quad \gcd(a, c) = 1 \quad \gcd(a, bc) = 1$$

$$G(3) \quad a|bc \quad \gcd(a, b) = 1 \Rightarrow a|c$$

$$G(4) \quad m|a \quad m|b \quad m|\gcd(a, b)$$

$$\gcd(a, b) = sa + tb \quad a \geq b$$

$$\rightarrow \gcd(a, b) = \gcd(b, \text{rem}(a, b))$$

$$a = bq + r \quad \begin{matrix} a/b \\ \downarrow \\ \text{rem}(a, b) = r = a - q \cdot b \end{matrix}$$

$$\gcd(a, b) = \gcd(b, a - qb)$$

The set of common divisors does not change!

$$= \gcd(b, a - b)$$

Euclid's algorithm

EA(a, b):

if  $b == 0$ : return  $a$

else: return ( $\text{rem}(a, b)$ ,  $b$ )

$$\gcd(a, b) = \gcd(b, \text{rem}(a, b))$$

$$\text{Ex } \quad \gcd(259, 70) = \gcd(70, 49)$$

$$= \gcd(49, 21)$$

$$= \gcd(21, 7)$$

$$= \gcd(7, 0)$$

$$= 7$$

$$\text{rem}(259, 70) = 49$$

$$\text{rem}(70, 49) = 21$$

$$\text{rem}(49, 21) = 7$$

$$\gcd(x, 0) = x$$

# Pulverizer aka Extended Euclidian Algorithm

$x$	$y$	$\text{rem}(x, y) = r = x - qy$
259	70	$\text{rem}(259, 70) = 49 = 259 - 3(70)$
70	49	$\text{rem}(70, 49) = 21 = 70 - 1(49)$ $= 70 - 1(259 - 3(70))$ $= -1 \cdot 259 + 4(70)$ $= 7 = 49 - 2(21)$
		$= 7 = 3 \cdot 259 - 11(70)$

$7 = 3(259) - 11(70)$   
 $d = sa + tb$

$$\gcd(a, b) = sa + tb \quad a > 0, b > 0$$

Keeps track of EA to find coefficients and the gcd

## Problems for Recitation 4

### 1 Problem: The Pulverizer! (Extended Euclidian Algorithms)

There is a pond. Inside the pond there are  $n$  pebbles, arranged in a cycle. A frog is sitting on one of the pebbles. Whenever he jumps, he lands exactly  $k$  pebbles away in the clockwise direction, where  $0 < k < n$ . The frog's meal, a delicious worm, lies on the pebble right next to his, in the clockwise direction.

- (a) Describe a situation where the frog can't reach the worm.
- (b) In a situation where the frog can actually reach the worm, explain how to use the Pulverizer to find how many jumps the frog will need.
- (c) Compute the number of jumps if  $n = 50$  and  $k = 21$ . Anything strange? Can you fix it?

## Modular Arithmetic + RSA!

Cryptography - The mathematics of hiding information

Encryption: Encoding a message so only authorized parties <sup>can</sup> read.

Decryption: The process of decoding data

Messages  $\equiv$  numbers

v i c t o r y

22 09 . . . . 25 = 2209032015182513

★ def: We say  $a$  is congruent to  $b$  modulo  $n$   
 denote  $a \equiv b \pmod{n}$   
 if  $n \mid (a-b)$

lemma  $a \equiv b \pmod{n}$  iff  $\text{rem}(a, n) = \text{rem}(b, n)$

Proof: We can use division algorithm:

$$a = q_1 n + r_1, \quad 0 \leq r_1 < n$$

$$b = q_2 n + r_2, \quad 0 \leq r_2 < n$$

Proove  
 $\Rightarrow n \mid (a-b)$      $a-b = (q_1-q_2)n + (r_1-r_2)$

$$r_1-r_2=0 \rightarrow (n-1) \leq (r_1-r_2) \leq n-1$$

Hence  $r_1=r_2$

$\Leftarrow a-b = (q_1-q_2)n + (\cancel{r_1-r_2})^0$

Hence  $n \mid (a-b) \Rightarrow a \equiv b \pmod{n}$



Example:  $29 \stackrel{?}{\equiv} 15 \pmod{7}$

$$\text{rem}(29, 7) = 1 = \text{rem}(15, 7)$$

$$29-15 = 14 \leftarrow \text{Multiple of mod } 7 = 2 \cdot 7$$

What is  $\{a \mid a \equiv 0 \pmod{3}\}$ ? partition of the integers into 3 sets

$$= \{ \dots, -6, -3, 0, 3, 6, \dots \} \leftarrow \text{all the multiples of } 3$$

$$\{a \mid a \equiv 1 \pmod{3}\} \quad (3k+1)$$

$$\{ \dots, -5, -2, 1, 4, 7 \}$$

Properties of Modular Arithmetic (more on handout)

$$a \stackrel{\text{congruent}}{\equiv} a \pmod{n}$$

$$a \equiv b \pmod{n} \text{ and } b \equiv c \pmod{n} \Rightarrow a \equiv c \pmod{n}$$

def: The multiplicative inverse of  $x \pmod{n}$  denoted by  $x^{-1}=y$  is a number s.t.

$$x \cdot y \equiv 1 \pmod{n}$$

$$(a) \quad x=7, n=5, y=3$$

$$7y \equiv 1 \pmod{5}$$

Looking for

$$(b) \quad x=3, n=6, y=$$

$$3y \equiv 1 \pmod{6}$$

What if  $y > 6$ ?

$$y = y + 6$$

$$2 \cdot 3 \equiv 4 \cdot 3 \pmod{6}$$

$$(c) \quad x=5, n=6$$

$$5y \equiv 1 \pmod{6}$$

$$y=5$$

Remember:  
 $a \equiv b \pmod{n}$   
 iff  
 $n \mid (a-b)$

Lemma: Any integer  $x$  not a multiple of a prime  $p$  has an inverse modulo  $p$ .

Proof: From last lecture (recitation)

There are integers  $s, t$

$$\cancel{sx + tp^0} = \text{GCD}(x, p) = 1 \pmod{p}$$

(a)  $\text{GCD}(7, 5) = 1$

<sup>no inverse</sup> (b)  $\text{GCD}(3, 6) = 3$

(c)  $\text{GCD}(5, 6) = 1$

Thm [Fermat]:

If  $p$  is prime  $x \equiv 0 \pmod{p}$

$$x^{p-1} \equiv 1 \pmod{p}$$

$$x^{p-2} x \equiv 1 \pmod{p}$$

## Cryptography

#1: Sender and receiver agree on large prime  $p$  <sup>(public)</sup> and secret key  $k \in \{1, 2, \dots, p-1\}$  <sup>(private)</sup>  $\rightarrow k$

Encryption:

$$m^* = \underset{\text{scrambled}}{\text{rem}}(mk, p) \quad m \in \{1, 2, \dots, p-1\}$$

so:  $m^* = mk \pmod{p}$

Decryption:

$$k^{p-2} \quad (\text{inverse of } k \text{ modulo } p)$$

$$m^* k^{p-2} = (mk) k^{p-2} = m \pmod{p}$$

Def: Euler's  $\phi(n) \equiv \# \text{m's} \in \{1, 2, \dots, n-1\}$  where  $\text{GCD}(m, n) = 1$

Example:  $\underset{\text{Prime}}{\phi(p)} = p-1$       Fact  $\underset{\text{Primes}}{\phi(pq)} = (p-1)(q-1)$

Thm: <sup>Euler</sup> Modulo  $n$  with  $\text{GCD}(x, n) = 1$

$$x^{\phi(n)} \equiv 1 \pmod{n}$$

## RSA

Each player generates a pair of primes  $(p_i, q_i)$

Let  $n_i = p_i q_i$ , select integer  $e_i$  s.t.

$$\text{GCD}(e_i, (p_i-1)(q_i-1)) = 1$$

Public Key  $(e_i, n_i)$

Private Key  $[d_i]^{e_i \text{ inverse}} \equiv 1 \pmod{(p_i-1)(q_i-1)}$

Encryption:

$$m^* = \text{rem}(m^{e_i}, n_i) \\ \equiv m^{e_i} \pmod{n_i}$$

Decryption:

$$(m^*)^{d_i} \equiv (m^{e_i})^{d_i} \equiv m^{e_i d_i} \pmod{n} \\ \equiv m^{k(p_i-1)(q_i-1)+1} \pmod{n} \\ = m \pmod{n_i}$$

# 6.042 Notes

## Chapter 4

### Cryptograph

Turing's Code

- 1) Turn message  $m$  into int
- 2) Encrypt message with secret key  $k$  prime  
 $m^* = m \cdot k$

both parties know  
 $k$  prime

- 3) Decrypt message using  $k$

$$\frac{m^*}{k} = m$$

- Works because it is very difficult to factor large prime #'s  
However if a second message is sent with same  $k$ ,  
 $\Rightarrow \gcd(m_1^*, m_2^*) = k$   
Easy to compute.

### Modular Arithmetic

$$a \equiv b \pmod{n} \Leftrightarrow n \mid (a - b) \Leftrightarrow \text{rem}(a, n) = \text{rem}(b, n)$$

$$a \equiv \text{rem}(a, n) \pmod{n}$$

New Turing Code

Encoding:  $m^* = \text{rem}(m \cdot k, p)$  key, public large prime

Decoding:

$$\begin{aligned} m^* \cdot k^{-1} &= \text{rem}(mk, p) \cdot k^{-1} \\ &\equiv (mk)k^{-1} \pmod{p} \\ &\equiv m \pmod{p} \quad \text{From Euler's theorem} \\ \Rightarrow m &= \text{rem}(m^* k^{-1}, p) \end{aligned}$$

\* Can't cancel multiplicative terms unless using modulo of prime

## Euler's Theorem

$\phi(n) = \# \text{ of int that are relatively prime to } n$   
 $\text{gcd}(a, b) = 1$

Thm 4.7.4

prime factors  $n = p_1 \dots p_a$   
 $\Rightarrow \phi(n) = n \left(1 - \frac{1}{p_1}\right) \left(1 - \frac{1}{p_2}\right) \dots \left(1 - \frac{1}{p_a}\right)$

Example

$$\begin{aligned}\phi(300) &= \phi(2^2 \cdot 3 \cdot 5^2) \\ &= 300 \left(1 - \frac{1}{2}\right) \left(1 - \frac{1}{3}\right) \left(1 - \frac{1}{5}\right) \\ &= 80\end{aligned}$$

If  $n = \prod_{\substack{\text{primes} \\ p_i}} p_i^q \Rightarrow \phi(n) = (p_1-1)(p_2-1)\dots(p_a-1)$

Euler's Theorem

$$k^{\phi(n)} \equiv 1 \pmod{n}$$

$\sim k^{\phi(n)-1}$  inverse of  $k \pmod{n}$

## 2 RSA: Let's try it out!

You'll probably need extra paper. *Check your work carefully!*

1. As a team, go through the **beforehand** steps.

- (a) Choose primes  $p$  and  $q$  to be relatively small, say in the range 5-15. In practice,  $p$  and  $q$  might contain several hundred digits, but small numbers are easier to handle with pencil and paper.

**Solution.** We choose  $p = 7$  and  $q = 11$  for our example. ■

- (b) Calculate  $n = pq$ . This number will be used to encrypt and decrypt your messages.

**Solution.** In our example,  $n = pq = 77$ . ■

- (c) Find an  $e > 1$  such that  $\gcd(e, (p-1)(q-1)) = 1$ .

The pair  $(e, n)$  will be your *public key*. This value will be broadcast to other groups, and they will use it to send you messages.

**Solution.** In our example,  $p-1 = 6 = 2 \cdot 3$  and  $q-1 = 10 = 2 \cdot 5$ . Therefore, any  $e$  that is odd and neither a multiple of 5 nor 3 would work. We choose  $e = 13$ . ■

- (d) Now you will need to find a  $d$  such that  $de \equiv 1 \pmod{(p-1)(q-1)}$ .

- Explain how this could be done using the Pulverizer. (Do not carry out the computations!)

**Solution.** We can rewrite the equation  $de \equiv 1 \pmod{(p-1)(q-1)}$  to read  $de - 1 = k(p-1)(q-1)$  for some integer value  $k$ . Rearranging this yields the equation  $de - k(p-1)(q-1) = 1$ . Because  $\gcd(e, (p-1)(q-1)) = 1$ , we know such a linear combination of  $e$  and  $(p-1)(q-1)$  exists! Using the Pulverizer will give us the coefficient  $d$ , and then we can adjust  $d$  to be positive using techniques from class. In this case  $d = -23$ , which can be adjusted to 37. ■

- Find  $d$  using Euler's Theorem given in yesterday's lecture.

The pair  $(d, n)$  will be your *secret key*. Do not share this with anybody!

**Solution.** Since  $e$  and  $(p-1)(q-1)$  are relatively prime, we can claim by Euler's Theorem that  $e^{\phi((p-1)(q-1))} \equiv 1 \pmod{(p-1)(q-1)}$  and hence  $e^{\phi((p-1)(q-1))-1} \cdot e \equiv 1 \pmod{(p-1)(q-1)}$ .

This means  $d = e^{\phi((p-1)(q-1))-1}$  is an *inverse* of  $e \pmod{(p-1)(q-1)}$ . To find the value of  $d$ , we first calculate  $\phi((p-1)(q-1))$ . In our example, the factorization of  $(p-1)(q-1)$  is  $2^2 \cdot 3 \cdot 5$ , so  $\phi((p-1)(q-1)) = (2^2 - 2^1)(3^1 - 3^0)(5^1 - 5^0) = 2 \cdot 2 \cdot 4 = 16$ . We substitute  $e$  and  $\phi((p-1)(q-1))$  into our equation to get  $d = 13^{16-1} = 13^{15}$ .

$13^{15}$  is a huge number! Therefore, we must reduce  $d$  to something more manageable using *repeated squaring*. In our example, we square 13 to get  $13^2 = 169 \equiv 49$

## RSA Public-Key Encryption

**Beforehand** The receiver creates a public key and a secret key as follows.

1. Generate two distinct primes,  $p$  and  $q$ .
2. Let  $n = pq$ .
3. Select an integer  $e$  such that  $\gcd(e, (p-1)(q-1)) = 1$ .  
The *public key* is the pair  $(e, n)$ . This should be distributed widely.
4. Compute  $d$  such that  $de \equiv 1 \pmod{(p-1)(q-1)}$ .  
The *secret key* is the pair  $(d, n)$ . This should be kept hidden!

**Encoding** The sender encrypts message  $m$  to produce  $m'$  using the public key:

$$m' = \text{rem}(m^e, n)$$

**Decoding** The receiver decrypts message  $m'$  back to message  $m$  using the secret key:

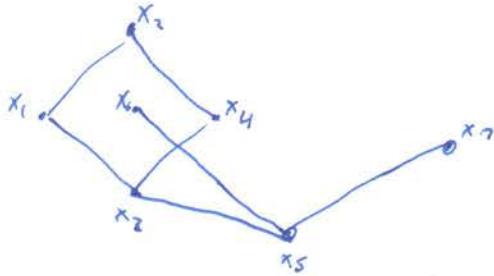
$$m = \text{rem}((m')^d, n).$$

## Graph Theory

Claim (U. Chicago): on average, men have 74% more opposite-gender partners than women

Claim (ABC news): 233%

### Graphs



Def: A graph  $G$  is a pair of sets  $(V, E)$ , where  $V$  is a non-empty set of vertices (or nodes) &  $E$  is a set of 2 element subsets of  $V$  called edges.

$$\text{Ex: } V: \{x_1, x_2, \dots, x_7\}$$

$$E: \{ \{x_1, x_2\}, \{x_1, x_3\}, \dots \}$$

aka, they are connected by an edge.

Def: Two nodes are adjacent if  $\{x_i, x_j\} \in E$

Def: An edge  $\{x_i, x_j\}$  is incident to  $x_i$  &  $x_j$

Def: # of edges incident to a node is the degree of node.

$$\text{Ex. } \deg(x_5) = 3$$



Def: A graph is simple if it has no loops or multiple edges.

Def:  $A_m$  = average # of opposite gender partners for men

$$A_w = \dots \quad \text{men} \qquad \dots \quad \text{women}$$

$$\text{What is } A_m/A_w? \quad 1.74? \quad 3.33?$$

$$A_m = \frac{\sum_{x \in V_m} \deg(x)}{|V_m|} = \frac{|E|}{|V_m|}$$

$$A_w = \frac{\sum_{x \in V_w} \deg(x)}{|V_w|} = \frac{|E|}{|V_w|}$$

$$\frac{A_m}{A_w} = \frac{|E| / |V_m|}{|E| / |V_w|} = \frac{|V_w|}{|V_m|} = \underline{1.0475}$$

### Graph Coloring Problem:

Given a graph  $G$  and  $k$  colors. Assign a color to each node so adjacent nodes get different colors.

Def: The min value of  $k$  for which coloring exists

- Chromatic Number  $\chi(G)$  of  $G$ .

↳ very difficult to calculate efficiently

P  $\neq$  NP problem

↑ Can check in polynomial time  
Can solve in polynomial time.

### "Basic" Coloring Algorithm for $G = (V, E)$

1) Order nodes  $V_1, V_2, V_3, \dots, V_n$  (Large deg first)

2) Order colors  $C_1, C_2, C_3$

"Greedy Algorithm"

3) For  $i=1, 2, \dots, n$ : Assign  $v_i$  the lowest legal color

in an  $n$  node graph

Thm: For all  $d$ , if every node  $v$  in  $G$  has degree  $\leq d$ , then  
Basic alg. uses at most  $d+1$  colors for  $G$ .

Proof: (by induction) ↗ induct on # of nodes

I.H.  $\text{flat } P(n)$

Base case :  $n=1 \Rightarrow 0 \text{ edges} \Rightarrow d=0$   
 $| \text{color}| = d+1 \quad \checkmark$

Inductive Step : Assume  $P(n)$  for purposes of induction.

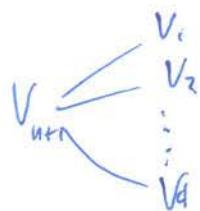
Let  $G = (V, E)$  be any  $(n+1)$ -node graph & let  $d$  be the max deg in  $G$ .

Order the nodes :  $V_1, V_2, \dots, V_n, V_{n+1}$

Remove  $V_{n+1}$  from  $G$  to create  $G' = (V', E')$

$G'$  has max deg  $d$  &  $n$  nodes so  $P(n)$  says that  
Basic Alg uses  $\leq d+1$  colors on  $V_1, V_2, \dots, V_n$

$V_{n+1}$  has  $\leq d$  neighbors



$\exists$  color in  $\{c_1, c_2, \dots, c_{d+1}\}$  not used  
by any neighbor

Give  $V_{n+1}$  that unused color

$\Rightarrow$  Basic Alg uses  $\leq d+1$  colors



## Graph coloring problem

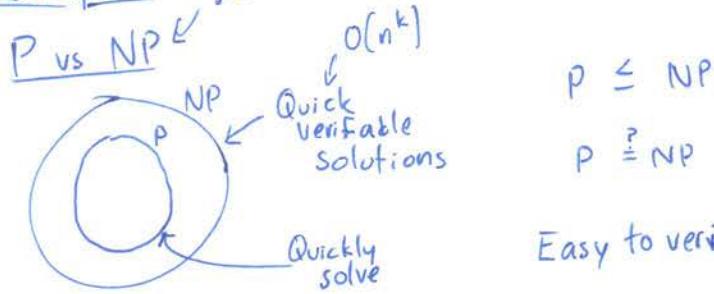
Given a graph  $G \in k$  colors assign colors to each  $v_i \in V$  s.t

$$\begin{array}{c} \textcircled{1} - \textcircled{2} \\ v_i \quad v_j \end{array} \quad c(v_i) \neq c(v_j)$$

min value of  $k$ (colors) you need to satisfy GCP. (chromatic #)

Handshake:  $\sum_{v \in V} \deg(v_i) = 2|E|$

NP-Complete Non-Deterministic Polynomial time



4  $\rightarrow$  P

Problem 1:

$P(n) = \text{any } n\text{-node } G \text{ with width } \leq w, (w+1)\text{-colorable}$

$$P(n) \Rightarrow P(n+1)$$

~~width~~  
v<sub>1</sub> v<sub>2</sub> ... v<sub>n</sub>

Base case    n=1     $\checkmark$     1     $\checkmark$   
                  w=0

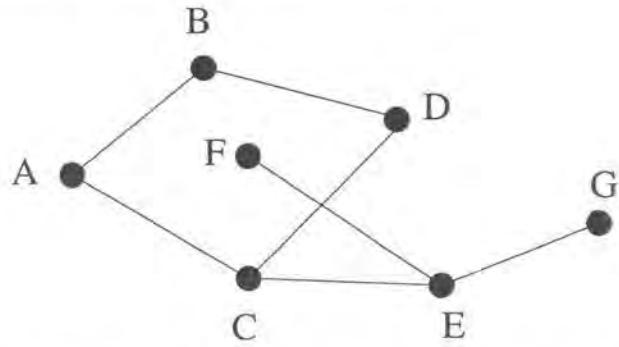
$P(n+1) \quad G, \text{ width } \leq w \Rightarrow (w+1)\text{-colorable}$

See solutions.

## Problems for Recitation 6

### 1 Graph Basics

Let  $G = (V, E)$  be a graph. Here is a picture of a graph.



Recall that the elements of  $V$  are called vertices, and those of  $E$  are called edges. In this example the vertices are  $\{A, B, C, D, E, F, G\}$  and the edges are

$$\{A-B, B-D, C-D, A-C, E-F, C-E, E-G\}.$$

Deleting some vertices or edges from a graph leaves a *subgraph*. Formally, a subgraph of  $G = (V, E)$  is a graph  $G' = (V', E')$  where  $V'$  is a nonempty subset of  $V$  and  $E'$  is a subset of  $E$ . Since a subgraph is itself a graph, the endpoints of every edge in  $E'$  must be vertices in  $V'$ . For example,  $V' = \{A, B, C, D\}$  and  $E' = \{A-B, B-D, C-D, A-C\}$  forms a subgraph of  $G$ .

In the special case where we only remove edges incident to removed nodes, we say that  $G'$  is the *subgraph induced on  $V'$*  if  $E' = \{(x-y | x, y \in V' \text{ and } x-y \in E)\}$ . In other words, we keep all edges unless they are incident to a node not in  $V'$ . For instance, for a new set of vertices  $V' = \{A, B, C, D\}$ , the induced subgraph  $G'$  has the set of edges  $E' = \{A-B, B-D, C-D, A-C\}$ .

### 2 Problem 1

An undirected graph  $G$  has *width w* if the vertices can be arranged in a sequence

$$v_1, v_2, v_3, \dots, v_n$$

such that each vertex  $v_i$  is joined by an edge to at most  $w$  preceding vertices. (Vertex  $v_j$  precedes  $v_i$  if  $j < i$ .) Use induction to prove that every graph with width at most  $w$  is  $(w + 1)$ -colorable.

(Recall that a graph is  $k$ -colorable iff every vertex can be assigned one of  $k$  colors so that adjacent vertices get different colors.)

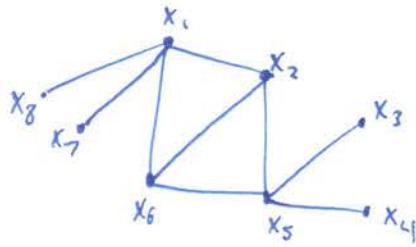
### 3 Problem 2

A **planar graph** is a graph that can be drawn without any edges crossing.

1. First, show that any subgraph of a planar graph is planar.
2. Also, any planar graph has a node of degree at most 5. Now, prove by induction that any graph can be colored in at most 6 colors.

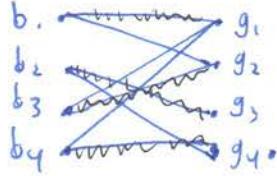
Def: Given a graph  $G = (V, E)$ , a matching is a subgraph of  $G$  where every node has degree 1.

$\{x_1-x_6, x_2-x_5\}$  is a matching of size 2



Def: A matching of a Graph  $G = (V, E)$  is perfect if it has size  $\frac{|V|}{2}$ . Aka everybody gets married.

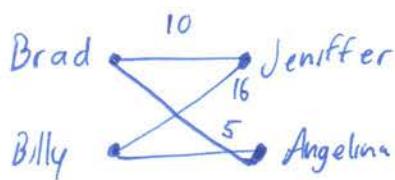
Ex.



Def: The weight of a matching  $M$  is the sum of the weights on the edges in  $M$

Def: A min-weight-matching for  $G$  is a perfect matching for  $G$  with minimum weight.

Ex.



Brad-Jenn  
Billy-Angelina } 20

Def: Given a matching  $M$ ,  $x \notin y$  are a rogue couple in  $M$  if  $x \leftrightarrow y$  prefer each other to their mates in  $M$ .

Def: A matching is stable if no rogue couples.

## The Mating Algorithm (TMA)

**Initial Condition:** Each of the  $N$  boys has an ordered list of the  $N$  girls according to his preferences. Each of the girls has an ordered list of the boys according to her preferences.

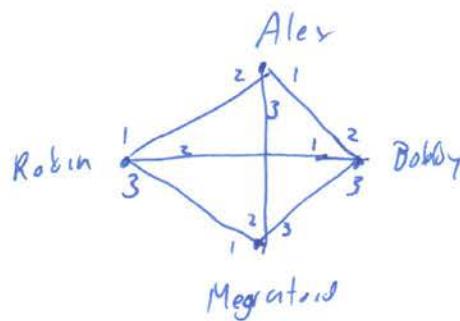
**Each Day:**

- Morning:
  - Each girl stands on her balcony
  - Each boy stands under the balcony of his favorite girl whom he has not yet crossed off his list and serenades. If there are no girls left on his list, he stays home and does 6.042 homework.
- Afternoon:
  - Girls who have at least one suitor say to their favorite from among the suitors that day: “Maybe, come back tomorrow.”
  - To the others, they say “No, I will never marry you!”
- Evening:
  - Any boy who hears “No” crosses that girl off his list.

**Termination Condition:** If there is a day when every girl has at most one suitor, we stop and each girl marries her current suitor (if any).

Goal: Find a perfect match that is stable

Thm:  $\neg \exists$  a stable match in



Proof: (By contradiction)

Assume (for purposes of contradiction) that  $\exists$  a stable matching  $M$ .

$\Rightarrow$  Megatronid is matched to someone in  $M$

$\Rightarrow$  Without loss of generality (WLOG) (by symmetry)

assume Megatronid matches to Alex in  $M$ .

$\Rightarrow$  Alex  $\in$  Robin are a rogue couple in  $M$ .

$\Rightarrow M$  is not stable  $\#$

□

## Stable Marriage Problems

•  $N$  boys &  $N$  girls

• Each boy has his own ranked list of all the girls

• " girl ... her ... - - - boys

Goal: Find stable matching

### Pref Lists

B  
1: C BEAD  
2: ABECD  
3: DCBAE  
4: ACDBE  
5: ABDEC

A  
A: 35214  
B: 52143  
C: 43512  
D: 12345  
E: 23415

### greedy Algorithm

1  $\rightarrow$  C  
2  $\rightarrow$  A  
3  $\rightarrow$  D  
4  $\rightarrow$  B  
5  $\rightarrow$  E

4, 5 rogue :-

girl	The Mating Algorithm (see handout)				Boys	Crossouts				
	Day 1	Day 2	Day 3	Day 4		Day 1	Day 2	Day 3	Day 4	Day 5
A	2, 4, 5	5	5	5	1	C				
B		2	2, 1	2	2	A				
C	1	1, 4	4	4	3					
D	3	3	3	3	4					
E				1	5	A				

$$\begin{array}{ll}
 A \rightarrow & 5 \\
 B \rightarrow & 2 \\
 C \rightarrow & 4 \\
 D \rightarrow & 3 \\
 E \rightarrow & 1
 \end{array}
 \quad \text{Stable!}$$

Need to show

- TMA terminates (quickly)
- Everyone gets married
- No rogue couples

Thm 1: TMA terminates within  $N^2 + 1$  days.

Proof: (By contradiction)

Suppose (for the purposes of contradiction) that the algorithm does not terminate after  $N^2 + 1$  days.

Claim: If we don't terminate on a day, then at least one boy crosses a girl off his list that evening.

$\Rightarrow \geq N^2 + 1$  crossouts

$N$  lists with  $N$  names  $\Rightarrow \leq N^2$  crossouts

#  $\square$

Lemma 1: If a girl rejects  $B$ , then henceforth  $g$  always has a suitor who she prefers to  $B$ .

Proof (By induction):

Thm 2: Everyone is married in TMA

Proof: (By contradiction). Assume for purposes of contradiction that Not everyone gets married at termination.

$\Rightarrow \exists$  boy  $B$  who is not married at termination

$\Rightarrow B$  was rejected by every girl

$\Rightarrow$  Every girl has a suitor at termination (by Lemma 1)

$\Rightarrow$  Every girl gets married

$\Rightarrow$  Every boy got married. Including  $B$ . #

□

Thm 3: TMA produces a stable matching

Proof (By contradiction). Assume (for the purposes of contradiction) TMA does not produce a stable matching

$\Rightarrow \exists$  rogue couple - call them Bob and Sail

Case 1: Bob serenaded Sail

$\Rightarrow$  Sail rejected Bob

$\Rightarrow$  Sail prefers husband to Bob (By Lemma 1)

$\Rightarrow$  Sail and Bob are not Rogue #

Case 2: Bob did not serenade Sail

$\Rightarrow$  Bob married someone he prefers to Sail

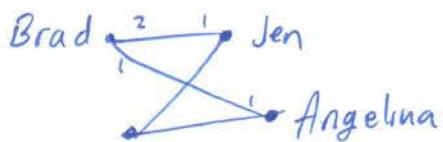
$\Rightarrow$  Bob and Sail are not rogue #

□

Let  $S$  = set of all stable matchings ( $S \neq \emptyset$ )

Def: For each person  $P$ , the realm of possibilities  
For  $P$  is  $\{Q \mid \exists m \in S, \{P, Q\} \in m\}$

Ex



Def: A person's optimal mate is his/her favorite from the realm of possibility.

Def: A person's pessimal mate is least favorite in ...

Thm 4: TMA pairs every boy with his optimal mate

Thm 5: TMA pairs every girl with her pessimal mate.

Proofs: In recitation

## Problems for Recitation 6

### Problem 1

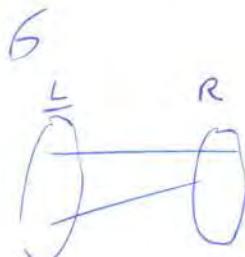
Recall that a graph is called  $d$ -regular if every vertex in the graph has degree exactly  $d$ . Let  $G = (V, E)$  be a  $d$ -regular bipartite graph, with the same number of vertices in the left part  $L$  as in the right part  $R$ .

Prove, using Hall's theorem and induction, that  $G$  can be partitioned into  $d$  perfect matchings. In other words, we can find  $E_1, E_2, \dots, E_d \subseteq E$ , all disjoint ( $E_i \cap E_j = \emptyset$ ) and which together form  $E$ , so that  $E_i$  is a perfect matching of  $G$  for each  $1 \leq i \leq d$ .

$$\begin{aligned} P(d) &= G, B_i, d\text{-reg}, L, R, \frac{|V|}{2} \\ &\Rightarrow E_1, \dots, E_d \subseteq E \end{aligned}$$

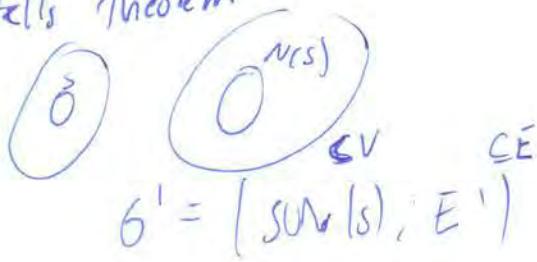
$$P(0) \checkmark$$

$$\begin{aligned} P(d+1) &= G, B_i, d+1\text{-reg}, L, R, \frac{|V|}{2} \\ &\stackrel{?}{\Rightarrow} E_1, \dots, E_d, E_{d+1} \end{aligned}$$



If we knew one matching  $\rightarrow$  we can remove it and be left with a  $d$ -regular.  $\checkmark$  By IH can then add the removed matching and done.

Hall's theorem



$$\begin{aligned} |E'| &= |S| (d+1) \\ &\Rightarrow |E'| = \sum_{v \in N(s)} \deg(v) \leq (d+1)|N(s)| \\ &\Rightarrow |S|(d+1) \leq |N(s)|(d+1) \\ &|S| \leq |N(s)| \end{aligned}$$

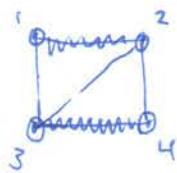
□

## Problem 3

Similarly to the previous problem, we say that the *pessimai mate* of girl  $j$  is her least favorite boy from the set  $P_j$  of boys she can be matched to in some stable matching.

Prove that The Mating Algorithm returns a matching where every girl is matched with her pessimai mate.

## Matching



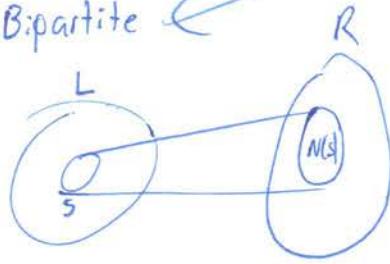
$$M = \{(1,2), (3,4)\} \Rightarrow \begin{matrix} 1 & 2 \\ 3 & 4 \end{matrix}$$

$\forall m \in M \quad m \leq E$

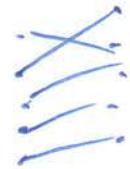
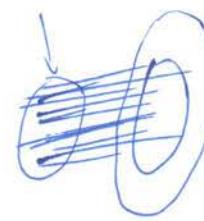
*everything connects from one side to the other.*

Halls Theorem

Bipartite



$\Rightarrow$



*everything has an edge.*

$$\forall s \subseteq L, |s| \leq |N(s)|$$

$\Rightarrow \exists \text{ matching } \subseteq E \text{ covers } L$

## Connectivity and Spanning Trees

Theme: Graphs can be simple but powerful objects for reasoning about discrete problems.

Connectivity: getting from  $u$  to  $v$  by traversing edges.

Question: What sorts of real-world graphs would we care about.

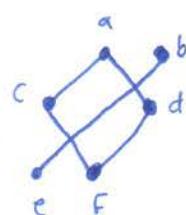
### Minimum Spanning Tree (MST)

What is the min cost set of edges we need to keep that still allow every vertex to reach every other vertex.

Def: A walk in  $G$  is a sequence of vertices  $v_0, v_1 \dots v_k$  and edges  $\underbrace{(v_0, v_1), (v_1, v_2) \dots (v_{k-1}, v_k)}$  where each edge is in  $G$

length- $k$  walk

(undirected)  $G = (V, E)$



ex (of walks): acFd  
acFdFcF

Def: A path is a type of walk where all the  $v_i$ 's are different.

Def: A cycle is walk where  $(v_0 = v_k)$  and  $v_0, v_1 \dots v_{k-1}$  are different

Ex: acFdFa

Lemma 1: IF there is a walk in  $G$  from  $u$  to  $v$   
there is a path from  $u$  to  $v$ .

Proof: Let  $u = v_0, v_1, \dots, v_k = v$  be the shortest walk from  $u$  to  $v$ .  
For the purposes of contradiction suppose it is not a path  
 $\Rightarrow$  Some vertex  $w$  is repeated.

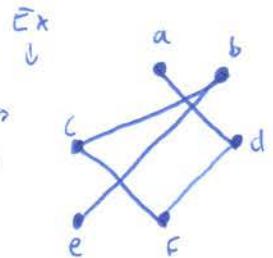
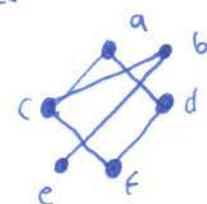
$$v_1, v_2, \dots, \boxed{v_j \dots v_j}, \dots, v_k$$

$\hookrightarrow$  can remove this and make a shorter path.

#yoloSWAG  $\square$

Def:  $u$  and  $v$  are connected in  $G$  if there is a path from  $u$  to  $v$ .

Def:  $G$  is connected if every pair of vertices is connected.



Def: A tree is a connected graph that has no cycles.

Def: A spanning tree of  $G$  is a subgraph that is a tree and has the same set of vertices.

Minimum Spanning Tree (MST): Given a connected, undirected graph  $G$  where each  $e = (u, v)$  has cost  $c(u, v)$ .

Goal: Find a spanning tree with minimum total cost ( $\triangleq$  sum of edge cost)

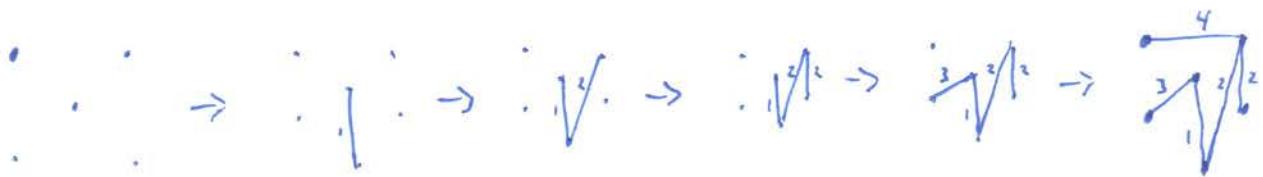
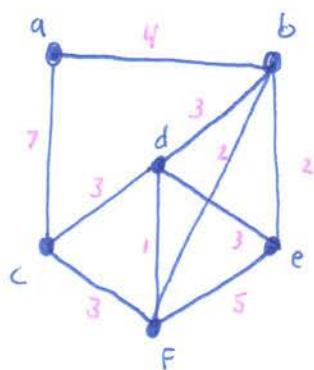
Kraskal's Algorithm

set  $S = \emptyset$   
while there is an edge that can be added to  $S$  w/o creating a cycle

Add the min cost (or tied for min) For such edge

End

$$G = (V, E)$$



Lemma 2: For each step  $t \geq 0$ , let  $S$  be the first  $t$  edges added by Kruskal's Algorithm, then there is some MST  $T = (V, E_T)$  s.t.  $S \subseteq E_T$

Proof (by induction)

$P(t)$  = "Let  $S$  be the first  $t$  edges, then there is some MST  $T = (V, E_T)$  s.t.  $S \subseteq E_T$ "

Base Case:  $P(0)$   $S \neq \emptyset$  for any MST,  $S \subseteq E_T$

Inductive Step:  $P(t) \Rightarrow P(t+1)$

Suppose  $t+1$  s.t. edge is  $x$

$S, P(t)$  holds  $\Rightarrow$  MST  $T = (V, E_T)$   $S \subseteq E_T$

Case #1:  $x \in E_T \Rightarrow S \cup S_{[x]} \subseteq E_T$   $P(t+1)$  is True

Case #2:  $x \notin E_T$

What happens if we add  $x$  to  $P$

There is some cycle (contains  $x$ )

Let  $T' = (V, E')$  by deleting  $(T') \subseteq c(T)$  and yet it is a spanning tree

## Some useful facts about connectivity/trees

Throughout this handout, we will assume that  $G = (V, E)$  is an *undirected* graph.

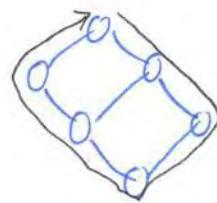
### Connectivity

- C1.  $u$  is connected to  $v$  if and only if  $v$  is connected to  $u$
- C2. If  $u$  is connected to  $w$  and  $v$  is connected to  $w$  then  $v$  is connected to  $u$ .

### Trees

- T1. If  $G$  is connected then it has a subgraph (obtained by deleting edges) on the same set of vertices that is a tree. This is called a *spanning tree*.
- T2. A graph on  $n$  nodes is a tree if and only if it is connected and has  $n - 1$  edges.
- T3. A graph on  $n$  nodes is a tree if and only if it has no cycles and has  $n - 1$  edges.
- T4. In a tree  $T$ , each pair of nodes  $u$  and  $v$  has a unique path from  $u$  to  $v$ .

- Def Hamiltonian cycle
  - touch every vertex

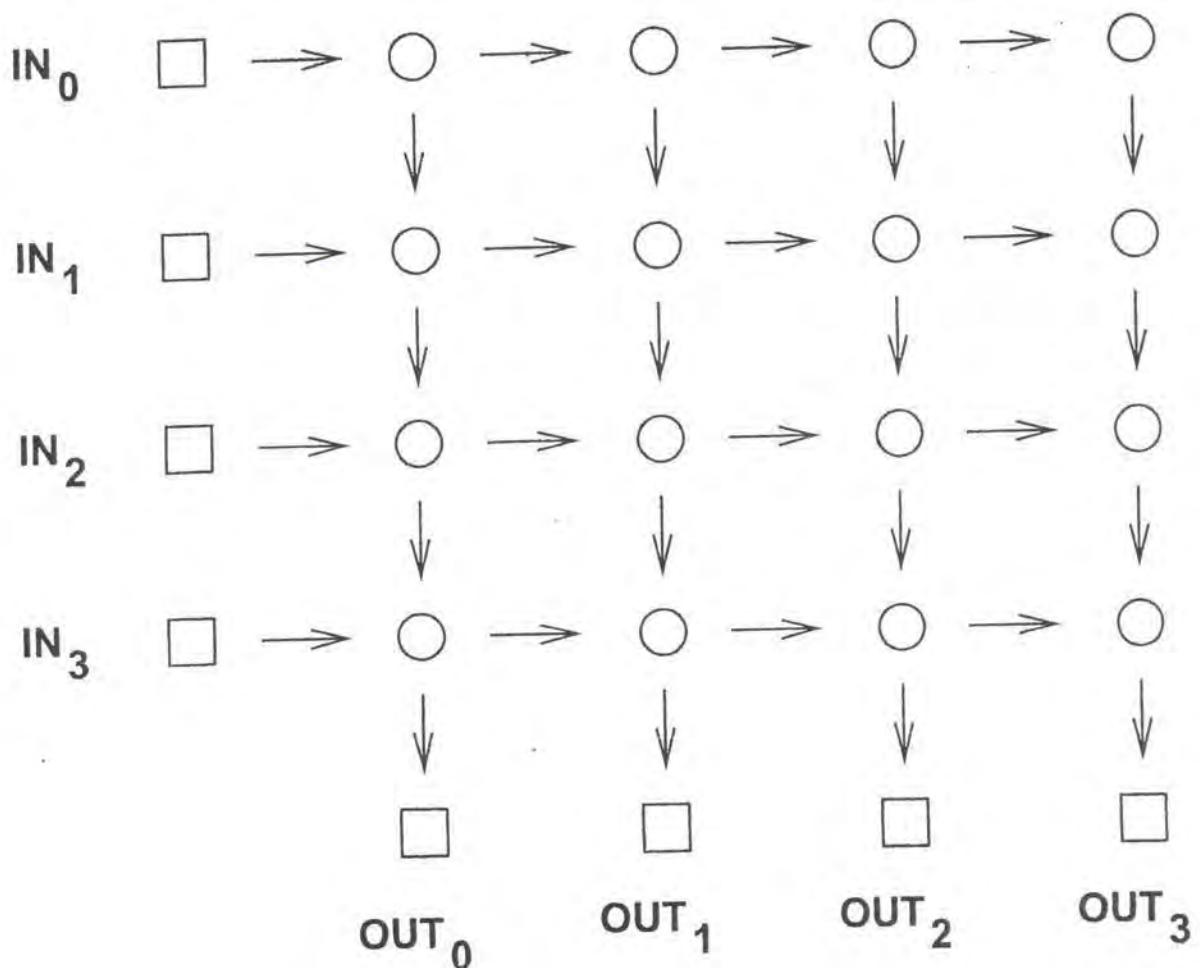


Def Euler walk Traverse every edge once

Euler Tour: end in the start

Fernando Trujano

**Two-Dimenional Array ( $N = 4$ )**  
(a.k.a. Grid, Crossbar)

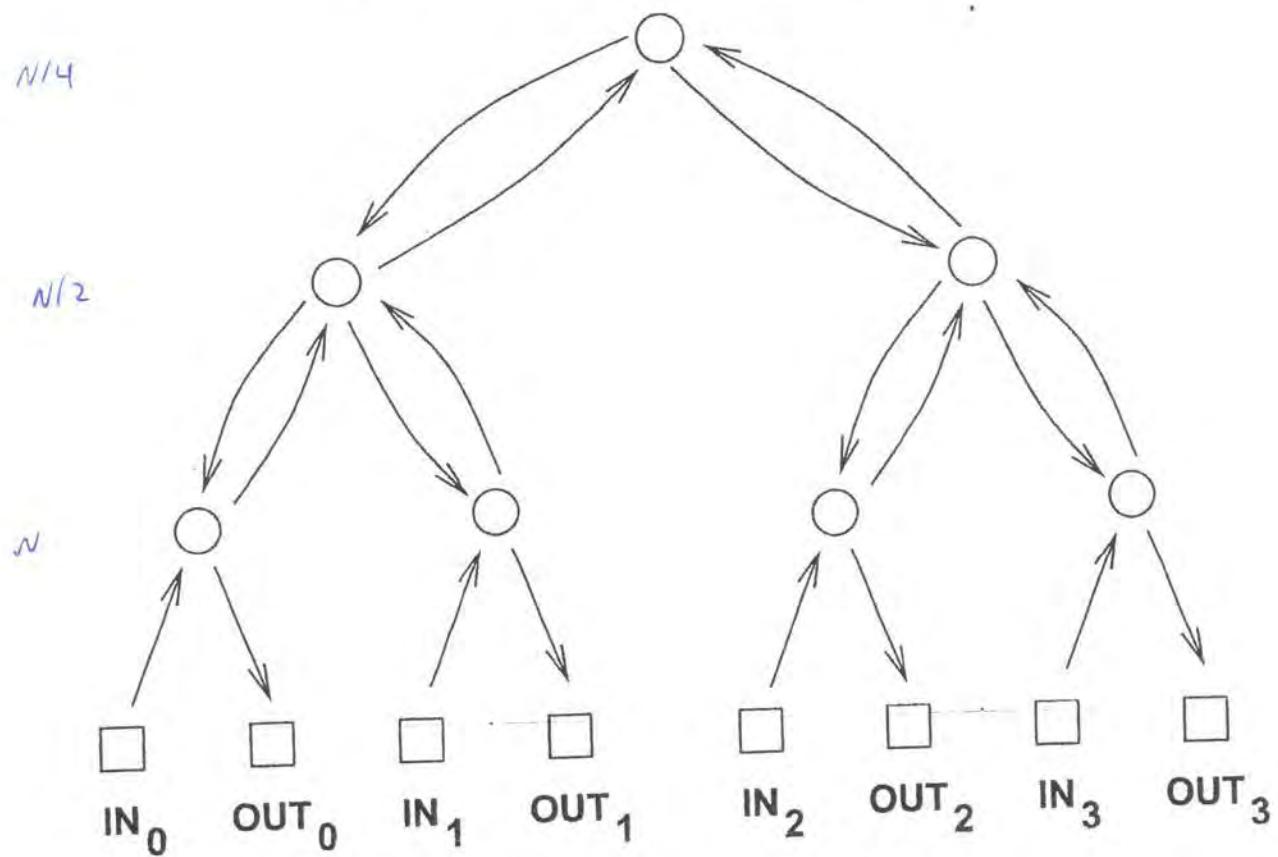


Diameter: worst case

$$\log n + 1$$

Going up  
Going down

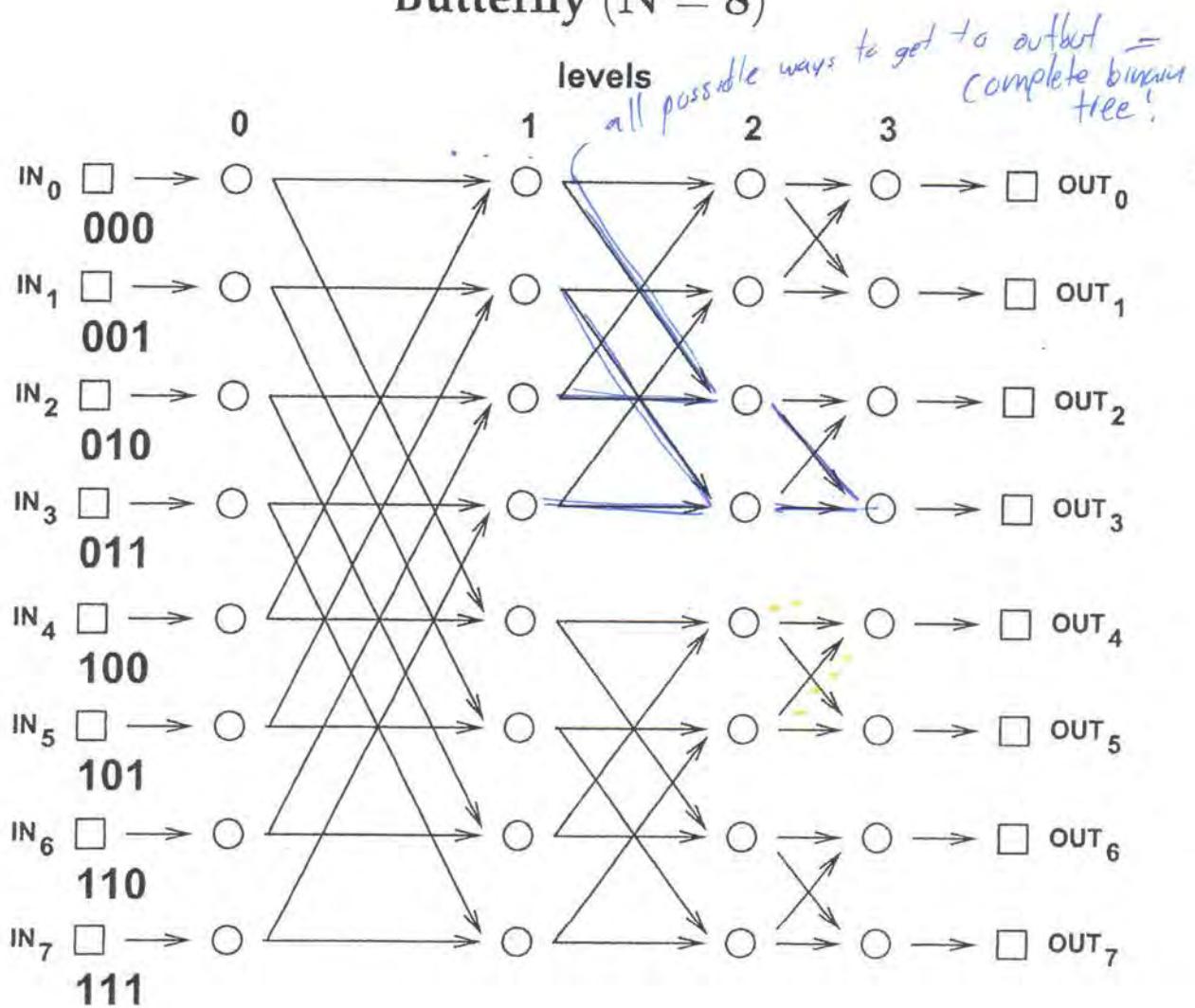
### Complete Binary Tree ( $N = 4$ )



$$N(1 + \frac{1}{2} + \frac{1}{4} + \dots + \frac{1}{n}) \\ = 2N - 1$$

can change bit at each level

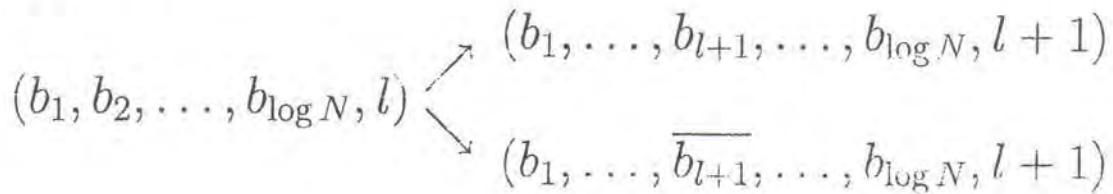
## Butterfly ( $N = 8$ )



Nodes:

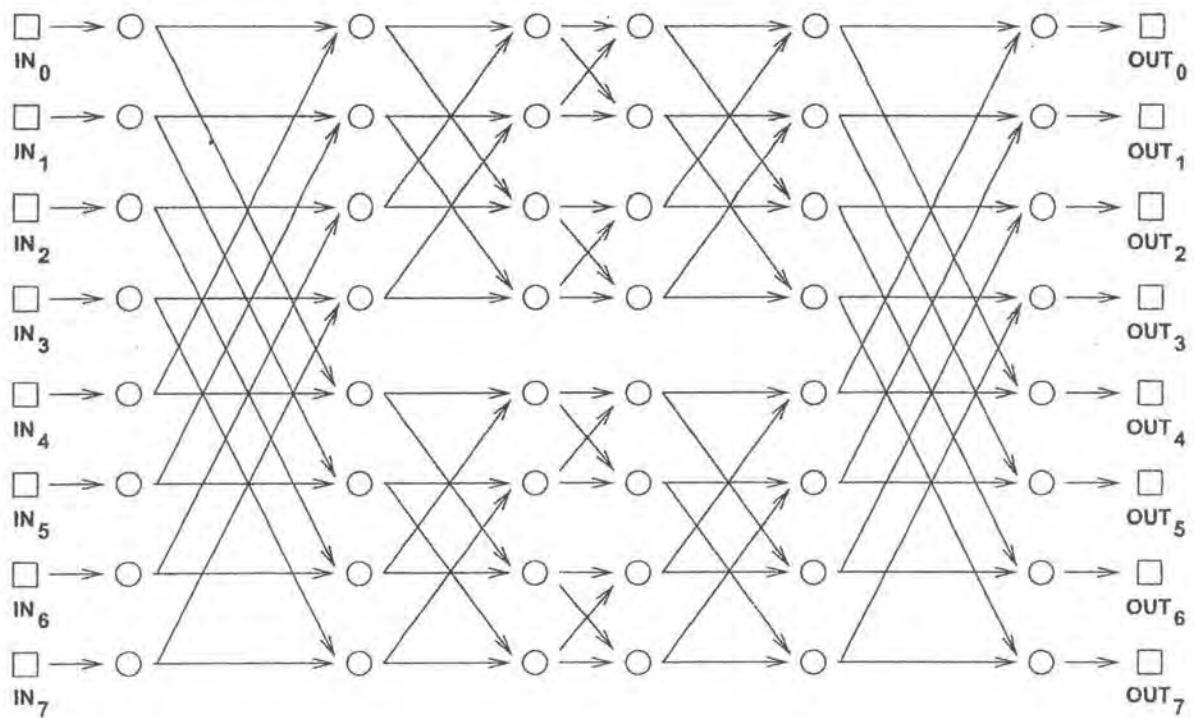
$$\{(b_1, b_2, \dots, b_{\log N}, l) \mid b_i = 0 \text{ or } 1, 0 \leq l \leq \log N\}$$

Edges:

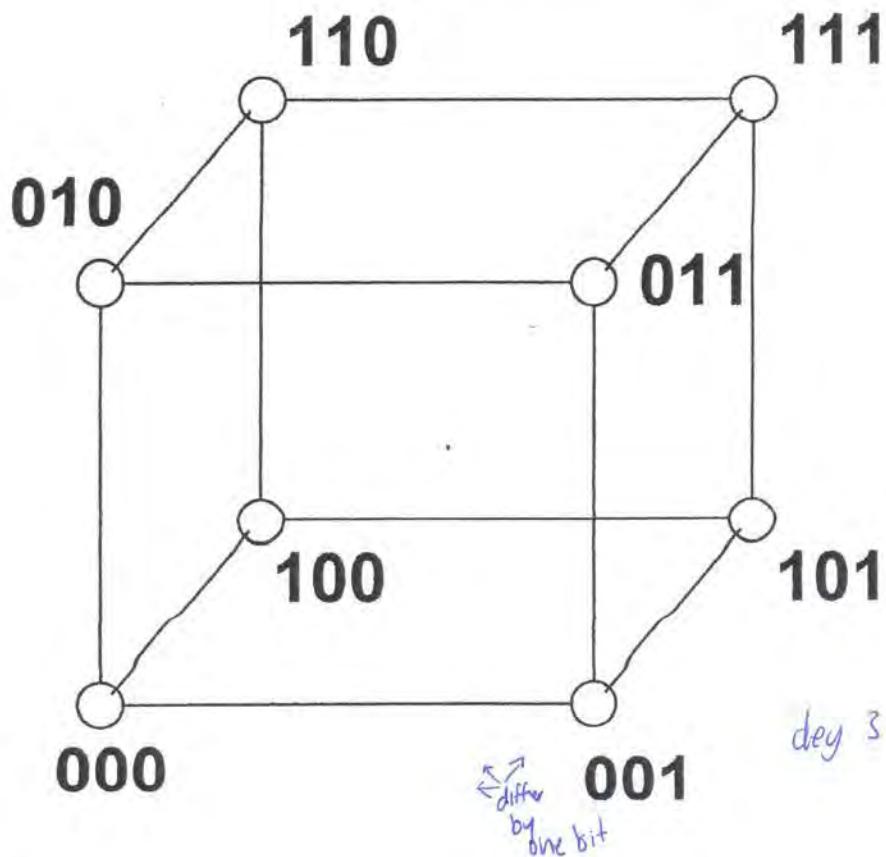


*Recursive structure*

## Beneš network ( $N = 8$ )



## Hypercube ( $N = 8$ )



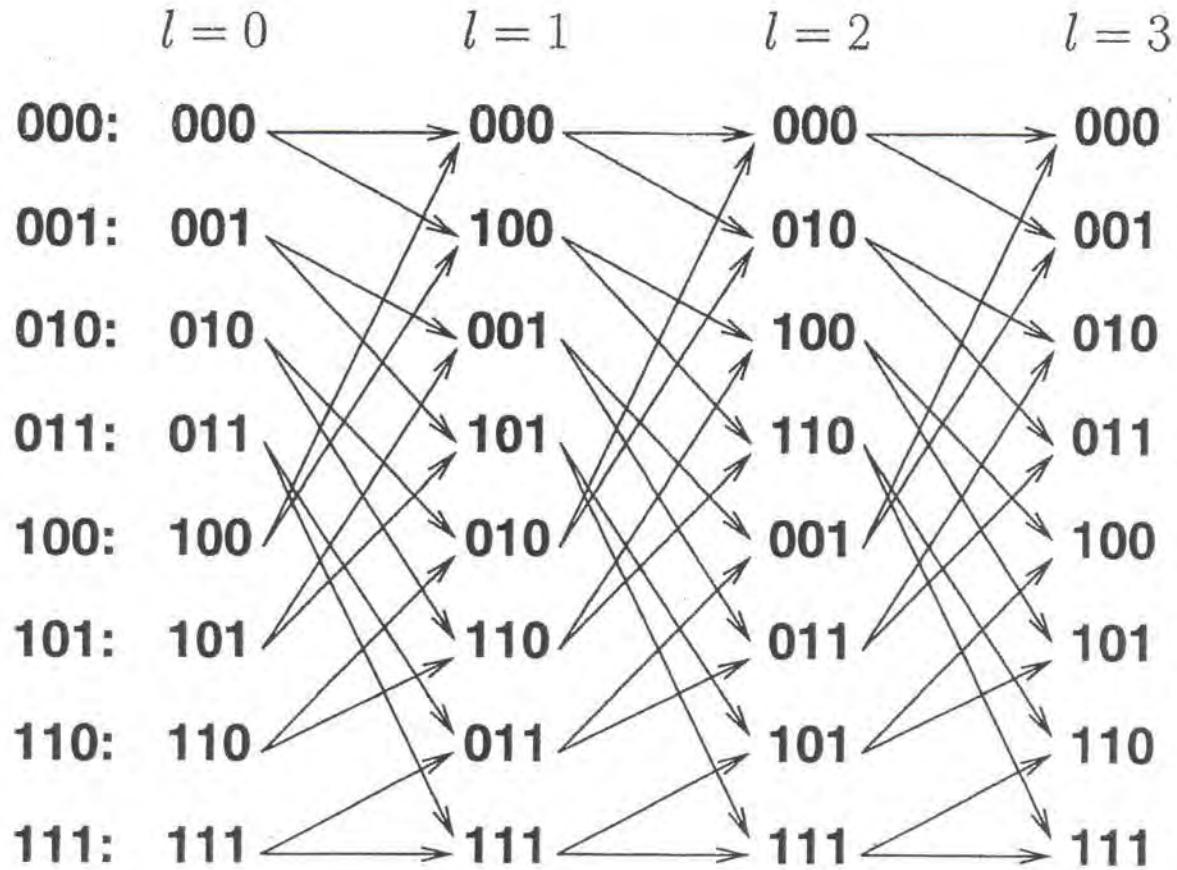
Nodes:

$$\{(b_1, b_2, \dots, b_{\log N}) \mid b_i = 0 \text{ or } 1\}$$

Edges:

$$\left\{ ((a_1, \dots, a_{\log N}), (b_1, \dots, b_{\log N})) \mid \vec{a} \text{ and } \vec{b} \text{ differ in precisely one bit} \right\}$$

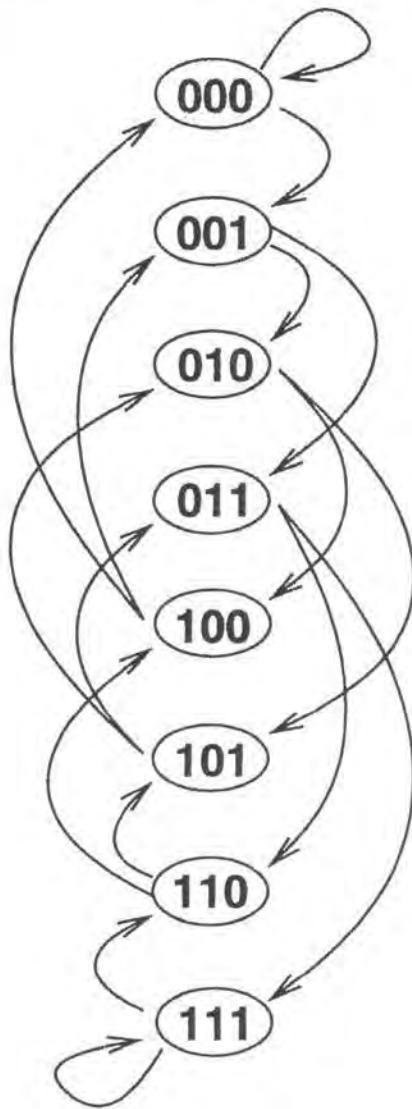
## Butterfly (redrawn)



Node  $(b_1, b_2, \dots, b_{\log N}, l)$  has been drawn in row:

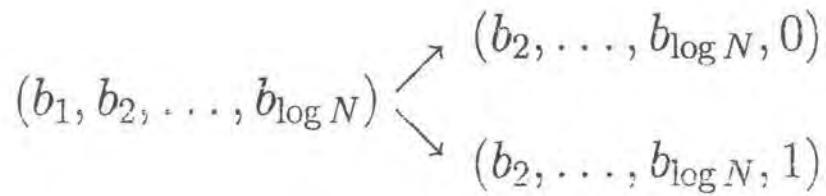
$$b_{l+1}, b_{l+2}, \dots, b_{\log N}, b_1, \dots, b_l$$

## De Bruijn Graph ( $N = 8$ )



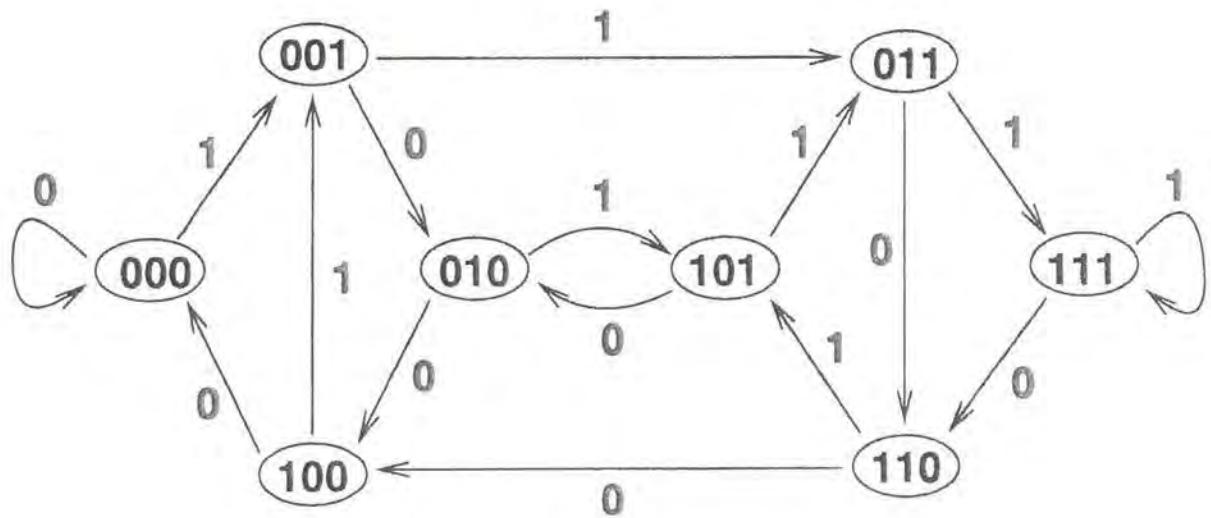
Nodes:  $\{(b_1, b_2, \dots, b_{\log N}) \mid b_i = 0 \text{ or } 1\}$

Edges:



*Briin*

## De Bruijn Graph (redrawn)



Edges labels:

$$(b_1, b_2, \dots, b_{\log N}) \xrightarrow{0} (b_2, \dots, b_{\log N}, 0)$$
$$(b_1, b_2, \dots, b_{\log N}) \xrightarrow{1} (b_2, \dots, b_{\log N}, 1)$$

## Communication Networks

<u>Network</u>	<u># switches</u>	<u>Diameter</u>	<u>congestion</u>
N-input 2D grid	$N^2$	$2N$	2
Complete Binary Tree	$2N-1$	$2\log_2 N + 2$	$\frac{N}{\sqrt{N}}$
ButterFly	$N(\log N + 1)$	$\log N + 2$	$\sqrt{N}$
Benes network	$2N(\log N + 1)$	$2\log N + 2$	1

Def: The distance between nodes  $u$  and  $v$  in a graph is the length of the shortest path from  $u$  to  $v$ .

Def: The diameter of a network is the distance between the input and output that are farthest apart.

Def: The congestion of a set of paths in a network is the max (over all nodes  $v$ ) of the number of paths that go thru  $v$ .

Def: The congestion of a routing problem is the congestion of the best set of paths for the routing problem.  
(that minimize congestion.)

Def: The congestion of a network is the congestion of the worst case routing problem.

Congestion of  $N \times N$  grid is 2

For routing problems,  $\exists$  set of paths,  $\forall$  nodes  $v$ ,  $\leq 2$  paths thru  $v$ .

Assuming 1 to one communication

## Properties of relations

Properties of a relation on  $A$ :

**Reflexivity**  $R$  is *reflexive* if

$$\forall x \in A. xRx.$$

“Everyone likes themselves.”

Every node in  $G$  has a loop.

**Irreflexivity**  $R$  is *irreflexive* if

$$\neg \exists x \in A. xRx.$$

“No one likes themselves.”

There are no loops in  $G$ .

**Symmetry**  $R$  is *symmetric* if

$$\forall x, y \in A. xRy \Rightarrow yRx.$$

“If  $x$  likes  $y$ , then  $y$  likes  $x$ .”

If there is an edge from  $x$  to  $y$  in  $G$ , then there is an edge from  $y$  to  $x$  in  $G$  as well.

**Antisymmetry**  $R$  is *antisymmetric* if

$$\forall x, y \in A. (xRy \wedge yRx) \Rightarrow x = y.$$

“No pair of distinct people like each other.”

There is at most one directed edge between any pair of distinct nodes.

**Transitivity**  $R$  is *transitive* if

$$\forall x, y, z \in A. (xRy \wedge yRz) \Rightarrow xRz.$$

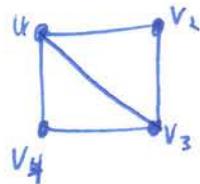
“If  $x$  likes  $y$  and  $y$  likes  $z$ , then  $x$  likes  $z$  too.”

For any walk  $v_0, v_1, \dots, v_k$  in  $G$  where  $k \geq 2$ ,  $v_0 \rightarrow v_k$  is in  $G$  (and, hence,  $v_i \rightarrow v_j$  is also in  $G$  for all  $i < j$ ).

# Matrices and Pagerank

Ways to represent a graph:

- a) Draw it
- b) List its edges (on vertices)



vertices  
 $v_1, v_2, v_3, v_4$

edges  
 $(v_1, v_2), (v_2, v_3), (v_1, v_3), (v_3, v_4), (v_1, v_4)$

- c) adjacency matrix!

Def: An  $n \times m$  matrix  $A$  is a rectangular table of numbers with  $n$  rows and  $m$  columns.

We will use  $A_{i,j}$  to denote the entry in row  $i$ , col  $j$

- c) adjacency matrix

Let's give a representation of  $G$

$$A = \begin{bmatrix} v_1 & | & 0 & 1 & 1 & 1 \\ v_2 & | & 1 & 0 & 1 & 0 \\ v_3 & | & 1 & 1 & 0 & 1 \\ v_4 & | & 1 & 0 & 1 & 0 \end{bmatrix} \quad \text{edge from } v_3 \text{ to } v_4$$

Def: The adjacency matrix of an  $n$ -vertex graph is

an  $n \times n$  where

$$A_{i,j} = \begin{cases} 1 & \text{if } (v_i, v_j) \in E \\ 0 & \text{else} \end{cases}$$

Def: If  $A$  and  $B$  are  $n \times m$  and  $m \times p$  then  $AB$  as an  $n \times p$  matrix

$$(AB)_{i,j} = \sum_{k=1}^m A_{i,k} B_{k,j}$$

Theorem: The number of walks of length  $K$  from  $V_i$  to  $V_j$  is exactly  $(A^k)_{ij}$

Proof: (By induction)  
on  $k$

Base Case: ( $k=1$ ) there is a walk iff there is an edge so the # of walks from  $V_i$  to  $V_j$  is  $A_{ij}$

Inductive Step defined as

Let  $Z_{ij}^k \triangleq \# \text{ walks of len } k \text{ From } V_i \text{ to } V_j$

want to prove  $Z_{ij}^{(k+1)} = (A^{k+1})_{ij}$

$$Z_{ij}^{(k+1)} = \sum_{V_l \text{ is adjacent } v_j} Z_{i,l}^k = \sum_l Z_{i,l}^k Z_{l,j} \xrightarrow{\substack{l \mapsto \\ (A^k)_{i,l}}} A_{i,j} = A_{ij}^{(k+1)}$$



What is the smallest  $k$  s.t.  $(A^k)_{ij} \neq 0$ ?

PageRank Claim: It is the length of the shortest path from  $V_i$  to  $V_j$

Abundance Problem: How can we filter a huge number of relevant web pages

Lesson #1: Wordcount can be misleading

Key: A link from  $P$  to  $q$  is an endorsement of  $q$  by  $P$

First Attempt: Return the highest degree

## Basic Page Rank

• Initially, each PageRank =  $1/n$

• Repeat  $i=1 \dots K$

• Each page divides its PR equally along outgoing edges, passes it along

• Set new PR to sum of received shares.

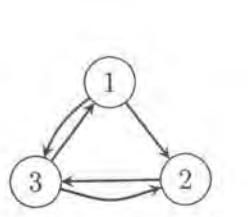
## Problems for Recitation 9

When you search for a broad term (e.g., “mathematics”) on Google, there are typically millions of matches; many webpages contain the word mathematics! In order to give useful results, Google needs to find a good way of ranking these results.

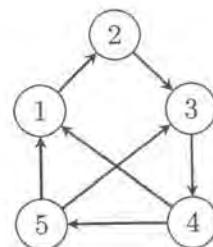
The success of Google is founded on a good way to do this, using the digraph structure of the web. The “PageRank” algorithm was invented by Larry Page and Sergey Brin. There are a lot more complications than we’ll discuss here, but we will be able to understand the core of the idea.

### The web graph

We can construct a digraph from the web very easily: the vertices of the graph are webpages, and there is a directed edge from webpage  $i$  to webpage  $j$  if  $i$  has a link to  $j$ . Here are two (very small) web graphs:



(a)



(b)

Just for today’s recitation, we will assume that (i) the web graph has no self loops (we could just remove them, anyway), and (ii) it is strongly connected. This is very far from being true. The real PageRank algorithm does not need this, but it will simplify things a bit for us.

### Problem 1: A first try

How can we use the web graph to determine the importance of a page? A natural idea is to look at the number of links to that page; the more incoming links, the higher the rank!

- What ranking does this give us in the two graphs above?
- Give some reasons why this ranking will not work well, in general. (How could you game the system?)

### Problem 2: PageRank

Let's try something more sophisticated. We would like links from "important" webpages to count more than links from unimportant ones. But this sounds self-referential, since "importance" is what we're trying to determine!

Let's try an iterative process. Let's give each webpage a million dollars (!). On the hour, each webpage does the following: it takes all its money, divides it equally amongst the webpages it links to, and sends it along to those pages. The process continues, and we hope that things settle down eventually, so that after a while the amount of money a given webpage has stays essentially constant. We then say that the importance (or PageRank) of the webpage is how much money it ends up with.

- (a) Consider graph (a) shown earlier. Suppose we have  $x_i$  millions of dollars at vertex  $i$ . Find a formula for the amount of money (in millions)  $x'_i$  at each node  $i$  after 1 hour.
- (b) Figure out a formula for the amount of money  $x_i^{(n)}$  (in millions) that node  $i$  has after  $n$  hours. Prove that your formula is correct using induction. Hint: the formula for  $x_3^{(n)}$  is  $\frac{1}{3}(4 - (-\frac{1}{2})^n)$ , this should help you get started.
- (c) What happens as  $n$  goes to infinity? Hence determine the PageRank of the webpages.
- (d) The formula for  $\vec{x}'$  in terms of  $\vec{x}$  can be written as a matrix product:  $\vec{x}' = W\vec{x}$ , for some matrix  $W$  (we'll call this the *update matrix*). Determine  $W$ .
- (e) Check that it satisfies the equation  $\vec{p} = W\vec{p}$ , where  $\vec{p}$  is the vector of PageRanks<sup>1</sup>. (In other words,  $p_i$  is the PageRank of page  $i$ .)

This last fact is true in general: for any strongly connected web graph  $G$ , with  $W$  being its update matrix, the equation

$$\vec{p} = W\vec{p} \tag{1}$$

is satisfied by the vector  $\vec{p}$  of PageRanks. We won't prove this, but assume this for the next question.

- (f) Determine the update matrix  $W$  for the web graph (b) shown earlier. Hence determine the PageRank vector  $\vec{p}$  by finding a non-zero solution to (1). (The solution is not unique; but if you add the requirement that, e.g.,  $p_1 = 1$ , then it will be unique).

---

<sup>1</sup>We snuck in some linear algebra; some of you may recognize that  $p$  is an eigenvector of  $W$ , with associated eigenvalue 1.

10/8/14

6.042 Recitation

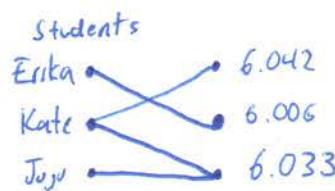
Fernando Trujano

## Relations, Posets, Scheduling. 7.1, 7.3-7.8

Def: Given  $A, B$  a relation  $R \subseteq A \times B$   
 $(a, b) \in R$  or  $a R b$  (pairs)

Relation  $\longleftrightarrow$  Bipartite Graph

Ex:  $A = \text{students}$   $R = \text{who's taking what}$   
 $B = \text{classes}$



Often:  $A = B$ . Relations of sets to themselves

Example:  $R = \text{"adjacency"}$ ,  $\text{"connected to"}$

$R = \text{"congruency"}$   $a \equiv b \pmod{5}$

$R = \text{"divisibility"}$   $a | b$

( $\rightarrow$  always Anti-Symmetric)

Properties:

Reflexive:  $\forall a, (a, a) \in R$

Symmetric:  $\forall a, b \quad (a, b) \in R \Rightarrow (b, a) \in R$

Anti-Symmetric  $\forall a, b \quad (a, b) \in R \wedge (b, a) \in R \Rightarrow b = a$

Transitivity  $\forall a, b, c \quad (a, b) \in R \wedge (b, c) \in R \Rightarrow (a, c) \in R$

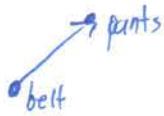
Hi :)

Equivalence Relation. reflexive, symmetric and transitive

(weak) Partial Order. reflexive, anti-symmetric, transitive

Poset  $\leq$

Hasse Diagram: Remove edges implied by transitivity

Ex. pants  $\leq$  belt       "belt depends on pants"

Lemma: The Hasse Diagram has no cycle with distinct verticies

Proof (By contradiction)

Assume we have a cycle  $a_1 \leq a_2 \leq \dots \leq a_n \leq a_1$

$a_1 \leq a_n \wedge a_n \leq a_1$  but  $a_n \neq a_1$   $\#$

□

Def: A total order is a partial order where every pair of elements is comparable -- ie  $x \leq y$  or  $y \leq x$

Def: A topological sort of  $(A, \leq) \cup (A, \leq)$  where  $\leq$  is total  
every set is comparable

$$a \leq b \Rightarrow a \leq_{+} b$$

Lemma: Every Finite poset has a minimal element (ie  $a \in A$  s.t  
"nothing depends on the minimal element"  
 $\forall b \neq a, b \not\leq a$ )

Proof: (By contradiction)

Let  $a_1 \leq a_2 \dots \leq a_n$  be the longest chain

(ie a sequence of elements  $a_1 \leq a_2 \dots$  where  $a_i$ 's are distinct)

Suppose  $a_1$  is not minimal

$\Rightarrow \exists b \neq a_1$  where  $b \leq a_1$  "something depends on  $a_1$ "

Case 1:  $b \neq a_1, a_2 \dots a_n$  then  $b \leq a_1 \dots \leq a_n$  is longer ✓

Case 2: cycle  $\#$

□

Theorem: Every finite poset  $(A, \leq)$  has a topographical sort.

Proof: (by induction)

$P(n) \triangleq$  every  $n$ -element poset has a top sort

Base Case  $n=1$  : It's already a total order

Inductive Step : Assume  $P(n)$  let  $(A, \leq)$  be  $n+1$  element poset.

Let  $a$  be minimal (by Lemma)

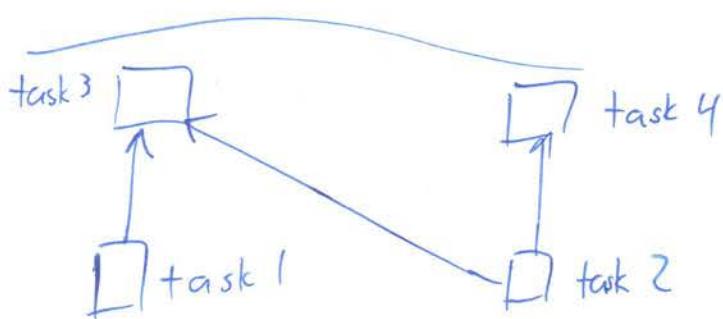
Set  $A' = A \setminus \{a\}$ ,  $P(n) \Rightarrow \leq'_T$  top sort

Set  $\leq_T = \leq'_T \cup \{(a, z) \mid \forall z \in A\}$

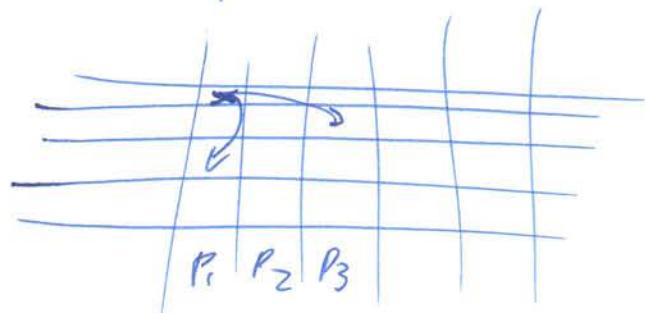
Case 1:  $x=a$  then  $a \leq_T z \quad \forall z \in A$

Case 2:  $x \neq a$  then  $y \neq a$

then  $x \leq y \Rightarrow x \leq'_T y \Rightarrow x \leq_T y$



Parallel processors



$\square \rightarrow \square \rightarrow \square \rightarrow \square$

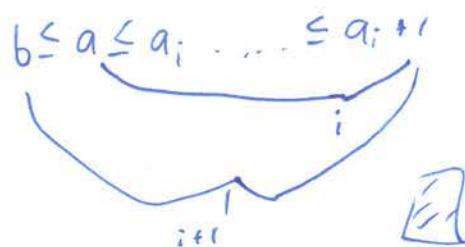
↪ can only run one task per time step so parallel processing would not be helpful.

Ihm: IF the longest chain is length  $t$  there is a partition  $A_1, \dots, A_t$  s.t.  $\forall a \in A_i$ , all  $b \leq a$   $b$  is in  $A_1, \dots, A_{i-1}$   $(b \neq a)$

Proof : For  $a \in A$ , suppose the longest chain ending at  $a$  has length  $i$  then put  $a$  in  $A_{t-i+1}$

Suppose  $\exists b \leq a$  ( $b \neq a$ ) and  $a \in A_{t-i+1}$  but  $b \notin A_1, \dots, A_{t-i}$  for contradiction.

$b \in A_{t-j+1} \quad j \geq i$



10/11/14

## 6.042 Recitation

### Binary Relations

Given sets  $A \text{ and } B$  a Binary Relation  $R: A \rightarrow B$

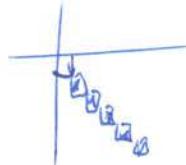
From  $A$  to  $B$  is a subset of  $A \times B$

$$R \subseteq A \times B \Rightarrow \forall a \in A \ \forall b \in B \ (a, b) \in A \times B$$

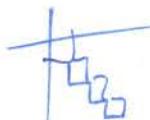
### Properties

Reflexive:  $aRa$

$$(a, a) \in R$$



Irreflexive:  $\nexists x \in A \times R_x$



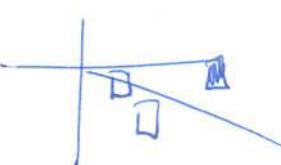
Symmetric:  $\forall x, y \in A \quad xRy \Rightarrow yRx$



Transitive:

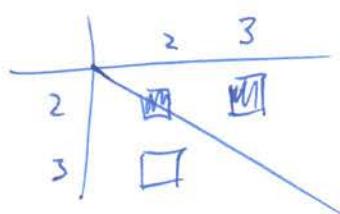


asymmetric:



Does not allow self loops.

antisym:



Allows self loops.

## Equivalence Relation

- Reflexive  $(a, a) \in R$   $aRa$   $a = a$
- Symmetric  $aRb, bRa$   $a = b$   $b = a$
- Transitive  $aRb, bRc$   $a = b$   $b = c \Rightarrow a = c$

## Equivalence Class

- All the elements that are considered the same.

If  $b \sim a$

## Poset $\preceq$

Partial order that is:

- Reflexive | 9/17/99
- Antisymmetric
- Transitive

## Topological Sort

$x \leq y \Rightarrow x \preceq_T y$

## Total Order

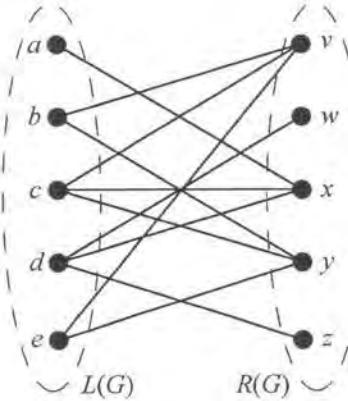
Everything in the set is comparable.

6.006 Bridges question!

## Solutions to Midterm Exam November 6

### Problem 1 (Bipartite Matching) (10 points).

Explain why the graph  $G$  below has no matching.



**Solution.** It is not possible because  $\{a, b, c, e\}$  is a bottleneck:  $|G(\{a, b, c, e\})| = |\{v, x, y\}| = 3 < 4 = |\{a, b, c, e\}|$ .

Since the left and right vertex sets of  $G$  are the same size, there is a bottleneck on the right iff there is one on the left. So an alternative answer is to observe that  $\{w, z\}$  is a bottleneck in the other direction:  $|G^{-1}(\{w, z\})| = |\{d\}| = 1 < 2 = |\{w, z\}|$ . ■

### Problem 2 (Big Oh) (10 points).

Verify that  $f \neq O(g)$  and  $g \neq O(f)$ , where

$$f(n) := n^n, \quad g(n) := \begin{cases} n^{n-(1/2)} & \text{if } n \text{ is odd,} \\ n^{n+(1/2)} & \text{if } n \text{ is even.} \end{cases}$$

**Solution.**

$$\limsup_{n \rightarrow \infty} \frac{f}{g} = \lim \frac{n^n}{n^{n-(1/2)}} = \lim \frac{n^n}{(1/\sqrt{n})n^n} = \sqrt{n} = \infty,$$

so  $f \neq O(g)$ .

$$\limsup_{n \rightarrow \infty} \frac{g}{f} = \lim \frac{n^{n+(1/2)}}{n^n} = \lim \frac{\sqrt{n} n^n}{n^n} = \sqrt{n} = \infty,$$

so  $g \neq O(f)$ . ■

**Base case:** ( $n = 1$ )  $G$  has one vertex, the degree of which is 0. Since  $G$  is 1-colorable,  $P(1)$  holds.

**Inductive step:** We may assume  $P(n)$ . To prove  $P(n + 1)$ , let  $G_{n+1}$  be a graph with  $n + 1$  vertices whose vertex degrees are all  $k$  or less. Also, suppose  $G_{n+1}$  has a vertex,  $v$ , of degree strictly less than  $k$ . Now we only need to prove that  $G_{n+1}$  is  $k$ -colorable.

To do this, first remove the vertex  $v$  to produce a graph,  $G_n$ , with  $n$  vertices. Let  $u$  be a vertex that is adjacent to  $v$  in  $G_{n+1}$ . Removing  $v$  reduces the degree of  $u$  by 1. So in  $G_n$ , vertex  $u$  has degree strictly less than  $k$ . Since no edges were added, the vertex degrees of  $G_n$  remain  $\leq k$ . So  $G_n$  satisfies the conditions of the induction hypothesis,  $P(n)$ , and so we conclude that  $G_n$  is  $k$ -colorable.

Now a  $k$ -coloring of  $G_n$  gives a coloring of all the vertices of  $G_{n+1}$ , except for  $v$ . Since  $v$  has degree less than  $k$ , there will be fewer than  $k$  colors assigned to the nodes adjacent to  $v$ . So among the  $k$  possible colors, there will be a color not used to color these adjacent nodes, and this color can be assigned to  $v$  to form a  $k$ -coloring of  $G_{n+1}$ . ■

**Solution.** The flaw is that if  $v$  has degree 0, then no such  $u$  exists. In such a case, removing  $v$  will not reduce the degree of any vertex, and so there may not be any vertex of degree less than  $k$  in  $G_n$ , as in the counterexample of part (a).

So the mistaken sentence is “Let  $u$  be a vertex that is adjacent to  $v$  in  $G_{n+1}$ .”

Alternatively, you could say that it’s OK to reason about a nonexistent  $u$ , and the only mistake is the claim that  $u$  exists. This claim is hidden in the phrase “So  $G_n$  satisfies the conditions of the induction hypothesis,  $P(n)$ ”. ■

(c) With a slightly strengthened condition, the preceding proof of the False Claim could be revised into a sound proof of the following Claim:

**Claim.** *Let  $G$  be a graph whose vertex degrees are all  $\leq k$ . If (statement inserted from below) has a vertex of degree strictly less than  $k$ , then  $G$  is  $k$ -colorable.*

Circle each of the statements below that could be inserted to make the proof correct.

- $G$  is connected and
- $G$  has no vertex of degree zero and
- $G$  does not contain a complete graph on  $k$  vertices and
- every connected component of  $G$
- some connected component of  $G$

**Solution.** Either the first statement “ $G$  is connected and” or the fourth statement “every connected component of  $G$ ” will work. ■

### Problem 6 (Digraph Induction) (25 points).

In a round-robin tournament, every two distinct players play against each other just once. For a round-robin tournament with no tied games, a record of who lost to whom can be described with a *tournament digraph*, where the vertices correspond to players, and there is an edge  $\langle x \rightarrow y \rangle$  iff  $x$  lost to  $y$  in their game.

A *ranking* is a path that includes all the players. So in a ranking, each player lost the game against the next player in the path, but may very well have won their games against other players further along the path.

(a) Give an example of a tournament digraph with more than one ranking.

**Solution.** Let  $n = 3$  with edges  $\langle u \rightarrow v \rangle$ ,  $\langle v \rightarrow w \rangle$  and  $\langle w \rightarrow u \rangle$ . Then both  $u, v, w$  and  $v, w, u$  are rankings. ■

## Solutions to Midterm Exam October 9

The following problem is a slight variant of a problem which appeared on the Spring13 midterm that was made available for review.

**Problem 1 (Predicates & Relations) (20 points).**

Five basic properties of binary relations  $R : A \rightarrow B$  are:

1.  $R$  is a surjection [ $\geq 1$  in]
2.  $R$  is an injection [ $\leq 1$  in]
3.  $R$  is a function [ $\leq 1$  out]
4.  $R$  is total [ $\geq 1$  out]
5.  $R$  is empty [= 0 out]

Below are some assertions about a relation  $R$ . For each assertion, write the numbers of all the properties above that the relation  $R$  must have; write "none" if  $R$  might not have any of these properties. For example, you should write "1, 4" next to the first assertion.

Variables  $a, a_1, \dots$  range over  $A$  and  $b, b_1, \dots$  range over  $B$ .

- (a)  $\forall a, b. a R b.$  1, 4
- (b)  $\text{NOT}(\forall a, b. a R b).$  ■

**Solution.** none ■

- (c)  $\forall a, b. \text{NOT}(a R b).$  ■

**Solution.** empty 5 ■

- (d)  $\forall a \exists b. a R b.$  ■

**Solution.** total 4 ■

- (e)  $\forall b \exists a. a R b.$  ■

**Solution.** surjection 1 ■

- (f)  $R$  is a bijection. ■

**Solution.** all the properties 1, 2, 3, 4 ■

- (g)  $\forall a \exists b_1. [a R b_1 \text{ AND } (\forall b. a R b \text{ IMPLIES } b = b_1)].$  ■

**Solution.** total function 3, 4 ■

- (a) One of the inclusions

$$\text{GoodCount} \subseteq \text{RecMatch}, \\ \text{RecMatch} \subseteq \text{GoodCount},$$

is easy to prove by structural induction, while the other inclusion follows easily by strong induction. Which inclusion is easy to prove by structural induction?

**Solution.**  $\text{RecMatch} \subseteq \text{GoodCount}$

- (b) State an induction hypothesis that allows an easy proof by structural induction of the inclusion from part (a). No proof is required.

**Solution.** The structural induction hypothesis is simply.

$$P(s) ::= s \in \text{GoodCount}.$$

The proof was not required, but here it is:

*Proof.* The proof is by structural induction on the recursive definition of RecMatch.

**Base Case:**  $P(\lambda)$  holds since the count of the empty string ends where it starts at zero.

**Inductive Step:** Assume  $P(s)$  and  $P(t)$  are true. We need to show that  $P([s]t)$  is true.

The count values for  $[s]t$  start with 0. Reading the initial left bracket yields 1 as the next count value. This 1 serves as the start of a series of count values exactly equal to the count values of  $s$ , with each value incremented by one. Since  $s \in \text{GoodCount}$  by induction hypothesis, these incremented count values begin with 1, always stay positive, and end with 1. The right bracket immediately after  $s$  reduces the ending count to 0. This 0 serves as the start of the remaining count values which are exactly the count values of  $t$ . Since by induction hypothesis  $t \in \text{GoodCount}$ , these remaining values never go negative and end at 0. Hence the entire sequence of count values for  $[s]t$  starts with 0, never goes negative, and ends with 0, which proves that  $[s]t \in \text{GoodCount}$ .

- (c) The other inclusion can be proved by strong induction. State a strong induction hypothesis that leads to a straightforward proof of this other inclusion. No proof is required.

**Solution.** A straightforward strong induction hypothesis that works is

$$Q(n) ::= \forall r \in \text{GoodCount}. |r| = n \text{ IMPLIES } r \in \text{RecMatch}.$$

The proof was not required, but here it is:

*Proof.* **Base Case**  $n = 0$ : There is only one string of length 0, namely the empty string, which is in RecMatch by definition, proving  $Q(0)$ .

**Inductive Step:** Assume that  $Q(k)$  is true for all  $k \leq n$ , we need to prove that  $Q(n+1)$  is also true.

So suppose  $r$  is a length  $n+1$  string that counts well. We must prove that  $r \in \text{RecMatch}$ .

**Solution. Inductive step:** Now we assume  $P(n)$  holds for some  $n \geq 1$  and prove  $P(n + 1)$ .

Let  $a := a_1 a_2 \cdots a_n a_{n+1}$  and suppose that  $p \mid a$ . We need to prove that  $p \mid a_i$  for some  $i \in [1, n + 1]$ .

Let  $b := a_1 \cdots a_n$ , so  $a = ba_{n+1}$ . Now since  $p \mid ba_{n+1}$  and  $p$  is prime, we know from the text Lemma 8.4.2 that  $p \mid b$  or  $p \mid a_{n+1}$ . If  $p \mid a_{n+1}$ , then  $P(n + 1)$  follows immediately by letting  $i = n + 1$ . If  $p \mid b$ , then the induction hypothesis  $P(n)$  implies that  $p \mid a_i$  for some  $i \in [1, n]$ , which also implies  $P(n + 1)$ . So in either case,  $P(n + 1)$  holds, which completes the inductive step.

By induction, the claim holds for all  $n \geq 1$ . ■

### Problem 5 (Preserved Invariant) (20 points).

There is a bucket containing more blue balls than red balls. As long as there continue to be more blues than reds, balls may be added and removed from the bucket according to the following rules:

- (i) Add a red ball.
- (ii) Remove a blue ball.
- (iii) Add two reds and one blue.
- (iv) Remove two blues and one red.

(a) Starting with 10 reds and 16 blues, what is the largest number of balls the bucket will contain by applying these rules?

**Solution.**

$$44 = 22 \text{ red balls} + 22 \text{ blue balls.}$$

By applying rule (iii) six times. ■

Let  $b$  be the number of blue balls and  $r$  be the number of red balls in the bucket at any given time.

(b) Prove that  $b - r \geq 0$  is a preserved invariant of the process of adding and removing balls according to rules (i)–(iv).

**Solution.** If  $b \leq r$ , then no rule applies, so  $b - r \geq 0$  is vacuously preserved. If a rule applies, we must have  $b - r > 0$ . Since each rule reduces the difference of  $b$  and  $r$  by one, we have  $b - r \geq 0$  after application of a rule. So again  $b - r \geq 0$  is preserved. ■

(c) Prove that no matter how many balls the bucket contains, repeatedly applying rules (i)–(iv) will eventually lead to a state where no further rule can be applied.

**Solution.** It is easy to verify that  $b - r$  is a strictly decreasing derived variable. For example, applying rule (iii) takes  $(b, r)$  to  $(b + 1, r + 2)$  and

$$b - r > b - r - 1 = (b + 1) - (r + 2).$$

Since  $b - r$  is also nonnegative by part (b), it follows that starting in state  $(b, r)$ , the rules can be applied at most  $b - r$  times. ■

# 1 Exponentiation and Modular Arithmetic

Recall that RSA encryption and decryption both involve exponentiation. To encrypt a message  $m$ , we use the following equation:

$$m' = \text{rem}(m^e, n) \equiv m^e \pmod{n}.$$

And to decrypt a message  $m'$ , we use

$$m = \text{rem}((m')^d, n) \equiv (m')^d \pmod{n}.$$

In practice,  $e$  and  $d$  might be quite large. But even for relatively small values of these variables, the quantities  $m^e$  and  $(m')^d$  can be very difficult to compute directly. Fortunately, there are tractable and efficient methods for carrying out exponentiation of large integer powers modulo a number.

Let's say we are trying to encrypt a message. First, note that:

$$\begin{aligned}\text{rem}(a \cdot b, c) &\equiv a \cdot b \pmod{c} \\ &\equiv \text{rem}(a, c) \cdot \text{rem}(b, c) \pmod{c} \\ &= \text{rem}(\text{rem}(a, c) \cdot \text{rem}(b, c), c)\end{aligned}$$

This principle extends to an arbitrary number of factors, such that:

$$a_1 \cdot a_2 \cdot \dots \cdot a_n \equiv \text{rem}(a_1, c) \cdot \text{rem}(a_2, c) \cdot \dots \cdot \text{rem}(a_n, c) \pmod{c}$$

We illustrate this point with an example:

**Example:** Find  $\text{rem}(23 \cdot 61 \cdot 19, 17)$ .

We could find the remainder of  $23 \cdot 61 \cdot 19 = 26657$  divided by 17, but that would be a lot of unnecessary work! Instead, we notice the fact that  $23 \equiv 6 \pmod{17}$ ,  $61 \equiv 10 \pmod{17}$ , and  $19 \equiv 2 \pmod{17}$ . Therefore,  $23 \cdot 61 \cdot 19 \equiv 6 \cdot 10 \cdot 2 \pmod{17}$ .

Similarly, we can reduce the remainder of  $6 \cdot 10 \cdot 2$  divided by 17. We notice the fact that  $10 \cdot 2 = 20 \equiv 3 \pmod{17}$ , so  $6 \cdot 10 \cdot 2 \equiv 6 \cdot 3 = 18 \equiv 1 \pmod{17}$ . We could have also calculated  $6 \cdot 10 = 60 \equiv 9 \pmod{17}$  to get the same answer  $6 \cdot 10 \cdot 2 \equiv 9 \cdot 2 = 18 \equiv 1 \pmod{17}$ . While both methods here were relatively simple to use, how you choose to associate your factors may sometimes greatly affect the difficulty of a calculation!

Let's return to RSA. Here's one way we might go about encrypting our message (though in a minute we'll consider a more efficient technique). We can compute  $m = \text{rem}(m^e, n)$  by breaking the exponentiation into a sequence of  $e - 1$  multiplications. We then take the remainder after dividing by  $n$  after each one of these multiplications.

**Example:** Encrypt the message  $m = 5$  with  $e = 6$  and  $n = 17$ .

# Number Theory

## Problem Solving

### Problem

\* Find the remainder of  $11^{11211}$  when divided by 113.

— aka Find  $y \in \{0, 1, \dots, 112\}$  such that 113 is prime

$$11^{11211} \equiv y \pmod{113}$$

By Fermat's Little Theorem  $\leftarrow$  b/c  $\gcd(11, 113) = 1$

$$\cancel{a^{\cancel{c}}} \quad 11^{112} \equiv 1 \pmod{113} \quad \begin{matrix} 113 \text{ is prime} \\ 112 = 113 - 1 \end{matrix}$$

Strategy: If  $a^r \equiv 1 \pmod{n} \Rightarrow a^{rk} \equiv 1 \pmod{n}$  any multiple

Take out as many  $112$  as possible, since they go to  $1 \pmod{113}$

So:

$$11^{11211} \equiv y \pmod{113}$$

$$11^{11211} \equiv 11^{112 \cdot 1000 + 111} \equiv 11^{112 \cdot 1000} \cdot 11^{111}$$

$$\equiv 1^{1000} \cdot 11^{111} \pmod{113}$$

Fermat's

$$\text{Note that: } 11^{111} \cdot 11 = 11^{112} \equiv 1 \pmod{113}$$

We know (from part a) that  $72 \cdot 11 \equiv 1 \pmod{113}$

$$\text{So: } 11^{111} \equiv 72 \pmod{113}$$

## Number Theory

GCD

- two #'s  $a, b$ , the smallest linear combination of  $a$  and  $b$  is the gcd.

Ex:

$$a = 12, b = 10 \Rightarrow \text{gcd} = 2$$

$$as + bt = 1 ?? \text{ NOT POSSIBLE}$$

- An integer is a linear combo of  $a$  and  $b$  iff it is a multiple of the gcd.

Euclid's Algorithm : Find the gcd

$$\text{gcd}(a, b) = \text{gcd}(b, \text{rem}(a, b))$$

repeatedly

Pulverizer (Extended Euclid's Algorithm)

↳ Find coefficients in linear combo that produces the gcd.  
( $s, t$ )

Modular Arithmetic

$$a \equiv b \pmod{n}$$

iff  $n \mid (a - b)$ 

$$\text{iff } \text{rem}(a, n) = \text{rem}(b, n)$$

Multiplicative inverse

$$k^{-1} k \equiv 1 \pmod{n}$$

exists iff  $\text{gcd}(k, n) = 1$ 

To get, use pulverizer:

$$qk + bn = 1$$

Multiplicative inverse of  $k \pmod{n}$ 

Or Euler's Theorem:

$$k^{\phi(n)} \equiv 1 \pmod{n}$$

+  $k$  and  $n$  relatively prime

## Totient Function

$\phi(n)$  ← # of ints  $[0, n]$  that are relatively prime to  $n$

$$\text{Ex: } \phi(12) = 4 \quad \text{sb} (1, 5, 7, 11)$$

If  $p$  and  $q$  are both prime:

$$\phi(pq) = (p-1)(q-1)$$

$$\phi(p) = p-1$$

$$\phi(n) = n \left(1 - \frac{1}{p_1}\right) \left(1 - \frac{1}{p_2}\right) \left(1 - \frac{1}{p_3}\right) \dots$$

↑      ↓      ↘  
Prime Factors

$$k \cdot \underbrace{k^{\phi(n)-1}}_{\substack{\text{multiplicative} \\ \text{Inverse!}}} \equiv 1 \pmod{n}$$

## RSA

receiver: pick two primes  $p$  and  $q$

$$\text{compute } n = pq$$

Select  $e$  such that

$$\gcd(e, (p-1)(q-1)) = 1$$

distribute this

compute  $d$  such that

$$ede \equiv 1 \pmod{(p-1)(q-1)}$$

public key is  $(e, n)$   
private key is  $(d, n)$

From Pulverizer:

$$ae + b(p-1)(q-1) = 1$$

↑      ↓      ↗  
d

sender: encrypt using public key

$$m^* = \text{rem}(m^e, n)$$

receiver: decrypt using private key

$$m = \text{rem}((m^*)^d, n)$$

Why does this work?

$$\text{We know: } m^* \equiv m^e \pmod{n}$$

So  
 $(m^*)^d \equiv m^{ed} \pmod{n}$

$$de \equiv 1 \pmod{(p-1)(q-1)}$$

$$de = 1 + r(p-1)(q-1) \text{ for some } r$$

So  
 $(m^*)^d \equiv m^{1+r(p-1)(q-1)} \pmod{n}$

$$\equiv m \cdot (m^{(p-1)(q-1)})^r \pmod{n}$$

$$\overbrace{\quad\quad\quad}^{1\quad\quad\quad} +$$

$$\begin{aligned} n &= pq \\ \phi(n) &\equiv (p-1)(q-1) \end{aligned}$$

$$m^{\phi(n)} \equiv 1 \pmod{n}$$

$$\equiv m \pmod{n}$$

## 6.042 Review

### Graph Theory I

Terminologysimple graph  $G = (V, E)$ 

Handshaking Lemma

$$\sum_{v \in V} \deg(v) = 2|E|$$

connected component - set of vertices where every vertex is connected to each other.

subgraph  $G' = (V', E')$   $V' \subseteq V$   $E' \subseteq E$ 

regular - all vertices have same degree

planar - can be drawn w/o intersecting edges

Coloringchromatic # : smallest  $k$  st graph is  $k$ -colorable

Need to do 2 things:

1) show possible to color  $G$  with  $k$  colors2) Show that  $G$  requires at least  $k$  colors

e.g. completely connected subgraph

n-sized graph takes  $n$  colorsBipartite Graph + Matching

Every regular graph has a perfect matching

matching - a set of pairs where each vertex appears at most once

Hall's Theorem:

S - set of vertices

N(S) - set of adjacent vertices

$\forall S \subseteq L, |S| \leq |N(S)| \Leftrightarrow \exists$  matching on  $G$  that covers  $L$ .

every vertex in  $L$   
gets matched.

## Isomorphisms

" $G_1$  and  $G_2$  if we can map  $V_{1s}$  to  $V_{2s}$  and have edges match"

Tricks: no actual way to easily find isomorphic

- All degree numbers maintained
- All path lengths maintained

## Trees + MST

up Tree = connected acyclic graph

$n$  vertices  $n-1$  edges

adding an edge = cycle

remove edge  $\Rightarrow$  disconnected

Spanning Tree: Remove edges from a  $G$  until  
 $n-1$  edges left  $\Rightarrow$  Tree

Minimum Spanning Tree: Spanning tree of  $G$  with  
smallest possible edge cost sum

Euler Tour: cycle, every edge visited exactly once.

## Induction Strategies

- Induction on size (# of vertices)

- Induction on degree

- common with regular graphs

### Build Up Error:

Assume  $P(n)$  to prove  $P(n+1)$

Should not add a node to a graph  $n!$       specific

Instead, start with size  $n+1$  graph, remove a node  
to get graph of size  $n$  that fits I.H.

10/20/14

## 6.042 Review

Fernando Troyano  
is a poopy  
head.

### DAG, Networks, PageRank

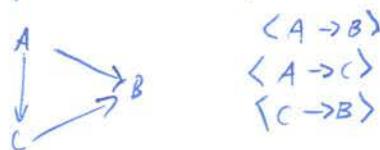
#### Networks

switches - # of vertices in network  
- not counting inputs or outputs

diameter - # of edges for shortest path that connects input output that are farthest apart.

congestion - Max congestion - largest # of packets that go through a switch with best solution to worst problem.

#### Directed Graphs aka Digraph



indegree : # of edges that point to vertex

outdegree : # of edges that start at vertex

#### DAGs (Directed Acyclic Graphs)

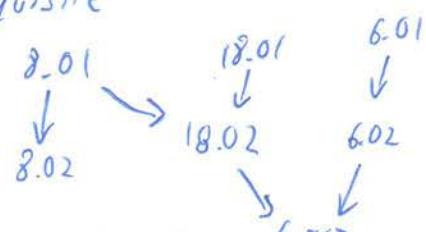
- Directed graphs with no cycles (except self loops)

#### Topological Sorts

Chains : Ordered Set of vertices where for a vertex, the vertex to its left is a pre requisite

Ex: 8.01 → 18.02 → 6.042

Max length = 3



Antichains : Set of vertices where none of the vertices are prerequisites of each other.

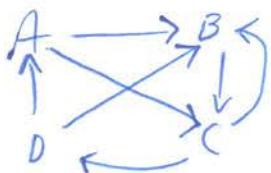
Ex: 8.01, 18.01, 6.01

- Dilworth's Lemma

Every DAG with  $n$  vertices needs either a chain of size greater than  $\lceil \frac{n}{2} \rceil$  or antichain of least  $\lceil \frac{n}{2} \rceil$

DAG == poset.

## PageRank



Unscaled PageRank:

give each vertex value of  $\frac{1}{n}$   $\checkmark$  total # of vertices

$$A: \frac{1}{4} \quad C: \frac{1}{4}$$

$$B: \frac{1}{4} \quad D: \frac{1}{4}$$

For each time step:

For each vertex:

Distribute it's previous PR evenly among vertexes it points to.

$$B: \frac{1}{2}(\frac{1}{4}) \quad C: \frac{1}{2}(\frac{1}{4})$$

Receive values from vertexes that point to it.

$$\begin{bmatrix} A' \\ B' \\ C' \\ D' \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} A \\ B \\ C \\ D \end{bmatrix}$$

Should add up to 1

$$So: P_A = \frac{1}{2} P_D$$

$P_i$ : Value  $A$  converges to.

$$P_B = \frac{1}{2} P_A + \frac{1}{2} P_C + \frac{1}{2} P_D$$

$$P_C = \frac{1}{2} P_A + P_B$$

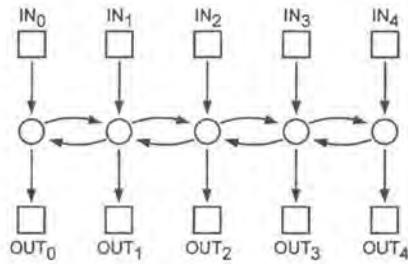
$$P_D = \frac{1}{2} P_C$$

$$P_A + P_B + P_C + P_D = 1$$

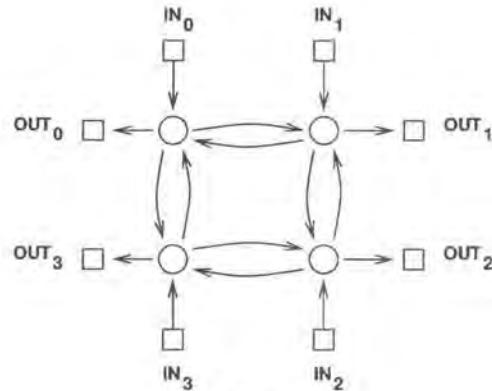
## Problems for Recitation 8

### Analysis of Two Networks

Two communication networks are shown below. Complete the table of properties and be prepared to justify your answers.



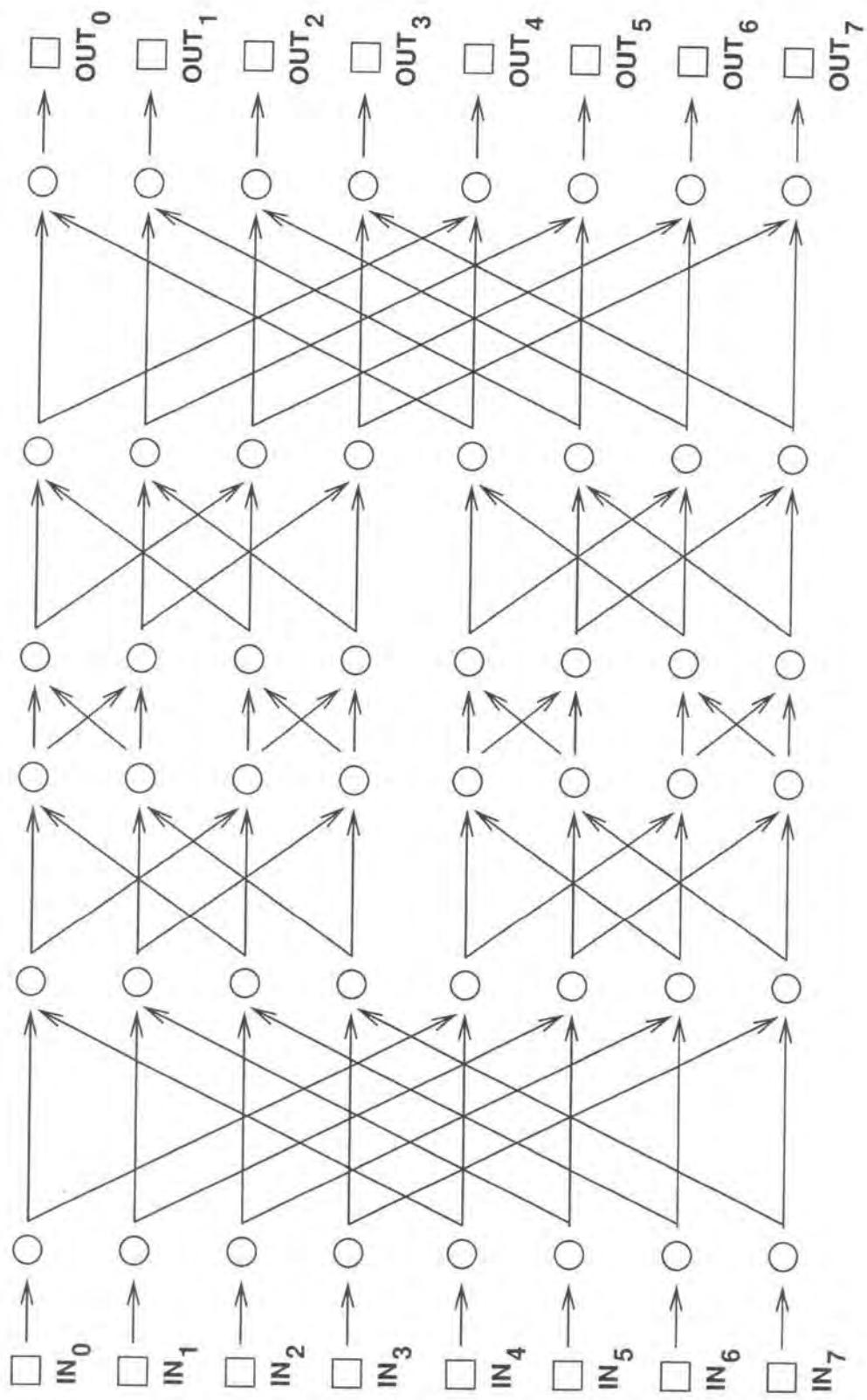
5-Path



4-Cycle

network	# switches	switch size	diameter	max congestion
5-path				
4-cycle				

Recall that the *diameter* of a communication network is the number of edges on the shortest path between the input and output that are farthest apart. The *max congestion* of a network is the largest number of packets that pass through any switch in the best solution to the hardest permutation routing problem. You might imagine that your enemy picks a permutation and then you pick the path taken by each packet. (Her goal is to cause congestion, and yours is to eliminate it.) Assuming you both do your best, the max congestion is then equal to the largest number of packets passing through a single switch.



## Notes for Recitation 1

### 1 Team Problem: Contrapositive

Prove by truth table that an implication is equivalent to its contrapositive.

Solution.

x	y	$x \rightarrow y$	$\neg y$	$\neg x$	$\neg y \rightarrow \neg x$	$(x \rightarrow y) \leftrightarrow (\neg y \rightarrow \neg x)$
T	T	T	F	F	T	T
T	F	F	T	F	F	T
F	T	T	F	T	T	T
F	F	T	T	T	T	T

In every row,  $x \rightarrow y$  is T precisely when  $\neg y \rightarrow \neg x$  is T. Thus, we conclude that an implication is equivalent to its contrapositive. ■

**Solution.** If either of Madalina or Nirvan is in the cabal, then Ankur is not. Equivalently, if Ankur is in the cabal, then neither Madalina nor Nirvan is. ■

So much for the translations. We now argue that the only cabal satisfying all seven propositions above is one whose members are exactly Madalina, Nirvan, and Brando.

We first observe that by (ii), there must be someone — either Chennah or Nirvan — who is not in the cabal. But if either Ashley or Jeffrey were in the cabal, then by (iii), everyone would be. So we conclude by contradiction that

$$\text{Ashley and Jeffrey are not in the cabal.} \quad (1)$$

Now consider that (v) implies its contrapositive: if Ashley is not in the cabal, then neither is Catherine. Therefore, since Ashley is not in the cabal,

$$\text{Catherine is not in the cabal.} \quad (2)$$

Next observe that if Chennah were in the cabal, then by (iv), Nirvan would be too, contradicting (ii). So by again contradiction, we conclude that

$$\text{Chennah is not in the cabal.} \quad (3)$$

Now suppose Tom is in the cabal. Then by (vi), Madalina and Brando are not. We already know Ashley, Jeffrey, Catherine, and Chennah are not in the cabal, leaving only three who could be — Tom, Ankur, and Nirvan. But by (i) the cabal must have at least three members, so it follows that the cabal must consist of exactly these three. This proves:

**Lemma 1.** *If Tom is in the cabal, then Ankur and Nirvan are in the cabal.*

But by (vii), if Nirvan is the cabal, then Ankur is not. That is,

**Lemma 2.** *Nirvan and Ankur cannot both be in the cabal.*

Now from Lemma 2 we conclude that the conclusion of Lemma 1 is false. So by contrapositive, the hypothesis of Lemma 1 must also be false, namely,

$$\text{Tom is not in the cabal.} \quad (4)$$

Finally, suppose Ankur is in the cabal. Then by (vii), Madalina and Nirvan are not, and we already know Ashley, Jeffrey, Catherine, Chennah, and Tom are not. So the cabal must consist of at most two people (Ankur and Brando). This contradicts (i), and we conclude by contradiction that

$$\text{Ankur is not in the cabal.} \quad (5)$$

So the only remaining people who could be in the cabal are Madalina, Nirvan, and Brando. Since the cabal must have at least three members, we conclude that

## Notes for Recitation 2

### 1 Problem: A Geometric Sum

Perhaps you encountered this classic formula in school:

$$1 + r + r^2 + r^3 + \dots + r^n = \frac{1 - r^{n+1}}{1 - r}$$

Use induction to prove that this formula is correct for all real values  $r \neq 1$ .

*Prepare a complete, careful solution. You'll be passing your proof to another group for "constructive criticism"!*

#### Solution. Proof by Induction

*Proof.* We use induction. Let  $P(n)$  be the proposition that the following equation holds for all  $r \neq 1$ :

$$1 + r + r^2 + r^3 + \dots + r^n = \frac{1 - r^{n+1}}{1 - r}$$

*Base case:*  $P(0)$  is true, because both sides of the equation are equal to 1.

*Inductive step:* We must show that  $P(n)$  implies  $P(n + 1)$  for all  $n \in \mathbb{N}$ . So assume that  $P(n)$  is true, where  $n$  denotes an arbitrary natural number. We can reason as follows:

$$\begin{aligned} 1 + r + r^2 + r^3 + \dots + r^n + r^{n+1} &= \frac{1 - r^{n+1}}{1 - r} + r^{n+1} \\ &= \frac{1 - r^{n+1} + (1 - r) \cdot r^{n+1}}{1 - r} \\ &= \frac{1 - r^{n+2}}{1 - r} \end{aligned}$$

The first equation follows from the assumption  $P(n)$ , and the remaining steps are simplifications. This proves that  $P(n + 1)$  is also true. Therefore,  $P(n)$  implies  $P(n + 1)$  for all  $n \in \mathbb{N}$ . By the principle of induction,  $P(n)$  is true for all  $n \in \mathbb{N}$ .  $\square$

## 2 Problem: Fibonacci

The Fibonacci sequence is defined by the equation  $F_n = F_{n-1} + F_{n-2}$ , with  $F_1 = 1$  and  $F_2 = 1$ . Prove that  $\sum_{i=1}^n F_i = F_{n+2} - 1$ .

**Solution.** *Proof.* The proof is by induction on  $n$ .

*Base case:* There are two things to check in the base, or we could think of it as having two base cases, with  $n = 1$  and  $n = 2$ . The first base case is  $\sum_{i=1}^1 F_i = 1 = F_3 - 1$ . The second base case is  $\sum_{i=2}^1 F_i = 1 + 1 = F_4 - 1$ .

*Induction step:* We must show that  $\sum_{i=1}^n F_i = F_{n+2} - 1$  implies that  $\sum_{i=1}^{n+1} F_i = F_{n+3} - 1$ .

$$\begin{aligned}\sum_{i=1}^{n+1} F_i &= \sum_{i=1}^n F_i + F_{n+1} \\&= F_{n+2} + F_{n+1} - 1 \\&= F_{n+3} - 1\end{aligned}$$

Thus, we conclude that  $\sum_{i=1}^n F_i = F_{n+2} - 1$ . □ ■

## 4 Problem: Surveyevor

In a new reality TV series called *Surveyevor*, a group of contestants is placed on a small island. Before the series begins, each contestant agrees to have a small purple or red tattoo, in the shape of an eye, applied to the middle of his or her forehead. In all, there are  $p \geq 1$  purple eyes and  $r \geq 0$  red eyes. However, none of the contestants knows the color of his or her third eye, nor how many total purple and red eyes there are. Furthermore, there are no mirrors and no one is allowed to discuss the tattoos ever. Therefore, everyone knows the colors of everyone else's third eye, but not their own. Good thing, because a contestant who learns that he or she has a purple eye must leave the island at the end of the show that day, and is therefore no longer eligible to win the \$1 million cash prize at the end of the show!

The contestants live in uneasy ignorance for several weeks. As time goes on, however, most of them lose their fear of being exiled, adapt to island living, and even make friends with one another. Things are going quite well for the islanders, but as you might suppose, the television audience grows bored, and the show's ratings plummet. When the network threatens to cancel the series, the producer decides she needs to do something, fast: on the next show, to the surprise of the happy islanders, the producer herself appears and convenes a meeting. Very loudly, she proclaims, "I see that at least one person here has a purple eye." Assuming that all the contestants are master logicians, what happens?

**Solution.** All the purple-tattooed contestants leave the island at the end of the  $p$ th day. ■

Use induction to prove that your conclusion is correct. We suggest a hypothesis  $P(n)$  that asserts all of the following are true on day  $n$ :

1. If  $p > n$ , then \_\_\_\_\_,
2. If  $p = n$ , then \_\_\_\_\_,
3. If  $p < n$ , then \_\_\_\_\_.

(We leave the task of filling in the blanks to you.)

**Solution.** Note that a red-eyed islander shouldn't ever conclude that she has a purple eye, since she doesn't, and we're assuming the contestants always reason correctly from what they know (and that what they know from the producer is also true). So no red-eyed contestant should ever leave the island.

**Theorem 2.** *All the purple-eyed people leave the island on day  $p$ .*

*Proof.* We use induction. Let  $P(n)$  be the proposition that all of the following are true on day  $n$ :

1. If  $p > n$ , then all purple-eyed people survive the day.

## Notes for Recitation 3

### 1 Problem: Breaking a chocolate bar

We are given a chocolate bar with  $m \times n$  squares of chocolate, and our task is to divide it into  $mn$  individual squares. We are only allowed to split one piece of chocolate at a time using a vertical or a horizontal break.

For example, suppose that the chocolate bar is  $2 \times 2$ . The first split makes two pieces, both  $2 \times 1$ . Each of these pieces requires one more split to form single squares. This gives a total of three splits.

Prove that the number of times you split the bar does not depend on the sequence of splits you make.

**Solution.** As with the “stacking game” from class, we approach the problem with experimentation, strong induction and a strong hypothesis. First we guess what the number of splits needed is:

**Theorem.** *To divide up a chocolate bar with  $m \times n$  squares, we need at most  $mn - 1$  splits.*

This theorem does not immediately lend itself to an induction or strong induction proof, since there are *two* variables. In general, propositions involving several natural-valued variables can often be proved by using a sort of nested induction (make sure to try that).

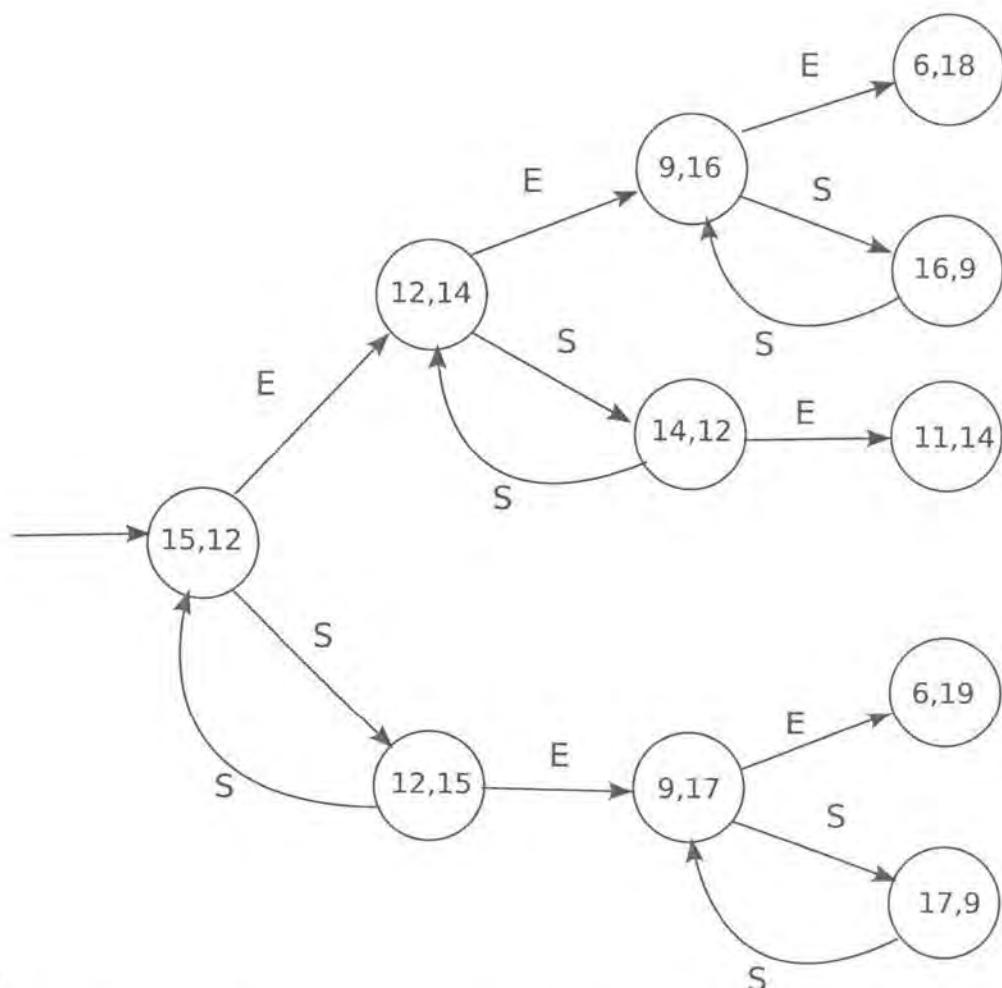
However, in this case, we can get by with a single-variable induction and a trick.

Intuitively, to break up a big chocolate bar, we need one split to make two pieces, and then we can break up the two pieces recursively. This suggests a proof using strong induction on the *size* of the chocolate bar, where size is measured in chocolate squares. Now instead of a problem involving two variables (the two dimensions), we have a problem in one variable (the size). With this simplification, we can prove the theorem using strong induction.

*Proof.* The proof is by strong induction on the size of the chocolate bar. Let  $P(k)$  be the proposition that a chocolate bar of size  $k$  requires at most  $k - 1$  splits.

*Base case,  $k = 1$ :*  $P(1)$  is true because there is only a single square of chocolate, and  $1 - 1 = 0$  splits are required.

*Induction step:* We suppose  $k \geq 1$  and any chocolate bar of size  $s$ , where  $1 \leq s \leq k$ , requires at most  $s - 1$  splits. We must now show there is a way to split a chocolate bar of size  $k + 1$  with at most  $k$  splits.



Solution.

Figure 1: State machine diagram for first few levels

*Inductive step:* Now assume that  $P(n)$  holds after  $n$  gong rings, where  $n \geq 0$ . Let  $r$  denote the number of red beads in the monk's bowl, and let  $g$  denote the number of green beads. In these terms, we are assuming that  $r - g$  is equal to  $5k + 2$  or  $5k + 3$  for some integer  $k$ . After  $n + 1$  gong rings, there are two cases to consider, depending on the monk's action:

1. If  $r \geq 3$ , then the monk may have exchanged 3 red beads for 2 green beads. Thus, the number of red beads minus the number of green becomes:

$$(r - 3) - (g + 2) = (r - g) - 5$$

This is equal to either  $5(k - 1) + 2$  or  $5(k - 1) + 3$ , so  $P(n + 1)$  is true.

2. Alternatively, the monk may have swapped every red bead for a green bead and vice versa. In this case, the number of reds minus the number of greens becomes  $g - r$ . If  $r - g = 5k + 3$ , then  $g - r = 5(-k) - 3 = 5(-k - 1) + 2$ . If  $r - g = 5k + 2$ , then  $g - r = 5(-k) - 2 = 5(-k - 1) + 3$ . Thus,  $P(n + 1)$  is again true.

Therefore,  $P(n)$  implies  $P(n + 1)$  for all  $n \geq 0$ .

By the induction principle,  $P(n)$  is true for all  $n \geq 0$ . Since the number of red beads minus the number of greens is always of the form  $5k + 2$  or  $5k + 3$  and the difference required to leave the temple does not match either form, no monk can ever leave the Temple of Forever.  $\square$

Now let's take a look at a different property of the Temple of Forever machine.

**Theorem 2.** *There is a finite number of reachable states in the Temple of Forever machine.*

Prove this theorem. (Hint: First find an invariant that suggests an upper bound on the number of reachable states. Be sure to prove the invariant.)

**Solution.** We begin by noting that the Temple of Forever machine exhibits the following invariant:

**Lemma 3.** *For all reachable states, the total number of red beads and green beads in the monk's bowl —  $r + g$  — is at most 27.*

*Proof.* We use induction on the number of gong rings. Let  $P(n)$  be the proposition that after  $n$  gong rings,  $r + g \leq 27$ .

*Base case:*  $P(0)$  is true because initially (after zero rings) the number of red beads plus the number of green beads is  $15 + 12 = 27$ .

*Inductive step:* Now assume that  $P(n)$  holds after  $n$  gong rings, where  $n \geq 0$ . Let  $r$  denote the number of red beads in the monk's bowl, and let  $g$  denote the number of green beads. In these terms, we are assuming that  $r + g$  is at most 27 after  $n$  gong rings. After  $n + 1$  gong rings, there are two cases to consider, depending on the monk's action:

**Solution.** *Proof.* The proof is by contradiction. Assume that it is possible for a monk to visit 108 unique states in some execution of the Temple of Forever machine, and consider the sequence of moves that the monk must have made to visit these states. Each move in the sequence must be either an *exchange* or a *swap*, since these are the only legal moves. Now, whenever the monk performs an *exchange* operation, the sum  $r + g$  decreases by one:

$$(r - 3) + (g + 2) = (r + g) - 1$$

In contrast, swaps do not have any effect on the sum. Furthermore, we know that the sum  $r + g$  must be at least 3 to perform an exchange operation. Therefore, there can be at most 25 exchange operations in the sequence.

Now consider swap operations: between each pair of exchanges in the sequence, there may be an unlimited number of swaps. However, only a single swap can take the monk to a new state: if at step  $k$  the monk is in state  $(r, g)$ , then at step  $k + 2$ , he will return to the same state. Therefore, an upper bound on the number of unique states in any execution of the machine is  $25 + 26 + 1 = 52$  (if swaps are inserted at both the beginning and end of the sequence). But then this contradicts the assumption that the monk visits 108 unique states, so no monk ever leaves the Temple of Forever.

□

■

What is the true maximal number of unique states a monk can visit in any execution of the Temple of Forever machine? How can this number be achieved?

**Solution.** The true maximum is 52. To achieve this number, the monk can perform sequential swaps and exchanges until he reaches the state  $(5, 2)$  via an exchange. At this point, the longest path goes to  $(2, 4)$ , via an exchange, instead of  $(2, 5)$ , via a swap. This is because the path leading to  $(2, 5)$  ends at  $(2, 5)$ , whereas the path leading to  $(2, 4)$  continues with swaps and exchanges, with the final state being  $(2, 0)$  (arrived at via a swap). However, the monk can reach 52 unique states if, at state  $(5, 2)$ , he performs two swap in a row to pick up state  $(2, 5)$ . Alternatively, the monk can perform two swaps at the start state, picking up state  $(12, 15)$ , and then continue with 25 pairs of sequential exchange, swap operations until he reaches  $(2, 0)$ . This also generates a path with 52 unique states.

■

## Notes for Recitation 4

### 1 Problem: The Pulverizer!

There is a pond. Inside the pond there are  $n$  pebbles, arranged in a cycle. A frog is sitting on one of the pebbles. Whenever he jumps, he lands exactly  $k$  pebbles away in the clockwise direction, where  $0 < k < n$ . The frog's meal, a delicious worm, lies on the pebble right next to his, in the clockwise direction.

- (a) Describe a situation where the frog can't reach the worm.

**Solution.** If  $k \mid n$  (say  $k = 3$  and  $n = 6$ ), then no number of jumps will lead the frog to the worm, as the frog will be returning to his original pebble ad infinitum. ■

(b) In a situation where the frog can actually reach the worm, explain how to use the Pulverizer to find how many jumps the frog will need.

**Solution.** Suppose the frog can reach the worm. When he actually reaches it, he has jumped a number of times, say  $j$ , and he has travelled around the cycle a number of times, call it  $c$ . Then, the distance that the frog has covered is both  $j \cdot k$  and  $c \cdot n + 1$ , so that

$$jk = cn + 1.$$

But this means that 1 can be written as a *linear combination* of  $n$  and  $k$ :

$$(-c)n + jk = 1,$$

Since 1 is positive, we conclude that it is a *positive linear combination* of  $n$  and  $k$ . And since it is the smallest positive integer, we also conclude that it is the *smallest positive linear combination* of  $n$  and  $k$ . But we have proved in lecture that the smallest positive linear combination of two integers is their GCD. So, the GCD of  $n$  and  $k$  is 1:

$$\gcd(n, k) = 1$$

and we can use the Pulverizer to find  $-c$  and  $j$ . ■

- (c) Compute the number of jumps if  $n = 50$  and  $k = 21$ . Anything strange? Can you fix it?

## Divisibility and modular Arithmetic

Divisibility: For all  $t, c$

If  $a|b$  and  $b|c \rightarrow a|c$

If  $a|b$  and  $a|c \rightarrow a|b+c$

For all  $c \neq 0$ ,  $a|b \Leftrightarrow ca|cb$

Greatest Common Divisor (GCD)

$$\gcd(ka, kb) = k \gcd(a, b)$$

$$\text{If } \gcd(a, b) = 1 \text{ and } \gcd(a, c) = 1 \Rightarrow \gcd(a, bc) = 1$$

$$\text{If } a|bc \text{ and } \gcd(a, b) = 1 \Rightarrow a|c$$

$$\text{If } m|a \text{ and } m|b \Rightarrow m|\gcd(a, b)$$

Modular Arithmetic

$$a \equiv a \pmod{n}$$

$$a \equiv b \pmod{n} \Rightarrow b \equiv a \pmod{n}$$

$$a \equiv b \pmod{n} \text{ and } b \equiv c \pmod{n} \Rightarrow a \equiv c \pmod{n}$$

$$a \equiv b \pmod{n} \Rightarrow a+c \equiv b+c \pmod{n}$$

$$a \equiv b \pmod{n} \text{ and } c \equiv d \pmod{n}$$

$$\Rightarrow a+c \equiv b+d \pmod{n}$$

$$a \equiv b \pmod{n} \text{ and } c \equiv d \pmod{n} \Rightarrow ac \equiv bd \pmod{n}$$

Multiplicative Inverse of  $x \pmod{n}$  ( $x^{-1} = y$ )  $x \cdot y \equiv 1 \pmod{n}$

$$a \equiv b \pmod{n} \Leftrightarrow n|(a-b)$$

$$\text{FF } \text{rem}(a, n) = \text{rem}(b, n)$$

## Properties of relations

Reflexivity:  $\forall x \in A. xRx$

"Everyone likes themselves"

"Every node has a loop"

Irreflexivity:  $\neg \exists x \in A. xRx$

"No one likes themselves, no loops"

Symmetry:  $\forall x, y \in A. xRy \Rightarrow yRx$

"If  $x$  likes  $y$ ,  $y$  likes  $x$ "

Antisymmetry:

$$\forall x, y \in A. (xRy \wedge yRx) \Rightarrow x=y$$

"No pair of distinct people can like each other"

Transitivity:  $\forall x, y, z \in A. (xRy \wedge yRz) \Rightarrow xRz$

Equivalence Relation: Reflexive, Symmetric

Weak Partial Order: Reflexive, Antisym and Transitive

Strong Partial Order: Irreflexive, Antisym and Transitive

## Well Ordering Principle

"Every non empty set of nonnegative ints has a smallest element"

Proof Template: (By contradiction)

i) Define set  $C$  of counterexamples

$$C := \{n \in \mathbb{N} \mid P(n) \text{ is False}\}$$

ii) Assume  $C$  is non empty

By WOP  $\rightarrow$  smallest element in  $C$

iii) Reach contradiction

iv) Conclude that  $C$  must be empty  $\square$

Greatest Common Denominator (GCD), repeat till done

Euclidean Algorithm:  $\gcd(a, b) = \gcd(b, \text{rem}(a, b))$

GCD can be written as a linear combination

$$\text{of } a, b. \quad \gcd(a, b) = sa + tb$$

Fulverizer - Find  $s$  and  $t$

$$x, y, \text{rem}(x, y) = x - q_1 y$$

$$259 \quad 70 \quad 49 = 259 - (3 \cdot 70)$$

$$70 \quad 49 \quad 21 = 70 - 1(49)$$

$$= 70 - 1(259 - 3(70))$$

$$= (-1)259 + 4(70)$$

$$49 \quad 21 \quad 7 = 49 - 2(21)$$

$$= (259 - 3 \cdot 70) - 2(-1 \cdot 259 + 4 \cdot 70)$$

$$= 3 \cdot 259 - 11(70)$$

$$21 \quad 7 \quad 0 \quad \text{so } a=3, t=-11$$

## Graph Theory

Chromatic #: min val of  $k$  for coloring problem

Basic Coloring Algorithm:

1) Order nodes by largest degree first

2) Order colors  $C_1, C_2, C_3$

3) For all nodes Assign  $v_i$  the lowest legal color will use at most  $\max(\text{degree})+1$  colors

Definitions:

Walk: Sequence of vertices in  $G$

Path: Walk with all different  $v_i$ 's

Cycle: Walk where  $v_0 = v_n$  and  $v_0, v_1, v_{k-1}$  are different.

Connectivity: Vertices : If there is a path from  $v$  to  $v'$

Graph: Every pair of vertices is connected

Acyclic connected graph

Spanning Tree: Subgraph (that is a tree) set of vertices,

A graph of  $n$ -nodes is a tree if:

connected via cycles and  $n-1$  edges.

Hamiltonian Cycle: Cycle that visits each node exactly once.

Euler Walk: Walk that traverses every edge once

Euler Tour: Euler walk that ends at start node

Networks Definitions:  $\nexists$  exists if a vertex has odd deg  $\Rightarrow$  indegree must = out degree + strongly connected

Distance b/w  $u-v$ : Shortest path from  $u$  to  $v$

Diameter of network: Distance b/w input/output apart

Congestion: of paths - max (of all nodes) paths

of Routing Problem Minimize through  $v$ .

congestion of best set of paths.

Network Congestion of worst case RP.

Network	# Switches	Diameter	Congestion
2D Grid	$N^2$	$2N$	2

(complete binary tree	$2N-1$	$2\log(N+2)$	$N$
-----------------------	--------	--------------	-----

Butterfly	$N(\log N+1)$	$\log(N+2)$	$\sqrt{N}$
-----------	---------------	-------------	------------

Benes Network	$2N(\log N+1)$	$2\log N+2$	1
---------------	----------------	-------------	---

Kraskals Algorithm: Find Minimum Spanning Tree (MST)

- start with an empty set

- add the minimum edge that does not create a cycle

- stop when there is no such edge.

$$\sum_{v \in V} \deg(v) = 2|E|$$

### Stable Marriage Problem

Find stable matching given ranked preferences for boy and girl.

Definitions:

Perfect matching - everyone gets married  
Rogue couple - when  $x$  and  $y$  prefer each other to their mates

Stable matching - no rogue couples

The Matching Algorithm (TMA)

Each Day:  
Morning: Boy serenades highest ranked girl still on list.

AFTERNOON: girls pick favorite among serenaders

EVENINGS: Boys cross girl off list if rejected.

Stop when every girl has at most one suitor.

- TMA Terminates within  $N^2+1$  days
- Everyone is married at end Stable matching
- Every boy paired with optimal mate
- Every girl paired with pessimal mate.
- Girls suitors only get better
- For guys it only gets worse.

### PageRank • Network with $n$ nodes

Initially every page =  $\frac{1}{n}$  PageRank

Every Update: each page distributes its PageRank equally along outgoing edges. And sets its new PR to sum of received shares.

Giving PR  $x$  can receive at most  $1-x$

### Multiplicative Inverse

$$k^{-1}k \equiv 1 \pmod{n} \quad \text{if } k \text{ relatively prime}$$

exists when  $\gcd(k, n) = 1$

2 ways to obtain:

1) Pulverizer

$$sk + tn = 1$$

$\uparrow$

Multiplicative inverse of  $k \pmod{n}$

2) Euler's Theorem

$$k^{\phi(n)} \equiv 1 \pmod{n}$$

Totient Function  $\phi(n)$

# of ints  $[0, n]$  relatively prime to  $n$

For prime  $p, q$

$$\phi(p) = p-1 \quad \phi(pq) = (p-1)(q-1)$$

$$\phi(n) = n \left(1 - \frac{1}{p_1}\right) \left(1 - \frac{1}{p_2}\right) \cdots$$

Prime Factors

$$k \cdot k^{\phi(n)-1} \equiv 1 \pmod{n}$$

$\overbrace{k}^{\text{Multiplicative inverse}}$

To show graph is  $k$ -colorable:

1) Show its possible to color  $G$  w/  $k$  colors.

2) Show that  $G$  requires at least  $k$  colors.

Eg: completely connected subgraph  
 $n$ -sized graph takes  $n$  colors

### Isomorphism

Map  $V_1$  of  $G_1$  to  $V_2$  of  $G_2$  and maintain edges

Tricks: All degree #s maintained  
All path lengths maintained

Trees: Add an edge  $\Rightarrow$  cycle

Remove edge  $e \Rightarrow$  disconnected

### Graph Theory Proofs

• Induct on size (# vertices)

• Induct on degree  $\rightarrow$  same deg tree  $V$   
common w/ regular graphs

Inductive step: start with size  $n+1$  graph,

remove a node to get size  $n$  that fits  $P(n)$ . Or else... buildup error!

## RSA

Receiver: pick two primes  $p, q$

compute  $n = pq$

select  $e$  such that

$$\gcd(e, (p-1)(q-1)) = 1$$

compute  $d$  such that

$$de \equiv 1 \pmod{(p-1)(q-1)}$$

Using pulverizer:

$$d \rightarrow ae + b(p-1)(q-1) = 1$$

Public Key:  $(e, n)$  private key  $(d, n)$

Sender: Encrypt using public key

$$m^* = \text{rem}(m^e, n)$$

Receiver: Decrypt using private key

$$M = \text{rem}(m^*)^d \pmod{n}$$

### Fermat's Little Theorem

For prime  $p$ :  $a^p \equiv a \pmod{p}$

$$a^{p-1} \equiv 1 \pmod{p}$$

### Halls Theorem:

Let  $G$  be a bipartite graph with  $(L, R)$

IF For every subset  $X \subseteq L$ ,  $L$  has more neighbors than its size  $|N(X)| \geq |X|$ .

$\Rightarrow$  Then, there is a matching for  $L$

## DAGs (Directed Acyclic Graphs)

Directed graphs with no cycles

Topological Sort! (think putting on clothes example)

Chains: Ordered set of vertices where, for a vertex  $x$  the vertex to its left is a prerequisite.

Antichains: Set of vertices where none of the vertices are prerequisites of each other.

Example Problem: Find remainder of  $38^{82248}$  divided by 83

1) 38 and 83 are relatively prime  
⇒ Use Euler's Theorem:

$$38^{\phi(83)} \equiv 1 \pmod{83}$$

↪ b/c prime

$$\phi(83) = 82$$

2) Try to remove as many 82's from power as possible (since they go to 1 (mod 83))

$$\begin{aligned} 38^{82248} &= 38^2 \cdot 38^{82 \cdot 1003} \\ &\equiv 38^2 \cdot 1^{1003} \pmod{83} = 144 \\ &\equiv 33 \pmod{83} \quad \therefore \boxed{\text{solution: 33}} \end{aligned}$$

Example Problem:  $\text{rem}(96^{123456789}, 97)$

$$\begin{aligned} 96 &\equiv -1 \pmod{97} \Rightarrow 96^{\text{odd power}} \equiv -1 \pmod{97} \\ &\Rightarrow \boxed{96} \end{aligned}$$

Find remainder of  $38^{82248}$  divided by 83

38 and 83 are relatively prime  
so we can use Euler's theorem

$$38^{\phi(83)} \equiv 1 \pmod{83}$$

and since 83 is prime

$$\phi(83) = 82$$

$\Rightarrow$  Try to remove as many of these  
from power as possible  
(since they go to 1 mod 83)

$$38^{82248} = 38^2 \cdot 38^{82 \cdot 1003}$$

$$\equiv 38^2 \cdot 1^{1003} \pmod{83}$$

$$= 1444$$

$$\equiv 33 \pmod{83}$$

so solution is  $\boxed{33}$

# Number Theory

## Divisibility:

for all  $t, c$

$$\text{If } a|b \text{ and } b|c \rightarrow a|c$$

$$\text{If } a|b \text{ and } a|c \rightarrow a|sb+tc$$

For all  $c \neq 0$ ,  $a|b \iff ca|cb$

## Greatest Common Denominator (GCD)

$$\gcd(ka, kb) = k \gcd(a, b)$$

$$\text{If } \gcd(a, b) = 1 \text{ and } \gcd(a, c) = 1$$

$$\rightarrow \gcd(a, bc) = 1$$

$$\text{If } a|bc \text{ and } \gcd(a, b) = 1 \rightarrow a|c$$

$$\text{If } m|a \text{ and } m|b \rightarrow m|\gcd(a, b)$$

## Modular Arithmetic

$$a \equiv a \pmod{n}$$

$$a \equiv b \pmod{n} \Rightarrow b \equiv a \pmod{n}$$

$$a \equiv b \pmod{n} \text{ and } b \equiv c \pmod{n} \Rightarrow a \equiv c \pmod{n}$$

$$a \equiv b \pmod{n} \Rightarrow a+c \equiv b+c \pmod{n}$$

$$a \equiv b \pmod{n} \text{ and } c \equiv d \pmod{n}$$

$$\Rightarrow a+c \equiv b+d \pmod{n}$$

$$a \equiv b \pmod{n} \text{ and } c \equiv d \pmod{n} \Rightarrow ac \equiv bd \pmod{n}$$

Multiplicative Inverse of  $x \pmod{n}$  ( $x^{-1} \equiv y$ )  $x \cdot y \equiv 1 \pmod{n}$

$$a \equiv b \pmod{n} \iff n|(a-b)$$

$$\iff \text{rem}(a, n) = \text{rem}(b, n)$$

## Greatest Common Denominator (GCD), repeat till done

$$\text{Euclidean Algorithm: } \gcd(a, b) = \gcd(b, \text{rem}(a, b))$$

GCD can be written as a linear combination of  $a, b$ .  $\gcd(a, b) = sa + tb$

Pulverizer - Find  $s$  and  $t$

$$x \dots y \dots \text{rem}(x, y) = x - qy$$

$$259 \quad 70 \quad 49 = 259 - (3 \cdot 70)$$

$$70 \quad 49 \quad 21 = 70 - 1(49) \\ = 70 - 1(259 - 3 \cdot 70) \\ = (-1)259 + 4(70)$$

$$49 \quad 21 \quad 7 = 49 - 2(21) \\ = (259 - 3 \cdot 70) - 2(-1 \cdot 259 + 4(70)) \\ = 3 \cdot 259 - 11(70)$$

$$21 \quad 7 \quad 0 \quad \Rightarrow \text{so } a=3, t=-11$$

## Multiplicative Inverse

$$k^{-1} \equiv 1 \pmod{n} \quad (\text{relatively prime})$$

exists when  $\gcd(k, n) = 1$

2 ways to obtain:

1) Pulverizer

$$sk + tn = 1$$

↑  
Multiplicative inverse of  $k \pmod{n}$

2) Euler's Theorem

$$k^{d(n)} \equiv 1 \pmod{n}$$

Totient Function  $\phi(n)$

# of ints  $[0, n]$  relatively prime to  $n$

For prime  $p, q$

$$\phi(p) = p-1 \quad \phi(pq) = (p-1)(q-1)$$

$$\phi(n) = n \left(1 - \frac{1}{p_1}\right) \left(1 - \frac{1}{p_2}\right) \dots$$

Prime Factors

$$k \cdot k^{\phi(n)-1} \equiv 1 \pmod{n}$$

↑ Multiplicative inverse

## Graph Theory

Chromatic #: min val of  $k$  for coloring problem

Basic Coloring Algorithm:

1) Order nodes by largest degree first

2) Order colors  $C_1, C_2, C_3$

3) For all nodes Assign  $V_i$  the lowest legal color will use at most  $\max(\text{degree})+1$  colors

Definitions: Odd length cycle  $\rightarrow$  not bipartite

Walk: Sequence of vertices in  $\tilde{G}$

Path: Walk with all different  $V_i$ 's

Cycle: Walk where  $V_0 = V_k$  and  $V_0, V_1, V_{k-1}$  are different.

Connectivity: Vertices: If there is a path from  $v$  to  $u$

Graph: Every pair of vertices is connected with the same

Spanning Tree: Subgraph (that is a tree) set of vertices.

A graph of  $n$ -nodes is a tree if:

connected w/o cycles and  $n-1$  edges.

Hamiltonian cycle: Cycle that visits each node exactly once.

Euler Walk: Walk that traverses every edge once

Euler Tour: Euler walks that ends at start node

won't exists if a vertex has odd deg

To show graph is  $k$ -colorable:

1) Show its possible to color  $G$  w/  $k$  colors.

2) Show that  $G$  requires at least  $k$  colors.

Eg: completely connected subgraph  
 $n$ -sized graph takes  $n$  colors

## Isomorphism

Map  $V_1$  of  $G_1$  to  $V_2$  of  $G_2$  and maintain edges

Tricks: All degree  $\#$ 's maintained

All path lengths maintained

Trees: Add an edge  $\Rightarrow$  cycle

Remove edge  $e \Rightarrow$  disconnected

## Graph Theory Proofs

• Induct on size (# vertices)

• Induct on degree  $\rightarrow$  same dg tree  $V$

• Contradiction w/ regular graphs

Inductive step: Start with size  $n+1$  graph,

remove a node to get size  $n$  that fits  $f(n)$ . Or else... buildup error!

Hall's Theorem:  
Let  $G$  be a bipartite graph with  $(L, R)$   
If for every subset  $X \subseteq L$ ,  $L$  has more neighbors than it's size  $N(X) \geq |N|$ .

⇒ Then, there is a matching for  $L$

## RSA

Receiver: pick two primes  $p, q$

compute  $n = pq$

select  $e$  such that

$$\gcd(e, (p-1)(q-1)) = 1$$

compute  $d$  such that

$$de \equiv 1 \pmod{(p-1)(q-1)}$$

\* Using pulverizer:

$$d \rightarrow ae + b(p-1)(q-1) = 1$$

Public Key:  $(e, n)$  private key  $(d, n)$

Sender: encrypt using public key

$$m^e = \text{rem}(m^e, n)$$

Receiver: Decrypt using private key

$$m = \text{rem}(m^d, n)$$

Fermat's Little Theorem

For prime  $p$ :  $a^p \equiv a \pmod{p}$

$$a^{p-1} \equiv 1 \pmod{p}$$

## Stable Marriage Problem

Find stable matching given ranked preferences for boy and girl

Definitions:

Perfect matching - everyone gets married

Rogue couple - when  $x$  and  $y$  prefer each other to their mates

Stable matching - no rogue couples

The Mating Algorithm (TMA)

Each day: Boy serenades highest ranked

girl still on list.

Afternoon: Girls pick favorite among serenaders

Evenings: Boys cross girl off list if rejected.

Stop when every girl has at most one suitor.

TMA Terminates within  $N^2+1$  days

Everyone is married at end Stable matching

Every boy paired with optimal mate

Every girl paired with optimal mate

Boys' suitors only get better

For guys: it only gets worse.

Kruskal's Algorithm: Find Minimum Spanning Tree (MST)  
Start with an empty set  
Add the minimum edge that does not create a cycle  
Stop when there is no such edge.

Hopcroft-Karp Lemma:  
 $|E| = |V| \leq N^2$

Kruskal's Algorithm: Find Minimum Spanning Tree (MST)  
Start with an empty set  
Add the minimum edge that does not create a cycle  
Stop when there is no such edge.

## Well Ordering Principle

\* Every non empty set of nonnegative ints has a smallest element

Proof Template: (by contradiction)

1) Define set C of counterexamples

$$C := \{ n \in \mathbb{N} \mid P(n) \text{ is False} \}$$

2) Assume C is non empty

By WOP  $\Rightarrow$  smallest element in C

3) Reach contradiction

4) Conclude that C must be empty  $\square$

## Properties of relations

Reflexivity:  $\forall x \in A, xRx$

"Everyone likes themselves"

"Every node has a loop"

Irreflexivity:  $\neg \exists x \in A, xRx$

"No one likes themselves, no loops"

Symmetry:  $\forall x, y \in A, xRy \Rightarrow yRx$

"If x likes y, y likes x"

Antisymmetry:

$\forall x, y \in A, (xRy \wedge yRx) \Rightarrow x = y$

"No pair of distinct people can like each other"

Transitivity:  $\forall x, y, z \in A, (xRy \wedge yRz) \Rightarrow xRz$

Equivalence Relation: Reflexive, Symmetric

Weak Partial Order: Reflexive, Antisym and Transitive

Strong Partial Order: Irreflexive, Antisym and Transitive

## Networks Definitions

Distance b/w u-v: shortest path from u to v

Diameter of network: distance b/w input/output apart

Congestion: # of paths - max (# of all nodes) paths

of Routing Problem: minimize congestion through v.

Congestion of best set of paths.

of Network: congestion of worst case RP.

N inputs # switches Diameter Congestion

2D Grid  $N^2$   $2N$  2

Complete Binary Tree  $2^{N-1}$   $2\log(N+1)$   $N$

Butterfly  $N(\log(N))$   $\log(N+2)$   $\sqrt{N}$

Benes Network  $2N(\log N)$   $2\log(N+2)$  1

## DAGs (Directed Acyclic Graphs)

Directed graphs with no cycles

Topological Sort: (think putting on clothes example)

Chains: Ordered set of vertices where, for a vertex x  
the vertex to its left is a prerequisite.

Antichains: Set of vertices where none of the  
vertices are prerequisites of each other.

PageRank • Network with n nodes

Initially every page =  $\frac{1}{n}$  PageRank

Every Update: each page distributes

its PageRank equally along outgoing

edges. And sets its new PR to

sum of received shares.

Giving PR X can receive at most  $1-x$

Example Problem: Find remainder of  $38^{82248}$   
divided by 83

1) 38 and 83 are relatively prime

$\Rightarrow$  Use Euler's Theorem:

$$38^{\phi(83)} \equiv 1 \pmod{83}$$

$\phi(83) = 82$   
 $\approx$  b/c prime

2) Try to remove as many  $82$ 's from power  
as possible (since they go to 1  $\pmod{83}$ )

$$38^{82248} = 38^2 \cdot 38^{82 \cdot 1003}$$

$$\equiv 38^2 \cdot 1^{1003} \pmod{83} = 144$$

$$\equiv 33 \pmod{83} \therefore \boxed{\text{Solution: 33}}$$

Example Problem:  $\text{rem}(96^{123456789}, 97)$

$$96 \equiv -1 \pmod{97} \Rightarrow 96^{\text{odd power}} \equiv -1 \pmod{97}$$

$\Rightarrow \boxed{96}$

## Well Ordering Principle

"Every non empty set of nonnegative integers has a smallest element"

Proof Template: (by contradiction)

1) Define set  $C$  of counterexamples

$$C := \{n \in \mathbb{N} \mid P(n) \text{ is False}\}$$

2) Assume  $C$  is non empty  
by WOP  $\Rightarrow$  smallest element in  $C$

3) Reach contradiction

4) Conclude that  $C$  must be empty  $\square$

## Properties of relations

Reflexivity:  $\forall x \in A, xRx$

"Everyone likes themselves"

"Every node has a loop"

Irreflexivity:  $\neg \exists x \in A, xRx$

"No one likes themselves, no loops"

Symmetry:  $\forall x, y \in A, xRy \Rightarrow yRx$

"If  $x$  likes  $y$ ,  $y$  likes  $x$ "

Antisymmetry:

$\forall x, y \in A, (xRy \wedge yRx) \Rightarrow x=y$

"No pair of distinct people can like each other"

Transitivity:  $\forall x, y, z \in A, (xRy \wedge yRz) \Rightarrow xRz$

Equivalence Relation: Reflexive, Symmetric and Transitive

Weak Partial Order: Reflexive, Antisymmetric and Transitive

Strong Partial Order: Irreflexive, Antisymmetric and Transitive

Truth Table:

x	y	$x \rightarrow y$
T	T	T
T	F	F
F	F	T
F	T	T

$$\text{contrapositive } x \rightarrow y = \bar{y} \rightarrow \bar{x}$$

## Networks Definitions

Distance b/w  $v-v$ : shortest path from  $v$  to  $v$   
Diameter of network: distance b/w input/output apart  
Congestion: of paths - max (of all nodes) paths of Routing Problem minimize congestion through  $v$ .  
longest of best set of paths.

(N input)	# switches	Diameter	Congestion
2D Grid	$N^2$	$2N$	2
Complete binary Tree	$2^{N-1}$	$2\log(N+1)$	$N$
Butterfly	$N(\log(N))$	$\log(N+2)$	$\sqrt{N}$
Benes Network	$2N(\log N)$	$2\log(N+2)$	1

## DAG's (Directed Acyclic Graphs)

Directed graphs with no cycles

Topological Sort: (think putting on clothes example)  
every finite poset has a topological sort

Chains: Ordered set of vertices where, for a vertex the vertex to its left is a prerequisite.

Antichains: set of vertices where none of the vertices are prerequisites of each other.

Dilworth's Lemma: Every DAG with  $n$  vertices needs either a chain of size  $\lceil n/2 \rceil$  or an antichain of at least  $n/2$

Any planar graph can be colored in at most 6 colors.  
 $\hookrightarrow$  none of the lines cross

Give a description of the equivalence class.

a) Integers  $x$  and  $y$  are equivalent if

$$x \equiv y \pmod{3}$$

$$A: \left\{ \dots, -6, -3, 0, 3, 6, \dots \right\}$$

$$\left\{ \dots, -5, -2, 1, 4, 7, \dots \right\}$$

$$\left\{ \dots, -4, -1, 2, 5, 8, \dots \right\}$$

PageRank • Network with  $n$  nodes!  
Initially every page =  $\frac{1}{n}$  Page Rank  
Every update, each page distributes its PR equally along outgoing edges. And sets its new PR to sum of received shares.

Giving PR  $x$  can receive at most  $1-x$

Sample Problem:



Receive PR from vertices that point to it.

$$\begin{bmatrix} A' \\ B' \\ C' \\ D' \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} A \\ B \\ C \\ D \end{bmatrix}$$

A B C D  
↓ ↓ add up to 1.

$$P_A = \frac{1}{2} P_0$$

$$P_B = \frac{1}{2} P_A + \frac{1}{2} P_C + \frac{1}{2} P_D$$

$$P_C = \frac{1}{2} P_A + P_B$$

$$P_D = \frac{1}{2} P_C$$

$$P_A + P_B + P_C + P_D = 1$$

Hasse Diagram - Directed acyclic graph. Represent a poset.

- Remove edges implied by transitivity

- Every finite poset has a topological sort.

Example Problem: Find remainder of  $38^{82248}$  divided by 83

1) 38 and 83 are relatively prime  
 $\Rightarrow$  Use Euler's Theorem:

$$\phi(83) \equiv 1 \pmod{83}$$

$\because 83$  prime

$$\phi(83) = 82$$

2) Try to remove as many 82's from power as possible (since they go to 1  $\pmod{83}$ )

$$38^{82248} = 38^2 \cdot 38^{82 \cdot 1003}$$

$$\equiv 38^2 \cdot 1^{1003} \pmod{83} = 144$$

$$\equiv 33 \pmod{83} \quad \therefore \boxed{\text{Solution: 33}}$$

Example Problem:  $\text{rem}(96^{123456789}, 97)$

$$96 \equiv -1 \pmod{97} \Rightarrow 96^{\text{odd power}} \equiv -1 \pmod{97}$$

$\Rightarrow \boxed{96}$

Any simple graph with  $n$  nodes and strictly more than  $\frac{1}{2}(n-1)(n-2)$  edges is connected.



This lecture is not on the First quiz  
Chapter 9, 10.0 - 10.2

Def: An  $n$ -year  $\$m$ -payment annuity pays  $\$m$  at the start of each year for a total of  $n$  years.

Assumption: Fixed annual interest rate  $p$

$$\$1 \text{ today} = \$1(1+p) \text{ in 1 year}$$

$$\$1 \text{ today} = \$1(1+p)^2 \text{ in 2 years}$$

$$\$ \frac{1}{1+p} \text{ today} = \$1 \text{ in 1 year}$$

$$\$ \frac{1}{(1+p)^2} \text{ today} = \$1 \text{ in 2 years}$$

<u>Current Value</u>	<u>Payments</u>
$\$m$	$\$m$ now
$\$ \frac{m}{(1+p)}$	$\$m$ in 1 yr
$\$ \frac{m}{(1+p)^2}$	... 2 yrs
$\vdots$	$\vdots$
$\$ \frac{m}{(1+p)^{n-1}}$	$\$m$ in $n-1$ years

$$V = \sum_{i=0}^{n-1} \frac{m}{(1+p)^i}; \quad = \text{Total current value.}$$

$$= m \sum_{i=0}^n x^i \quad \text{where } x = \frac{1}{1+p}$$

can prove by induction  
but how do we derive?

$$\text{Thm: if } n \geq 1, x \neq 1, \sum_{i=0}^{n-1} x^i = \frac{1-x^n}{1-x}$$

Perturbation Method:  $S = 1 + x + x^2 + x^3 + \dots + x^{n-1}$

$$\frac{-xS = x + x^2 + x^3 + \dots + x^{n-1} + x^n}{(1-x)S = 1 - x^n} \Rightarrow S = \frac{1-x^n}{1-x}$$

$$\text{So, } V = m \sum_{i=0}^n x^i \text{ where } x = \frac{1}{1+p}$$

$$V = m \left( \frac{1-x^n}{1-x} \right) \\ = m \frac{1 - \left( \frac{1}{1+p} \right)^n}{1 - \frac{1}{1+p}} = m \left( \frac{1+p - \frac{1}{(1+p)^{n+1}}}{p} \right)$$

$$\text{So if } n=\infty, \text{ then } V = m \frac{1+p}{p} \quad \text{so if } m=50k \ p=.06 \\ V = 883,333$$

Therefore, if the interest rate is .06%. you are better off getting \$1M today than \$50k every year forever.

Corollary

$$\sum_{i=0}^{n-1} x^i = \frac{1-x^n}{1-x} \Rightarrow \sum_{i=0}^{\infty} x^i = \frac{1}{1-x}$$

What would you pay for a company that grows by \$m every year.

$$\text{Todays value: } \frac{m}{1+p} + \frac{2m}{(1+p)^2} + \frac{3m}{(1+p)^3} + \dots = \sum_{i=1}^{\infty} \frac{im}{(1+p)^i}$$

$$= m \sum_{i=1}^{\infty} i x^i \text{ where } x = \frac{1}{1+p}$$

$$\sum_{i=1}^n i x^i = x + 2x^2 + 3x^3 + \dots + n x^n$$

Perturbation Method:

$$S = x + 2x^2 + 3x^3 + \dots + nx^n \\ - xS = x^2 + 2x^3 + \dots + (n-1)x^n + nx^{n+1}$$

$$(1-x)S = \underbrace{x + x^2 + x^3 + \dots + x^n}_{\frac{1-x^{n+1}}{1-x}-1} - nx^{n+1} \Rightarrow S = \frac{x - (n+1)x^{n+1} + nx^{n+2}}{(1-x)^2}$$

## Derivative Method

$$\text{For } x \neq 1 \quad \sum_{i=0}^n x^i = \frac{1-x^{n+1}}{1-x}$$

Take derivative of both sides

$$\Rightarrow \sum_{i=0}^n i x^{i-1} = \frac{-(1-x)(n+1)x^n - (1-x^{n+1})(-1)}{(1-x)^2}$$

$$= \frac{1 - (n+1)x^n + nx^{n+1}}{(1-x)^2}$$

Multiply by  $x$   
 $\Rightarrow \sum_{i=0}^n i x^i$

$$|x| < 1$$

So if  $n \rightarrow \infty$

$$S = \frac{x}{(1-x)^2}$$

$$\sum_{i=1}^{\infty} i 2^{-i} = \frac{1}{2} + \frac{2}{4} + \frac{3}{8} + \dots = \frac{1/2}{(1-1/2)^2} = 2$$

$$\text{Company Value} = m \frac{\frac{1}{1+p}}{\left(1 - \frac{1}{1+p}\right)^2} = \frac{m(1+p)}{p^2}$$

$$\text{So if } m = 50K, p = 0.06 \quad V = 1,222,222$$

Another example of Perturbation Method

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

$$S = 1 + 2 + 3 + \dots + n-1 + n \quad \text{Reversed}$$

$$+ S = n + n-1 + n-2 + \dots + 2 + 1$$

$$2S = (n+1) + (n+1) + (n+1) + \dots + (n+1)$$

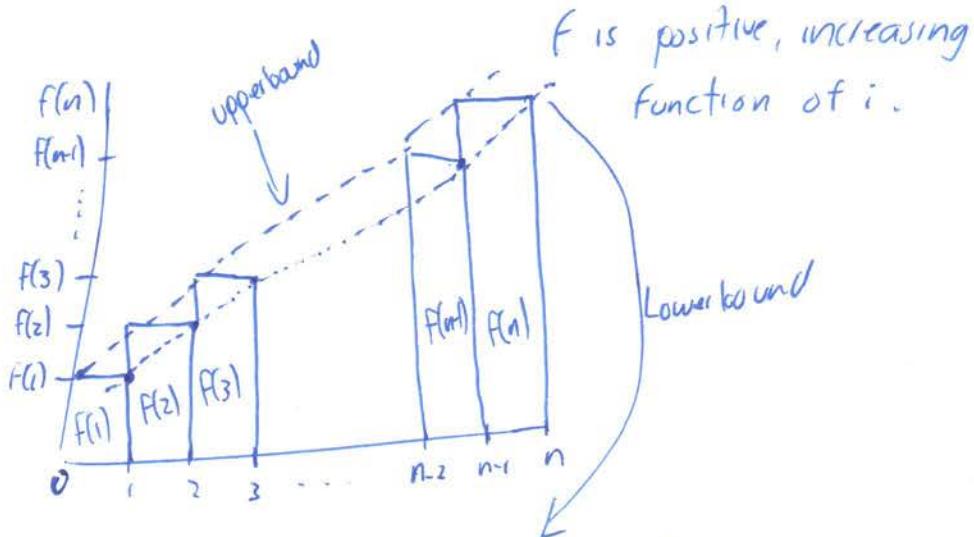
$$2S = n(n+1)$$

$$S = \frac{n(n+1)}{2}$$

"Nasty One"

$$\sum_{i=0}^n \sqrt{i}$$

Integration Bounds for  $\sum_{i=1}^n f(i)$  where



$$f(1) + \int_1^n f(x) dx \leq \sum_{i=1}^n f(i) \leq \int_1^n f(x) dx + f(n)$$

Ex:  $f(i) = \sqrt{i}$

$$\int_1^n \sqrt{x} dx = \frac{x^{3/2}}{3/2} \Big|_1^n = \frac{2}{3} (n^{3/2} - 1)$$

$$\frac{2}{3} (n^{2/3} - 1) + 1 \leq \sum_{i=1}^n \sqrt{i} \leq \frac{2}{3} (n^{2/3} - 1) + \sqrt{n}$$

$$\frac{2}{3} n^{2/3} + 1/3 \leq \sum_{i=1}^n \sqrt{i} \leq \frac{2}{3} n^{3/2} + \sqrt{n} - 2/3$$

So sum of first 100  $\sqrt{i}$ :  $n=100$

$$667 \leq \sum_{i=1}^{100} \sqrt{i} \leq 676$$

Don't know exactly, but within range.

$$\sum_{i=1}^n \sqrt{i} = \frac{2}{3} n^{3/2} + \delta(n) \text{ where } \frac{1}{3} \leq \delta(n) \leq \sqrt{n} - 2/3$$

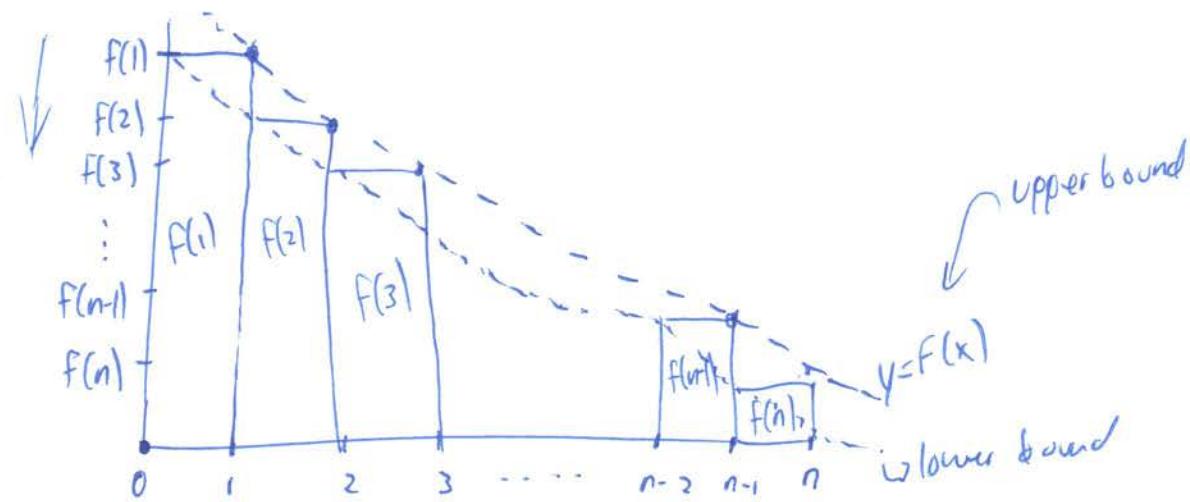
OR

$$\sum_{i=1}^n \sqrt{i} \sim \frac{2}{3} n^{3/2}$$

Def:  $g(x) \sim h(x)$  means  $\lim_{x \rightarrow \infty} \frac{g(x)}{h(x)} = 1$

Integration Bounds for  $\sum_{i=m}^n f(i)$  when  $f$  is positive and decreasing

Ex:  $\sum_{i=1}^n \sqrt{i}$



$$\int_1^n f(x) dx + f(n) \leq \sum_{i=1}^n f(i) \leq \int_1^n f(x) dx + f(1)$$

Same as increasing with these two switched.

Ex:  
 $f(i) = \sqrt{i}$      $\int_1^n \frac{dx}{\sqrt{x}} = \frac{\sqrt{x}}{1/2} \Big|_1^n = 2(\sqrt{n} - 1) = 2\sqrt{n} - 2$

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

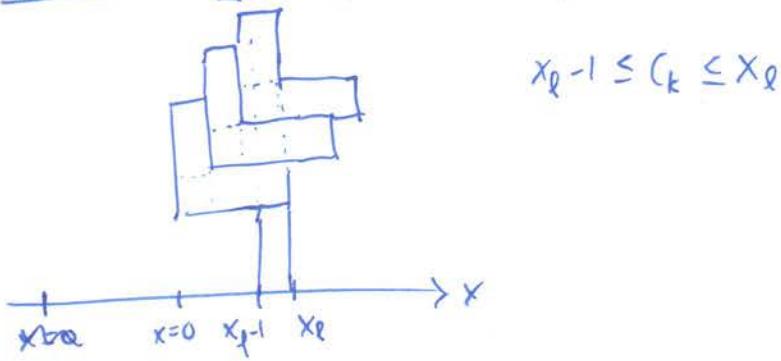
# Counting ↗ Sums Asymptotics

Key to Counting: STRUCTURE

## Problems

## L-tower

Proof :  $\exists k$  .  $k$  L-shaped blocks stacks on a stand is stable



Condition 1: The top block has to be stable on the 2nd block

Condition 2: The top 2 blocks have to rest on the bottom block

Condition 3: 3 blocks have to rest on the stand

So, 10 blocks  $\rightarrow$  10 conditions

Center of gravity of a stack of 1 block :  $C_1 = \frac{1}{2} + \frac{x\ell}{2}$

$$C_h = \frac{C_1 + (C_1 + 1) + (C_1 + 2) + \dots + (C_1 + (h-1))}{h}$$

$$= \frac{h(1 + (1+2 + \dots + (h-1)))}{1}$$

$$= \frac{h\left(1 + \frac{h(h-1)}{2}\right)^h}{h} = C_1 + \frac{(h-1)}{2}$$

$$C_h = C_1 + \frac{(h-1)}{z}$$

## Problems for Recitation 11

### 1 The L-tower problem

Observe the structures shown in Figure 1. One has 2 L-shapes, the other 5 L-shapes. Consider a tower with  $k$  L-shapes. Assume that the blocks are all of size  $x \times 1$  where  $x > 1$ . As the picture indicates, if  $k$  is too small then the tower falls to the left. On the other hand, if  $k$  is too large the tower would fall to the right. Will the tower be stable for some  $k$ ? Prove there is at least one value of  $k$  for which the L-tower is stable. Assume that a structure is stable if and only if its center of gravity is not hanging in the air horizontally. The L-tower is stable if and only if each of its subparts is stable.

*Hint:* Show the tower is stable if and only if  $\frac{3x-3}{2} \leq k \leq \frac{3x-1}{2}$ .

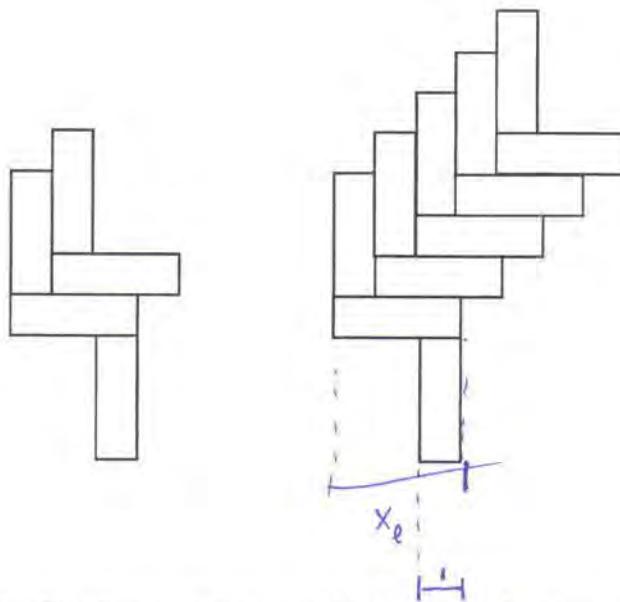


Figure 1: Too few or too many L shapes make the tower unstable

- (b) If we think about the pairs  $(k, j)$  over which we are summing, they form a triangle in the table below. The values in the cells of the table correspond to the terms in the double summation. Complete this table to see the pattern.

$k \setminus j$	1	2	3	4	$\dots$	$n$
1						
2						
3						
4						
$\vdots$						
$n$						

- (c) The summation above is summing each row and then adding the row sums. But we can tame this beast if, instead, we first sum the columns and then add the column sums. Use the table to rewrite the double summation. The inner summation should sum over  $k$ , and the outer summation should sum over  $j$ .
- (d) Now simplify the summation to derive a closed formula for  $S_n$ .

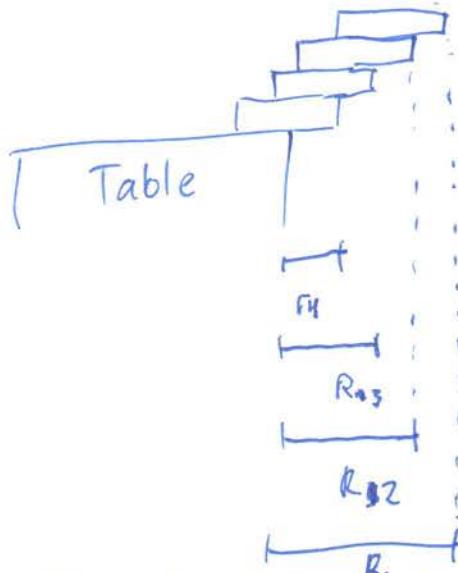
Now try your hand at another double sum.

- (e) Find a formula for  $F_n$ , for  $n$  a positive integer:

$$F_n = \sum_{k=1}^n k \sum_{j=k}^n 2^j / (j+1).$$

Given:  $n$  blocks of length 1 ( $L=1$ )

Def:  $r_i$  = amount by which the  $i$ th block extends beyond table



Stability Constraint: The center of mass ( $k$  of the top  $k$  blocks) must lie on the  $(k+1)$ st block. (Table =  $n+1$  blocks)

For Greedy Strategy:  $C_k = r_{k+1}$

Center of mass of  $i$ th block is at  $r_i - 1/2$

$$C_k = \frac{(r_1 - 1/2) + (r_2 - 1/2) + (r_k - 1/2)}{K}$$

$$r_{k+1} = \frac{r_1 + r_2 + r_3 + \dots + r_k - k/2}{K}$$

$$Kr_{k+1} = r_1 + r_2 + \dots + r_k - k/2$$

Perturbation

$$-(k-1)r_k = r_1 + r_2 + \dots + r_{k-1} - \frac{k-1}{2}$$

$$Kr_{k+1} - (k-1)r_k = r_k - 1/2$$

$$k(r_{k+1} - r_k) = -\frac{1}{2}$$

$$\Rightarrow r_k - r_{k+1} = \frac{1}{2k}$$

$$r_1 - r_2 = \frac{1}{2}$$

$$r_2 - r_3 = \frac{1}{4}$$

$$r_3 - r_4 = \frac{1}{6}$$

$$r_n - r_{n+1} = \frac{1}{2n}$$

$$r_1 - r_{n+1} = \frac{1}{2} + \frac{1}{4} + \frac{1}{6} + \dots + \frac{1}{2n}$$

$$r_1 = \sum_{i=1}^n \frac{1}{2i}$$

$$= \frac{1}{2} \sum_{i=1}^n \frac{1}{i}$$

$= \frac{1}{2} H_n$  where  $H_n$  = nth Harmonic number

$$H_1 = 1$$

$$H_2 = 1 + \frac{1}{2} = \frac{3}{2}$$

$$H_3 = \frac{3}{2} + \frac{1}{3} = \frac{11}{6}$$

$$H_4 = \frac{25}{12} > 2$$

So 4 blocks is enough to get one block over the table.

$$H_{100,000,000} = 1439$$

With 102,029,000 blocks the last one could be 7 blocks over.

$$\frac{1}{2} H_n$$

## Integration Bounds for decreasing sum

$$f(n) + \int_1^n F(x) dx \leq \sum_{i=1}^n f(i) \leq f(1) + \int_1^n f(x) dx$$

$f(i) = 1/i$   
 $\int_1^n \frac{dx}{x} = |\ln x|_1^n$   
 $= \ln n$

$$\frac{1}{n} + \ln n \leq \sum_{i=0}^n 1/i \leq 1 + \ln n$$

$$H_n \sim \ln n \quad t \ln = \ln n + \underbrace{\frac{1}{2n} - \frac{1}{12n^2} + \frac{E(n)}{120n^4}}_{\text{don't need to know this.}}$$

Euler's constant = 0.57721

$$n! = n(n-1) \cdots 2 \cdot 1 = \prod_{i=1}^n i$$

$$\begin{aligned} \ln(n!) &= \ln(1 \cdot 2 \cdot 3 \cdots n) = \ln(1) + \ln(2) + \cdots + \ln(n) \\ &= \sum_{i=1}^n \ln(i) \end{aligned}$$

## Integration Bounds for increasing function

$$f(1) + \int_1^n f(x) dx \leq \sum_{i=1}^n f(i) \leq f(n) + \int_1^n f(x) dx$$

$$f(i) = \ln(i)$$

$$0 + n \ln n - n + 1 \leq \sum_{i=1}^n \ln(i) \leq \ln n + n \ln n - n + 1$$

$$\int_1^n \ln x dx$$

$$x \ln x - x \Big|_1^n$$

$$\begin{aligned} e^{n \ln(n) - n + 1} &\stackrel{\text{Exponentiate}}{\leq} n! \leq e^{(n+1) \ln(n+1)} \\ \frac{n^n}{e^{n-1}} &\leq n! \leq \frac{n^{n+1}}{e^{n-1}} \end{aligned}$$

\* Stirling's Formula :  $n! = \left(\frac{n}{e}\right)^n \sqrt{2\pi n} e^{\varepsilon(n)}$

where  $\frac{1}{12n+1} \leq \varepsilon(n) \leq \frac{1}{12n}$

Ex:

$$100! \geq \left(\frac{100}{e}\right)^{100} \sqrt{200\pi} e^{\frac{1}{1201}}$$

$$n! \sim \left(\frac{n}{e}\right)^n \sqrt{2\pi n}$$

$$100! \leq \dots \dots e^{\frac{1}{1200}}$$

## Asymptotic Notation

"tilde"  $f(x) \sim g(x)$  if  $\lim_{x \rightarrow 0} \frac{f(x)}{g(x)} = 1$

"Oh", "big O"  $f(x) = \underbrace{O(g(x))}_{1}$  if  $\lim_{x \rightarrow 0} \frac{f(x)}{g(x)} \leq \infty$  (finite)

Upper Bound.

Thm Let  $f(x) = x$ ,  $g(x) = x^2$ . Then  $f(x) = O(g(x))$

Proof

$$\lim_{x \rightarrow 0} \frac{f(x)}{g(x)} = \lim_{x \rightarrow 0} \frac{x}{x^2} = 0 \leq \infty$$

$$x = O(x^2)$$

So Is  $x^2 = O(10^6 x)$ ?

No  $\lim := \infty \times$

$100x^2 = O(x^2)$ ?

Yes  $\lim \frac{100x^2}{x^2} = 100 < \infty \checkmark$

ignore constants

Example:

Time to multiply  $n \times n$  matrices is  $T(n) = O(n^3)$

$$T_n = \ln n + 5 + O(\ln n) \quad T_n - \ln n - 5 = O(\ln n)$$

WARNING: BAD THINGS COMING UP:

$f(n) \geq O(g(n)) \Rightarrow_{\text{BAD}} \text{meaningless}$

$O$  is upperbound not lower bound

" $\Omega$ "  $f(x) = \Omega(g(x))$  if  $\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} > 0$

Lower bound:  $f(x)$  grows at least as fast as  $g$

-Opposite of  $O$

" $\Theta$ "  $f(x) = \Theta(g(x))$  if  $\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} < \infty, > 0$

grow at the same rate.

Both  $O$  and  $\Omega$

Then:  $f(x) = \Theta(g(x))$  iff  $f(x) = O(g(x))$  and  $f(x) = \Omega(g(x))$

Ex  $10x^3 - 20x^2 + 1 = \Theta(x^3)$

$x/\log x = \Theta(x)??$  Nope X

## Summary

$O$  means  $\leq$

$\Omega$  "  $\geq$

$\Theta$  "  $=$

$o$  "  $<$

$w$  "  $>$

"little oh"  $f(x) = o(g(x)) \Leftrightarrow \lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = 0$

"little omega"  $f(x) = w(g(x)) \Leftrightarrow \lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = \infty$

$$\ln x = o(x) \quad x^2 = w(x)$$

Thm (NOT) Let  $f(n) = \sum_{i=1}^n i$   
Then  $f(n) = O(n)$ .  $f(n) = \Theta(n^2)$   
 $\frac{n(n+1)}{2}$

False Proof (By induction on  $n$ )

I.H  $P(n) : f(n) = O(n)$

Base Case  $f(1) = 1 = O(1) \checkmark$

Inductive Step : Assume  $P(n)$  to prove  $P(n+1)$

$$P(n) \Rightarrow f(n) = O(n)$$

$$P(n+1) = f(n) + n+1$$

$$= O(n) + O(n)$$

$$= O(n)$$

$$O(n+1)$$

$$\Rightarrow P(n+1)$$

□

~ What went wrong?

$$f(n) \text{ Fixes } n \text{ so } \lim_{n \rightarrow \infty} \frac{f(n)}{n} < \text{Finite}$$

0 only makes sense if n goes to infinity.

NEVER PUT 0 IN  
A PREDICATE OR  
USE IT IN AN  
INDUCTION  
HYPOTHESIS!

Asymptotic Notation

$O$  ( $\leq$ )  $\rightarrow$  growth of functions

$\Omega$  ( $\geq$ )

$\Theta$  ( $=$ )

$\circ$  ( $<$ )

$\sim$  ( $>$ )

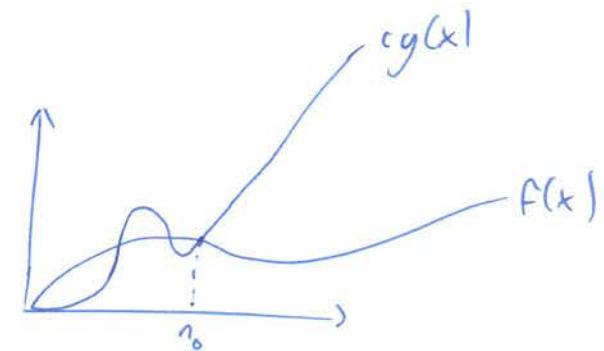
$\sim$  tilde

" $=$ "

Big-Oh

$$f(x) = O(g(x))$$

" $f(x)$  is upperbounded by  $g(x)$ "



So:

$$\lim_{n \rightarrow \infty} \left| \frac{f(x)}{g(x)} \right| < \infty \text{ (Finite)}$$

growing Faster than  $f(x)$

$$n^2 = O(n^2)$$

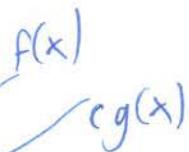
$$n = O(n^2)$$

Little Oh

$$\lim_{n \rightarrow \infty} \left| \frac{f(x)}{g(x)} \right| = 0$$

$$n^2 \neq o(n^2)$$

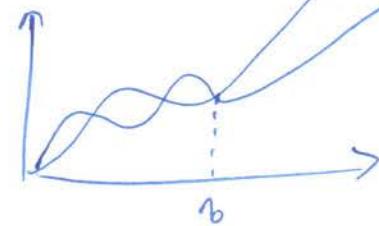
$$n = o(n^2)$$

Big Omega

$$f(x) = \Omega(g(x))$$

" $g(x)$  is a lower bound"

$$\lim_{n \rightarrow \infty} \left| \frac{f(x)}{g(x)} \right| > 0$$



$$n^2 = \Omega(n^2)$$

$$n^3 = \Omega(n^2)$$

### Little Omega

$$f(x) = \omega(g(x))$$

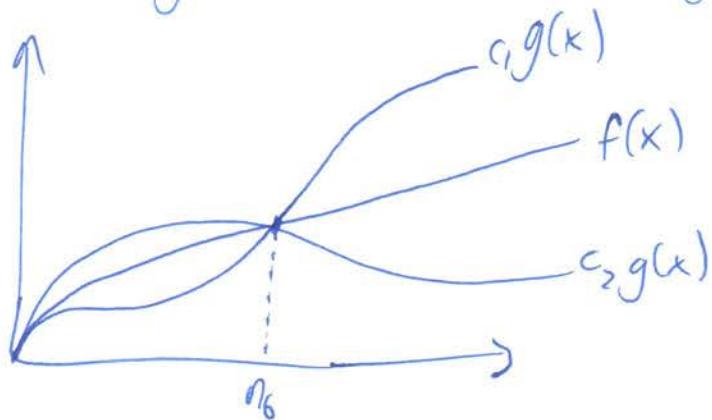
$$\lim_{n \rightarrow \infty} \left| \frac{f(x)}{g(x)} \right| = \infty$$

$$n^2 \neq \omega(n^2)$$

$$n^3 = \omega(n^2)$$

### Theta

$$f(x) = \Theta(g(x)) \Rightarrow f(x) = \Omega(g(x)) \wedge f(x) = O(g(x))$$



$$n^2 = \Theta(n^2)$$

$$\lim_{n \rightarrow \infty} \left| \frac{f(x)}{g(x)} \right| = L$$

$$0 < L < \infty$$

### Tilda

$$f(x) \sim g(x) \text{ if } \lim_{n \rightarrow \infty} \frac{f(x)}{g(x)} = 1$$

## Problems for Recitation 12

### 1 Asymptotic Notation

Which of these symbols

$\Theta$      $O$      $\Omega$      $o$      $\omega$

can go in these boxes? (List all that apply.)

$$2n + \log n = \boxed{O \ \Theta \ \Omega} (n)$$

$$\log n = \boxed{O, o} (n)$$

$$\sqrt{n} = \boxed{\Omega, \omega} (\log^{300} n)$$

$$n2^n = \boxed{\Omega} (n)$$

$$n^7 = \boxed{O} (1.01^n)$$

# Divide and Conquer Recurrences

10.1 - 10.4

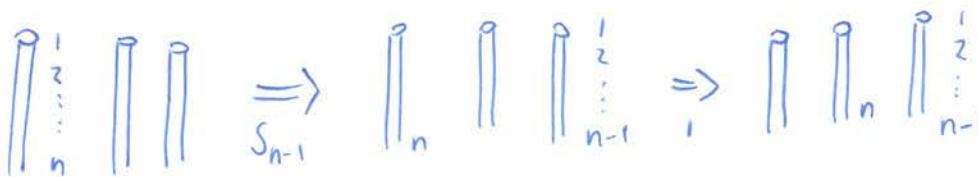
Game: Towers of Hanoi

Def:  $S_n = \min \# \text{ of moves to solve the } n\text{-disk problem.}$

$$S_1 = 1$$

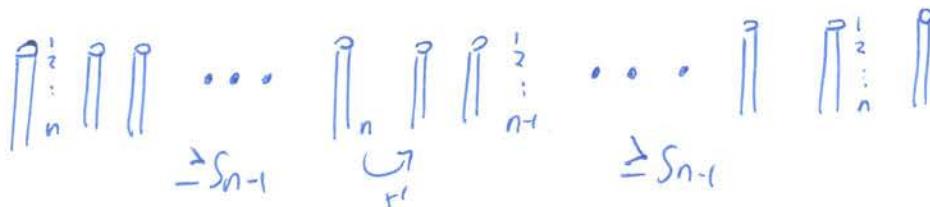
$$S_2 = 3$$

$$S_3 = 7$$



$$\Rightarrow S_{n-1} \quad \begin{array}{c} \text{ } \\ \text{ } \\ \text{ } \end{array} \quad \Rightarrow S_n \leq 2S_{n-1} + 1$$

## Lower Bound



lower bound = upper bound so equality ✓

Guess and check

$$\text{Thm: } S_n = 2^n - 1$$

Proof (By induction)

$$\text{I.H: } S_n = 2^n - 1$$

$$\text{Base Case } \underset{n=1}{\text{ }} S_1 = 2^1 - 1 = 1 \quad \checkmark$$

### Inductive Step

$$\begin{aligned}
 S_{n+1} &= 2S_n + 1 \\
 &= 2(2^n - 1) + 1 \\
 &= 2^{n+1} - 1 \quad \square
 \end{aligned}$$

Plug and Chug:

$$\begin{aligned}
 S_n &= 1 + 2S_{n-1} \\
 &= 1 + 2(1 + 2S_{n-2}) = 1 + 2 + 4S_{n-2} \\
 &= 1 + 2 + 4(1 + 2S_{n-3}) = 1 + 2 + 4 + 8S_{n-3} \\
 &\vdots \\
 S_n &= 1 + 2 + 4 \dots 2^{n-1} S_1 = 2^n - 1
 \end{aligned}$$

Merge Sort: Goal: Sort  $x_1 \dots x_n$ ,  $n = \text{power of 2}$

1. Sort  $x_1 \dots x_{\frac{n}{2}}$  and Sort  $x_{\frac{n}{2}+1} \dots x_n$
2. Merge

Def  $S(n) = \# \text{ of comparisons used by merge sort.}$

(worst case)

$$S(1) = 0$$

$$S(n) = \underbrace{2S\left(\frac{n}{2}\right)}_{\text{Sorting}} + \underbrace{n-1}_{\text{Merge}}$$

Plug and Chug

$$\begin{aligned} S(n) &= n-1 + 2S\left(\frac{n}{2}\right) \\ &= n-1 + 2\left(\frac{n}{2}-1 + 2S\left(\frac{n}{4}\right)\right) \\ &= n-1 + n-2 + 4S\left(\frac{n}{4}\right) \\ &= n-1 + n-2 + 4\left(\frac{n}{4}-1 + 2S\left(\frac{n}{8}\right)\right) \end{aligned}$$

$$\begin{aligned} S(n) &= n-1 + n-2 + n-4 \dots n-2^{\log n-1} + 2^{\log n} \\ &= \sum_{i=0}^{\log n-1} n - \sum_{i=0}^{\log n-1} 2^i = n \log n - n + 1 \end{aligned}$$

Nasty Recurrence

$$S(x) = \begin{cases} 2S\left(\frac{x}{2}\right) + \frac{8}{9}S\left(\frac{3}{4}x\right) + x^2 & \text{for } x \geq 1 \\ 0 & \text{for } x < 1 \end{cases}$$

def

A divide and conquer recurrence has the form:

$$S(x) = a_1 S(b_1 x + \varepsilon_1(x)) + a_2 S(b_2 x + \varepsilon_2(x)) \dots$$

$$\dots a_k S(b_k x + \varepsilon_k(x)) + g(x)$$

where  $a_i > 0$ ,  $0 < b_i < 1$ ,  $k = O(1)$ ,  $|\varepsilon_i(x)| \leq O\left(\frac{x}{\log^2 x}\right)$ ,  $g(x) \in \Theta(n)$

$$|g(x)| = O(x)$$

Ihm [Akra-Bazzi 96'] Set p s.t  $\sum a_i b_i^p = 1$

$$S(x) = \Theta\left(x^p + x^p \int_1^x \frac{g(u)}{u^{p+1}} du\right)$$

Ex  $S(x) = 2S\left(\frac{x}{2}\right) + x - 1$  (Merge Sort)

Good guess  $a_1 = 2$   $b_1 = \frac{1}{2}$   $k = 1 \Rightarrow p = 1$

$$\begin{aligned} S(x) &= \Theta\left(x + x \int_1^x \frac{u^{-1}}{u^2} du\right) \\ &= \Theta\left(x + x \int_1^x \frac{1}{u} - \frac{1}{u^2} du\right) = \Theta\left(x + x(\ln u - \frac{1}{u})\right)_1^x \\ &= \Theta(x \ln x) \end{aligned}$$

Ex  $S(x) = \begin{cases} 2S\left(\frac{x}{2}\right) + \frac{8}{9}S\left(\frac{3}{4}x\right) + x^2 & x \geq 1 \\ 0 & \text{else} \end{cases}$

$$2\left(\frac{1}{2}\right)^p + \frac{8}{9}\left(\frac{3}{4}\right)^p = 1 \Rightarrow p = 2$$

$$\begin{aligned} S(x) &= \Theta\left(x^2 + x^2 \int_1^x \frac{u^2}{u^3} du\right) \\ &= \Theta(x^2 + x^2 \ln u \Big|_1^x) = \Theta(x^2 \ln x) \end{aligned}$$

$$S(x) = 3S\left(\frac{x}{3}\right) + 4S\left(\frac{x}{4}\right) + x^2$$

$$3\left(\frac{1}{3}\right)^p + 4\left(\frac{1}{4}\right)^p = 1$$

## Problems for Recitation 13

### 1 Plug & Chug

Suppose you put \$1000 in a bank account. At the end of each month, you earn 1% interest and then you immediately withdraw \$5. Let  $M_n$  be the amount of money in the account after  $n$  months.

1. Express the amount in the account after  $n$  months with a recurrence and base cases.

2. Now we're going to find a closed form for  $M_n$  using the "plug and chug" method. Rewrite  $M_n$  in terms of smaller and smaller  $M_i$  by applying the recurrence equation over and over. Stop when you uncover a pattern. Simplify enough to keep the expressions manageable, but not so much that you destroy the pattern!

3. Based on the pattern you observed, what expression would you have after  $k$  rounds of plug-and-chug?

## Recitation 13

3. Define a divide-and-conquer recurrence for this algorithm. Let  $T(n)$  be the number of comparisons to sort a list of  $n$  items.

4. We could analyze the running time of this using plug-and-chug, but let's try Akra-Bazzi. First, what is  $p$ ?

5. Does the condition  $|g'(x)| = O(x^c)$  hold for some  $c \in N$ ?

6. Determine the theta bound on  $T(n)$  by integration.

7. Turns out that any equal partition of the list into a constant number of sublists  $c > 1$  will yield the same theta bound. Can you see why?

1/24/14

## 6.042 Recitation

Fernando Irujo and

Recurrences

Ex: Mergesort

$$T(n) = 2T\left(\frac{n}{2}\right) + g(n)$$

Merge

- Techniques : Plug and Chug  
 Guess and Check  
 Akra - Buzz  
 (Master Theorem)

Plug and Chug

$$\begin{aligned} T(n) &= T\left(\frac{n}{2}\right) + 1 \\ &= \left(T\left(\frac{n}{4}\right) + 1\right) + 1 \\ &= T\left(\frac{n}{16}\right) + 1 + 1 \\ &\quad \vdots \\ &= T(1) + 1 + \dots + 1 \\ &= \log n \end{aligned}$$

Akra - Buzz (see Appendix)

$$T_a(x) = 3T\left(\frac{x}{3}\right) + n$$

$$k=1$$

$$a=3$$

$$b=\frac{1}{3}$$

$$b/a = h_i(x) = 0$$

$$p=1$$

$$\sum_{i=1}^k a_i b_i^{p-1} = 3 \left(\frac{1}{3}\right)^0 = 1$$

$$g(n) = n$$

$$T_b(n) = 3T\left(\left[\frac{n}{3}\right]\right) + n$$

Floor

$$a=3$$

$$\frac{n}{3} + \left\lfloor \frac{n}{3} \right\rfloor - \frac{n}{3}$$

$$b/x$$

$$h_i(x)$$

$$b = \frac{1}{3} \quad p = 1$$

$$g(n) = n$$

$$T(n) = \Theta\left(x\left(1 + \int_1^x \frac{u}{\sqrt{u}} du\right)\right)$$

$$T(n) = \Theta(n \log n)$$

Def: A linear recurrence has the form

$$f(n) = a_1 f(n-1) + a_2 f(n-2) + \dots + a_d f(n-d) + g(n)$$

where  $a_1, a_2, \dots, a_d$  and  $d$  are constant and  $a_d \neq 0$

Ex:  $f(n) = 2f(n-1) + 1$   
 $d=1, a_1=2, g(n)=1$

↳ Sub problems get smaller by a constant.

Def: IF  $g(n)=0$ , the recurrence is homogeneous

IF  $g(n) \neq 0$       "      inhomogeneous

The order of recurrence is  $d$ .

### Grad Student Job Problem

- Total # of jobs =  $m$  (Fixed over time)
- Each professor generates 1 next prof each year
- Except 1<sup>st</sup> year profs produce 0
- No retirements

Question: When are all  $m$  jobs filled.

Boundary Condition: 1<sup>st</sup> professor is hired in year 1

### Solution

Let  $F(n) = \# \text{ of professors during year } n$

$$f(0) = 0 \quad f(3) = 1+1 = 2$$

$$f(1) = 1 \quad f(4) = 2+1 = 3$$

$$f(2) = 1 \quad f(5) = 3+2 = 5$$

$$\text{So } F(n) = \underbrace{F(n-1)}_{\text{old}} + \underbrace{F(n-2)}_{\text{new}}$$

$d=2$  Order: 2

$$a_1 = a_2 = 1$$

$$g(n) = 0$$

Fibonacci !!

$$f(n) = f(n-1) + f(n-2)$$

Solution: Try  $f(n) = \alpha^n$  for constant  $\alpha$

$$\alpha^n = \alpha^{n-1} + \alpha^{n-2}$$

$$\Rightarrow \alpha^2 = \alpha + 1$$

$$\alpha^2 - \alpha - 1 = 0$$

$$\Rightarrow \alpha = \frac{1 \pm \sqrt{1+4}}{2} = \frac{1 \pm \sqrt{5}}{2}$$

$$\alpha_1 = \frac{1+\sqrt{5}}{2} \quad \alpha_2 = \frac{1-\sqrt{5}}{2}$$

Golden Ratio!

Fact: IF  $f(n) = \alpha_1^n$  and  $f(n) = \alpha_2^n$  are solutions to a linear recurrence.  
(w/o boundary conditions), then  $f(n) = c_1\alpha_1^n + c_2\alpha_2^n$  is also a solution  
for any constants  $c_1$  and  $c_2$ . Linear combination

$$\Rightarrow f(n) = c_1 \left( \frac{1+\sqrt{5}}{2} \right)^n + c_2 \left( \frac{1-\sqrt{5}}{2} \right)^n$$

### Determining the constant factors

$$f(0) = 0 = c_1(1)^0 + c_2(1)^0 = c_1 + c_2 \Rightarrow c_2 = -c_1$$

$$f(1) = 1 \quad c_1 \left( \frac{1+\sqrt{5}}{2} \right)^1 + c_2 \left( \frac{1-\sqrt{5}}{2} \right)^1 = c_1 \left( \frac{2\sqrt{5}}{2} \right) \Rightarrow c_1 = \frac{1}{\sqrt{5}}, c_2 = -\frac{1}{\sqrt{5}}$$

### Solution

$$\boxed{f(n) = \frac{1}{\sqrt{5}} \left( \frac{1+\sqrt{5}}{2} \right)^n - \frac{1}{\sqrt{5}} \left( \frac{1-\sqrt{5}}{2} \right)^n}$$

So,  $f(n) = \frac{1}{\sqrt{5}} \left( \frac{1+\sqrt{5}}{2} \right)^n + \delta(n)$  approaches  $\rightarrow 0$  as  $n \rightarrow \infty$

$$\delta(n) \rightarrow 0$$

$$f(n) = \frac{1}{\sqrt{5}} \left( \frac{1+\sqrt{5}}{2} \right)^n + o(n)$$

So when are all the jobs filled?



$$f(n) \geq m$$

$$\frac{1}{\sqrt{5}} \left( \frac{1+\sqrt{5}}{2} \right)^n \geq m$$

$$n \log \left( \frac{1+\sqrt{5}}{2} \right) = \log \sqrt{5} m$$

$$n = \Theta(\log m)$$

$$\text{So if } m = 10,000$$

$$n = 20.8$$

### Solving General Linear Recurrence (homogeneous)

$$f(n) = \sum_{i=1}^d a_i f(n-i)$$

$$f(0) = b_0, f(1) = b_1, \dots, f(d-1) = b_{d-1}$$

$$\text{Try } f(n) = \alpha^n$$

$$\alpha^n = a_1 \alpha^{n-1} + a_2 \alpha^{n-2} + \dots + a_d \alpha^{n-d}$$

$$\alpha^d = a_1 \alpha^{d-1} + a_2 \alpha^{d-2} + \dots + a_d$$

$$\Rightarrow \alpha^d - a_1 \alpha^{d-1} - \dots - a_d = 0 \leftarrow \text{Characteristic Equation}$$

Simple Case: All d roots are different

Solution:  $f(n) = c_1 \alpha_1^n + c_2 \alpha_2^n + \dots + c_d \alpha_d^n$   
(w/o boundary conditions)

Solve for  $c_1, c_2, \dots, c_d$  From  $f(i) = b_i$  for  $i=0, 1, \dots, d-1$

Ex:  $f(0) = c_1\alpha_1^0 + c_2\alpha_2^0 + \dots + c_d\alpha_d^0 = b_0$  \*can always be solved

$$f(1) = c_1\alpha_1^1 + c_2\alpha_2^1 + \dots + c_d\alpha_d^1 = b_1$$

Tricky Case : Repeated roots

Thm: IF  $\alpha$  is a root of the characteristic Equation and  $\alpha$  is repeated  $r$  times, then

$$\alpha^n, n\alpha^n, n^2\alpha^n, \dots, n^{r-1}\alpha^n$$

are all solutions to the recurrence.

Then proceed as before... (take linear combinations)

Example

- plant reproduces one for one during 1<sup>st</sup> year of life and never again.
- plant lives forever

Solution

Let  $f(n) = \#$  of plants in year  $n$

$$f(0) = 0, f(1) = 1$$

$$f(n) = f(n-1) + \underbrace{(f(n-1) - f(n-2))}_{\text{old} + \text{new}}$$

$$f(n) = 2f(n-1) - f(n-2)$$

$$f(n) = 2f(n-1) + f(n-2) = 0$$

So,  $\alpha^2 - 2\alpha + 1 = 0$  characteristic polynomial

Solve,  $(\alpha-1)^2 = 0 \Rightarrow \alpha=1$ , double root

$$(\lambda - 1)^2 = 0 \quad \lambda = 1, \text{ double root}$$

$$\Rightarrow f(n) = c_1 \frac{(1)^n}{\lambda} + c_2 n \frac{(1)^n}{\lambda^2}$$

$$f(n) = c_1 + c_2 n$$

now we need to use boundary conditions to find  $c_1$  and  $c_2$

$$0 = f(0) = c_1 \Rightarrow c_1 = 0$$

$$1 = f(1) = c_1 + c_2 \Rightarrow c_2 = 1$$

so

$$f(n) = c_1 + c_2 n \Rightarrow \boxed{f(n) = n}$$

## Solving Inhomogeneous Recurrences

$$f(n) - a_1 f(n-1) - \dots - a_d f(n-d) = g(n)$$

- Step 1: Replace  $g(n)$  by 0 and solve the homogeneous recurrence w/o boundary conditions
- Step 2: Restore  $g(n)$  and find a particular solution w/o boundary condition
- Step 3: Add homogeneous and particular solution together and then use boundary condition to produce the general solution

Example:  $f(n) = 4f(n-1) + 3^n$ ,  $f(1) = 1$        $g(n) = 3^n$

Step 1 :  $f(n) - 4f(n-1) = 0$       homogeneous solution:  $f(n) = c_1 4^n$

$$\lambda - 4 = 0 \Rightarrow \lambda = 4$$

Step 2 : Find a particular solution

$$f(n) - 4f(n-1) = 3^n$$

Guess that  $f(n) = c 3^n$

$$\Rightarrow c 3^n - 4c 3^{n-1} = 3^n$$

$$c3^n - 4c3^{n-1} = 3^n$$

divide by  $3^{n-1}$

$$3c - 4c = 3 \\ \Rightarrow c = -3 \Rightarrow \text{Particular Solution is } f(n) = -3^{n+1}$$

Step 3: The general solution =

$$f(n) = c_1 4^n - 3^{n+1}$$

Use boundary conditions  $f(1) = 1$  to find  $c_1$

$$f(1) = 1 = c_1 4^1 - 9$$

$$10 = 4c_1 \Rightarrow c_1 = 5/2$$

Final Solution:

$$\boxed{f(n) = \frac{5}{2} 4^n - 3^{n+1}}$$

Check:

$$f(2) = 4f(1) + 3^2 = 13$$

$$f(2) = \frac{5}{2} 4^2 - 3^3 = 40 - 27 = 13 \quad \checkmark$$

### Guessing A Particular Solution

• If  $g(n)$  is an exponential, guess an exponential of same type

$$\text{Ex: If } g(n) = 2^n + 3^n \text{ guess } f(n) = a2^n + b3^n$$

• If  $g(n)$  is a polynomial, guess a polynomial of the same degree

$$\text{Ex: If } g(n) = n^2 \text{ guess } f(n) = an^2 + bn + c$$

$$\text{Ex: If } g(n) = 2^n + n \text{ guess } f(n) = a2^n + bn + c$$

If guesses fail, multiply guess by  $n$

$$\text{Ex: if } g(n) = 2^n, f(n) = a2^n \text{ fails } \rightarrow \text{guess } f(n) = (an + b)2^n \\ \text{then } f(n) = (an^2 + bn + c)2^n, \dots$$

Ex

$$f(n) = 2f(n-1) + 2^n, \quad f(0) = 1 \quad g(n) = 2^n$$

Homogeneous solution:

$$\lambda - 2 = 0 \Rightarrow \lambda = 2$$

$$f(n) = c_1 2^n$$

Particular Solution

Guess  $f(n) = a 2^n$

$$a 2^n = 2a 2^{n-1} + 2^n$$

$$\text{So, } a = a + 1 \Rightarrow 0 = 1 \quad \text{FAIL!}$$

try  $f(n) = (an+b) 2^n \rightsquigarrow$

$$(an+b) 2^n = 2(a(n-1)+b) 2^{n-1} + 2^n$$

divide by  $2^n$

$$an+b = an - a + b + 1$$

$$\Rightarrow a=1 \quad b=0$$

So particular Solution  $f(n) = n 2^n$

General Solution

$$f(n) = c_1 2^n + n 2^n$$

Plug in boundary conditions

$$1 = f(0) = c_1 \Rightarrow c_1 = 1$$

So,

$$f(n) = 2^n + n 2^n = \boxed{(n+1) 2^n}$$

Check:

$$f(1) = 2f(0) + 2^1 = 4$$

$$f(1) = (1+1) 2^1 = 4 \checkmark$$

Linear Recurrences

$$T(n) = \sum_{i=1}^d a_i T(n-i) \quad (\text{homogenous})$$

What is a solution to a homogenous Linear Recurrence?

$$f(n) = \sum_{i=1}^d a_i f(n-i)$$

If  $f(n)$  and  $g(n)$  are solutions  $\Rightarrow h(n) = sf(n) + tg(n)$   
is also a solution

Goal: Show that  $h(n) = \sum_{i=1}^d a_i h(n-i)$

So, Proof

$$h(n) = sf(n) + tg(n)$$

$$f(n) = \sum_{i=1}^d a_i f(n-i), \quad g(n) = \sum_{i=1}^d a_i g(n-i) \quad \text{Since } f(a), g(a) \text{ are solutions}$$

$$h(n) = s \sum_{i=1}^d a_i f(n-i) + t \sum_{i=1}^d a_i g(n-i) \quad \text{Should be true...}$$

$$= \sum_{i=1}^d [s a_i f(n-i) + t a_i g(n-i)] = \sum_{i=1}^d a_i [s f(n-i) + t g(n-i)]$$

$\underbrace{s f(n-i) + t g(n-i)}_{h(n-i)}$

$$= \sum_{i=1}^d a_i h(n-i)$$

□

# Solving Homogeneous Linear Recurrences

\* See handout

If our guess  $f(n) = x^n$

$$x^n = \sum_{i=1}^d a_i x^{n-i}$$

$$x^n = a_1 x^{n-1} + \dots +$$

Each root is a solution (By theorem above)

$\Rightarrow$  Every linear combination of roots is a solution

$$\sum_{i=1}^d r_i^n$$

$$T(n) = \sum_{i=1}^d a_i T(n-i) + g(n)$$

+

homog  
 $f_h(n)$

$\Rightarrow$

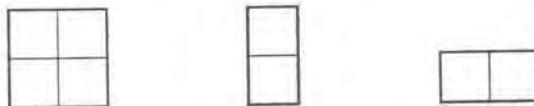
$$f_g(n) = f_h(n) + f_p(n)$$

$f_p(n)$

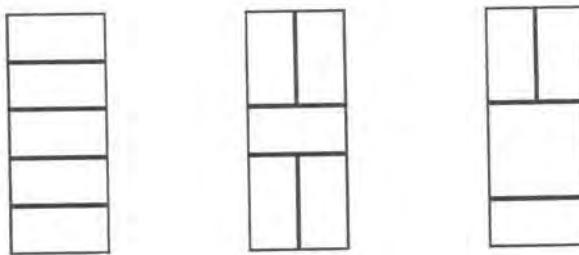
## Problems for Recitation 14

### Mini-Tetris

A *winning configuration* in the game of Mini-Tetris is a complete tiling of a  $2 \times n$  board using only the three shapes shown below:



For example, here are several possible winning configurations on a  $2 \times 5$  board:



1. Let  $T_n$  denote the number of different winning configurations on a  $2 \times n$  board. Determine the values of  $T_1$ ,  $T_2$ , and  $T_3$ .
2. Find a recurrence equation that expresses  $T_n$  in terms of  $T_{n-1}$  and  $T_{n-2}$ .
3. Find a closed-form expression for the number of winning configurations on a  $2 \times n$  Mini-Tetris board

## Short Guide to Solving Linear Recurrences

A *linear recurrence* is an equation

$$\underbrace{f(n) = a_1 f(n-1) + a_2 f(n-2) + \dots + a_d f(n-d)}_{\text{homogeneous part}} \quad + \underbrace{g(n)}_{\text{inhomogeneous part}}$$

together with boundary conditions such as  $f(0) = b_0$ ,  $f(1) = b_1$ , etc.

1. Find the roots of the *characteristic equation*:

$$x^n = a_1 x^{n-1} + a_2 x^{n-2} + \dots + a_k$$

2. Write down the *homogeneous solution*. Each root generates one term and the homogeneous solution is the sum of these terms. A nonrepeated root  $r$  generates the term  $c_r r^n$ , where  $c_r$  is a constant to be determined later. A root  $r$  with multiplicity  $k$  generates the terms:

$$c_{r_1} r^{n_1}, \quad c_{r_2} n r^{n_2}, \quad c_{r_3} n^2 r^{n_3}, \quad \dots, \quad c_{r_k} n^{k-1} r^{n_k}$$

where  $c_{r_1}, \dots, c_{r_k}$  are constants to be determined later.

3. Find a *particular solution*. This is a solution to the full recurrence that need not be consistent with the boundary conditions. Use guess and verify. If  $g(n)$  is a polynomial, try a polynomial of the same degree, then a polynomial of degree one higher, then two higher, etc. For example, if  $g(n) = n$ , then try  $f(n) = bn+c$  and then  $f(n) = an^2+bn+c$ . If  $g(n)$  is an exponential, such as  $3^n$ , then first guess that  $f(n) = c3^n$ . Failing that, try  $f(n) = bn3^n + c3^n$  and then  $an^23^n + bn3^n + c3^n$ , etc.

4. Form the *general solution*, which is the sum of the homogeneous solution and the particular solution. Here is a typical general solution:

$$f(n) = \underbrace{c2^n + d(-1)^n}_{\text{homogeneous solution}} + \underbrace{3n+1}_{\text{particular solution}}$$

5. Substitute the boundary conditions into the general solution. Each boundary condition gives a linear equation in the unknown constants. For example, substituting  $f(1) = 2$  into the general solution above gives:

$$\begin{aligned} 2 &= c \cdot 2^1 + d \cdot (-1)^1 + 3 \cdot 1 + 1 \\ \Rightarrow -2 &= 2c - d \end{aligned}$$

Determine the values of these constants by solving the resulting system of linear equations.

# Counting Methods I

7.2, 11.1-11.4, 11.10

"Counting one thing by counting another"

Ex1: There are Five types of doughnuts; how many ways are there to select a dozen.  
glazed, creme, cider, jelly, powder

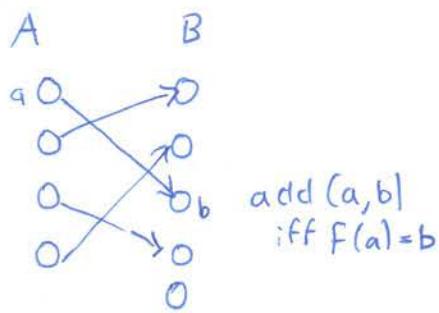
Def: The size (or cardinality) denoted by  $|A|$  is the number of elements of  $A$ .

Def: Given  $f: A \rightarrow B$  ( $f$  maps  $A$  to  $B$ )

• Injective if every element of B is mapped to at most once

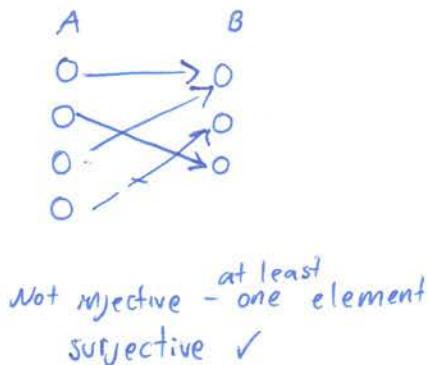
• Surjective      ↗      ↗      ↗  
least

Bijection if Both Injective and surjective



Not surjective - not all elements of B have incoming edge.

Injective ✓



Not injective - at least one element  
surjective ✓

Theorem Let  $f: A \rightarrow B$

- (a) if  $f$  is injective :  $|A| \leq |B|$
  - (b) if  $f$  is surjective :  $|A| \geq |B|$
  - (c) if  $f$  is bijection :  $|A| = |B|$

Back to Ex 1 ...

$A \triangleq$  all ways to select a dozen doughnuts

$B \triangleq$  all 16-bit strings with exactly 4 ones

$$a = (2, 0, 6, 2, 2) \mapsto 00110000001001000$$

#glazed

$$a = (a_1, a_2, a_3, a_4, a_5) \mapsto \underbrace{0 \dots 0}_1 \underbrace{1 \dots 1}_1 \dots \underbrace{0 \dots 0}_1 \dots \underbrace{1 \dots 1}_1 \dots \underbrace{0 \dots 0}_1$$

12 zeroes  
4 ones

$$\text{So, } |A| = |B|$$

and  $B$  should be easier to count than  $A$

YAY!

Recall  $A_1, A_2, A_3, \dots, A_n$  (sets)

$A_1 \times A_2 \times A_3 \times \dots \times A_n$  is the set of seq

$(a_1, a_2, \dots, a_n)$  where  $a_i \in A_i$

Product Rule:  $|A_1 \times A_2 \times \dots \times A_n| = |A_1| \cdot |A_2| \cdot \dots \cdot |A_n|$

Ex 2

Let  $|x|=n$ . Then,  $A \triangleq 2^x$  (powerset)  $\triangleq \{Y \mid Y \subseteq x\}$  ( $Y$  can be empty)

$x = \{1, 2, 3\}$   $2^x = \emptyset, \{1\}, \{2\}, \dots, \{1, 2, 3\}$  all valid subsets of  $x$

$B \triangleq$  length  $n$  binary strings

$$f(Y) = \underline{\quad} - \underline{\quad} - \underline{\quad} -$$

$\begin{matrix} x_i \\ \uparrow \text{is } x_i \text{ in } Y? \\ \text{True} = 1 \end{matrix}$

Ex:

$$X = \{1, 2, 3\} \quad Y = \{2\}$$

$$f(Y) = \emptyset \text{ or } \{0\}$$

Because there is a bijection between A and B.  
 $|A| = |B| = 2^n$   $\hookrightarrow$  different ways to name to make it easier to count  
 share no elements in common

Sum Rule: If  $A_1, \dots, A_n$  are disjoint (ie  $\forall i \neq j A_i \cap A_j = \emptyset$ )  
 $\hookrightarrow$  share no elements in common

$$|A_1 \cup A_2 \cup \dots \cup A_n| = |A_1| + |A_2| + \dots + |A_n|$$

Ex 3: Passwords must be 6-8 chars (upper/lowercase or digits)  
 but must start with a letter.

$\Sigma$  = all valid passwords

Break up  $\Sigma$  into disjoint sets that are easier to count

$$\Sigma = \{a, b, \dots, z, A, B, \dots, Z\}$$

$$\Phi = \{a, b, \dots, z, A, B, \dots, Z, 0, 1, \dots, 9\}$$

$$\Sigma = \underbrace{\Sigma \times \Phi \times \Phi \times \Phi \times \Phi \times \Phi}_5 \cup \underbrace{\Sigma \times \Phi \times \Phi \times \Phi \times \Phi}_6 \cup \underbrace{\Sigma \times \Phi \times \Phi \times \Phi \times \Phi}_7$$

Disjoint sets

$$|\Sigma| = |\Sigma \times \Phi^5| + |\Sigma \times \Phi^6| + |\Sigma \times \Phi^7| \quad (\text{By sum rule})$$

$$= |\Sigma| (|\Phi|^5 + |\Phi|^6 + |\Phi|^7) \quad (\text{By product rule})$$

## Generalized Product Rule

Let  $S$  be a set of length  $k$  sequences

- $n_1$  choices for First entry
- $n_2$  choices for second entry, for every choice of First entry
- $n_3$  " third " " First and second "

$$|S| = n_1 \cdot n_2 \cdot n_3 \dots n_k$$

### Ex 4

Def: A defective dollar bill has a repeated digit in the serial.

What fraction are non defective

$$\frac{10}{n_1} \quad \frac{9}{n_2} \quad \frac{8}{n_3} \quad - \quad - \quad -$$

$$\text{So, } |S| = 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3$$

$$B \triangleq \text{all bills, } 10^8 \quad \frac{|S|}{|B|} = .018$$

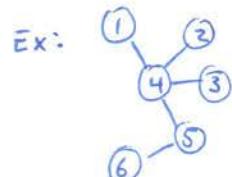
So only 1.8% of bills are non-defective

### Ex 5 :

$A \triangleq$  all labelled trees on  $n$ -vertices

$B \triangleq$  all length  $n-2$  sequences whose values are from  $\{1, 2, \dots, n\}$

$$|B| = n^{n-2}$$



## Prüfer Code

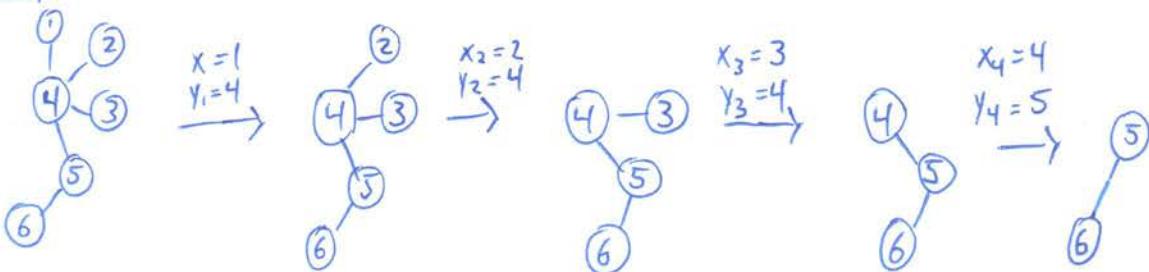
For  $i = 1 \dots n-2$

Set  $x_i$  smallest label of a leaf, let  $y_i$  be its neighbor

Delete  $x_i$

Return  $y_1, y_2, y_3, \dots, y_{n-2}$

### Example



$$x_i = 12345$$

$$y_i = 44456$$

Can reconstruct tree from this.

## Pigeonhole Principle (PHP)

IF  $|A| > |B|$   
    ↑ pigeons      ↑ holes

collision!

$f: A \rightarrow B$ , then there are  $a \neq a'$   $f(a) = f(a')$   
    ↳ generalized

6 PHP: IF  $|A| > k|B|$ , then there is some  $b \in B$  with more than  $k$  elements repeated to it.

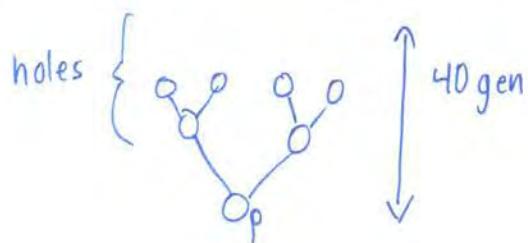
Proposition: Every person on Earth has a pair  $(x, y)$  of ancestors where both pairs of  $x$  have  $y$  as an ancestor

Facts: (1) no one has children after 100 yrs

(2) Human race is 4000 years old

(3) At most  $10^{12}$  humans lived during that period

For person p, its Family tree (For 4K years) has at least 40 generations



$$\begin{matrix} \circ & \circ & \circ \\ \circ & \circ & \circ \\ & \circ & \circ \end{matrix} \} 10^{12}$$
$$2^{40} > 10^{12}$$

There must be a collision

1/31/14

# 6.042 Recitation

Fernando Trujillo

Set = collection of distinct unordered elements  
 $\{A, B, C\} = \{B, C, A\}$

Cardinality =  $|S|$  = "size"

Sequence = collection of ordered elements (not necessarily distinct)  
 $a, a, b$

Function = mapping  $f: X \rightarrow Y$

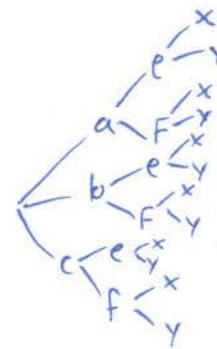
Permutation = sequence built from a set using each element exactly once.

Rule 1: Bijection

$$A \xrightarrow{\sim} B \quad |A|=|B|$$

Rule 4: Product

$$\{a, b, c\} \quad \{e, f\} \quad \{x, y\}$$



$$= |P_1||P_2|\dots|P_k|$$

Rule 6: Sum

$$A_1 \quad A_2 \quad A_3$$

$A_1, A_2, A_3$  disjoint sets

$$|A_1 \cup A_2 \cup A_3| = \sum_i |A_i|$$

Ex

How many subsets from a set  $X$ ?

$$X = \{x_1, \dots, x_n\}$$

$$\{x_1, x_2\}$$

$$\{x_1, x_3\}$$

$$\{x_1, x_2, x_3\}$$

.....

bit sequence  
of len n       $\xrightarrow{\text{Bijectron}}$  subset

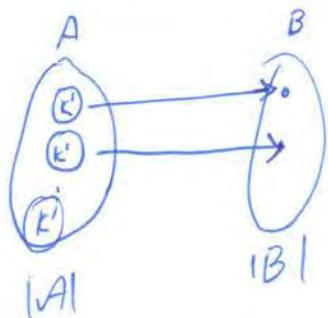
$$1, 1, 0 \dots 0 \rightarrow \{x_1, x_2\}$$

$$0, 1, 0 \dots 0 \rightarrow \{x_2\}$$

—————  
1

$$|\text{size}| = 2^n$$

Rule 3: Division



$$|A| = k^1 |B|$$

## Problems for Recitation 15

### 1 The Tao of BOOKKEEPER

In this problem, we seek enlightenment through contemplation of the word *BOOKKEEPER*.

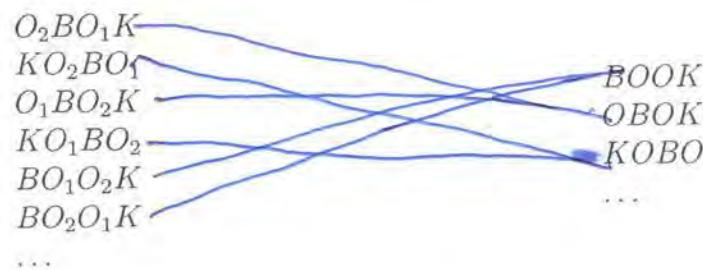
1. In how many ways can you arrange the letters in the word *POKE*?

$$4!$$

2. In how many ways can you arrange the letters in the word  $BO_1O_2K$ ? Observe that we have subscripted the O's to make them distinct symbols.

$$4!$$

3. Suppose we map arrangements of the letters in  $BO_1O_2K$  to arrangements of the letters in *BOOK* by erasing the subscripts. Indicate with arrows how the arrangements on the left are mapped to the arrangements on the right.



4. What kind of mapping is this, young grasshopper?

$$2 \rightarrow 1$$

5. In light of the Division Rule, how many arrangements are there of *BOOK*?

$$\frac{4!}{2}$$

15. (IMPORTANT) How many  $n$ -bit sequences contain  $k$  zeros and  $(n - k)$  ones?

$$\binom{n}{k} = \frac{n!}{(k!)(n-k)!}$$

This quantity is denoted  $\binom{n}{k}$  and read “ $n$  choose  $k$ ”. You will see it almost every day in 6.042 from now until the end of the term.

*Remember well what you have learned: subscripts on, subscripts off.*

*This is the Tao of Bookkeeper.*

### 3 More Counting Problems

Solve the following counting problems. Define an appropriate mapping (bijective or  $k$ -to-1) between a set whose size you know and the set in question.

1. (IMPORTANT) In how many ways can  $k$  elements be chosen from an  $n$ -element set  $\{x_1, x_2, \dots, x_n\}$ ?

$$\binom{n}{k}$$

2. How many different ways are there to select a dozen donuts if five varieties are available? (We discussed a bijection for this set in Recitation 15. Now use that bijection to give a count.)

$$\binom{16}{4}$$

16 bit sequence w/ 4 ones

3. An independent living group is hosting eight pre-frosh, affectionately known as  $P_1, \dots, P_8$  by the permanent residents. Each pre-frosh is assigned a task: 2 must wash pots, 2 must clean the kitchen, 1 must clean the bathrooms, 1 must clean the common area, and 2 must serve dinner. In how many ways can  $P_1, \dots, P_8$  be put to productive use?

~~$$\frac{8!}{2!2!1!1!2!}$$~~

4. Suppose that two identical 52-card decks of are mixed together. In how many ways can the cards in this double-size deck be arranged?

$$\frac{104!}{(2!)^{52}}$$

6. How many  $n$ -phoneme words are there in Hawaiian? (You don't have to find a closed form for your expression.)

## Counting Methods II

counting  $\leftrightarrow$  algebra

Rule: Let  $t_1, t_2, \dots, t_n$  are distinct elements,

the # of distinct sequences with  $n_i$  occurrences of  $t_i$ ,

$$\frac{(n_1 + n_2 + n_3 + \dots + n_m)!}{n_1! n_2! \dots n_m!} = \binom{n_1 + n_2 + n_3 + \dots + n_m}{n_1, n_2, \dots, n_m} \text{ multinomial}$$

Ex: How many distinct ways are there to arrange the letters in BOOKKEEPER

$$\begin{aligned} B: n_1 &= 1 & O: n_2 &= 2 & K: n_3 &= 2 & E: n_4 &= 3 & P: n_5 &= 1 & R: n_6 &= 1 \\ &\Rightarrow \frac{(10!)}{1! 2! 2! 3! 1! 1!} \end{aligned}$$

$$(a+b)^2 = a^2 + 2ab + b^2$$

$$(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$$

$$(a+b)^4 = ?$$

Thm: Binomial:  $(a+b)^n = \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k$

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \binom{n}{k, n-k}$$

$$(a+b)^4 = (a+b)(a+b)(a+b)(a+b)$$

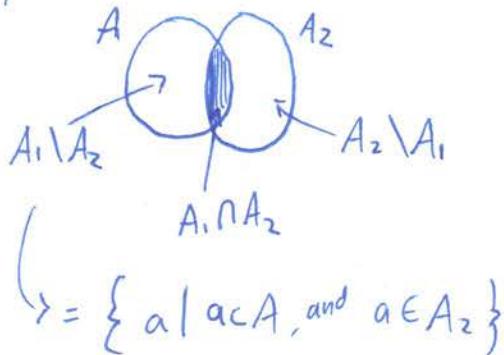
$$= aaaa + aaab + aaba + aabb + abaa + abab + abba + abbb + baaa \dots$$

Proof:  $(a+b)^n = \underbrace{(a+b) \dots (a+b)}_n$

The only terms are of the form  $a^{n-k} b^k$

The coefficient of  $a^{n-k} b^k$  is exactly the # of seq's  
of length n w/ n-k a's and k b's  $\square$

Beyond the Sum Rule



$$|A_1 \cup A_2| = |A_1 \setminus A_2| + |A_1 \cap A_2|$$

$$+ |A_2 \setminus A_1|$$

$$|A_1 \cup A_2| + |A_1 \cap A_2| = |A_1 \setminus A_2| + |A_1 \cap A_2|$$

$$+ |A_2 \setminus A_1| + |A_2 \cap A_1|$$

$$= |A_1| + |A_2|$$

$$\underline{|A_1 \cup A_2| = |A_1| + |A_2| - |A_1 \cap A_2|}$$

$$|A_1 \cup A_2 \cup A_3| = |A_1| + |A_2| + |A_3| - |A_1 \cap A_2| - |A_1 \cap A_3|$$

$$- |A_2 \cap A_3| + |A_1 \cap A_2 \cap A_3|$$

$\nwarrow$  compensates for  
underestimate

## Inclusion-Exclusion

$$|\bigcup_{i=1}^n A_i| = \sum_{i=1}^n |A_i| - \sum_{1 \leq i < j \leq n} |A_i \cap A_j| + \sum_{1 \leq i < j < k \leq n} |A_i \cap A_j \cap A_k|$$

Add and subtract . . .

$$\dots (-1)^{n+1} \left| \bigcap_{i=1}^n A_i \right|$$

Ex: A permutation on  $n$  numbers  $\pi: \{1, \dots, n\} \rightarrow \{1, \dots, n\}$   $\leftarrow$  # of students at graduation

$n=4$

$$\begin{array}{ccc} 1 & \xrightarrow{\pi} & 0_1 \\ 2 & \cancel{\xrightarrow{\pi}} & 0_2 \\ 3 & \xrightarrow{\pi} & 0_3 \\ 4 & \xrightarrow{\pi} & 0_4 \end{array} \quad \begin{array}{l} \pi(2)=1 \\ \pi(1)=2 \\ \pi(3)=3 \\ \pi(4)=4 \end{array}$$

How many  $\pi$ 's have at least 1  $i$  where  $\pi(i)=i$

Let  $A_i = \{\pi \mid \pi(i)=i\}$ : what is  $|\bigcup_{i=1}^n A_i|$

$$|A_1| \mapsto \frac{1}{n!} + \frac{1}{n-1!} - \frac{1}{n-2!} + \dots - \frac{1}{1!} = (n-1)!$$

$$|A_2| = (n-1)! \quad \vdots$$

$$|A_1 \cap A_2| = (n-2)! \mapsto \frac{1}{n!} - \frac{1}{n-1!} + \dots - \frac{1}{2!} + \frac{1}{1!}$$

$$S \subseteq \{1, \dots, n\} \quad S = \{1, 2, 3, \dots, k\} \quad |S|=k$$

$$|\bigcap_{i=k}^n A_i| = (n-k)!$$

$$|A_1 \cup A_2 \dots \cup A_n| = \sum_{i=1}^n |A_i| - \sum_{\substack{1 \leq i < j \leq n \\ (n-1)!}} |A_i \cap A_j|$$

$$= \frac{n(n-1)! - \binom{n}{2}(n-2)! + \binom{n}{3}(n-3)!}{n!}$$

$$1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} \dots := e^{-1}$$

Recall:  $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} \dots$

$$\binom{n}{k} = \frac{n!}{(n-k)! k!}$$

Not always so easy to verify.

### Outline

- Define some set  $S$

- Show that  $|S|=m$  by counting one way

- Show that  $|S|=n$  by counting another way

- (conclude that  $m=n$ )

### Pascal's Identity

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$$

IF we chose Jay, how many ways are there to complete the team?  
 If we don't " "

candidates :  $\binom{n-1}{k-1}$   
 spot :  $\binom{n-1}{k}$

~ Recall:

Inclusion Exclusion Formula

$$|\bigcup_{i=1}^n A_i| = \sum_{i=1}^n |A_i| - \sum_{1 \leq i < j \leq n} |A_i \cap A_j| + \sum_{1 \leq i < j < k \leq n} |A_i \cap A_j \cap A_k|$$

Proof:

Consider  $x \in \bigcup_{i=1}^n A_i$

Left Hand side  $|\bigcup_{i=1}^n A_i| = \sum_{x \in \bigcup_{i=1}^n A_i} 1$  ; Right Hand side  $\sum_{S \subseteq \{1, \dots, n\}} (-1)^{|S|+1} \left| \bigcap_{i \in S} A_i \right|$

$$= \sum_{x \in \bigcup_{i=1}^n A_i} \sum_{\substack{S \subseteq T \\ S \neq \emptyset}} \binom{|T_x|}{k} (-1)^{k+1}$$

$$T_x = \{i \mid x \in A_i\}$$

Binomial Formula!

(somehow...)

$$\cancel{(1-1)^{|T_x|} - 1}$$

$$0 = (1-1)^n = \sum_{k=0}^n \binom{n}{k} (-1)^k \cancel{(1+1)^{n-k}}$$

## "Counting with algebra and calculus"

12.1-12.4, 12.6

Counting Methods

$$\begin{array}{ccc} \text{Sequences} & \longleftrightarrow & \text{Generating Functions} \\ (a_0, a_1, \dots) & & a_0 + a_1 x + a_2 x^2 + \dots \end{array}$$

"A generating function is a clothesline on which we hang up a sequence for display!"

$$\begin{aligned} (1, 1, 1, 1, \dots) &\longleftrightarrow 1 + x + x^2 + x^3 + \dots = \sum_{i=0}^{\infty} x^i = \frac{1}{1-x} \\ \sum_{i=0}^n x^i &= \frac{1-x^{n+1}}{1-x} \quad \text{for } x \neq 1 \quad \sum_{i=0}^{\infty} x^i = \frac{1}{1-x} \quad |x| < 1 \end{aligned}$$

Dictionary

$$(a_0, a_1, a_2, \dots) \longleftrightarrow A(x) = \sum_{i=0}^{\infty} a_i x^i$$

$$(b_0, b_1, b_2, \dots) \longleftrightarrow B(x) = \sum_{i=0}^{\infty} b_i x^i$$

$$\text{Scaling } (c a_0, c a_1, c a_2, \dots) \quad c A(x) = \sum_{i=0}^{\infty} c a_i x^i$$

$$\text{Addition } (a_0 + b_0, a_1 + b_1, a_2 + b_2, \dots) \quad A(x) + B(x) = \sum_{i=0}^{\infty} (a_i + b_i) x^i$$

$$\begin{array}{c} \frac{1}{2} \text{ Ex} \\ \frac{1}{2} + \frac{1}{2} \end{array} \quad (1, 1, 1, \dots) \longleftrightarrow \frac{1}{1-x} \quad \sum_{i=0}^{\infty} y^i = \frac{1}{1-y}$$

$$\frac{1}{2} - \frac{1}{2} \quad (-1, -1, -1, \dots) \longleftrightarrow \sum_{i=0}^{\infty} (-1)^i x^i = \sum_{i=0}^{\infty} (-x)^i = \frac{1}{1+x}$$

$$\frac{1}{2} + \frac{1}{2} \quad (1, 0, 1, 0, \dots) \longleftrightarrow \frac{1/2}{1-x} + \frac{1/2}{1+x} = \frac{\frac{1}{2}(1+x) + \frac{1}{2}(1-x)}{(1-x^2)} = \frac{1}{1-x^2}$$

Right Shift  $(0, 0, 0, \dots, 0, a_0, a_1, \dots) \longleftrightarrow \sum_{i=k}^{\infty} a_{i-k} x^i = x^k A(x)$

$$x \sum_{i=k}^{\infty} a_{i-k} x^{i-k}$$

Differentiation  $(a_1, 2a_2, 3a_3, \dots) \longleftrightarrow A'(x) = \sum_{i=1}^{\infty} a_i x^{i-1}$

Ex  $(0, 1, 4, 9, 16, \dots) \longleftrightarrow x + 4x^2 + 9x^3, \dots$

$$(1, 1, 1, 1) \longleftrightarrow \frac{1}{1-x}$$

Deriv  $(1, 2, 3, 4, \dots) \longleftrightarrow \frac{1}{(1-x)^2}$

$$(0, 1, 4, 9, 16, \dots)$$

Right Shift  $(0, 1, 2, \dots) \longleftrightarrow \frac{x}{(1-x)^2}$

Deriv  $(1, 4, 9, 16, \dots) \longleftrightarrow \frac{1+x}{x^3}$

Right Shift  $(0, 1, 4, 9, 16, \dots) \longleftrightarrow \frac{(1+x)}{(1-x)^3}$

$(c_0, c_1, c_2, \dots) \longleftrightarrow C(x) = A(x) B(x)$

$c_n := a_0 b_n + a_1 b_{n-1} + a_2 b_{n-2} + \dots + a_n b_0$   
(convolution)

	$b_0$	$b_1 x$	$b_2 x^2$
$a_0$	$a_0 b_0$	$a_0 b_1 x$	$a_0 b_2 x^2$
$a_1 x$	$a_1 b_0 x$	$a_1 b_1 x^2$	
$a_2 x^2$	$a_2 b_0 x^2$		

$$(A(x) = \left( \sum_{i=0}^{\infty} a_i x^i \right) \left( \sum_{i=0}^{\infty} b_i x^i \right)$$

$$(a_0 + a_1 x + \dots)(b_0 + b_1 x + \dots)$$

Ex: Fruit Salad

- # of apples must be even
- # of bananas must be a mult of 5
- At most 4 oranges
- At most 1 pear

How many ways for n fruit

Convolution Rule Let  $A(x)$  be generating function for selecting items from  $A$  according to some rules. Let  $B(x)$  be " "  $B$  " " some other rules.

Generating Function for selecting items from  $A \cup B$  is  $A(x)B(x)$

$$A(x) = 1 + 0 \cdot x + 1 \cdot x^2 + 0 \cdot x^3 + 1 \cdot x^4$$

# of ways of selecting zero apples

$$\text{Apples} = \frac{1+x^2+x^4}{1-x^2}$$

$$B(x) = 1 + x^5 + x^{10} = \frac{1}{1-x^5}$$

$$O(x) = 1 + x + x^2 + x^3 + x^4 = \frac{1-x^5}{1-x}$$

$$P(x) = 1 + x$$

$$F(x) = A(x) B(x) O(x) P(x) = \frac{1}{1-x^2} \cdot \frac{1}{1-x^5} \cdot \frac{1-x^5}{1-x} 1+x$$

↑  
Frut Salads

$$= \frac{1}{(1-x)^2}$$

$$C(x) = \frac{x(1+x)}{(1-x)^4} \longleftrightarrow \left( \sum_{i=0}^{\infty} i^2, \sum_{i=0}^{\infty} i^2, \sum_{i=0}^{\infty} i^2, \dots \right)$$

$$C(x) = S_0 + S_1 x + S_2 x^2$$

$$S_n = \sum_{i=0}^n i^2$$

$$C(0) = S_0$$

Taylor Series!

$$C'(x) = S_1 + 2S_2 x + 3S_3 x^2 \dots | C'(0) = S_1$$

$$C''(x) = 2S_2 + 6S_3 x \dots | C''(0) = 2S_2$$

Taylor Series Rule

$$S_n = \frac{C^{(n)}(0)}{n!} \quad \text{n}^{\text{th}} \text{ derivative } \circ$$

## Problems for Recitation 17

The (*ordinary*) *generating function* for a sequence  $\langle a_0, a_1, a_2, a_3, \dots \rangle$  is the power series:

$$a_0 + a_1x + a_2x^2 + a_3x^3 + \dots$$

Find closed-form generating functions for the following sequences. Do not concern yourself with issues of convergence.

- (a)  $\langle 2, 3, 5, 0, 0, 0, 0, \dots \rangle$
- (b)  $\langle 1, 1, 1, 1, 1, 1, 1, \dots \rangle$
- (c)  $\langle 1, 2, 4, 8, 16, 32, 64, \dots \rangle$
- (d)  $\langle 1, 0, 1, 0, 1, 0, 1, 0, \dots \rangle$
- (e)  $\langle 0, 0, 0, 1, 1, 1, 1, 1, \dots \rangle$
- (f)  $\langle 1, 3, 5, 7, 9, 11, \dots \rangle$

$$F(x) = \frac{1}{1-x} \leftrightarrow (1, 1, 1, \dots)$$

$$F(0) = 1$$

$$F'(x) = \frac{1}{(1-x)^2}, \quad F'(0) = 1 \Rightarrow F^{(n)}(0) = n!$$

$$F''(x) = \frac{2}{(1-x)^3}, \quad F''(0) = \frac{2}{n!} = 1$$

$$(1+x) = \frac{x(1+x)}{(1-x)^4} = \frac{x}{(1-x)^4} + \frac{x^2}{(1-x)^4}$$

$$\underbrace{\qquad\qquad\qquad}_{\substack{\text{$n$th coeff}}} = (n-1)^{\text{th}} \text{coeff} \left( \frac{1}{(1-x)^4} \right) + (n-2)^{\text{th}} \text{coeff} \left( \frac{1}{(1-x)^4} \right)$$

$$F(x) = \frac{1}{(1-x)^4} \Rightarrow \frac{F^{(n)}(x)}{n!} = \frac{(n+3)(n+2)(n+1)}{6}$$

$$S_n = \frac{(n+2)(n+1)n}{6} + \frac{(n+1)n(n-1)}{6}$$

Generating Function

Closed form expression related to infinite sum of polynomials.

$$G(x) = \sum_{n=0}^{\infty} g_n x^n \longleftrightarrow (g_0, \dots)$$

Scaling

$$cG(x) = \sum_{n=0}^{\infty} c g_n x^n \longleftrightarrow (cg_0, \dots)$$

Addition

$$g(x) + F(x) \longleftrightarrow (\dots g_n + F_n \dots)$$

Derivative

$$F(x) = \sum_{n=0}^{\infty} F_n x^n$$

$$F'(x) = \sum_{n=0}^{\infty} n F_n x^{n-1} \longleftrightarrow (F_1, 2F_2, 3F_3, \dots)$$

$x_0, x_1, x_2$

Right Shift

$$(0, \dots 0, \underbrace{1, 1, 1, 1}_k) \longleftrightarrow \sum_{n=k}^{\infty} f_n x^n = F_x x^k + F_{k+1} x^{k+1} + \dots$$

$$x^k (1 + \dots) = \frac{x^k}{1-x}$$

$$\frac{\text{Product}}{(x)} = G(x) F(x) = \sum_{n=0}^{\infty} g_n x^n \sum_{n=0}^{\infty} f_n x^n$$

$c_n = \underbrace{\sum_{j=0}^n}_{\text{convolution}} \overbrace{F_j g_{n-j}}$   
n<sup>th</sup> coefficient

Taylor Series  $\curvearrowleft$   $n^{\text{th}}$  derivative

$$f_n = \frac{\overbrace{F}^{(n)}(0)}{n!}$$

Generating Functions can be used for counting!

$$(1+x)^k \Rightarrow \left( \binom{k}{0}, \binom{k}{1}, \dots, \binom{k}{k}, 0, \dots \right)$$

$\binom{k}{n}$  = selecting subsets of length  $n$  from  $k$ .

$$\begin{array}{ll} \{a_1\} & 1+x \\ \{a_2\} & 1+x \end{array}$$

$$(1+x)(1+x) = 1+2x+x^2$$

\* The generating function for choosing elements from a union of disjoint sets is the product of the generating functions from choosing from each set.

Thm  $\binom{3n}{n} = \sum_{r=0}^n \binom{n}{r} \binom{2n}{n-r}$  Combinatorial Proof

$$S = \boxed{\begin{matrix} \text{red} & \text{black} \\ n & 2n \end{matrix}} \quad \text{select } n \text{ from } 2n+n$$

$$|S| = \binom{3n}{n}$$

$r$  cards  $n-r$  # blacks  $\leftarrow$  same way to count.

$$\binom{n}{r} \binom{2n}{n-r}$$

## Problems for Recitation 16

### 1 Combinatorial Proof

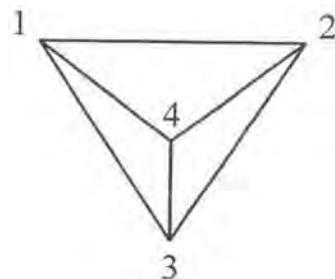
A *combinatorial proof* is an argument that establishes an algebraic fact by relying on counting principles. Many such proofs follow the same basic outline:

1. Define a set  $S$ .
2. Show that  $|S| = n$  by counting one way.
3. Show that  $|S| = m$  by counting another way.
4. Conclude that  $n = m$ .

### 2 Triangles

Let  $T = \{X_1, \dots, X_t\}$  be a set whose elements  $X_i$  are themselves sets such that each  $X_i$  has size 3 and is  $\subseteq \{1, 2, \dots, n\}$ . We call the elements of  $T$  “triangles”. Suppose that for all “edges”  $E \subseteq \{1, 2, \dots, n\}$  with  $|E| = 2$  there are exactly  $\lambda$  triangles  $X \in T$  with  $E \subseteq X$ .

For example, if we might have the triangles depicted in the following diagram, which has  $\lambda = 2$ ,  $n = 4$ , and  $t = 4$ :



In this example, each edge appears in exactly two of the following triangles:

$$\{1, 2, 3\}, \{1, 2, 4\}, \{1, 3, 4\}, \{2, 3, 4\}$$

4. Find the number of 5-card hands in which every suit appears at most twice.
5. There are 15 sidewalk squares in a row. Suppose that a ball is thrown down the row so that it bounces on 0, 1, 2, or 3 distinct sidewalk squares. How many different throws are possible? Two throws are considered to be equivalent if they bounce on the same squares in a different order.
6. In how many different ways can the numbers shown on a red die, a green die, and a blue die total up to 15? Assume that these are ordinary, 6-sided dice.
7. In how many ways can 20 indistinguishable pre-frosh be stored in four different crates if each crate must contain an *even* number of pre-frosh?
8. How many paths are there from point  $(0,0)$  to  $(50,50)$  if every step increments one coordinate and leaves the other unchanged and there are impassable boulders sitting at points  $(10,10)$  and  $(20,20)$ ?
9. In how many ways can the 180 students in 6.042 be divided into 36 groups of 5?

11/12/14

# 6.042 Recitation

Fernando Irujano

## Probability chapter 14

- Do not rely on intuition

Def: The sample space for an experiment is the set of all possible outcomes.

Def: an outcome (aka sample point) consists of all the information about the experiment after it's been performed. Including the values of all random choices

### Monty Hall

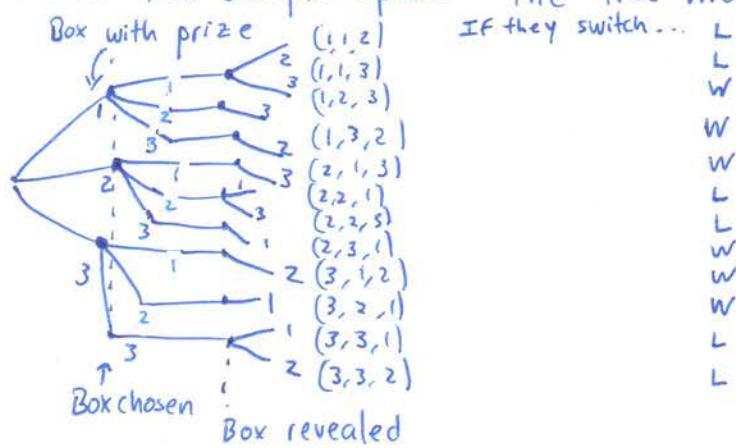
Def: an outcome of the monty Hall game (when the player switches) consists of:

1. Box with price
2. Box chosen by player
3. Box revealed

Ex: Sample point  $(2, 1, 3)$  is the outcome  
where:  
 2 has price  
 1 chosen by player  
 3 revealed by monty

$(1, 1, 1)$  is not a sample point  
 $(1, 2, 1)$  " " " "

Construct the sample space - The tree method



Def: A probability space consists of a sample space  $S$  and a probability function  $\Pr: S \rightarrow \mathbb{R}$  s.t

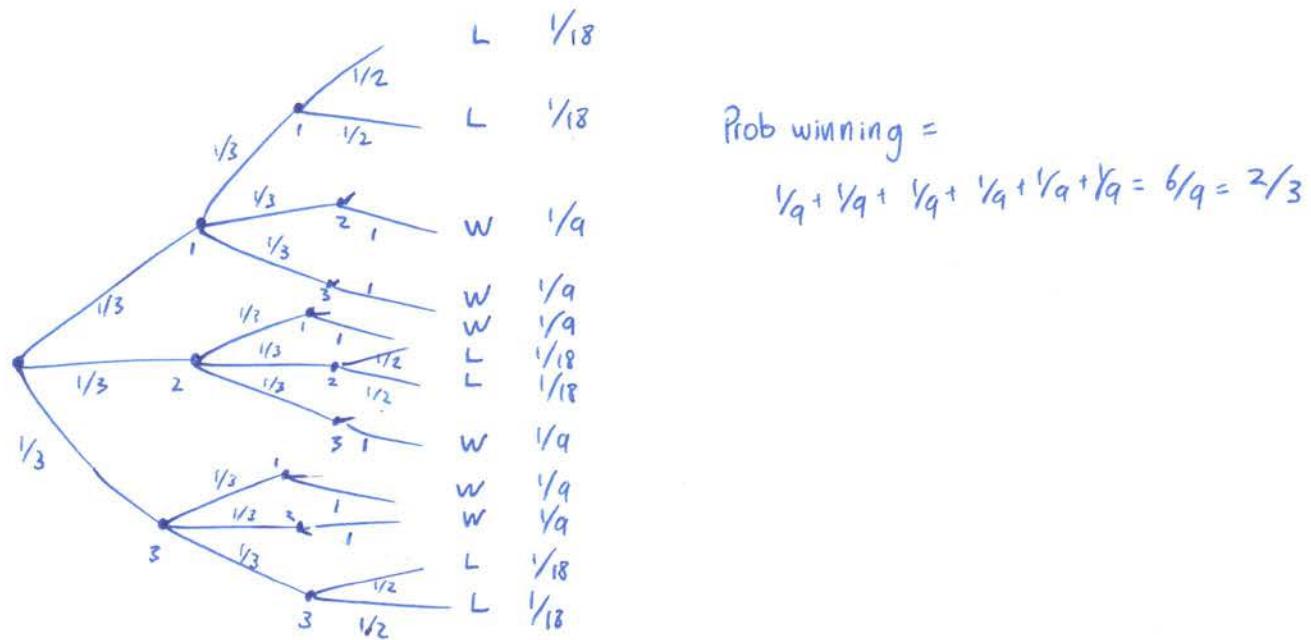
- 1)  $\forall w \in S, 0 \leq \Pr(w) \leq 1$

- 2)  $\sum_{w \in S} \Pr(w) = 1$

~ Interpretation :  $\forall w \in S, \Pr(w)$  = probability that  $w$  will be the outcome

### Assumptions

- 1) The prize is in each box with prob  $1/3$
- 2) The player picks each box with prob  $1/3$
- 3) IF Monty has a choice, he reveals each box with prob  $1/2$



Def: An event is a subset of the sample space

Ex.  $E_L$  = event that player loses

Def: The probability that an event  $E \subseteq S$  occurs is  $\sum_{w \in E} \Pr(w)$

$$\text{Ex: } \Pr(E_L) = 6/18 = 1/3$$

Def: A sample space  $S$  is uniform if every sample point has the same prob :  $\Pr(w) = 1/|S|$

# 6.042 HANDOUT



[marilynvossavant.com](http://marilynvossavant.com)

[»home](#)

[»ask a question](#)

[»discussions](#)

[»about marilyn](#)

[»idea box](#)

## Game Show Problem

(This material in this article was originally published in PARADE magazine in 1990 and 1991.)

Suppose you're on a game show, and you're given the choice of three doors. Behind one door is a car, behind the others, goats. You pick a door, say #1, and the host, who knows what's behind the doors, opens another door, say #3, which has a goat. He says to you, "Do you want to pick door #2?" Is it to your advantage to switch your choice of doors?

Craig F. Whitaker  
Columbia, Maryland

## Marilyn's Response:

Yes; you should switch. The first door has a  $1/3$  chance of winning, but the second door has a  $2/3$  chance. Here's a good way to visualize what happened. Suppose there are a million doors, and you pick door #1. Then the host, who knows what's behind the doors and will always avoid the one with the prize, opens them all except door #777,777. You'd switch to that door pretty fast, wouldn't you?

## Marilyn's Response:

Good heavens! With so much learned opposition, I'll bet this one is going to keep math classes all over the country busy on Monday.

My original answer is correct. But first, let me explain why your answer is wrong. The winning odds of  $1/3$  on the first choice can't go up to  $1/2$  just because the host opens a losing door. To illustrate this, let's say we play a shell game. You look away, and I put a pea under one of three shells. Then I ask you to put your finger on a shell. The odds that your choice contains a pea are  $1/3$ , agreed? Then I simply lift up an empty shell from the remaining other two. As I can (and will) do this regardless of what you've chosen, we've learned nothing to allow us to revise the odds on the shell under your finger.

The benefits of switching are readily proven by playing through the six games that exhaust all the possibilities. For the first three games, you choose #1 and "switch" each time, for the second three games, you choose #1 and "stay" each time, and the host always opens a loser. Here are the results.

	DOOR 1	DOOR 2	DOOR 3	RESULT
GAME 1	AUTO	GOAT	GOAT	Switch and you lose.
GAME 2	GOAT	AUTO	GOAT	Switch and you win.
GAME 3	GOAT	GOAT	AUTO	Switch and you win.
GAME 4	AUTO	GOAT	GOAT	Stay and you win.
GAME 5	GOAT	AUTO	GOAT	Stay and you lose.
GAME 6	GOAT	GOAT	AUTO	Stay and you lose.

When you switch, you win  $2/3$  of the time and lose  $1/3$ , but when you don't switch, you only win  $1/3$  of the time and lose  $2/3$ . You can try it yourself and see. Alternatively, you can actually play the game with another person acting as the host with three playing cards—two jokers for the goat and an ace for the prize. However, doing this a few hundred times to get statistically valid results can get a little tedious, so perhaps you can assign it as extra credit—or for punishment! (That'll get their goats!)

## Marilyn's Response:

Gasp! If this controversy continues, even the postman won't be able to fit into the mailroom. I'm receiving thousands of letters, nearly all insisting that I'm wrong, including the Deputy Director of the Center for Defense Information and a Research Mathematical Statistician from the National Institutes of Health! Of the letters from the general public, 92% are against my answer, and of the letters from universities, 65% are against my answer. Overall, nine out of ten readers completely disagree with my reply.

Now we're receiving far more mail, and even newspaper columnists are joining in the fray! The day after the second column appeared, lights started flashing here at the magazine. Telephone calls poured into the switchboard, fax machines churned out copy, and the mailroom began to sink under its own weight. Incredulous at the response, we read wild accusations of intellectual irresponsibility, and, as the days went by, we were even more incredulous to read embarrassed retractions from some of those same people!

So let's look at it again, remembering that the original answer defines certain conditions, the most significant of which is that the host always opens a losing door on purpose. (There's no way he can always open a losing door by chance!) Anything else is a different question.

The original answer is still correct, and the key to it lies in the question, "Should you switch?" Suppose we pause at that point, and a UFO settles down onto the stage. A little green woman emerges, and the host asks her to point to one of the two unopened doors. The chances that she'll randomly choose the one with the prize are  $1/2$ , all right. But that's because she lacks the advantage the original contestant had—the help of the host. (Try to forget any particular television show.)

When you first choose door #1 from three, there's a  $1/3$  chance that the prize is behind that one and a  $2/3$  chance that it's behind one of the others. But then the host steps in and gives you a clue. If the prize is behind #2, the host shows you #3, and if the prize is behind #3, the host shows you #2. So when you switch, you win if the prize is behind #2 or #3. You win either way! But if you don't switch, you win only if the prize is behind door #1.

And as this problem is of such intense interest, I'm willing to put my thinking to the test with a nationwide experiment. This is a call to math classes all across the country. Set up a probability trial exactly as outlined below and send me a chart of all the games along with a cover letter repeating just how you did it so we can make sure the methods are consistent.

One student plays the contestant, and another, the host. Label three paper cups #1, #2, and #3. While the contestant looks away, the host randomly hides a penny under a cup by throwing a die until a 1, 2, or 3 comes up. Next, the contestant randomly points to a cup by throwing a die the same way. Then the host purposely lifts up a losing cup from the two unchosen. Lastly, the contestant "stays" and lifts up his original cup to see if it covers the penny. Play "not switching" two hundred times and keep track of how often the contestant wins.

Then test the other strategy. Play the game the same way until the last instruction, at which point the contestant instead "switches" and lifts up the cup not chosen by anyone to see if it covers the penny. Play "switching" two hundred times, also.

## Problems for Recitation 19

### The Four-Step Method

This is a good approach to questions of the form, “What is the probability that ——?” Intuition *will* mislead you, but this formal approach gives the right answer every time.

1. Find the sample space. (Use a tree diagram.)
2. Define events of interest. (Mark leaves corresponding to these events.)
3. Determine outcome probabilities:
  - (a) Assign edge probabilities.
  - (b) Compute outcome probabilities. (Multiply along root-to-leaf paths.)
4. Compute event probabilities. (Sum the probabilities of all outcomes in the event.)

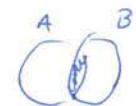
## Conditional Probability

$\Pr(A|B)$  = Prob of A given B

Ex: A = event player chooses box 1

B = prize is in box 1

$$\Pr(A|B) = \frac{1}{3}$$



Def: IF  $\Pr(B) \neq 0$ ,  $\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}$

↑  
and  
normalize

Product Rule:  $\Pr(A \cap B) = \Pr(B) \Pr(A|B) = \Pr(B|A) = \Pr(A) \Pr(B|A)$

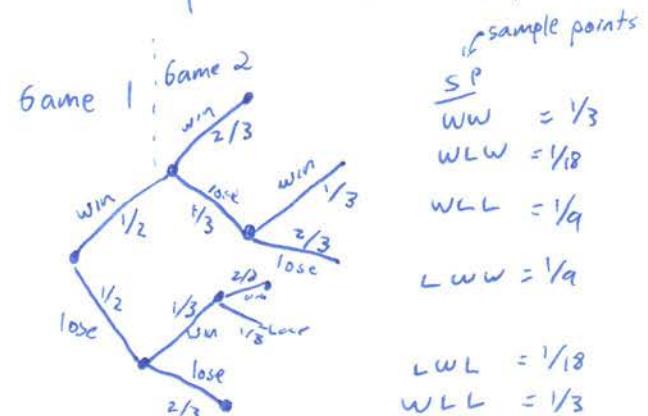
General Product Rule:  $\Pr(A_1 \cap A_2 \cap \dots \cap A_n) = \Pr(A_1) \Pr(A_2|A_1) \Pr(A_3|A_1 \cap A_2) \dots \Pr(A_n|A_1 \cap \dots \cap A_{n-1})$

Ex: In a best 2 out of 3 series, suppose prob of winning 1st game =  $1/2$ , the prob of winning a game after a win is  $2/3$ . So prob of winning after loss =  $1/3$ .

IF you win the first game, what is the probability that you win series.

(compute  $\Pr(A|B)$ )

↑  
win series      ↑  
win First game



$$\Pr(WW) = \Pr(W^{1st}) \Pr(W^{2nd} | W^{1st}) = \frac{1}{3}$$

$$\Pr(WLW) = \Pr(W^{1st}) \Pr(L^{2nd} | W^{1st}) \Pr(W^{3rd} | W^{1st} \wedge L^{2nd}) = \frac{1}{18}$$

$\Pr(W^3 | L^2)$

<u>SP</u>	<u>A</u>	<u>B</u>	<u><math>A \wedge B</math></u>
$\frac{1}{3}$ : WW	✓	✓	✓
$\frac{1}{18}$ : WLW	✓	✓	✓

$\frac{1}{9}$ : WLL

$\frac{1}{9}$ : LWW

LWL

LL

$$\Pr(A|B) = \frac{\Pr(A \wedge B)}{\Pr(B)}$$

$$= \frac{\frac{1}{3} + \frac{1}{18}}{\frac{1}{3} + \frac{1}{18} + \frac{1}{9}} = \frac{\frac{7}{18}}{\frac{9}{18}} = \frac{7}{9}$$

### a' postieri probability

$\Pr(B|A)$  where B precedes A in time  
 Pr winning 1st game given you won the series.

$$\hookrightarrow \frac{\Pr(B \wedge A)}{\Pr(A)} = \frac{\left(\frac{1}{3} + \frac{1}{18}\right)}{\left(\frac{1}{3} + \frac{1}{18} + \frac{1}{9}\right)} = \frac{\frac{7}{18}}{\frac{9}{18}} = \frac{7}{9}$$

$$\Pr(A|B) = \Pr(B|A) ; \text{FF}$$

$$\frac{\Pr(A \wedge B)}{\Pr(A)}$$

$$\frac{\Pr(A \wedge B)}{\Pr(A)}$$

$$\Pr(A \wedge B) = 0 \quad \text{or} \\ \Pr(A) = \Pr(B)$$

$\wedge$  and is commutative

| given is NOT

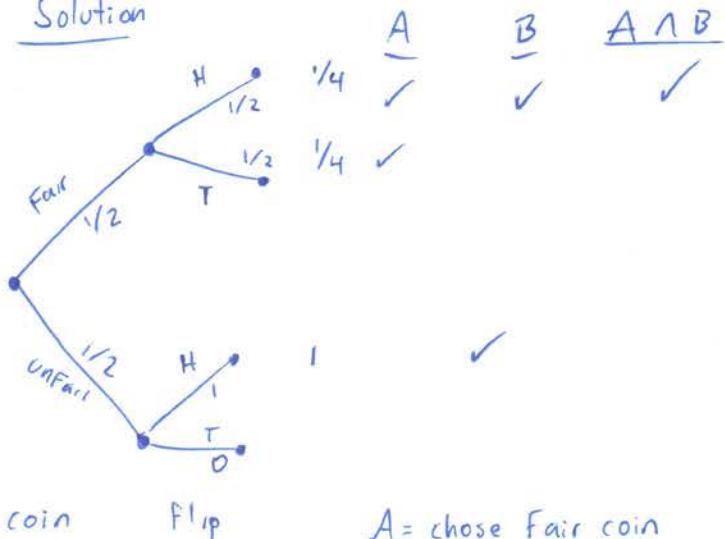
Ex:

Fair coin :  $\Pr(H) = \Pr(T) = 1/2$

Unfair coin :  $\Pr(H) = 1$ ,  $\Pr(T) = 0$

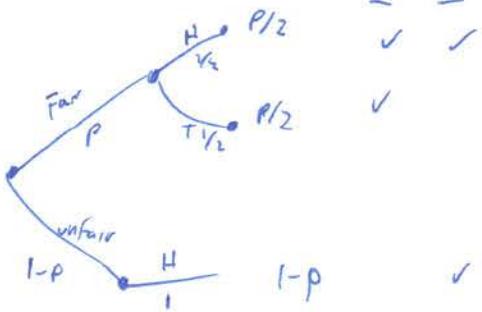
Pick coin randomly, flip, see result.  $\rightarrow$  Heads. What's the  $\Pr$  that chosen coin was fair.

Solution



coin      flip       $A = \text{choose Fair coin}$   
                 $B = \text{result is heads}$

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{1}{3}$$



$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{p/2}{1-p/2} = \frac{p}{2-p}$$

\* Ex: Medical Testing

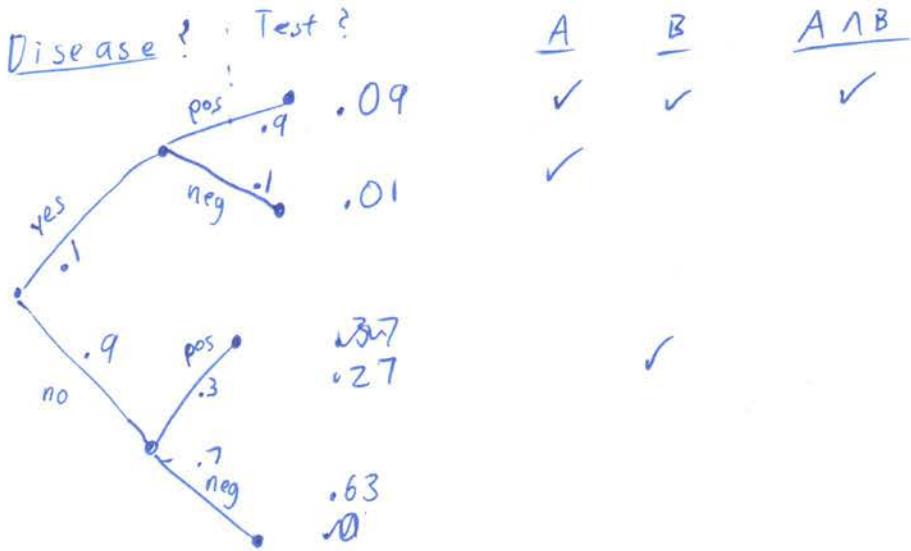
10% of people have decease X

IF you have the decease, 10% chance test says you dont. False negative  
" don't " , 30%      " do . False positive.

Problem: Pick random person  $\rightarrow$  Tests positive  $\rightarrow$   $\Pr$  Person has decease

$A$  = event person has disease

$B$  = event person tests positive



$$\Pr(A|B) = \frac{\Pr(A \wedge B)}{\Pr(B)} = \frac{.09}{(.09+.27)} = \frac{1}{4} \quad !$$

$$\Pr(\text{test correct}) = .09 + .63 = .72$$

### Ex Carnival Dice

player picks a #  $N \in [1, 6]$  and then roll 3 Fair dice

player wins if  $N$  matches  $\geq 1$  die.

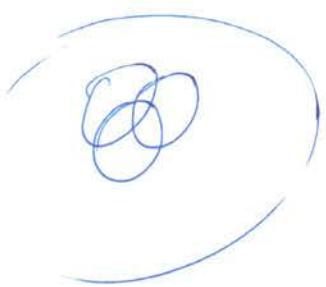
$$\text{Claim? } \Pr(\text{win}) = \frac{1}{2}$$

Proof? Let  $A_i$  = event  $i$ th die is  $N$

$$\begin{aligned}
 \Pr(\text{win}) &= \Pr(A_1 \cup A_2 \cup A_3) \\
 &= \Pr(A_1) + \Pr(A_2) + \Pr(A_3) \\
 &= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} \\
 &= \frac{1}{2} \quad \square
 \end{aligned}$$

This step only true if  
A<sub>i</sub>'s are disjoint!

Fact :  $\Pr(A_1 \cup A_2 \cup A_3) = \Pr(A_1) + \Pr(A_2) + \Pr(A_3)$



$$\Pr(A_1 \cup A_2 \cup A_3) = \Pr(A_1) + \Pr(A_2) + \Pr(A_3) - \Pr(A_1 \cap A_2) - \Pr(A_1 \cap A_3) - \Pr(A_2 \cap A_3) + \Pr(A_1 \cap A_2 \cap A_3)$$

to account for double counting!

$$\Pr(\text{Wm}) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} - \frac{1}{36} - \frac{1}{36} - \frac{1}{36} + \frac{1}{216} = .421$$

$$\Pr(A \cup B | C) = \Pr(A | C) + \Pr(B | C) - \Pr(A \cap B | C)$$

Claim? If  $C \neq D$  are disjoint then

$$\Pr(A | C \cup D) = \Pr(A | C) + \Pr(A | D)$$



$| \neq | + |$

NO

NEVER

### Sex Discrimination

#### Events

$A$  = applicant admitted

$F_{cs}$  = "Female and CS"

$F_{EE}$  = "Female and EE"

$M_{cs}$  = "Male and CS"

$M_{EE}$  = "Male and EE"

#### Lawyer's claim

$$\begin{cases} \Pr(A | M_{cs}) < \Pr(A | F_{cs}) \\ \Pr(A | M_{EE}) < \Pr(A | F_{EE}) \end{cases} \quad \text{"No discrimination"}$$

$$\begin{cases} \Pr(A | M_{cs} \cup M_{EE}) > \Pr(A | F_{cs} \cup F_{EE}) \end{cases} \quad \text{"Sexual discrimination"}$$

professor claims

Is someone lying?

No. This can happen!

## Airline On Time Rates

### Alaska Air

		on time
LA	500 / 560	89%
Pho	220 / 230	95%
SD	210 / 230	92%
JF	500 / 600	83%
Seattle	1900 / 2200	86%
Total	3330 / 3820	87%

### Am West

700 / 800	87%
4900 / 5300	92%
400 / 450	89%
320 / 450	71%
200 / 260	77%

Conclusion: Am West is better! (more on time)

but Alaska Air is better in each city....

gets weighting on different city

Seattle vs pho

## Problems for Recitation 20

### 1 Nerditosis

There is a rare and deadly disease called *Nerditosis* which afflicts about 1 person in 1000. One symptom is a compulsion to refer to everything— fields of study, classes, buildings, etc.— using numbers. It's horrible. As victims enter their final, downward spiral, they're awarded a degree from MIT. Two doctors claim that they can diagnose Nerditosis.

1. Doctor  $X$  received his degree from Harvard Medical School. He practices at Massachusetts General Hospital and has access to the latest scanners, lab tests, and research. Suppose you ask Doctor  $X$  whether you have the disease.
  - If you have Nerditosis, he says “yes” with probability 0.99.
  - If you don’t have it, he says “no” with probability 0.97.

Let  $D$  be the event that you have the disease, and let  $E$  be the event that the diagnosis is erroneous. Use the Total Probability Law to compute  $\Pr\{E\}$ , the probability that Doctor  $X$  makes a mistake.

2. “Doctor”  $Y$  received his genuine degree from a fully-accredited university for \$49.95 via a special internet offer. He knows that Nerditosis strikes 1 person in 1000, but is a little shaky on how to interpret this. So if you ask him whether you have the disease, he’ll helpfully say “yes” with probability 1 in 1000 regardless of whether you actually do or not.

Let  $D$  be the event that you have the disease, and let  $F$  be the event that the diagnosis is faulty. Use the Total Probability Law to compute  $\Pr\{F\}$ , the probability that Doctor  $Y$  made a mistake.

3. Which doctor is more reliable?

### 3 Prisoners

There are three prisoners in a maximum-security prison for fictional villains: the Evil Wizard Voldemort, the Dark Lord Sauron, and Little Bunny Foo-Foo. The parole board has declared that it will release two of the three, chosen uniformly at random, but has not yet released their names. Naturally, Sauron figures that he will be released to his home in Mordor, where the shadows lie, with probability  $\frac{2}{3}$ .

A guard offers to tell Sauron the name of one of the other prisoners who will be released (either Voldemort or Foo-Foo). However, Sauron declines this offer. He reasons that if the guard says, for example, “Little Bunny Foo-Foo will be released”, then his own probability of release will drop to  $\frac{1}{2}$ . This is because he will then know that either he or Voldemort will also be released, and these two events are equally likely.

Using a tree diagram and the four-step method, either prove that the Dark Lord Sauron has reasoned correctly or prove that he is wrong. Assume that if the guard has a choice of naming either Voldemort or Foo-Foo (because both are to be released), then he names one of the two uniformly at random.

## Chapter 16

Def: An event A is independent of an event B if

$$\Pr(A|B) = \Pr(A) \text{ or if } \Pr(B) = 0$$

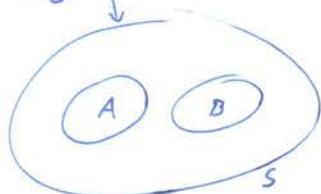
Ex: Flip 2 fair "independent" coins  $\Pr(A|B) = \Pr(A) = 1/2$

$B = \text{event 1st coin is heads } \Pr(B) = 1/2$

$$P_1(A) = \frac{1}{2}$$

disjoint  $\neq$  independent

disjoint



$$\Pr(A|B) = 0 \neq \Pr(A)$$

independent

Thm (Product rule for independent events)

IF A is independent of B, then  $\Pr(A \cap B) = \Pr(A)\Pr(B)$

## Proof

Case 1  $\Pr(B) = 0$ . Then  $\Pr(A \cap B) = 0 = \Pr(A)\Pr(B)$

□

Thm (Symmetry of independence)

IF  $A$  is independent of  $B$ , then  $B$  is independent of  $A$

Ex: Flip 2 unbiased coins (independent)

$A$  = event coins match

$B$  = event 1<sup>st</sup> coin Heads

Are  $A$  and  $B$  independent?

$$\Pr(A|B) = \Pr(\text{coin 2 is H}) = 1/2$$

$$\Pr(A) = \Pr(HH) + \Pr(TT) = \left(\frac{1}{2}\right)\left(\frac{1}{2}\right) + \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)$$

$$\Rightarrow \Pr(A|B) = \Pr(A)$$

$\Rightarrow A$  and  $B$  are independent

Ex: Flip 2 biased coins (independent)

$$\Pr(H) = p$$

$$\Pr(T) = 1-p$$

$A$  and  $B$  same from last example.

Are  $A$  and  $B$  independent?

$$\Pr(A|B) = \Pr(\text{coin 2 is H}) = p$$

$$\Pr(A) = \Pr(HH) + \Pr(TT) = p^2 + (1-p)^2$$

$A$  and  $B$  are independent iff  $p=0$  or  $p = p^2 + (1-p)^2$

$$0 = 1 - 3p + 2p^2$$

$$(1-2p)(1-p)$$

$$p = 1/2$$

$$p = 1$$

Def: Events  $A_1, A_2, \dots, A_n$  are mutually independent if

$$\forall i \in \mathbb{N}, \forall J \subseteq [1, n] - i$$

$$\Pr(A_i \cap_{j \in J} A_j) = \Pr(A_i) \quad \text{or} \quad \Pr(\bigcap_{j \in J} A_j) = 0$$

Equivalent Def (Product rule form)

$A_1, A_2, \dots, A_n$  are mutually independent if

$$\forall J \subseteq [1, n]$$

$$\Pr(\bigcap_{j \in J} A_j) = \prod_{j \in J} \Pr(A_j)$$

Ex:  $n=3$ ,  $A_1, A_2 \in A_3$  are mutually independent if

$$\Pr(A_1 \cap A_2) = \Pr(A_1) \Pr(A_2)$$

$$\Pr(A_1 \cap A_3) = \Pr(A_1) \Pr(A_3) \quad * \text{check every combination}$$

$$\Pr(A_2 \cap A_3) = \Pr(A_2) \Pr(A_3)$$

$$\Pr(A_1 \cap A_2 \cap A_3) = \Pr(A_1) \Pr(A_2) \Pr(A_3) \quad \Rightarrow \text{don't forget to check!}$$

Ex: Blood Matching

$\frac{1}{10}$  people match type O

$$\Pr(O \cap + \cap XYZ) = \Pr(O) \Pr(+) \Pr(XYZ)$$

$\frac{1}{5}$  " " rh factor +

$$= \left(\frac{1}{10}\right) \left(\frac{1}{5}\right) \left(\frac{1}{4}\right)$$

$\frac{1}{4}$  " " marker XYZ

$$= \frac{1}{200}$$

$\Rightarrow \frac{1}{200}$  " all 3

What if  $\Pr(+ \cap O) > \Pr(+)$

can't know these 3 are independent...

unless you test a lot of people

Ex: Flip 3 Fair mutually independent coins

$A_1$  = event coin 1 matches coin 2

$A_2$  = " 2 " 3

$A_3$  = " 1 " 3

Are  $A_1, A_2$  and  $A_3$  independent?

$$\forall i \Pr(A_i) = \Pr(HH) + \Pr(TT) = (\frac{1}{4}) + (\frac{1}{4}) = \frac{1}{2}$$

$$\forall i, j \Pr(A_i \wedge A_j) = \Pr(HHH) + \Pr(TTT) = (\frac{1}{8})(\frac{1}{8}) = \frac{1}{64} = \Pr(A_i)\Pr(A_j)$$

yes?  
not yet ... need to check all

$$\Pr(A_1 \wedge A_2 \wedge A_3) = \Pr(HHH) + \Pr(TTT) = \frac{1}{8} \neq \Pr(A_1)\Pr(A_2)\Pr(A_3)$$

NOT Independent

Def:

Events  $A_1, A_2, A_3, \dots, A_n$  are pairwise independent if  $\forall i \neq j$   
 $A_i \wedge A_j$  are independent

pairwise independence  $\not\Rightarrow$  mutual independence

mutual independence  $\Rightarrow$  pairwise independence

Ex Birthday Problem

$N$  birthdays ex  $N = 365$

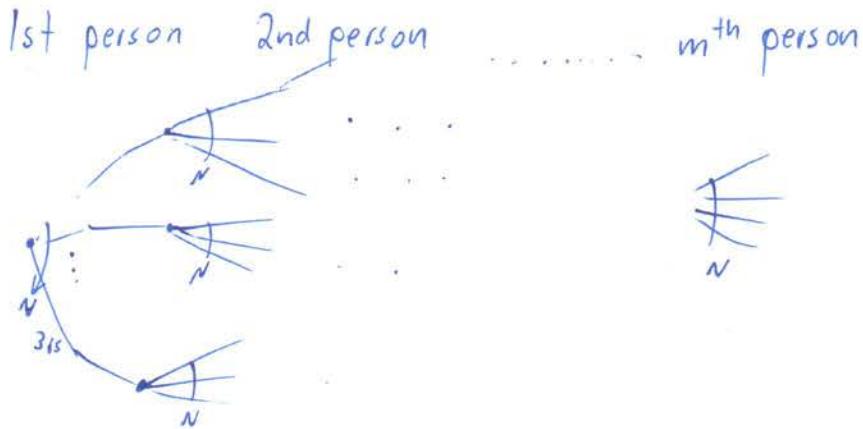
$M$  people ex  $M = 100$

What is probability 2 or more people have same birthday

Assumptions: Birthdays are equally likely and mutually independent

not really true ↴

## Birthday Sample Space



$$S = \{ \langle b_1, b_2, \dots, b_m \rangle \mid 1 \leq b_i \leq N \}$$

$$|S| = N^m$$

$$\Pr(\langle b_1, b_2, \dots, b_m \rangle) = \frac{1}{N^m}$$

# sample points w/ all b-days different

$$1 \cdot (N-1) \cdot (N-2) \cdots (N-M+1) = \frac{N!}{(N-M)!}$$

$$\Pr(\text{all b-days differ}) = \frac{N!}{(N-M)! N^m}$$

Recall Stirling's Formula:  $N! \sim \sqrt{2\pi n} \left(\frac{N}{e}\right)^N$

$$\sim e^{(N-M+1/2)/n} \left(\frac{N}{N-M}\right)^{N-M}$$

For  $N=365$ ,  $M=100$

$$\Pr(\text{all b-days diff}) = 3.07 \cdot 10^{-7}$$

For  $N=365$ ,  $M=23$

$$\Pr(\text{all b-days diff}) = .49$$

## Problems for Recitation 21

### 1 Bayes' Rule

*Bayes' Rule* says that if  $A$  and  $B$  are events with nonzero probabilities, then:

$$\Pr\{A | B\} \cdot \Pr\{B\} = \Pr\{B | A\} \cdot \Pr\{A\}$$

- a. Prove Bayes' Rule.
  
  
  
  
  
  
- b. A weatherman walks to work each day. Some days it rains:  
$$\Pr\{\text{rains}\} = 0.30$$
Sometimes the weatherman brings his umbrella. Usually this is because he predicts rain, but he also sometimes carries it to ward off bright sunshine.  
$$\Pr\{\text{carries umbrella}\} = 0.40$$
As a weatherman, he usually doesn't get caught out in a storm without protection:  
$$\Pr\{\text{carries umbrella} | \text{rains}\} = 0.80$$
Suppose you see the weatherman walking to work, carrying an umbrella. What is the probability of rain? Use Bayes' Rule.

### 3 The Immortals

There were  $n$  Immortal Warriors born into our world, but in the end *there can be only one*. The Immortals' original plan was to stalk the world for centuries, dueling one another with ancient swords in dramatic landscapes until only one survivor remained. However, after a thought-provoking discussion of probabilistic independence, they opt to give the following protocol a try:

1. The Immortals forge a coin that comes up heads with probability  $p$ .
2. Each Immortal flips the coin once.
3. If *exactly one* Immortal flips heads, then he or she is declared The One. Otherwise, the protocol is declared a failure, and they all go back to hacking each other up with swords.
  - a. One of the Immortals (the Kurgan from the Russian steppe) argues that as  $n$  grows large, the probability that this protocol succeeds must tend to zero. Another (McLeod from the Scottish highlands) argues that this need not be the case, provided  $p$  is chosen *very carefully*. What does your intuition tell you?
  - b. What is the probability that the experiment succeeds as a function of  $p$  and  $n$ ?
  - c. How should  $p$ , the bias of the coin, be chosen in order to maximize the probability that the experiment succeeds? (You're going to have to compute a derivative!)

## Expectation

Chpt 18, sec 19.1, 19.5.1-19.5.2

Def: The expectation (or average or mean) of a random variable  $R$  over sample space  $S$

$$Ex(R) = \sum_{w \in S} R(w) Pr(w)$$

Ex: Roll a 6-sided Fair dice.

$R$  = outcome  
weighted

$$Ex(R) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = \left(\frac{1}{6}\right)\left(\frac{7 \cdot 6}{2}\right) = 7/2$$

Def: The median of a r.v  $R$  is  $x \in Range(R)$

$$\Pr(R < x) \leq \frac{1}{2} \text{ and } \Pr(R > x) \leq \frac{1}{2}$$

Ex: So Median in 6sided Fair dice = 4

Thm:  $Ex[R] = \sum_{x \in Range(R)} x \Pr(R=x)$

$$\begin{aligned}
 \text{Proof: } Ex(R) &= \sum_{w \in S} R(w) Pr(w) = \sum_{x \in Range(R)} \sum_{\substack{w \in S \\ R(w)=x}} R(w) Pr(w) \\
 &= \sum_{x \in Range(R)} \sum_{\substack{w \in S \\ R(w)=x}} x R(w) Pr(w) = \sum_{x \in R} x \overbrace{\sum_{\substack{w \in S \\ R(w)=x}} Pr(w)}^{\Pr(R=x)}
 \end{aligned}$$

□

Corollary: IF  $R: S \rightarrow N$ ,  $Ex(R) = \sum_{i=0}^{\infty} i \Pr(R=i)$

Another rearrangement

Thm: IF  $R: S \rightarrow N$ ,  $Ex(R) = \sum_{i=0}^{\infty} \Pr[R > i]$

$$= \sum_{i=1}^{\infty} \Pr[R \geq i]$$

Proof

$$\begin{aligned} \sum_{i=0}^{\infty} \Pr[R > i] &= \Pr[R > 0] + \Pr[R > 1] + \Pr[R > 2] \\ &= \Pr[R=1] + \Pr[R=2] + \Pr[R=3] \\ &\quad + \Pr[R > 2] \\ &= \Pr[R=2] + \Pr[R=3] \\ &\quad + \Pr[R=3] \end{aligned}$$

$$\begin{aligned} &\Pr[R=1] + 2\Pr[R=2] + 3\Pr[R=3] \\ &= \sum_{i=0}^{\infty} i \Pr[R=i] = Ex(R) \end{aligned}$$

See online Notes

## Expectations Problems for Recitation 23

### 1 Expected Payoff

Here's yet another fun 6.042 game! You pick a number between 1 and 6. Then you roll three fair, independent dice.

- If your number never comes up, then you lose a dollar.
- If your number comes up once, then you win a dollar.
- If your number comes up twice, then you win two dollars.
- If your number comes up three times, you win *four* dollars!

What is your expected payoff? Is playing this game likely to be profitable for you or not?

3. To simplify the analysis, suppose that we always roll the dice three times, but may ignore the second or third rolls if we didn't previously get doubles. Let the random variable  $X_i$  be the sum of the dice on the  $i$ -th roll, and let  $E_i$  be the event that the  $i$ -th roll is doubles. Write the expected number of squares a piece advances in these terms.
  
4. What is the expected number of squares that a piece advances in Monopoly?

Expectation

$$\mathbb{E}[R] = \sum_{w \in S} R(w) \Pr[w]$$

$$= \sum_{r \in \text{range}(R)} r \Pr[R=r]$$

Indicator

$$I_E(w) = \begin{cases} 1 & w \in E \\ 0 & w \notin E \end{cases}$$

$$\mathbb{E}[I_E] = \Pr[I_E = 1] = \Pr[E]$$

$$\mathbb{E}[I_E] = 1 \Pr[R=1] + 0 \Pr[\cancel{R=0}] \xrightarrow{0}$$

Linearity of Expectation

$$\mathbb{E}[X_1 + X_2] = \mathbb{E}[X_1] + \mathbb{E}[X_2]$$

Conditional Expectations

$$\mathbb{E}[R|E] = \sum_{\substack{w \in S \\ w \in E}} R(w) \Pr[w|E]$$

$R = \# \text{ on Fair die}$      $E = \text{number is even}$

$$\mathbb{E}[R|E] = \sum_{w \in \{1, \dots, 6\}} R(w) \Pr[w|E]$$

$$= 1 \cdot 0 + 2 \frac{1}{3} + 3 \cdot 0 + 4 \frac{1}{3} + 5 \cdot 0 + 6 \frac{1}{3} = 4$$



$$\mathbb{E}[R_1 + R_2 | E] = \mathbb{E}[R_1 | E] + \mathbb{E}[R_2 | E]$$



Total Expectation

$$\mathbb{E}[R] = \mathbb{E}[\mathbb{E}[R|E]]$$

$$\mathbb{E}[R] = \mathbb{E}[R|E_1] \Pr[E_1] + \dots + \mathbb{E}[R|E_n] \Pr[E_n]$$

Lemma 1

$$\mathbb{E}[R|E] = \frac{\mathbb{E}[R \cdot I_E]}{\Pr[E]}$$

Proof

$$\sum_{s \in E^W} R(s) \Pr[s|E] = \frac{\Pr[s \cap E]}{\Pr[E]} = \begin{cases} 0 & I_E(s)=0 \\ 1 & I_E(s)=1 \end{cases}$$

See notes

Thm: Given a probability space  $S$  and events  $A_1, A_2 \dots A_n \subseteq S$

the expected #'s of events to occur is  $\sum_{i=1}^n \Pr(A_i)$

Proof: Let  $T_i(w) = \begin{cases} 1 & \text{if } w \in A_i \\ 0 & \text{otherwise} \end{cases}$

$T_i = 1$  iff  $A_i$  occurs

Let  $\#T = T_1 + T_2 + \dots + T_n$

$$\mathbb{E}(T) = \sum_{i=1}^n \Pr(T_i = 1)$$

$$= \sum_{i=1}^n \Pr(A_i)$$

□

Ex: Flip Fair coins

$A_i$  = event  $i^{\text{th}}$  coin is Heads

$T = \# \text{heads}$

independence not necessary

$$\mathbb{E}(T) = \Pr(A_1) + \Pr(A_2) + \dots + \Pr(A_n) \quad \swarrow$$

$$= \frac{1}{2} + \frac{1}{2} + \dots + \frac{1}{2} = n/2$$

or

$$\mathbb{E}(T) = \sum_{i=0}^n i \Pr(T=i) = \sum_{i=0}^n i \binom{n}{i} 2^{-n}$$

Assume mutual independence

$$\Rightarrow \sum_{i=0}^n i \binom{n}{i} 2^{-n} = n/2$$

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

+ (combinatorial / Probabilistic proof)

Thm 2

$$\Pr(T \geq 1) \leq E(T)$$

"upper bounded"

Proof

$$\begin{aligned} E(T) &= \sum_{i=1}^{\infty} \Pr(T \geq i) \\ &\geq \Pr(T \geq 1) \quad \square \end{aligned}$$

Corollary

$$\Pr(T \geq 1) \leq \sum_{i=1}^n \Pr(A_i)$$

Thm 3 Murphy's Law

Given mutually independent events,  $A_1, \dots, A_n$

then  $\Pr(T=0) \leq e^{-E(T)}$

$\nearrow$   
none of the  
events happen.

Proof

$$\begin{aligned} \Pr(T=0) &= \Pr(\bar{A}_1 \wedge \bar{A}_2 \wedge \bar{A}_3 \dots \bar{A}_n) \\ &= \prod_{i=0}^n \Pr(\bar{A}_i) \quad \text{Since mutually independent} \\ &= \prod_{i=0}^n (1 - \Pr(A_i)) \\ &\leq \prod_{i=1}^n e^{-\Pr(A_i)} = e^{-\sum_{i=1}^n \Pr(A_i)} = e^{-E(T)} \end{aligned}$$

Thm 4 : Product Rule For Expectation

For any independent random variables  $R_1 \in R_2$

$$E(R_1 R_2) = E(R_1) E(R_2)$$

Ex: Roll 2 6-sided fair independent dice

$R_1$  = value 1<sup>st</sup> die,  $R_2$  = value 2<sup>nd</sup> die

$$E(R_1 R_2) = E(R_1) E(R_2)$$

$$= \frac{7}{2} \cdot \frac{7}{2} = \frac{49}{4} = 12\frac{1}{4}$$

$$E(R_1^2) = \sum_{i=1}^6 i^2 P_i(R_1 = i) = \frac{1}{6}(1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2) = \frac{91}{6} = 15\frac{1}{6} \neq 12\frac{1}{4}$$

\* Need independence for product rule

Corollary 4.1

If  $R_1, R_2, \dots, R_n$  are mutually independent then

$$E(R_1 R_2 \dots R_n) = E(R_1) E(R_2) \dots E(R_n)$$

Corollary 4.2

For any constants  $a, b$  & any r.v.  $R$

$$E(aR + b) = aE(R) + b$$

Note

$$E(\frac{1}{R}) \neq \frac{1}{E(R)}$$

$$\text{Ex } R = \begin{cases} 0 & \text{w/prob } \frac{1}{2} \\ 1 & \end{cases}$$

$$E(R) = \frac{1}{2}$$

$$E(\frac{1}{R}) = \infty \neq \frac{1}{2}$$

~~Corollary ??~~

$$E(R/T) > 1$$

$$\Rightarrow E(R/T)E(T) > E(T)$$

not independent!

We also don't know that

$$E(T) > 0$$

$$\Rightarrow E(RT) > E(T)$$

$$\Rightarrow E(R) > E(T) \quad \square$$

YOU ARE  
THE COOL  
GUY I'VE EVER  
HAD THE UNFORTUNE  
TO BE ACCAINTED  
WITH

\* Never take averages of ratios!

... unless your name is Erika Lu.

Def The variance of a r.v  $R$  is  $\text{Var}(R) = E((R - E(R))^2)$

Variance is high if I'm likely to be far from expectation

$$\Pr(R=1000) = 1/2$$

$$\Pr(R=-1000) = 1/2$$

$$E(R) = 0$$

$$R - E(R) = \begin{cases} 1000 & w/ 1/2 \\ -1000 & w/ 1/2 \end{cases}$$

$$(R - E(R))^2 = 1M w/ prob 1$$

$$\text{Var}(R) = 1,000,000$$

$$\Pr(R=1) = 1/2$$

$$\Pr(R=-1) = 1/2$$

$$E(S) = 0$$

$$\text{Var}(S) = 1$$

Fun Fact:

$$\begin{aligned} E(R - E(R)) &= E(R) - E(E(R)) \\ &= (E(R) - E(R)) = 0 \end{aligned}$$

This is why we square  $(R - E(R))$

Def For any r.v  $R$ , the standard deviation of  $R$  is

$$\begin{aligned}\sigma(R) &= \sqrt{\text{Var}(R)} \\ &= \sqrt{E(\text{dev}^2)} \\ &= \text{root-mean square}\end{aligned}$$

## Problems for Recitation 24

### 1 Properties of Variance

In this problem we will study some properties of the variance and the standard deviation of random variables.

- (a) Show that for any random variable  $R$ ,  $\text{Var}[R] = E[R^2] - E^2[R]$ .

$$\begin{aligned}
 \text{Var}(R) &= E((R - E[R])^2) \\
 &= \sum_{\text{all } r} P(r) (r - E[R])^2 \\
 &= \sum_r P(r) \left( r^2 - 2rE[R] + E^2[R] \right) \\
 &\quad \begin{array}{c} \downarrow \\ E[R^2] \end{array} \quad \begin{array}{c} \downarrow \\ -E^2[R] \end{array} \quad \begin{array}{c} \Rightarrow \\ = [E[R^2] - E^2[R]] \end{array} \\
 &\quad \begin{array}{c} \text{Factor out} \\ -2E[R] \sum P(r)(r) \end{array} + E[R]
 \end{aligned}$$

- (b) Show that for any random variable  $R$  and constants  $a$  and  $b$ ,  $\text{Var}[aR + b] = a^2 \text{Var}[R]$ . Conclude that the standard deviation of  $aR + b$  is  $a$  times the standard deviation of  $R$ .

$$\begin{aligned}
 \text{Var}[aR + b] &= \text{Var}[aR] + \text{Var}[b] \\
 &= a^2 \text{Var}[R]
 \end{aligned}$$

- (c) Show that if  $R_1$  and  $R_2$  are independent random variables, then

$$\text{Var}[R_1 + R_2] = \text{Var}[R_1] + \text{Var}[R_2].$$

## 2 Law of Total Expectation

Suppose that an unpredictable friend has gone to buy you candy, and in eager wait, you're trying to imagine how many pieces of candy you might be about to receive. Your friend might buy Swedish Fish, say with 0.3 probability, and on average 25 individual fish in the package; or your friend might buy Sour Patch Kids, with 0.7 probability, and on average 43 individual pieces in the package. You might reasonably compute that the expected number of pieces of candy is

$$0.3 \cdot 25 + 0.7 \cdot 43.$$

The Law of Total Expectation justifies this:

**Rule** (Law of Total Expectation). *Let  $X$  and  $Y$  be two random variables, not necessarily independent. Then*

$$\mathbb{E}[X] = \mathbb{E}_Y [\mathbb{E}_{X|Y} [X|Y]].$$

In our candy scenario,  $Y$  represents the type of candy bought – Swedish Fish or Sour Patch Kids – and  $X$  represents the number of pieces of candy.

For Problem 2, prove the Law of Total Expectation. (You may assume that each random variable has only finitely many outcomes.)

## Expected Value for Function of a Random Variable

Let  $R$  be a random variable and  $f(R)$  is a function of  $R$ .  
 Then the expected value of a random variable  $f(R)$  is.

$$E[f(R)] = \sum_{x \in \text{Range}(R)} f(x) \cdot \Pr\{R=x\}$$

### Variants

$$\text{Var}[R] = E[(R - E[R])^2] = \sum_{r \in \text{range}} (r - \mu_R)^2 \Pr(R=r)$$

$$V = E[R^2] - E[R]^2 \geq 0 \Rightarrow E[R^2] \geq E[R]^2$$

$$\text{Var}[aR+b] = a^2 \text{Var}[R]$$

$$E[R_1 R_2] = E[R_1] E[R_2]$$

Independent

### Linearity of Var

\* Need pair-wise independence

$$\text{Var}\left[\sum_{i=1}^n R_i\right] = \sum_{i=1}^n \text{Var}[R_i]$$

### Binomial Distribution      $k$ heads in $n$ tosses

$$H_{n,p} = \sum_{i=1}^n I_i$$

$$f_n(k) = \text{distribution} = \binom{n}{k} p^k (1-p)^{n-k}$$

- toss  $n$  coins w/ bias  $p$
- $k$  of them heads
- Find distribution

$$\text{Variance of an indicator random variable} = p - p^2$$

Linearity of Expectation      no independence required

$$E[x_1 + x_2] = E[x_1] + E[x_2]$$

||

$$\sum_{s \in S} (x_1(s) + x_2(s)) P(s)$$

$$= \sum_{s \in S} x_1(s) p(s) + \sum_{s \in S} x_2(s) p(s)$$

||

$$E[x_1] + E[x_2]$$

Iterated Expectation / Total expectation

$$E[x] = E_{\Sigma} [f(\Sigma)]$$

$$f(\Sigma) = E[x | \Sigma = y]$$

constant that is random with respect to  $y$ .

$$E_{\Sigma} [$$

$$= \sum_{y \in \text{range}(\Sigma)} \Pr[\Sigma = y] \cdot f(y) = \sum \Pr[\Sigma = y] E[x | \Sigma = y]$$

(Chapter 18 and 19)

Markov's TheoremIF  $R$  is a non-neg random variable, then  $\forall x > 0$ 

$$\Pr(R \geq x) \leq \frac{Ex(R)}{x}$$

Proof

$$Ex(R) = \underbrace{Ex(R|R \geq x)}_{\geq x} \Pr(R \geq x) + \underbrace{Ex(R|R < x)}_{\geq 0} \Pr(R < x)$$

$$\geq x \Pr(R \geq x)$$

$$\Rightarrow \Pr(R \geq x) = \frac{Ex(R)}{x} \quad \square$$

Corollary : IF  $R$  is a non-neg r.v., then  $\forall c > 0$ 

$$\Pr(R \geq c Ex(R)) \leq \frac{1}{c}$$

Proof: set  $x = c Ex(R)$ Ex $R$  = weight of random personsuppose  $Ex(R) = 100$ 

$$\Pr(R \geq 200) = ?? \leq \frac{1}{2}$$

$c=2$

Markow with Negative <sup>rv</sup> numis?

$$\Pr(R=1000) = \frac{1}{2}$$

$$\Pr(R=-1000) = \frac{1}{2}$$

$$Ex(R) = 0$$

$$\text{Markov } \Rightarrow \Pr(R \geq 1000) \leq \frac{Ex(R)}{1000} = 0$$

Does NOT work.

Ex:

N-person Chinese Appetizer

Let  $R = \# \text{ people who get right appetizer back.}$

$$Ex(R) = 1$$

$$\Pr(R=N) = \Pr(R \geq N) \leq \frac{1}{N} \quad \leftarrow \text{Good Bound with Markov}$$

(but this might not work that well always..)

### Chebyshov's Theorem

For any rv  $R \in \mathbb{R} \times > 0$

$$\Pr(|R - Ex(R)| \geq x) \leq \frac{\text{Var}(R)}{x^2}$$

Proof

$$\begin{aligned} \Pr(\underbrace{|R - Ex(R)|}_{\text{deviation}} \geq x) &= \Pr(\underbrace{(|R - Ex(R)|)^2}_{\text{Always non-negative}} \geq x^2) \\ &\leq \frac{Ex((R - Ex(R))^2)}{x^2} = \frac{\text{Var}(R)}{x^2} \end{aligned}$$

□

standard dev.

Corollary:  $\Pr(|R - \bar{E}(R)| \geq c \sigma(R)) \leq \frac{1}{c^2}$

↑ off by  $c$  standard dev

Proof: Set  $x = c \sigma(R)$  in Chebychev

$$\Pr(|R - \bar{E}(R)| \geq c \sigma(R)) \leq \frac{\text{Var}(R)}{c^2 \sigma^2(R)} = \frac{1}{c^2} \quad \square$$

Example:  $R = \text{IQ of random person}$

Assume  $R \geq 0$ ,  $\bar{E}(R) = 100$ ,  $\sigma(R) = 15$

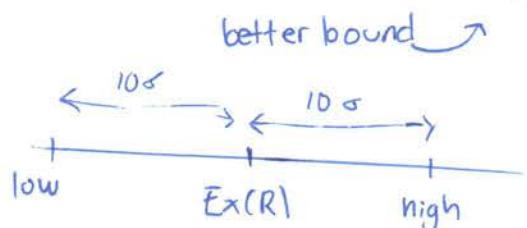
$$\Pr(R \geq 250) ??$$

Markov:  $\Pr(R \geq 250) \leq \frac{100}{250} = 2/5$

Chebychev:  $\Pr(R \geq 250) = \Pr(R - 100 \geq 150)$

$$\leq \Pr(|R - 100| \geq 150) = \Pr(|R - 100| \geq 10\sigma(R))$$

$$\leq \frac{1}{100} \text{ by corollary}$$



$$\Pr(|R - \bar{E}(R)| \geq 10\sigma(R))$$

We don't know if distribution is symmetric

Thm:

For any r.v  $R$   $\Pr(R - \bar{E}(R) \geq c \sigma(R)) \leq \frac{1}{c^2 + 1}$

II                   $\Pr(R - \bar{E}(R) \leq -c \sigma(R)) \leq \frac{1}{c^2 + 1}$

## Thm (Chernoff Bound)

Let  $T_1, T_2, \dots, T_n$  be any mutually independent r.v

such that  $\forall j \quad 0 \leq T_j \leq 1 \quad \text{Let } T = \sum_{j=1}^n T_j$

Then For any  $c > 1$

$$\Pr(T \geq c \mathbb{E}x(T)) \leq e^{-z \mathbb{E}x(T)}$$

where

$$z = c \ln c + 1 - c > 0$$

Proof :

$$\begin{aligned} \Pr(T \geq c \mathbb{E}x(T)) &= \Pr(c^T \geq c^{c \mathbb{E}x(T)}) \\ &\leq \frac{\mathbb{E}x(c^T)}{c^{c \mathbb{E}x(T)}} \quad \text{Markov} \end{aligned}$$

Lemma 1

$$\mathbb{E}x(c^T) \leq e^{(c-1) \mathbb{E}x(T)}$$

Proof

Math

$$\begin{aligned} \text{So} \quad &\leq \frac{e^{(c-1) \mathbb{E}x(T)}}{c^{c \mathbb{E}x(T)}} = e^{(c-1) \mathbb{E}x(T) - c \ln c \mathbb{E}x(T)} \\ &= e^{-(c \ln c - c + 1) \mathbb{E}x(T)} \\ &= e^{-z \mathbb{E}x(T)} \end{aligned}$$

Ex: Suppose  $E(X(T)) = 100$ ,  $c = 2$

$$\Rightarrow z = 2 \ln 2 + 1 - 2 > .38$$

$$(\text{Chernoff: } \Pr(T \geq z E(X(T))) \leq e^{-z E(X(T))} = e^{-38})$$

Markov:  $\dots = 1/2$   $\leftarrow$  not very good bound

Ex: 10 million people play "Pick 4"

$$\Pr(\text{win}) = 1/10,000$$

$$E(X(\#\text{winners})) = \frac{10,000,000}{10,000} = 1000$$

$$T_j = \begin{cases} 1 & \text{if } j^{\text{th}} \text{ player wins} \\ 0 & \dots \text{ loses} \end{cases}$$

$$T = T_1 + T_2 + \dots + T_{10,000,000}$$

$$\Pr(2000 \text{ winners}) \leq e^{-38 \cdot 1000} = e^{-380} \quad \text{Assuming mutual independence}$$

$$\Pr(1100 \text{ winners}) \leq e^{-0.048 \cdot 1000} = e^{-4.8} \leq .01$$

$$c = 1.1$$

$$z = 1.1 \ln 1.1 + 1 - 1.1 \geq .0048$$

Ex:  $N$  jobs  $B_1, B_2, \dots, B_N$   $M$  servers  $S_1, S_2, \dots, S_M$

Goal: Balance the load

$$N = 100,000$$

$$M = 10$$

$B_j$  takes  $L_j$  time ( $0 \leq L_j \leq 1$ )

$$\text{Let } L = \sum_{j=1}^N L_j \quad \text{Optimal time is } L/M$$

Let  $R_{ij}$  be the load on  $S_i$  from  $B_j$

$$R_{ij} = \begin{cases} L_j & \text{if } B_j \text{ assigned to } S_i, \text{ prob } 1/M \\ 0 & \text{if "not", } 1 - 1/M \end{cases}$$

Let  $R_i$  be the total load on  $S_i \Rightarrow R_i = \sum_{j=1}^N R_{ij}$

Linearity of Expectation

$$E(R_i) = \sum_{j=1}^N E(R_{ij})$$

$$= \sum_{j=1}^N L_j / M \quad P_i(R_i \geq \frac{cL}{M}) \leq e^{-z \frac{L}{M}} \quad z = c \ln n + 1 - c$$

$$= L / M$$

$$\Pr(\text{worst server takes } \geq \frac{cL}{M}) = \Pr(R_1 \geq \frac{cL}{M} \cup R_2 \geq \frac{cL}{M} \cup \dots \cup R_m \geq \frac{cL}{M})$$

$$\leq \Pr(R_1 \geq \frac{L}{M}) + \Pr(R_2 \geq \frac{cL}{M}) + \dots + \Pr(R_m \geq \frac{cL}{M})$$

$$\leq M e^{-z \frac{L}{M}}$$

$$c. 1.1 (10\% \text{ above opt}) \leq 10e^{-0.0017 \cdot 2500} < \frac{1}{160000}$$

$$z \geq .0048$$

$$L/M = \frac{L \cdot 0.0017}{10} \approx 2,500$$

## Problems for Recitation 25

**Theorem 1.** Let  $E_1, \dots, E_n$  be events, and let  $X$  be the number of these events that occur. Then:

$$\text{Ex}(X) = \Pr\{E_1\} + \Pr\{E_2\} + \dots + \Pr\{E_n\}$$

**Theorem 2** (Markov's Inequality). Let  $X$  be a nonnegative random variable. If  $c > 0$ , then:

$$\Pr\{X \geq c\} \leq \frac{\text{Ex}(X)}{c}$$

**Theorem 3** (Chebyshev's Inequality). For all  $x > 0$  and any random variable  $R$ ,

$$\Pr\{|R - \text{Ex}(R)| \geq x\} \leq \frac{\text{Var}[R]}{x^2}$$

**Theorem 4** (Union Bound). For events  $E_1, \dots, E_n$ :

$$\Pr\{E_1 \cup \dots \cup E_n\} \leq \Pr\{E_1\} + \dots + \Pr\{E_n\}$$

**Theorem 5** (“Murphy’s Law”). If events  $E_1, \dots, E_n$  are mutually independent and  $X$  is the number of these events that occur, then:

$$\Pr\{E_1 \cup \dots \cup E_n\} \geq 1 - e^{-\text{Ex}(X)}$$

**Theorem 6** (Chernoff Bounds). Let  $E_1, \dots, E_n$  be a collection of mutually independent events, and let  $X$  be the number of these events that occur. Then:

$$\Pr\{X \geq c \text{Ex}(X)\} \leq e^{-(c \ln c - c + 1) \text{Ex}(X)} \quad \text{when } c > 1$$

- e. Use Theorem 5 to lower bound the probability that I forget one or more items.
- g. I'm supposed to remember many other items, of course: clothing, watch, backpack, notebook, pencil, kleenex, ID, keys, etc. Let  $X$  be the total number of items I remember. Suppose I remember items mutually independently and  $\text{Ex}(X) = 36$ . Use Chernoff's Bound to give an upper bound on the probability that I remember 48 or more items.
- h. Give an upper bound on the probability that I remember 108 or more items.

12/10/14

# 6.042 Recitation

Fernando Trujano

### Basic Counting Notions

$f: A \rightarrow B$

Injective: Every elmt of B mapped to at most one.  $|A| \leq |B|$

Surjective: " least  $|A| \geq |B|$

Bijective: Both injective and surjective.  $|A| = |B|$  bijection rule

Generalized Pigeon Hole Principle:

If  $|X| > K \cdot |Y| \Rightarrow$  For any  $F: X \rightarrow Y$   $\exists$   $K+1$  diff elmts of X mapped to the same element in Y.

Division rule:  $F: A \rightarrow B$   $k \rightarrow 1 \Rightarrow |A| = K \cdot |B|$

Product rule:  $P_1, P_2, P_3 \dots$  sets  $|P_1 \times P_2 \times \dots| = |P_1| \cdot |P_2| \cdot \dots$

Sum rule:  $A_1, \dots, A_n$  disjoint sets  $|A_1 \cup A_2 \cup \dots \cup A_n| = \sum_{k=1}^n |A_k|$

Binomial Theorem:  $(a+b)^n = \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k$

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \binom{n}{k, n-k} \quad \binom{n}{k} = \binom{n}{n-k}$$

$\in n$  choose  $k$

Exclusion-Inclusion: count non-disjoint sets

$$\begin{aligned} A_1 \cup A_2 \cup A_3 &= |A_1| + |A_2| + |A_3| - |A_1 \cap A_2| \\ &\quad - |A_1 \cap A_3| - |A_2 \cap A_3| + |A_1 \cap A_2 \cap A_3| \\ \text{General: } |\bigcup_{i=1}^n A_i| &= \sum_{i=1}^n |A_i| - \sum_{1 \leq i < j \leq n} |A_i \cap A_j| \\ &\quad + \sum_{1 \leq i < j < k \leq n} |A_i \cap A_j \cap A_k| - \dots + (-1)^{n+1} |\bigcap_{i=1}^n A_i| \end{aligned}$$

### Generating Functions

$$(a_0, a_1, a_2, \dots) \leftrightarrow A(x) = \sum_{i=0}^{\infty} a_i x^i \quad (1, 1, 1, \dots) \leftrightarrow 1+x+x^2+\dots$$

$$\text{Right shift: } (0, 0, \dots, 0, a_0, a_1, \dots) \leftrightarrow x^k A(x) \quad = \sum_{i=0}^{\infty} x^i = \frac{1}{1-x}$$

$$\text{Derivative: } (1, 2, 3, 4, \dots) \rightarrow 1+2x^2+3x^3+\dots = (1+x+x^2+x^3+\dots)' = \left(\frac{1}{1-x}\right)' = \frac{1}{(1-x)^2}$$

The generating function for choosing elements from a union of disjoint sets is the product of the generating function from choosing each set.

Convolution Rule: # apples even

Ex: Fruit Salad # bananas mult of 5 At most 5 oranges At most 1 pear

$$A(x) = 1+0x+x^2+0x^3+\dots = \frac{1}{1-x^2} \quad B(x) = 1+x^5+x^{10}+\dots = \frac{1}{1-x^5}$$

# ways of selecting zero Apples

$$D(x) = 1+x+x^2+x^3+x^4 = \frac{1-x^5}{1-x} \quad P(x) = 1+x$$

$$F(x) = A(x)B(x)D(x)P(x) = \frac{1}{(1-x)^3} \quad (\text{By convolution rule})$$

Taylor Series Rule

$$\langle f_0, f_1, f_2, \dots \rangle \quad f_n = \frac{f^{(n)}(0)}{n!} \quad \text{Allows you to find coefficient of } x^n$$

Bookkeeper Rule:  $k_1, \dots, k_m$  distinct elements with  $k_i$  occurrences of  $k_i$ .

$$\# \text{ sequences} = \frac{(k_1+k_2+\dots+k_m)!}{k_1! k_2! \dots k_m!}$$

Combinatorial Proofs: 1) Define set S 2) Show  $|S|=n$  by counting

Pascals Identity 3) Show  $|S|=m$  by counting another way

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$$

4) Conclude that  $m=n$

## Counting

### Finding closed Form For sums

Perturbation method:

$$\sum_{i=0}^n i = ? \quad S = 1 + x + x^2 + x^3 + \dots + x^{n-1}$$

$$-xS = x + x^2 + x^3 + \dots + x^{n-1} + x^n$$

$$(1-x)S = 1 - x^n \Rightarrow S = \frac{1-x^n}{1-x}$$

Derivative Method.

$$\text{We know: } \sum_{i=0}^n i x^{i-1} = \frac{1-x^{n+1}}{1-x} \quad \sum_{i=0}^n i x^{i-1} = ??$$

Take derivative of both sides,

$$\sum_{i=0}^n i x^{i-1} = \frac{(n+1)x^n + nx^{n+1}}{(1-x)^2}$$

### Integration Bounds: Approximate a sum

f-positive, increasing  $\Rightarrow$

$$f(1) + \int_1^n f(x) dx \leq \sum_{i=1}^n f(i) \leq \int_1^n f(x) dx + f(n)$$

f-positive, decreasing  $\Rightarrow$

$$\int_1^n f(x) dx + f(n) \leq \sum_{i=1}^n f(i) \leq \int_1^n f(x) dx + f(1)$$

Stirling's Formula:  $n! \sim \left(\frac{n}{e}\right)^n \sqrt{2\pi n}$

Asymptotic Notation:  $f(x) = O(g(x))$

$$O(\leq) \Rightarrow \lim_{n \rightarrow \infty} \left| \frac{f(x)}{g(x)} \right| < \infty$$

$$o(<) \Rightarrow \lim_{n \rightarrow \infty} \left| \frac{f(x)}{g(x)} \right| = 0$$

$$\tilde{\mathcal{L}}(\geq) \Rightarrow \lim_{n \rightarrow \infty} \left| \frac{f(x)}{g(x)} \right| > 0$$

$$w(>) \Rightarrow \lim_{n \rightarrow \infty} \left| \frac{f(x)}{g(x)} \right| = \infty$$

$$\Theta(=) \Rightarrow \lim_{n \rightarrow \infty} \left| \frac{f(x)}{g(x)} \right| = L \quad 0 < L < \infty$$

$$\sim C^{(=)} \Rightarrow \lim_{n \rightarrow \infty} \left| \frac{f(x)}{g(x)} \right| = 1$$

### Solving Recurrences:

Plug and chug:  $S_n = 1 + 2S_{n-1}$

$$= 1 + 2[1 + 2S_{n-2}] = 1 + 2 + 4S_{n-2}$$

$$= 1 + 2 + 4(1 + 2S_{n-3}) = 1 + 2 + 4 + 8S_{n-3}$$

$$S_n = 1 + 2 + 4 + \dots + 2^{n-1} S_0 = 2^{n-1} \underbrace{S_0}_{\text{closed form}}$$

Akra-Bazzi: Recurrence in form:

$$\sum_{i=1}^k a_i T(b_i x + h_i(x)) + g(x)$$

$$T(x) = \Theta\left(x^p \left(1 + \int_1^x \frac{g(u)}{u^{p+1}} du\right)\right)$$

$$\text{where } \sum_{i=1}^k a_i b_i^p = 1$$

$$\text{Example: } T(n) = 2T\left(\frac{n}{2}\right) + \frac{8}{n} T\left(\frac{3n}{4}\right) + n^2$$

$$1) a_1 = 2, b_1 = \frac{1}{2}, a_2 = \frac{8}{3}, b_2 = \frac{3}{4}$$

$$2) \text{Find } p: 2\left(\frac{1}{2}\right)^p + \left(\frac{8}{3}\right)\left(\frac{3}{4}\right)^p \Rightarrow p = 2$$

$$3) \text{Magic: } T(n) = \Theta\left(n^2 \left(1 + \int_1^n \frac{u^2}{u^3} du\right)\right)$$

$$= \Theta(n^2(1 + \log n)) = \Theta(n^2 \log n)$$

## Solving Linear Recurrences

- 1) Rearrange linear recurrence

$$f(n) = \underbrace{a_1 f(n-1) + a_2 f(n-2) + \dots + a_d f(n-d)}_{\text{homogeneous part}} + g(n)$$

-Largest "n" on the left

- 2) Find the roots of the characteristic equation

$$x^n = a_1 x^{n-1} + a_2 x^{n-2} + \dots + a_k$$

- 3) Write the homogeneous solution

Non repeated root  $\Rightarrow c_1 r_1^n$   
 repeated root  $r$  times  $\Rightarrow c_1 r^n + c_2 n r^n + \dots + c_k n^{k-1} r^n$   
 sum up all the terms. are constants for later

- 4) Find a particular solution - Guess and verify

If  $g(n)$  is a polynomial try  $f(n) = \text{poly of same degree}$  If it doesn't work try with a degree higher

so if  $g(n) = n$  try  $f(n) = bn + c$  then  $f(n) = an^2 + bn + c$

If  $g(n)$  is an exponential like  $3^n$ . guess  $f(n) = c3^n$   
 then  $f(n) = bn3^n + c3^n$  then  $f(n) = an^23^n + bn3^n + c3^n$

- 5) Form general solution by adding homogeneous and particular solution.

- 6) Substitute boundary conditions into general solution to solve for unknown constants

## Linear Re

Linear Recurrence:  $t(n) = \frac{t(n+1)}{3} + \frac{2+t(n-1)}{3} + 1$        $t(0) = t(T) = 0$

Example:  $t(n) = 3t(n+1) - 3t(n) - 2t(n-1) - 3$

1) Rearrange:  $t(n+1) = 3t(n) + 2t(n-1) + 3$

2) Characteristic Equation:  $\lambda^{n+1} = 3\lambda^n + 2\lambda^{n-1}$

$$0 = \lambda^2 - 3\lambda + 2 \Rightarrow \lambda = 2 \quad \lambda = 1$$

3) Homogeneous solution:  $a\lambda^n + b\lambda^{-n}$

4) Particular solution: Since  $g(n)$  is degree 0  
guess  $t(n) = c$   $\Leftarrow$  For all  $n$

check:  $c = 3c - 2c - 1 \Rightarrow c = -1 \quad X$

guess degree higher:  $t(n) = bn + c$

check:  $b(n+1) + c = 3(bn + c) - 2(b(n-1) + c) - 3$

$$bn + b + c = 3bn + 3c - 2bn - 2b - 2c - 3$$

$b = 3 \Rightarrow$  Particular Solution =  $3n$

5) General Solution:  $t(n) = a\lambda^n + b\lambda^{-n} + 3n$

6) Boundary conditions: Recall  $t(0) = t(T) = 0$

$t(0) = a + b \Rightarrow b = -a$

$t(T) = a2^T + b2^{-T} = 0 = a2^T - a + 3T$

$= a(2^T - 1) + 3T \Rightarrow a = \frac{3T}{1-2^T}$

$\Rightarrow b = -\frac{3T}{1-2^T}$

$t(n) = \frac{-3T}{1-2^T} + \frac{3T}{1-2^T} 2^n + 3n$

## Random Variables

↳ Total function whose domain is the sample space  
RV maps each outcome in sample space to a #

Indicator RV: Maps outcomes to 0 or 1  
Partition the sample space  $\text{mr. } E(R) = A(A)$

so  $\Pr(R=1) \Rightarrow \Pr(R \text{ maps to } 1)$  for all events  
Probability density Function (pdf) in sample space

$\text{PDF}_R := \begin{cases} \Pr(R=x) & \text{if } x \in \text{range}(R) \\ 0 & \text{else} \end{cases}$

Cumulative Distribution Function (cdf)

$$\begin{aligned} \text{CDF}_R(x) &= \Pr(R \leq x) \\ &= \sum_{y \leq x} \Pr(R=y) = \sum_{y \leq x} \text{PDF}_R(y) \end{aligned}$$

$$\text{Variance} = \text{Var}[R] = E[(R - E(R))^2]$$

$$= E(R^2) - E^2(R)$$

↳ indicator rv where  $\Pr(B=1) = p$

$$\star \text{Var}(B) = p(1-p)$$

$$\text{Var}[aR+b] = a^2 \text{Var}[R]$$

IF  $R_1$  and  $R_2$  independent (pairwise)

$$\text{Var}[R_1 + R_2] = \text{Var}[R_1] + \text{Var}[R_2]$$

Markov's Theorem - Find probability that rv exceeds set threshold

$$\Pr(R \geq x) \leq \frac{E(R)}{x} \quad \forall x > 0$$

no independence assumptions

With Bounds: IF  $R > x \in R$

$$\Pr(R \geq x) \leq \frac{E(R) - x}{x - x} \quad \forall x \geq 1$$

Below:  $a \in R$   $R \leq a$

$$\Pr(R \leq x) \leq \frac{a - E(R)}{a - x} \quad \forall x < a$$

Chebychev's Theorem

$$\Pr[|R - E(R)| \geq x] \leq \frac{\text{Var}[R]}{x^2}$$

$$\Pr[|R - E(R)| \geq c \sigma_R] \leq \frac{1}{c^2}$$

↳ so Prob that  $R$  will deviate more than  $c$  standard devs from  $E(R) \leq 1/c^2$

Probability <sup>?</sup> Do NOT rely on intuition!

Sample space: set of all possible outcomes  
uniform if every pt. equal prob.

Conditional Probability:  $\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}$

Product Rule:  $\Pr(A \cap B) = \Pr(B) \Pr(A|B)$

A is independent of B:  $\Pr(A|B) = \Pr(A)$   
or  $\Pr(B) = 0$

A is disjoint of B:  $\Pr(A \cap B) = 0$

Product Rule for independent events:

IF A indep B:  $\Pr(A \cap B) = \Pr(A)\Pr(B)$

Symmetry of Independence:

If A indep B  $\rightarrow$  B indep A

Mutual Independence

Probability of any event unaffected by  
knowledge of other events.

$$\Pr(E_i) = \Pr\left(E_i \mid \bigwedge_{j \neq i} E_j\right)$$

$$\Pr\left[\bigwedge_{j \in E} E_j\right] = \prod_{j \in E} \Pr(E_j)$$

To check mutual independence check that  $\Pr$  of every combination of events  $= \prod \Pr(E_j)$

Pairwise Independence

All events are independent from each other

$$PW \nRightarrow M \quad M \Rightarrow PW$$

Bayes Rule:  $\Pr(A|B) \Pr(B) = \Pr(B|A) \Pr(A)$

Law of total probability:  $\Pr(A) = \Pr(A|E) \Pr(E) + \Pr(A|\bar{E}) \Pr(\bar{E})$

4) Step Method:

1) Find Sample space (tree diagram)

2) Define events of interest

3) Determine outcome probabilities

4) Sum prob of all outcomes for events

Expectation aka average, mean

$$E(R) = \sum_{w \in S} R(w) \Pr(w)$$

$$= \sum_{x \in \text{Range}(R)} x \Pr(R=x)$$

If  $\text{range}(R) = \mathbb{N}$

$$E(R) = \sum_{i=0}^{\infty} i \Pr(R=i) = \sum_{i=1}^{\infty} \Pr(R>i)$$

Conditional Expectation

$$E(R|A) := \sum_{r \in \text{Range}(R)} r \Pr(R=r|A)$$

Law of Total Expectation

$$E(R) = \sum_i E(R|A_i) \Pr(A_i)$$

Collusion - can't always assume everything is random. Recall coin flip game.

Linearity of Expectation

$$E(R_1 + R_2) = E(R_1) + E(R_2)$$

no independence required

$$E(a_1 R_1 + a_2 R_2) = a_1 E(R_1) + a_2 E(R_2)$$

$$E\left(\sum_{i=1}^k a_i R_i\right) = \sum_{i=1}^k a_i E(R_i)$$

IF  $R_1$  and  $R_2$  independent

$$E(R_1 R_2) = E(R_1) E(R_2)$$

Expected #events to occur =  $\sum_{i=1}^n \Pr(A_i)$

## Practice Final Exam

- The exam is **closed book**, but you may have three 8.5" × 11" sheet with notes (either printed or in your own handwriting) on both sides.
- Calculators and electronic devices (including cell phones) are not allowed.
- You may assume all of the results presented in class. This does **not** include results demonstrated in practice quiz material.
- Please show your work. Partial credit cannot be given for a wrong answer if your work isn't shown.
- Write your solutions in the space provided. If you need more space, write on the back of the sheet containing the problem. Please keep your entire answer to a problem on that problem's page.
- Be neat and write legibly. You will be graded not only on the correctness of your answers, but also on the clarity with which you express them.
- If you get stuck on a problem, move on to others. The problems are not arranged in order of difficulty.

NAME: \_\_\_\_\_

TA: \_\_\_\_\_

Problem	Value	Score	Grader
1	10		
2	12		
3	12		
4	10		
5	15		
6	15		
7	11		
8	14		
9	16		
10	15		
11	15		
12	15		
Total	160		

**Problem 1. [10 points]** It is well-known that *elitotis* is a common disease amongst Harvard students: in fact, 1 in 10 of their students have it. Fortunately, MIT has developed a reliable test for the presence of elevated *arrogentes* levels, which is helpful for testing for elitotis because:

- a student with elitotis has elevated arrogentes levels with probability 4/5, and
- a student with no elitosis has elevated arrogentes levels with probability 1/3.

What is the probability that a student selected uniformly at random from Harvard has elitosis, given that he or she has elevated arrogentes levels?

$$\Pr(E|A) = \frac{\Pr(A|E)\Pr(E)}{\Pr(A)} \quad \text{Bayes Rule}$$

$$\hookrightarrow \Pr(A) = \Pr(A|E)\Pr(E) + \Pr(A|\bar{E})\Pr(\bar{E}) \quad \text{Total Probability}$$

$$\left(\frac{4}{5}\right)\left(\frac{1}{10}\right) + \left(\frac{1}{3}\right)\left(\frac{9}{10}\right)$$

$$\Pr(E|A) = \frac{\left(\frac{4}{5}\right)\left(\frac{1}{10}\right)}{\left(\frac{4}{5}\right)\left(\frac{1}{10}\right) + \left(\frac{1}{3}\right)\left(\frac{9}{10}\right)}$$

**Problem 3. [12 points]** Let  $T$  be a positive integer. Consider the following recurrence equation:

$$t(n) = \frac{t(n+1)}{3} + \frac{2t(n-1)}{3} + 1 \text{ for } 1 \leq n \leq T-1; \quad t(0) = t(T) = 0.$$

Find a closed form solution for  $t(n)$  for  $0 \leq n \leq T$  as a function of  $T$ .

**Problem 4. [10 points]** Determine a closed form formula for the following sum (here,  $n$  is a positive integer):

$$\sum_{i=1}^n \sum_{j=i}^n \frac{1}{j}.$$

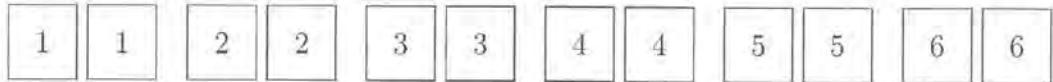
**Problem 6. [15 points]**

Let  $T$  be a tree with  $n$  nodes, where  $n$  is a positive integer. Suppose we color each node of  $T$  in a random way: each node is red with probability  $1/3$ , green with probability  $1/3$  and blue with probability  $1/3$ , and the colors of distinct nodes are mutually independent.

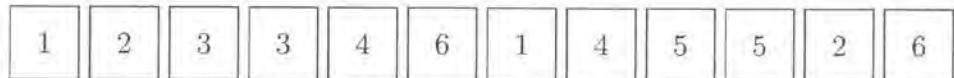
Find a formula for the probability that the resulting coloring is a proper coloring. You do not need to give a full proof of your answer, but do include your reasoning.

(*Hint: you may use the fact that any tree on at least 2 nodes has at least one leaf.*)

**Problem 7. [11 points]** You have twelve cards:



You shuffle them well, and deal them in a row (so the ordering will be a uniformly random permutation). For example, you might get:



What is the expected number of adjacent pairs with the same value? In the example, there are two adjacent pairs with the same value, the 3's and the 5's.

**Problem 8. [14 points]**

T-Pain is planning an epic boat trip and he needs to decide what to bring with him.

- He *definitely* wants to bring burgers, but they only come in packs of 6.
- He and his two friends can't decide whether they want to dress formally or casually. He'll either bring 0 pairs of flip flops or 3 pairs.
- He doesn't have very much room in his suitcase for towels, so he can bring at most 2 (and might not bring any!)
- In order for the boat trip to be truly epic, he has to bring at least 1 nautical-themed pashmina afghan.

(a) [7 pts] Let  $g_n$  be the number of different ways for T-Pain to bring  $n$  items (burgers, pairs of flip flops, towels, and/or afghans) on his boat trip, satisfying the restrictions above. Express the generating function  $G(x) := \sum_{n=0}^{\infty} g_n x^n$  as a quotient of polynomials.

(b) [7 pts] Let  $H(x) := \sum_{n=0}^{\infty} h_n x^n$  be the generating function for the sequence  $h_0, h_1, h_2, \dots$  representing the number of ways T-Pain could write an epic book of  $n$  chapters about his boat trip. It turns out that

$$H(x) = \frac{3 - 2x}{(1 - x)^2} - 2.$$

Using this information, determine a closed formula for  $h_n$ .

**Problem 9. [16 points]** Clumsy Clarke is rather injury prone:

- Every time he enters his car, which he does 72 times a month, he bumps his head with probability  $1/6$ .
- Each time he enters his house, which he does 32 times a month, he nicks his finger with probability  $1/2$ .
- Every time he does some shopping, which he does 25 times a month, he drops a bag on his foot with probability  $1/5$ .

All of these events are mutually independent.

(a) [4 pts] What is the expected number of injuries Clumsy Clarke experiences in a month?

$$E(X) = \left(\frac{1}{6}\right) 72 + \left(\frac{1}{2}\right) 32 + \left(\frac{1}{5}\right) 25$$

$$= 33$$

(b) [4 pts] What is the variance in the number of injuries Clarke has in a month?

$$\Pr(B=1) = 1/6 \text{ hits head}$$

$$\text{Var}(B) = \left(\frac{1}{6}\right)\left(\frac{5}{6}\right) 72$$

$$\text{Var}(N) = \left(\frac{1}{2}\right)\left(\frac{1}{2}\right) 32$$

$$\text{Var}(F) = \left(\frac{1}{5}\right)\left(\frac{4}{5}\right) 25$$

$$\text{Var}(X) = \text{Var}(B) + \text{Var}(N) + \text{Var}(F)$$

$$= 22$$

(c) [4 pts] What would the Markov bound be on the probability that Clarke has 100 or more injuries in a month?

$$P(R \geq 100) \leq \frac{E(X)}{100}$$

(d) [4 pts] What would the Chebyshev bound be on the probability that Clarke has 100 or more injuries in a month?

**Problem 10. [15 points]** Alyssa, an industrious software developer, writes 100 lines of code each hour, with probability  $p$  of making a mistake each hour (she never makes more than one mistake in an hour, though). Assuming Alyssa works  $n$  hours in a day and the mistakes she makes in each hour are independent, give formulas for the following:

(a) [4 pts] The probability of exactly  $k$  mistakes in a day.

(b) [5 pts] The probability of at least one mistake in a day.

(c) [6 pts] The expected number of hours until either the first mistake, or the end of the work day, whichever comes first. (Assume that if Alyssa makes a mistake in some hour, she makes it at the beginning of that hour).

**Problem 11. [15 points]** Consider the following game. You have the following grid:

1	2	3	4	5
6	7	8	9	10
11	12	13	14	15
16	17	18	19	20
21	22	23	24	25

There are 25 balls in a bucket, numbered from 1 to 25. 7 of these balls are randomly chosen from the bucket. You cross out the 7 numbers on your grid corresponding to the selected balls. If you cross out an entire row, an entire column, or either of the diagonals, then you win! E.g., if the draw yields balls 2, 6, 7, 12, 15, 17 and 22, then you win.

What is the probability of winning?

**Problem 12. [15 points]** Consider the following tennis tournament. There are  $n$  players, and every pair of players play against each other once. Moreover, all the players are equally matched, and so the winner of each matchup is uniformly random; the game outcomes are also mutually independent.

Call a player *awesome* if they win all their games, and *terrible* if they lose all of them.

- (a) [5 pts] What is the probability that there will be an awesome player?

(b) [5 pts] What is the probability that there will be both an awesome player and a terrible player?

(c) [5 pts] What is the probability that there will be neither an awesome player nor a terrible player?

## Counting

Finding closed form for sums:

Permutation method:  
 $\sum_{i=0}^n = ? \quad S = 1 + x + x^2 + x^3 + \dots + x^{n+1}$   
 $xS = x + x^2 + x^3 + \dots + x^{n+1}x^n$   
 $(1-x)S = 1 - x^{n+1} \Rightarrow S = \frac{1-x^{n+1}}{1-x}$

Derivative method:  
 $\frac{d}{dx} y = 1 = \frac{(1-x)^{n+1}}{(1-x)^2} \sum_{i=0}^n i x^{i-1}$

We know:  $\sum_{i=0}^n i x^{i-1} = \frac{(1-x)^{n+1}}{(1-x)^2} \sum_{i=0}^n i x^{i-1} = ?$

Integration bounds: approximate a sum  
 Take derivative of both sides  
 $\int_0^1 f(x) dx \leq \sum_{i=1}^n f(i) \leq \int_1^0 f(x) dx + f(0)$

Integrating, decreasing  $\Rightarrow$   
 $\int_0^1 f(x) dx + f(0) \leq \sum_{i=1}^n f(i) \leq \int_1^0 f(x) dx + f(0)$

Stirling's Formula:  $n! \sim \left(\frac{n}{e}\right)^n \sqrt{2\pi n}$

Asymptotic Notation:  $R(x) = O(g(x))$   
 $O(f(x)) \Rightarrow \lim_{n \rightarrow \infty} \left| \frac{f(x)}{g(x)} \right| < \infty$

$O(f(x)) \Rightarrow \lim_{n \rightarrow \infty} \left| \frac{f(x)}{g(x)} \right| = 0 \quad f = o(g) \Rightarrow f = O(g)$

$\sum_{i=1}^n (i-1) \Rightarrow \lim_{n \rightarrow \infty} \left| \frac{\sum_{i=1}^n (i-1)}{g(x)} \right| > 0$

$\ln(n!) \Rightarrow \lim_{n \rightarrow \infty} \left| \frac{\ln(n!)}{g(x)} \right| = \infty$

$O(\ln(n)) \Rightarrow \lim_{n \rightarrow \infty} \left| \frac{\ln(n)}{g(x)} \right| = 1$

$\sum_{i=1}^n (i-1) \Rightarrow \lim_{n \rightarrow \infty} \left| \frac{\sum_{i=1}^n (i-1)}{g(x)} \right| > 0$

$\ln(n!) \Rightarrow \lim_{n \rightarrow \infty} \left| \frac{\ln(n!)}{g(x)} \right| = \infty$

$O(\ln(n)) \Rightarrow \lim_{n \rightarrow \infty} \left| \frac{\ln(n)}{g(x)} \right| = 1$

Solving Recurrences:

Plot and check:  $S_n = 1 + 2S_{n-1}$   
 $= 1 + 2(1 + 2S_{n-2}) = 1 + 2 + 4S_{n-2}$   
 $+ 1 + 2 + 4(1 - 2S_{n-3}) = 1 + 2 + 4 + 8S_{n-3}$

$S_n = 1 + 2 + 4 + \dots + 2^n S_0 = 2^n - 1$  (closed form)

Akin-Bazzi: Recurrence in form:  
 $\sum_{i=1}^k a_i T(b_i x + b_i)$

$T(x) = \theta\left(\frac{x^k}{k!} \left(1 + \int_0^x \frac{g(u)}{u^{k+1}} du\right)\right)$

where  $\sum_{i=1}^k a_i b_i = 1$

Example:  $T(n) = 2T\left(\frac{n}{2}\right) + \frac{3}{4}T\left(\frac{3n}{4}\right) + n^2$   
 $\begin{cases} a_1 = 2, b_1 = \frac{1}{2}, a_2 = \frac{3}{4}, b_2 = \frac{3}{4} \\ \text{Find } p: z^1(a_1) + (z^2(a_2))^2 = p^2 = 2 \\ \text{Magic: } T(n) = \theta\left(n^2 \left(1 + \int_0^1 \frac{u^2}{u^3} du\right)\right) \\ = \theta(n^2(1 + \log n)) = \theta(n^2 \log n) \end{cases}$

Counting: How many different paths from  $(0,0,0)$  to  $(20,20,30)$ ?  
 point  $(0,0,0) \rightarrow (10,10,20) \rightarrow (20,20,30)$ , each step increments one coordinate.

Step increments are total direct P.  $\binom{20+20+30}{10,10,20}$   
 # strings,  $\binom{60}{30}$

A: Abstraction is a set of strings

## Basic Counting Notations

$f: A \rightarrow B$   
 Injective: Every element of  $B$  mapped to at most one.  $|A| \leq |B|$   
 Subjective: " "  
 Bijective: Both injective and surjective.  $|A| = |B|$ . Bijection rule

Generalized Pigeon Hole Principle:  
 If  $|X| > |Y| \rightarrow$  for any  $f: X \rightarrow Y \exists k+1$  diff elems of  $X$  mapped to the same element in  $Y$ .

Division rule:  $f: A \rightarrow B$   $k+1 \Rightarrow |A| = k \cdot |B|$   
 Product rule:  $P_1, P_2, P_3 \dots$  sets,  $|P_1|, |P_2|, \dots = |P_1| \cdot |P_2| \cdot \dots$

Sum rule:  $A_1, \dots, A_n$  disjoint sets,  $|A_1, U A_2, U \dots, U A_n| = \sum_{k=1}^n |A_k|$

Exclusion-Inclusion: Count non-disjoint sets

$\bigcup_{i=1}^n A_i = |A_1| + |A_2| + |A_3| - |A_1 \cap A_2| - |A_1 \cap A_3| - |A_2 \cap A_3| + |A_1 \cap A_2 \cap A_3|$

Binomial Theorem:  $(a+b)^n = \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k$

$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \binom{n}{n-k}$

$\binom{n}{k} = \binom{n}{n-k}$

Exclusion-Inclusion: Count non-disjoint sets

$\bigcup_{i=1}^n A_i = |A_1| + |A_2| + |A_3| - |A_1 \cap A_2| - |A_1 \cap A_3| - |A_2 \cap A_3| + |A_1 \cap A_2 \cap A_3|$

General:  $\bigcup_{i=1}^n |A_i| = \sum_{i=1}^n |A_i| - \sum_{1 \leq i < j \leq n} |A_i \cap A_j| + \sum_{1 \leq i < j < k \leq n} |A_i \cap A_j \cap A_k| - \dots$

Generating Functions

Convolution Rule:  $\# \text{ applies even} \quad \# \text{ binomials mult of 5} \quad \# \text{ at most 5 ones}$

Ex: Food Salad

Right shift:  $\{D, D, \dots, D, a_0, a_1, \dots\} \rightarrow \{y^k a(x)\}$

Derivative:  $\{1, 2, 3, 4, \dots\} \rightarrow \{1 + 2x + 3x^2 + \dots\} = (1+x+x^2+x^3+\dots)' = \left(\frac{1}{1-x}\right)' = \frac{1}{(1-x)^2}$

The generating function for choosing elements from a union of disjoint sets is the product of the generating function from choosing each set.

Ex: Food Salad

$\# \text{ sequences} = \frac{(k_1+k_2+\dots+k_m)!}{k_1! k_2! \dots k_m!}$

Combinatorial Proofs: 1) Define sets  $S$  & show  $|S| = n$  by counting exactly  $m$  cases. 2) Show  $|S| = m$  by counting another way.

Parity Identity: 1) conclude that  $m = n$

$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$

Picks: rest of cards

Ordering of three groups

Ex: # 5 card hands with exactly 3 aces

Ex: In how many ways divide 30 students into 36 groups of 5?

Total combinations:  $\binom{180}{36}$   
 Ordering of students:  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # 5 card hands with exactly 3 aces

Ex: In how many ways divide 30 students into 36 groups of 5?

Total combinations:  $\binom{180}{36}$   
 Ordering of students:  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Total combinations:  $\binom{180}{36}$   
 Ordering of students:  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Total combinations:  $\binom{180}{36}$   
 Ordering of students:  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Total combinations:  $\binom{180}{36}$   
 Ordering of students:  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Total combinations:  $\binom{180}{36}$   
 Ordering of students:  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Total combinations:  $\binom{180}{36}$   
 Ordering of students:  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Total combinations:  $\binom{180}{36}$   
 Ordering of students:  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Total combinations:  $\binom{180}{36}$   
 Ordering of students:  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Total combinations:  $\binom{180}{36}$   
 Ordering of students:  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Total combinations:  $\binom{180}{36}$   
 Ordering of students:  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Total combinations:  $\binom{180}{36}$   
 Ordering of students:  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Total combinations:  $\binom{180}{36}$   
 Ordering of students:  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Total combinations:  $\binom{180}{36}$   
 Ordering of students:  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Total combinations:  $\binom{180}{36}$   
 Ordering of students:  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Total combinations:  $\binom{180}{36}$   
 Ordering of students:  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Total combinations:  $\binom{180}{36}$   
 Ordering of students:  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Total combinations:  $\binom{180}{36}$   
 Ordering of students:  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Total combinations:  $\binom{180}{36}$   
 Ordering of students:  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Total combinations:  $\binom{180}{36}$   
 Ordering of students:  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Total combinations:  $\binom{180}{36}$   
 Ordering of students:  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Total combinations:  $\binom{180}{36}$   
 Ordering of students:  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Total combinations:  $\binom{180}{36}$   
 Ordering of students:  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Total combinations:  $\binom{180}{36}$   
 Ordering of students:  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Total combinations:  $\binom{180}{36}$   
 Ordering of students:  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Total combinations:  $\binom{180}{36}$   
 Ordering of students:  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Total combinations:  $\binom{180}{36}$   
 Ordering of students:  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Total combinations:  $\binom{180}{36}$   
 Ordering of students:  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Total combinations:  $\binom{180}{36}$   
 Ordering of students:  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Total combinations:  $\binom{180}{36}$   
 Ordering of students:  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Total combinations:  $\binom{180}{36}$   
 Ordering of students:  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Total combinations:  $\binom{180}{36}$   
 Ordering of students:  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Ex: # sequences =  $\binom{180}{36}$

Total combinations:  $\binom{180}{36}$   
 Ordering of students: <



# Number Theory

## Divisibility:

For all  $t, c$

$$\text{If } a|b \text{ and } b|c \rightarrow a|c$$

$$\text{If } a|b \text{ and } a|c \rightarrow a|sb+tc$$

For all  $c \neq 0$ ,  $a|b \Leftrightarrow ca|cb$

## Greatest Common Divisor (GCD)

$$\gcd(ka, kb) = k \gcd(a, b)$$

$$\text{If } \gcd(a, b) = 1 \text{ and } \gcd(a, c) = 1 \rightarrow \gcd(a, bc) = 1$$

$$\text{If } a|bc \text{ and } \gcd(a, b) = 1 \rightarrow a|c$$

$$\text{If } m|na \text{ and } m|b \rightarrow m|\gcd(a, b)$$

## Modular Arithmetic

$$a \equiv a \pmod{n}$$

$$a \equiv b \pmod{n} \Rightarrow b \equiv a \pmod{n}$$

$$a \equiv b \pmod{n} \text{ and } b \equiv c \pmod{n} \Rightarrow a \equiv c \pmod{n}$$

$$a \equiv b \pmod{n} \Rightarrow a+c \equiv b+c \pmod{n}$$

$$a \equiv b \pmod{n} \Rightarrow ac \equiv bc \pmod{n}$$

$$a \equiv b \pmod{n} \text{ and } c \equiv d \pmod{n} \Rightarrow a+c \equiv b+d \pmod{n}$$

$$a \equiv b \pmod{n} \text{ and } c \equiv d \pmod{n} \Rightarrow ac \equiv bd \pmod{n}$$

Multiplicative inverse of  $x \pmod{n}$  ( $x^{-1}$ )  $x \cdot y \equiv 1 \pmod{n}$

$$a \equiv b \pmod{n} \Leftrightarrow n|(a-b)$$

$$\Leftrightarrow \text{rem}(a, n) = \text{rem}(b, n)$$

## Greatest Common Divisor

Egyptian Algorithm:  $\gcd(a, b) = \gcd(b, \text{rem}(a, b))$

GCD can be written as a linear combination of  $a, b$ :  $\gcd(a, b) = sa + tb$

## Pulverizer - Find s and t

$$x \dots y, \text{rem}(x, y) = x - qy$$

$$259 \quad 70 \quad 49 = 259 - 3(70)$$

$$70 \quad 49 \quad 21 = 70 - 1(49)$$

$$= 70 - 1(259 - 3(70))$$

$$= (-1)259 + 4(70)$$

$$49 \quad 21 \quad 7 = 49 - 2(21)$$

$$= (259 - 3 \cdot 70) - 2(-1 \cdot 259 + 4 \cdot 70)$$

$$= 3 \cdot 259 - 11(70)$$

$$21 \quad 7 \quad 0 \quad 450 \quad S = 3, t = -11$$

## Multiplicative Inverse

$$k^{-1} \equiv 1 \pmod{n} \quad \text{relatively prime}$$

exists when  $\gcd(k, n) = 1$

2 ways to obtain:

### 1) Pulverizer

$$sk + tn = 1$$

$\uparrow$   
Multiplicative inverse of  $k \pmod{n}$

### 2) Euler's Theorem

$$k^{\phi(n)} \equiv 1 \pmod{n}$$

### Totient Function $\phi(n)$

# of ints  $[0, n]$  relatively prime to  $n$

For prime  $p, q$

$$\phi(p) = p-1 \quad \phi(pq) = (p-1)(q-1)$$

$$\phi(n) = n \left(1 - \frac{1}{p_1}\right) \left(1 - \frac{1}{p_2}\right) \dots$$

\* Prime Factors

$$k \cdot k^{\phi(n)-1} = 1 \pmod{n}$$

\* Multiplicative inverse

## Graph Theory

Chromatic #: min val of  $k$  for coloring problem

Basic Coloring Algorithm:

1) Order nodes by largest degree first

2) Order colors  $C_1, C_2, C_3$

3) For all nodes Assign  $V_i$  the lowest legal color  
will use at most  $\max(\text{degree})+1$  colors.

Definitions:

Odd length cycle  $\rightarrow$  not bipartite

Walk: Sequence of vertices in  $G$

Path: Walk with all different  $V_i$ 's

Cycle: Walk where  $V_0 = V_k$  and  $V_0, V_1, V_{k-1}$  are different.

Connectivity: Vertices: If there is a path from  $v$  to  $v'$

Graph: Every pair of vertices is connected

Spanning Tree: Subgraph (that is a tree) of vertices, with the same

A graph of  $n$ -nodes is a tree if:

connected, no cycles and  $n-1$  edges.

Hamiltonian cycle: Cycle that visits each node exactly once.

Euler Walk: Walk that traverses every edge once

Euler tour: Euler walk that ends at start node

(can't exists if a vertex has odd deg)

(undirected graph)

To show graph is  $k$ -colorable:

1) Show it's possible to color  $G$  w/  $k$  colors.

2) Show that  $G$  requires at least  $k$  colors.

Eg: completely connected subgraph

$n$ -sized graph takes  $n$  colors

Kashtan's Algorithm: Find Minimum Spanning Tree (MST)

- > start with an empty set
- > add the minimum edge that does not create a cycle
- > stop when there is no such edge.

Halls' Theorem:

Let  $G$  be a bipartite graph with  $(L, R)$

If for every subset  $X \subseteq L$ ,  $|X| \leq |N(X)| \geq |X|$

$\Rightarrow$  Then, there is a matching for  $L$

## Isomorphism

Map  $V_i$  of  $G_1$  to  $V_j$  of  $G_2$  and maintain edges

Tricks: All degree  $\ell$ 's maintained

All path lengths maintained

Trees: Add an edge  $\Rightarrow$  cycle

Remove edge  $e \Rightarrow$  disconnected

## Graph Theory Proofs

Induct on size (# vertices)

Induct on degree  $\rightarrow$  same deg tree  $V$

Common w/ regular graphs

Inductive step: Start with size  $n+1$  graph,

remove a node to get size  $n$  that

fits  $P(n)$ . Or else... buildup error!

## Stable Marriage Problem

Find stable matching given ranked preferences for boy and girl.

Definitions:

Perfect matching - everyone gets married

Rogue couple - when  $x$  and  $y$  prefer each other to their mates

Stable matching - no rogue couples

The Mating Algorithm (TMA)

Each day:  
Morning: Boy serenades highest ranked girl still on list.

Afternoon: Girls pick favorite among serenaders

Evenings: Boys cross girl off list if rejected.  
Stop when every girl has at most one suitor.

TMA Terminates within  $N^2+1$  days

Everyone is married at end. Stable matching

Every boy paired with optimal mate

Every girl paired with optimal mate.

Boys' suitors only get better

For girls, it only gets worse.

## RSA

Receiver: pick two primes  $p, q$

Compute  $n = pq$

Select  $e$  such that

$$\gcd(e, (p-1)(q-1)) = 1$$

compute  $d$  such that

$$de \equiv 1 \pmod{(p-1)(q-1)}$$

\* Using pulverizer:

$$d \rightarrow ae + b(p-1)(q-1) = 1$$

Public Key:  $(e, n)$  private key  $(d, n)$

Sender: encrypt using public key

$$m^e = \text{rem}(m^e, n)$$

Receiver: Decrypt using private key

$$M = \text{rem}(m^e)^d \pmod{n}$$

Fermat's Little Theorem

$$\text{For prime } p: a^p \equiv a \pmod{p}$$

$$a^{p-1} \equiv 1 \pmod{p}$$

## Well Ordering Principle

"Every non empty set of nonnegative integers has a smallest element"

Proof Template: (by contradiction)

1) Define set  $C$  of counterexamples

$$C := \{n \in \mathbb{N} \mid P(n) \text{ is False}\}$$

2) Assume  $C$  is non empty  
by WOP  $\Rightarrow$  smallest element in  $C$

3) Reach contradiction

4) Conclude that  $C$  must be empty  $\square$

## Properties of relations

Reflexivity:  $\forall x \in A, xRx$

"Everyone likes themselves"

"Every node has a loop"

Irreflexivity:  $\neg \exists x \in A, xRx$

"No one likes themselves, no loops"

Symmetry:  $\forall x, y \in A, xRy \Rightarrow yRx$

"If  $x$  likes  $y$ ,  $y$  likes  $x$ "

Antisymmetry:

$\forall x, y \in A, (xRy \wedge yRx) \Rightarrow x=y$

"No pair of distinct people can like each other"

Transitivity:  $\forall x, y, z \in A, (xRy \wedge yRz) \Rightarrow xRz$

Equivalence Relation: Reflexive, Symmetric and Transitive

Weak Partial Order: Reflexive, Antisymmetric and Transitive

Strong Partial Order: Irreflexive, Antisymmetric and Transitive

Truth Table:

x	y	$x \rightarrow y$
T	T	T
T	F	F
F	F	T
F	T	T

$$\text{contrapositive } x \rightarrow y = \bar{y} \rightarrow \bar{x}$$

## Networks Definitions

Distance b/w  $v-v$ : shortest path from  $v$  to  $v$   
Diameter of network: distance b/w input/output apart  
Congestion: of paths - max (of all nodes) paths of Routing Problem minimize congestion through  $v$ .  
longest of best set of paths.

(N input)	# switches	Diameter	Congestion
2D Grid	$N^2$	$2N$	2
Complete binary Tree	$2^{N-1}$	$2\log(N+1)$	$N$
Butterfly	$N(\log(N))$	$\log(N+2)$	$\sqrt{N}$
Benes Network	$2N(\log N)$	$2\log(N+2)$	1

## DAG's (Directed Acyclic Graphs)

Directed graphs with no cycles

Topological Sort: (Think putting on clothes example)  
every finite poset has a topological sort

Chains: Ordered set of vertices where, for a vertex the vertex to its left is a prerequisite.

Antichains: Set of vertices where none of the vertices are prerequisites of each other.

Dilworth's Lemma: Every DAG with  $n$  vertices needs either a chain of size  $\lceil n/2 \rceil$  or an antichain of at least  $n/2$

Any planar graph can be colored in at most 6 colors.  
 $\hookrightarrow$  none of the lines cross

Give a description of the equivalence class.

a) Integers  $x$  and  $y$  are equivalent if

$$x \equiv y \pmod{3}$$

$$A: \left\{ \dots, -6, -3, 0, 3, 6, \dots \right\}$$

$$\left\{ \dots, -5, -2, 1, 4, 7, \dots \right\}$$

$$\left\{ \dots, -4, -1, 2, 5, 8, \dots \right\}$$

PageRank • Network with  $n$  nodes!  
Initially every page =  $\frac{1}{n}$  Page Rank  
Every update, each page distributes its PR equally along outgoing edges. And sets its new PR to sum of received shares.

Giving PR  $x$  can receive at most  $1-x$

Sample Problem:



Receive PR from vertices that point to it.

$$\begin{bmatrix} A' \\ B' \\ C' \\ D' \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} A \\ B \\ C \\ D \end{bmatrix}$$

A B C D  
↓ ↓ add up to 1.

$$P_A = \frac{1}{2} P_0$$

$$P_B = \frac{1}{2} P_A + \frac{1}{2} P_C + \frac{1}{2} P_D$$

$$P_C = \frac{1}{2} P_A + P_B$$

$$P_D = \frac{1}{2} P_C$$

$$P_A + P_B + P_C + P_D = 1$$

Hasse Diagram - Directed acyclic graph. Represent a poset.

- Remove edges implied by transitivity

- Every finite poset has a topological sort.

Example Problem: Find remainder of  $38^{82248}$  divided by 83

1) 38 and 83 are relatively prime  
 $\Rightarrow$  Use Euler's Theorem:

$$\phi(83) \equiv 1 \pmod{83}$$

$\phi$  b/c prime

$$\phi(83) = 82$$

2) Try to remove as many 82's from power as possible (since they go to 1  $\pmod{83}$ )

$$38^{82248} = 38^2 \cdot 38^{82 \cdot 1003}$$

$$\equiv 38^2 \cdot 1^{1003} \pmod{83} = 144$$

$$\equiv 33 \pmod{83} \quad \therefore \boxed{\text{Solution: 33}}$$

Any simple graph with  $n$  nodes and strictly more than  $\frac{1}{2}(n-1)(n-2)$  edges is connected.

Example Problem:  $\text{rem}(96^{123456789}, 97)$

$$96 \equiv -1 \pmod{97} \Rightarrow 96^{\text{odd power}} \equiv -1 \pmod{97}$$

$\Rightarrow \boxed{96}$

(20)

48/100 ✓

## Problem 1

state prop.

$$w^2 + x^2 + y^2 = z^2$$

w x y  
Even, odd, odd

Because of the commutative property of addition this is equivalent to

$$(2i)^2 + (2j+1)^2 + (2k+1)^2 = z^2$$

w x y  
odd even odd and  
odd odd even

$$4i^2 + 4j^2 + 4j + 1 + 4k^2 + 4k + 1 = z^2$$

$$4i^2 + 4j^2 + 4j + 4k^2 + 4k + 2 = z^2$$

$$\underbrace{4(i^2 + j^2 + j + k^2 + k) + 2}_{P_1} = z^2$$

in order for  $z$  to be even,  $P_1$  must have an even square root.  
Even square root examples.

$$\begin{aligned}\sqrt{4} &= 2 \\ \sqrt{16} &= 4 \\ \sqrt{36} &= 6 \\ \sqrt{64} &= 8 \\ \sqrt{100} &= 10\end{aligned}$$

$$4, 16, 36, 64, 100, \dots$$

$(2i)^2$  for any integer  $i$  is a multiple of 4 because  
 $(2i)^2 = 4i^2$

So while  $P_1$  may be even, it is not a multiple of 4  
so  $\sqrt{P_1}$  will not be even.

x y z  
Even, Even, odd

$$(2i)^2 + (2j)^2 + (2k+1)^2$$

$$= 4i^2 + 4j^2 + 4k^2 + 4k + 1$$

$P_2$  is not even, or a multiple of 4 so  
 $\sqrt{P_2} = z$  will not be even

$$\underbrace{4(i^2 + j^2 + k^2 + k) + 1}_{P_2}$$

x y z  
Even, Even, Even

$$(2i)^2 + (2j)^2 + (2k)^2 = z^2$$

$$\underbrace{4i^2 + 2j^2 + 2k^2 - z^2}_{4(i^2 + j^2 + k^2) = z^2} \rightarrow P_3$$

$P_3$  is a multiple of 4 and  $4(i^2 + j^2 + k^2) = \frac{z^2}{2^2}$   
 $= (2n)^2 = 4(n^2)$

Fernando Trujano

## 6.042 Pset #1

Ria Brando M  
collab: Erika Lu  
Kate Fairis

Problem 2

14/14

1)

x	y	z	$x \rightarrow y$	$y \rightarrow z$	$z \rightarrow x$	$(x \rightarrow y) \wedge (y \rightarrow z) \wedge (z \rightarrow x)$
T	T	T	T	T	T	T
T	T	F	T	F	T	F
T	F	T	F	T	T	F
T	F	F	F	T	T	F
F	T	T	T	T	F	F
F	F	T	T	T	F	F
F	T	F	T	F	T	T
F	F	F	T	T	T	T

2)

x	y	z	$\bar{x}$	$\bar{y}$	$\bar{z}$	$\bar{x} \wedge \bar{y} \wedge \bar{z}$
T	T	T	F	F	F	F
T	T	F	F	F	T	F
T	F	T	F	T	F	F
T	F	F	F	T	T	F
F	T	T	T	F	F	F
F	F	T	T	T	F	F
F	T	F	T	F	T	F
F	F	F	T	T	T	T

3)

x	y	z	$\bar{x}$	$y \wedge z$	$\bar{y} \wedge \bar{z}$	$(\bar{x} \vee (y \wedge z))$	$x \vee (\bar{y} \wedge \bar{z})$	$(\bar{x} \vee (y \wedge z)) \wedge x \vee$
T	T	T	F	T	F	T	T	T
T	T	F	F	F	T	F	F	F
T	F	T	F	F	F	F	T	F
T	F	F	F	F	T	F	F	F
F	T	T	T	F	F	T	F	F
F	F	T	T	F	F	T	F	F
F	T	F	T	F	F	T	F	F
F	F	F	T	F	T	T	T	T

#2 is not like the others because the combination x,y,z = True = F

(13)

## Problem 3

$$a) ((p \rightarrow q) \wedge (q \rightarrow r) \wedge (r \rightarrow p)) \rightarrow p \wedge q \wedge r + 5$$

b)

<u>p</u>	<u>q</u>	<u>r</u>	<u><math>p \rightarrow q</math></u>	<u><math>q \rightarrow r</math></u>	<u><math>r \rightarrow p</math></u>	<u><math>(p \rightarrow q) \wedge (q \rightarrow r) \wedge (r \rightarrow p)</math></u>	<u><math>p \wedge q \wedge r</math></u>
T	T	T	T	T	T	T	T
T	T	F	T	F	T	F	F
T	F	T	F	T	T	F	F
T	F	F	F	T	T	F	F
F	T	T	T	T	F	F	F
F	F	T	T	F	F	F	F
F	T	F	T	T	T	F	F
F	F	F	T	T	T	T	F

inconsistent statement

4

c) No, ~ ... ... ... ~ ...

when  $p \wedge q \wedge r$  is False ' $(p \rightarrow q) \wedge (q \rightarrow r) \wedge (r \rightarrow p)$ ' is True.

"if part" is True and "then-part" is False

∴ The proposition is false.

Fernando Trujano

6.042 Pset #1

R19 Brando II.  
Collaboration: Erika Lu,  
Kate Farris

s.t. =

$$\frac{+19}{24}$$

Problem 4

a)

$$\forall n \exists a, b, c, d \in \mathbb{N} : n = a^2 + b^2 + c^2 + d^2$$

Domain is  $\mathbb{N}$  containing 0

$$\checkmark + \frac{4}{4}$$

b)

$$\forall n > 1 \exists a, b \in \text{prime}(a) \wedge \text{prime}(b) \text{ and } 2n = a + b$$

$$\checkmark + \frac{4}{4}$$

Domain is  $\mathbb{Z}^+$  ~~not correct translation~~ ~~+2~~ ~~+4~~

c)  $\forall x, y \exists \delta > 0. |x - y| \leq \delta$

needs to apply to f  $\checkmark + \frac{4}{4}$

d)  $\forall n > 0 \exists (x, y, z \in \mathbb{Z}^+ : x^n + y^n = z^n)$  Domain is  $\mathbb{Z}^+$

need  $n > 2$   $\frac{+3}{4}$

e)

$$\forall n \exists p \text{ prime}(p) \wedge n < p < 2n$$

Domain is  $\mathbb{Z}^+$

$$\forall n \exists m \text{ prime}(m) \wedge m > n \quad \checkmark + \frac{4}{4}$$

f)  $\forall n > 1 \exists p \text{ prime}(p) \wedge n < p < 2n$

$$\checkmark + \frac{4}{4}$$

Domain is  $\mathbb{Z}^+$



## Problem 5

Lemma 1: The four chosen stones  $w, x, y, z$  must add up to at least 39

$$w + x + y + z \geq 39$$

This is the only way to find an unknown weight of 39 or 40.

Since the max unknown weight is 40, a total of 39 would suffice.  
If it is heavier than 39, it must be 40.

Lemma 2: We can add or subtract stone values to find the unknown weight.

By placing a stone on either side of the scale it can either add or subtract from the total.

I selected ①, ③, ⑨, ⑯. These satisfies Lemma 1 because  $1+3+9+26 \geq 39$

By Lemma 2, I can use combinations of these stones to create any integer from 1 to 29.

$1 = 1$	$20 = 26 - 9 + 3$
$2 = 3 - 1$	$21 = 26 - 9 + 3 + 1$
$3 = 3$	$22 = 26 - 3 - 1$
$4 = 3 + 1$	$23 = 26 - 3$
$5 = 9 - 3 - 1$	$24 = 26 - 3 + 1$
$6 = 9 - 3$	$25 = 26 - 1$
$7 = 9 - 3 - 1$	$26 = 26$
$8 = 9 - 1$	$27 = 26 + 1$
$9 = 9$	$28 = 26 + 3 - 1$
$10 = 9 + 1$	$29 = 26 + 3$
$11 = 9 + 1 - 3$	$30 = 26 + 3 + 1$
$12 = 9 + 3$	$31 = 26 + 9 - 3 - 1$
$13 = 9 + 3 + 1$	$32 = 26 + 9 - 3$
$14 = 26 - 9 - 3$	$33 = 26 + 9 - 3 + 1$
$15 = 26 - 9 - 3 + 1$	$34 = 26 + 9 - 1$
$16 = 26 - 9 + 1$	$35 = 26 + 9$
$17 = 26 - 9$	$36 = 26 + 9 + 1$
$18 = 26 - 9 + 1$	$37 = 26 + 9 + 3 - 1$
$19 = 26 - 9 + 3 - 1$	$38 = 26 + 9 + 3$
	$39 = 26 + 9 + 3 + 1$

## Problem 6



(2)

In a six person group,

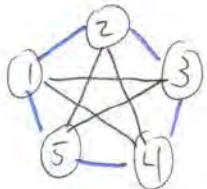
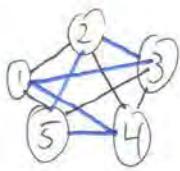
Lemma 1. There exists at least three other people that person 1 shook hands with or did not shake hands with.  $P \vee Q$

Since there are only six people this condition will always hold regardless of the # of handshakes. Why?

Initially, no one has shaken hands so  $P$  is true for any person  $X$  in a 6 person group.  $P$  holds until person  $X$  has shaken 3 hands, at this point  $Q$  becomes true and Lemma 1 still holds.

This (Because a triangle is defined as 3 people that shake or did not shake hands. By Lemma 1, Every group of 6 will have a triangle you're trying to prove. — 6)  $\square$

b)



m Shaken hands

m Not shaken

No, Lemma 1 does not hold. See pictures above for contradictions

 $\square$

Problem 1

(6/6) JK

89/100

Thm:  $a^b$  where  $a, b$  are irrational can be rational

Def: Irrational number  $a \neq c/d$  where  $c, d$  are integers

$$\begin{aligned} a &= \sqrt{3} && \leftarrow \text{irrational} \\ b &= \sqrt{2} \end{aligned}$$

Case 1  $a^b = (\sqrt{3})^{\sqrt{2}}$  is rational.  $\square$

Case 2

$$\begin{aligned} a^b &= (\sqrt{3})^{\sqrt{2}} \text{ is irrational} \\ b &= \sqrt{2} \leftarrow \text{also irrational} \end{aligned}$$

$$(a^b)^b = ((\sqrt{3})^{\sqrt{2}})^{\sqrt{2}} = (\sqrt{3})^2 = 3 \text{ is rational} \quad \square$$

18/18 HL

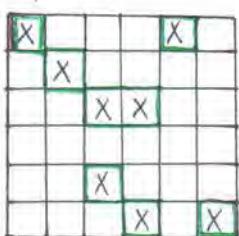
## Problem 2

Thm: IF Fewer than  $n$  students in class are initially infected, the whole class will never be completely infected.

Proof:

Def: Infected perimeter: The perimeter of all infected squares on the grid (including edges).

For example:



The infected perimeter (marked in green) of this grid is 30

Lemma 1. (invariant) After a time step, the total infected perimeter stays the same or decreases.

Proof:

Since the edges of adjacent infected cells don't count for the infected perimeter total, the <sup>total</sup> perimeter cannot increase and thus must stay the same or decrease.

Proof of Theorem

A Fully infected class would cover the whole  $nxn$  grid  $\Rightarrow$  infected perimeter =  $4n$

The initial infected perimeter can be, at most  $4i$ , where  $i$  is the number of initially infected students.

↑ This is the case where all infected cells are not adjacent to each other and thus contribute 4 sides to the total infected perimeter.

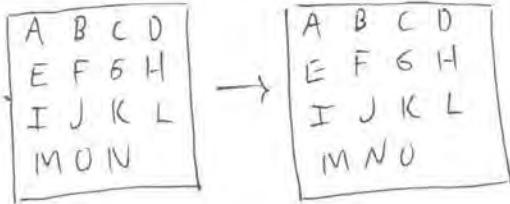
By Lemma 1, the total infected perimeter stays the same or decreases after each time step  $\Rightarrow$  in order to infect the whole class ( $4n$  perimeter) the initial number of <sup>initial</sup> infected students must be at least  $n \Rightarrow$  IF Fewer than  $n$  students in class are initially infected, the <sup>infected</sup> perimeter will never be equal to  $4n$  and the whole class will never be completely infected.

□

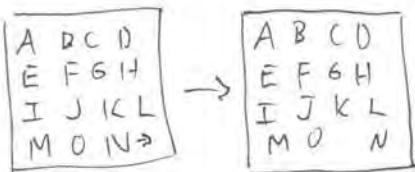
Fernando Trujano

+20/20

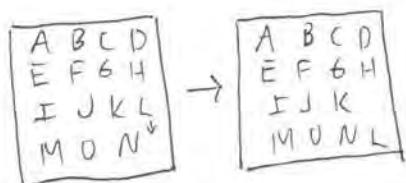
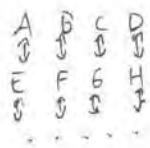
## Problem 3

Problem: Find a sequence of legal moves to goTheorem: The puzzle is not solvableDef: Order. Sequence of tiles on board reading from top row to bottom row and from left to right within a row

a)

Row move:Fact: A row move effectively swaps the empty cell with an adjacent (left or right) and does not change the order of the tiles.

b)

Column move:Lemma 1: A column move changes the relative order of precisely 3 pairs of tiles.Proof: In a column move, we move a letter from cell  $i$  to a blank spot in cell  $i+4$  or  $i-4$ .

When a letter moves 4 positions  
it changes relative order with  
three tiles in between  $\square$

Def:  $L_1 \& L_2$  are inverted if  $L_1$  precedes  $L_2$  in the alphabet but  $L_1$  appears after  $L_2$  in the puzzle.

c)

Fact 2: A row move does not change the parity of the number of inversions.

↳ Since a row move does not change the order of the letters, the parity of the number of inversions will remain the same.

d)

Lemma 2: A column move always changes the parity of the number of inversions.

Proof

By Lemma 1, a column move changes the relative order of 3 cells.

Case 1: There was an inverted pair: The change uninverts the pair

Case 2: There was an uninverted pair: The change inverts the pair

→ The parity of the number of inversions will always change with a col move  $\square$

e)

Lemma 3: (invariant)

In every configuration reachable from the position shown below, the parity of the # of inversions is different from the parity of the row containing the blank square.

row1	A	B	C	D
row2	E	F	G	H
row3	I	J	K	L
row4	M	O	N	

Proof: (by induction)

①  $P(n)$ : The parity of the # of inversions is different from the parity of the row containing the blank square after  $n$  moves.

② Base case

$P(0)$ :	A	B	C	D
	E	F	G	H
	I	J	K	L
	M	O	N	

# of inversions: 1  
row of blank square: 4  $\rightarrow$  different parity ✓

③ Inductive step:

Assume  $P(n)$  is true for the purposes of induction to show that  
 $P(n) \rightarrow P(n+1)$

Case 1: Row move (for move  $n+1$ )

By  $P(n)$ , the parity of the # of inversions is different from the row containing the blank square.  
 By Fact 2, the parity of the # of inversions stays the same.

A row move cannot change the row of the blank square → stays the same  
 $\rightarrow$  Both parities remain the same ✓

### Case 2 : column move (For move $n+1$ )

By  $P(n)$ , the parity of the # of inversions is different from the row containing the blank square.

By Lemma 2, the parity of the # of inversions will change with a column move.

Fact  
A column move changes the row of the blank square by one  $\rightarrow$  switching parity.  
 $\Rightarrow$  Both parities change  $\Rightarrow P(n+1)$  holds

$$\begin{aligned}\Rightarrow P(n) &\Rightarrow P(n+1) \\ \Rightarrow P(n) &\neq n > 0\end{aligned}$$

F)

### Proof of Theorem

The desired end position has an even number of inversions (0) with the blank square on an even row (4).

By Lemma 3, it is not reachable.





## Problem 4

a)

Thm: Every nonempty set of nonnegative integers has a smallest element. (WOP)

Proof (By contradiction and strong induction)

Assume For the purposes of contradiction that there exists a non-empty set of nonnegative integers that does not have a smallest element.

$P(n)$ :  $n$  cannot be in the set because it would be the smallest element

Base case:

$P(0)$ : cannot be in set because it would be the smallest one  
↳ smallest non negative integer

Inductive step:

Assume  $P(0), \dots, P(n)$  are not in the set for the purpose of induction. Then  $P(n+1)$  cannot be in the set, otherwise  $P(n+1)$  would be the smallest element.

$P(0) \dots P(n) \Rightarrow P(n+1) \Rightarrow P(n)$  cannot be in set  $\Rightarrow$  The set is empty  
→ contradiction

□

b) Thm:  $\sum_{i=0}^n i^3 = \left(\frac{n(n+1)}{2}\right)^2$

Proof: (by contradiction and WOP)

$C ::= \{n \in \mathbb{N} \mid \sum_{i=0}^n i^3 \neq \left(\frac{n(n+1)}{2}\right)^2\}$   
counterexamples

By WOP there exists  $c$ , a minimum element in  $C$ .

Since  $c$  is the smallest counterexample then the Theorem is False for  $n=c$   
but holds for all non negative integers  $n < c$ .

$P(0)$  is True so  $c > 0 \Rightarrow c-1$  is non negative  $\rightarrow P(c-1)$  is True

$\Rightarrow$  Contradiction

$$\Rightarrow \sum_{i=0}^n i^3 = \left(\frac{n(n+1)}{2}\right)^2$$

□

## Problem 5

Ihm:  $27a^4 + 9b^4 + 3c^4 = d^4$  has no positive integer solutions

Proof: By contradiction

Lemma 1:  $a, b, c$  and  $d$  must be divisible by 3, that is  
 $3 | a, b, c, d$

Proof:

Every coefficient in  $27a^4 + 9b^4 + 3c^4$  is divisible by 3  $\Rightarrow d^4$  is also divisible by 3  $\Rightarrow 3 | d$

Assume there exists a solution for the purposes of contradiction.

$\Rightarrow$  There must be a solution with a min value for  $a$ , however by Lemma one  $3 | a, b, c, d$  so another solution could be found by dividing  $a, b, c, d$  by 3.  $\rightarrow$  Contradiction  $\rightarrow 27a^4 + 9b^4 + 3c^4 = d^4$  has no positive integer solutions.

□

(48) On You only proved

$3 | d$

## Problem 6

- a) Thm 1. IF the Nim sum is 0, any move will result in a Nim sum that is not 0.

Fact: Nim sum is defined to be the binary xor of the number of stones in each pile.

Refer to the following table:

A	B	C	A xor B xor C
1	1	1	1
1	1	0	0
1	0	1	0
1	0	0	1
0	1	1	0
0	1	0	1
0	0	1	1
0	0	0	0

We can see that the nim sum = 1 if the parity of 1's in  $A \Delta B \Delta C$  is odd.

Similarly, if the nim sum = 0 the parity of 1's in  $A \Delta B \Delta C$  must be even.

IF 00000, parity of 1's is even and any legal move would change this  $\Rightarrow$  new nim sum will not be 0.

- b) Thm 2: IF the nim sum  $\neq 0$ , there is always a way to make the Nim sum 0 with one move.

Proof: There will always be a move to change the parity of 1's to be even  $\Rightarrow$  nim sum will be 0.

- c) Thm: IF the game begins with a non-zero Nim sum, then the first player has a winning strategy.

Proof: By Thm 1, the first player can make a move that makes the nim sum 0,  
 $\Rightarrow$  the other person cannot take the last piece  $\Rightarrow$  must make a move.  
 By Thm 2 any move that the second player does will result in a nim sum  $\neq 0$ .

This cycle is repeated until the first player wins.  $\square$

## Problem 7

8/8  
IB

Thm: A group of  $n \geq 1$  people can be divided into teams each containing either 4 or 7 people with all possible values of  $n$  being 4, 7, 8, 11, 12, 14, 15, 16 and  $\geq 18$

ProofLemma 1

By using a combination of 4's and 7's, we can create teams of 4, 7, 8, 11, 12, 14, 15, and 16

Proof.

This is done with simple arithmetic, each case is calculated below.

$$\begin{array}{ll} 4 = 4 & 14 = 7 + 7 \\ 7 = 7 & 15 = 4 + 4 + 7 \\ 8 = 4 + 4 & 16 = 4 + 4 + 4 + 4 \\ 11 = 4 + 7 & \\ 12 = 4 + 4 + 4 & \end{array}$$

Lemma 2: Every group  $n \geq 18$  can be divided into teams each containing either 4 or 7 people.

Proof: (by strong induction)

①  $P(n)$ : A group of  $n \geq 18$  people can be divided into teams each containing either 4 or 7 people

② Base case

$$P(18): 18 = 7 + 7 + 4$$

$$P(20) \text{ is } 4 + 4 + 4 + 4$$

$$P(19): 19 = 4 + 4 + 4 + 7$$

$$P(21) \text{ is } 7 + 7 + 7$$

③ Inductive step.

Assume  $P(18), P(19) \dots, P(n)$  are true to prove  $P(n+1)$  is True.

$$n+1 = 4 + (n-3) \Rightarrow (n+1)-4 = (n-3)$$

$\Rightarrow$  If a team of 4 is removed from the set of  $n+1$  people, we are left with  $n-3 \geq 18$  people.

By  $P(18), P(19) \dots, P(n)$  this group of  $n-3$  people can be divided into teams each containing either 4 or 7 people  $\therefore P(n) \Rightarrow P(n+1)$



## Problem 8

8/8 ✓

Thm.: The pirates can safely attack the Piñata

Proof (by strong induction)

$$\textcircled{1} \quad P(n) : n = 5x + 7y + 9z \quad x, y, z \in \mathbb{Z}^+$$

\textcircled{2} Multiple base cases:

$$P(20) : \begin{matrix} x=4 \\ y=0 \\ z=0 \end{matrix} \Rightarrow 5(4) + 7(0) + 9(0) = 20 \quad \checkmark$$

$$P(21) : \begin{matrix} x=1 \\ y=1 \\ z=1 \end{matrix} \Rightarrow 5(1) + 7(1) + 9(1) = 21 \quad \checkmark$$

$$P(22) : \begin{matrix} x=3 \\ y=1 \\ z=0 \end{matrix} \Rightarrow 5(3) + 7(1) + 9(0) = 22 \quad \checkmark$$

$$P(23) : \begin{matrix} x=0 \\ y=2 \\ z=1 \end{matrix} \Rightarrow 5(0) + 7(2) + 9(1) = 23 \quad \checkmark$$

$$P(24) : \begin{matrix} x=2 \\ y=2 \\ z=0 \end{matrix} \Rightarrow 5(2) + 7(2) + 9(0) = 24 \quad \checkmark$$

$$P(25) : \begin{matrix} x=5 \\ y=0 \\ z=0 \end{matrix} \Rightarrow 5(5) + 7(0) + 9(0) = 25 \quad \checkmark$$

\textcircled{3} Inductive step

For purposes of induction, assume  $P(20), \dots, P(n-1)$  are true when  $n \geq 25$ .  $\Rightarrow$  By  $P(20) \dots P(n-1)$ ,  $P(n-5)$  is True.

This means we can successfully divide  $n-5$  pieces.

We can take  $n$  pieces and apply  $\nearrow \Rightarrow P(n)$

$$\Rightarrow P(n) \quad \forall n \geq 20$$

□

## Problem 1

(S) 16  
B

-12

a)  $\gcd(139, 61) = 139s + 61t$

~~75/92~~

$$\begin{array}{ccccccc}
 x & y & \text{rem}(x, y) & = & x - q \cdot y \\
 \hline
 139 & 61 & 17 & = & 139 - 2 \cdot 61 \\
 61 & 17 & 10 & = & 61 - 3 \cdot 17 \\
 & & & = & 61 - 3 \cdot (139 - 2 \cdot 61) \\
 & & & = & 61 - 3(139) + 6(61) \\
 & & & = & 61 - 3(139) + 7(61) \\
 17 & 10 & 7 & = & 17 - 1 \cdot 10 \\
 & & & = & 17 - (-3(139) + 7(61)) \\
 & & & = & (139 - 2 \cdot 61) + 3(139) - 7(61) \\
 10 & 7 & 3 & = & 10 - 7(1) \\
 & & & = & -3(139) + 7(61) - (139 - 2 \cdot 61) + 3(139) - 7(61) \\
 & & & = & -3(139) + 7(61) - (4(139) - 9(61)) \\
 & & & = & -7(139) + 16(61) \\
 7 & 3 & 1 & = & 7 - 2 \cdot 3 \\
 & & & = & 4(139) - 9(61) - 2(-7(139) + 16(61)) \\
 & & & = & 4(139) - 9(61) + 14(139) - 32(61) \\
 & & & = & 18(139) - 41(61) \\
 3 & 1 & 0 & = & 3 - 3(1)
 \end{array}$$

$$\boxed{s = 18, t = -41}$$

b)  $n = 139$   
 $x = 61$   
 $y = ??$

$$61y \equiv 1 \pmod{139}$$

$$\boxed{y = -41} \quad \text{wrong}$$

$$\begin{array}{c|cc}
 & 139 & 61y - 1 \\
 \hline
 & 139 & 2502
 \end{array}$$

c) Inverse of  $19 \pmod{37} = y$

Euler's Theorem

$$k^{\phi(n)} \equiv 1 \pmod{n}$$

$\Rightarrow k^{\phi(n)-1}$  is inverse of  $(k \pmod{n})$

$$k=19, n=37$$

$$19^{\phi(37)-1} = y$$

$$\phi(37) = \\ \text{prime}$$

$$\phi(37) = 37-1 = 36$$

$$19^{35} = y$$

Compute

$$\boxed{19^{35}}$$

d)  $\text{rem}(38^{82248}, 83)$

$$\phi(83) = 82 \\ \text{prime}$$

$$38^{8248} \equiv 1 \pmod{83} \Rightarrow 38^{82000} \equiv 1 \pmod{83}$$

$$3^{248} \equiv 1 \pmod{83}$$

$$\Rightarrow \underbrace{3^{248} \pmod{83}}_9$$

~~9~~

5  
—  
16

JH

## Problem 2

a) IF  $a|b$ , then  $\nexists c$ .  $a|bc$ Assume  $a|b \Rightarrow$  there exists an integer  $k_1$  s.t.  $ak_1 = b$ Since  $c$  is an integer  $a(k_1 \cdot c) = bc \Rightarrow ak_2 = bc$  still an int  $\Rightarrow a|bc$ 

□

b) IF  $a|b$  and  $a|c$ , then  $a|sb+tc$ 

$$\begin{aligned} ak_1 &= b \\ ak_2 &= c \quad \nexists s, t \quad sb+tc = sak_1 + tak_2 \\ &\quad = a(sk_1 + tk_2) \\ &\quad \text{both integers} \Rightarrow a|sb+tc \end{aligned}$$

c)  $a|b \Leftrightarrow ca|cb \quad \nexists c$ 

$$\begin{aligned} \Rightarrow a|b && \Leftarrow ca|cb \\ \Rightarrow ak_1 &= b & \Rightarrow ak_1 = cb \\ \text{ca}k_1 &= c \cdot b \Rightarrow ca|cb & ak_1 = b \Rightarrow a|b \end{aligned}$$

□

d)  $\gcd(ka, kb) = k \gcd(a, b)$

$\downarrow$   
 $= s(ka) + t(kb)$

$= k(sa + tb)$  By Theorem 4.2.1 in Text

$= k \gcd(a, b)$

□

 $\Rightarrow$  show  $\gcd$ 

(1)

## Problem 3

a) Thm:  $x^2 \equiv y^2 \pmod{p} \iff x \equiv y \pmod{p} \vee x \equiv -y \pmod{p}$

Proof: By algebra

$\Rightarrow$

$$x^2 \equiv y^2 \pmod{p}$$

$$x^2 \pmod{p} = y^2 \pmod{p}$$

$$x^2 \pmod{p} - y^2 \pmod{p} = 0$$

$$(x^2 - y^2) \pmod{p} = 0$$

$$x^2 - y^2 \equiv 0 \pmod{p}$$

$$(x+4)(x-y) \equiv 0 \pmod{p}$$

$$(x+y) \not\equiv 0 \pmod{p}$$

$$x \equiv -y \pmod{p} \checkmark$$

$6|3^4$ , but  $6|3^3, 6|3^2$

②

$$(x-y) \equiv 0 \pmod{p}$$

$$x \equiv y \pmod{p} \checkmark$$

$\Leftarrow$

$$(x+y) \equiv 0 \pmod{p}$$

$$(x-y) \equiv 0 \pmod{p}$$

$$(x+y)(x-y) \equiv 0 \pmod{p}$$

$$x^2 - y^2 \equiv 0 \pmod{p}$$

$$x^2 \equiv y^2 \pmod{p} \checkmark$$



b) Thm

$$n \equiv x^2 \pmod{p} \Rightarrow n^{\frac{p-1}{2}} \equiv +1 \pmod{p}$$

Fermat's Little Theorem:

$$k^{p-1} \equiv 1 \pmod{p}$$

$$\begin{aligned} n^{\frac{p-1}{2}} &\equiv (x^2)^{\frac{p-1}{2}} \pmod{p} \\ &\equiv x^{p-1} \pmod{p} \stackrel{\text{By FLT}}{\equiv} 1 \pmod{p} \end{aligned}$$

□

c)

$$p \equiv 3 \pmod{4}$$

$$p = 4k + 3$$

$n \equiv x^2 \pmod{p} \rightarrow$  Square modulo

$$n^{\frac{p-1}{2}} \equiv 1 \pmod{p}$$

$$n^{\frac{(4k+3)-1}{2}} \equiv 1 \pmod{p}$$

$$n^{2k+1} \equiv 1 \pmod{p}$$

$$n \cdot n^{2k+1} \equiv n \pmod{p}$$

$$n^{2k+2} \equiv \underbrace{n \pmod{p}}_{\times^2 \pmod{p}} \quad \frac{p-3}{4} = k$$

$$(n^{k+1})^2 \equiv x^2 \pmod{p}$$

$\Rightarrow$

$$x \equiv n^{k+1} \pmod{p} \Rightarrow x = ap + n^{k+1}$$

$$\Rightarrow x = n^{k+1} = \underbrace{n^{\frac{p-3}{4} + \frac{1}{4}}}_{= n^{\frac{p+1}{4}}}$$

Use  
more  
words!  
→

10  
10

## Problem 4

Thm:For any prime,  $p$  and integer  $k \geq 1$ 

$$\phi(p^k) = p^k - p^{k-1}$$

Proof:

There are  $\frac{p^k}{p} = p^{k-1}$  #'s between 0 and  $\overbrace{p^{k-1}}$  that are divisible by  $p$ .  $\checkmark$   $\Rightarrow$  none of these

are relatively prime to  $p^k \Rightarrow$ 

$$\phi(p^k) = p^k - \cancel{p^{k-1}}$$

## Problem 5

2 distinct primes, hidden



a) Decoding

$$n = \overbrace{pq}$$

$$s = \text{rem}((m)^d, n) \equiv m^d \pmod{n}$$

Private key:  $(d, n)$ Public key:  $(e, n)$ 

$$s^e \equiv m^{de} \pmod{n}$$

$$de \equiv 1 \pmod{(p-1)(q-1)}$$

$$de = 1 + r \stackrel{\text{unknown}}{\pmod{(p-1)(q-1)}}$$

$$\Rightarrow s^e \equiv m \cdot m^{r(p-1)(q-1)} \pmod{n}$$

$$s^e \equiv m \cdot 1^r \pmod{n}$$

$$\text{b/c } \phi(n) = (p-1)(q-1)$$

(Corollary 4.7.5)

and Euler's Theorem:

$$k^{\phi(n)} \equiv 1 \pmod{n}$$

$$\Rightarrow s^e = m \pmod{n}$$

$$b) m^* = 1535$$

Public key:  $(7, 7613) \rightarrow [e, n]$ 

$$m \equiv s^e \pmod{n} \equiv 1535^7 \pmod{7613}$$

$$= 4229 \\ \begin{array}{cccc|c} & 6 & 6 & 1 & \\ & 6 & 0 & 0 & k \\ \hline & & & & \end{array} \\ \boxed{\text{book}}$$

Problem 6

(10/10 AC)

Thm:  $L_n$  and  $L_{n+1}$  in the Lucas Series are relatively prime.Def: Lucas Series

$$L_1 = 2$$

$$L_2 = 1$$

$$L_n = L_{n-1} + L_{n-2} \text{ for } n \geq 3$$

Ex: 2, 1, 3, 4, 7, 11, 18

Def: Relatively Prime (a, b)a and b relatively prime if  $\gcd(a, b) = 1$ Proof (By induction) $P(n)$ :  $L_n$  and  $L_{n+1}$  are relatively prime

Base Case:  $L_1 = 2$      $\gcd(2, 1) = 1 \Rightarrow$  relatively prime ✓  
 $L_2 = 1$      $\gcd(1, 3) = 1 \Rightarrow$  relatively prime ✓

$$L_3 = 3 \quad \gcd(1, 3) = 1 \Rightarrow$$
 relatively prime ✓

Inductive Step: Show that  $\forall n \geq 0 \quad P(n) \rightarrow P(n+1)$  Assume  $P(n)$  for induction.  
 To show that  $F_{n+1}$  and  $F_{n+2}$  are relatively prime.

Suppose  $F_{n+1}$  and  $F_{n+2}$  are not relatively prime. $\Rightarrow$  They have a common divisor  $d > 1$ 

$$d \mid \frac{F_{n+2} - F_{n+1}}{F_n} \text{ by } P(n)$$

 $\Rightarrow d > 1$  divides both  $F_n$  and  $F_{n+1}$  $\Rightarrow F_n, F_{n+1}$  not relatively prime.  
(contradiction) $\Rightarrow F_{n+1}$  and  $F_{n+2}$  are relatively prime  
 $P(n) \rightarrow P(n+1)$ 

□

(10/10) JK

Ihma) IF  $sm + tn = 1$ , for integers  $a, b$ 

$$x = \text{rem}(bsm + atan, mn)$$

is the unique solution to

$$x \equiv a \pmod{m}$$

$$x \equiv b \pmod{n}$$

(i) It is a solution

$$x = \text{rem}(bsm + atan, mn)$$

$$\Rightarrow x \equiv bsm + atan \pmod{mn} \quad x \pmod{m}$$

$$x = bsm + atan + mnk \quad \text{where } k \text{ is an integer}$$

$$\begin{array}{c} x \pmod{n} \\ x \equiv bsm \pmod{n} \end{array}$$

$$sm = 1 - tn$$

$$x \equiv b(1 - tn) \pmod{n}$$

$$\equiv (b - bta) \pmod{n}$$

$$x \equiv b \pmod{n}$$

$$x \equiv atan \pmod{m}$$

$$tn = 1 - sm$$

$$x \equiv a(1 - sm) \pmod{m}$$

$$x \equiv a - asm \pmod{m}$$

$$x \equiv a \pmod{m}$$

x solves both equations

ii) Show that this solution is unique: Ihm: The solution is uniqueProof (By Contradiction)Assume  $\exists a, y \mid y \neq x$ 

$$\Rightarrow y \equiv a \pmod{m} \quad x \equiv a \pmod{m}$$

$$\Rightarrow y \equiv b \pmod{n} \quad x \equiv b \pmod{n}$$

Subtracting both solutions to show that they are the same.

$$y - x \equiv 0 \pmod{m} \quad y - x \equiv 0 \pmod{n}$$

$$\Rightarrow y - x \equiv 0 \pmod{mn}$$

$$y \equiv x \pmod{mn}$$

$$\text{but } y < mn \wedge x < mn \Rightarrow y = x$$

(contradiction)

 $\Rightarrow$  The solution is unique

b)

Thm:

$$x = \sum_{i=1}^k a_i \frac{s_i N}{n_i}$$

*is a solution to*

$$x \equiv a \pmod{m}$$

$$x \equiv b \pmod{N}$$

Proof:*variable*

$$\sum_{i=1}^k a_i \frac{s_i N}{n_i} \pmod{n_j} \equiv a_j \frac{s_i N}{n_j} \pmod{n_j}$$

$$\sum_{i=1}^k a_i \frac{s_i N}{n_i} \equiv a_j (1 - r_j n_j) \pmod{n_j}$$

*because*  $r_i n_i + \frac{s_i N}{n_i} = 1 \Rightarrow \frac{s_i N}{n_i} = 1 - r_i n_i$

*From Chinese Remainder Theorem*

$$\begin{aligned} &\equiv a_j - a_j r_j n_j \pmod{n_j} \\ &\equiv a_j (1 - r_j n_j) \end{aligned}$$

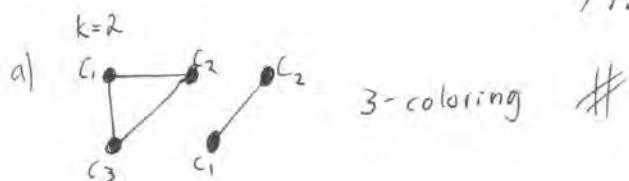
$$\sum_{i=1}^k a_i \frac{s_i N}{n_i} \equiv a_j \pmod{n_j}$$

*works for all values of n<sub>j</sub>*

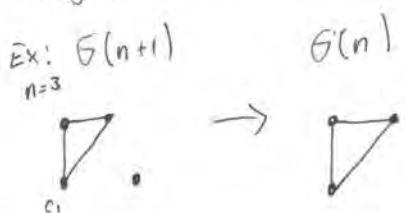
□

## Problem 1

+15/15



- 10) Wrong sentence: "So  $g(n)$  satisfies the conditions of the induction hypothesis"<sup>11</sup>  
 The proof does not take into account the possibility of vertex  $v$  having degree 0 and therefore no adjacent neighbors.



$g(n)$  does not satisfy the conditions of IH because it does not have a vertex of degree less than  $k$ .

16/100

DC

(15/15)

## Problem 2

Claim: For some  $n \geq 3$  ( $n$  boys and  $n$  girls), there exists a set of boys' and girls' preferences such that every dating arrangement is stable

Thm: This claim is false. For every set of boys' and girls' preferences there exists a rogue couple

Proof

Lemma 1: In any set of preferences, there exists a pair that did not rank each other last.

Proof: Consider a boy, let's call him boy A, and how he could be ranked. There are 3 cases.

Case 1: A girl did not rank him last and he did not rank her last ✓

Case 2: Every girl ranks boy A last. Then, consider another boy, let's call him boy B. He was not ranked last by any of the girls so we can pair him with anyone that is not his last choice ✓

Case 3: Consider a girl  $g_1$ . Everyone but  $g_1$  ranks  $B_1$  last but  $B_1$  ranks  $g_1$  last. Consider a boy, Boy B, he was not ranked last by anyone except maybe  $g_1$ , so there must be a girl in that set<sup>1</sup> that is not his last preference. Pair them up ✓

These 3 cases show all combinations that boy A could have been ranked

□

Proof of Theorem

By Lemma 1, there exists a pair that did not rank each other last in every combination of preferences. Now consider the matching where boy A is matched with his least favorite girl and girl 1 is matched with his least favorite boy. This means that they are a rogue couple and the arrangement is unstable

□

## Problem 3

 $\mu$ 

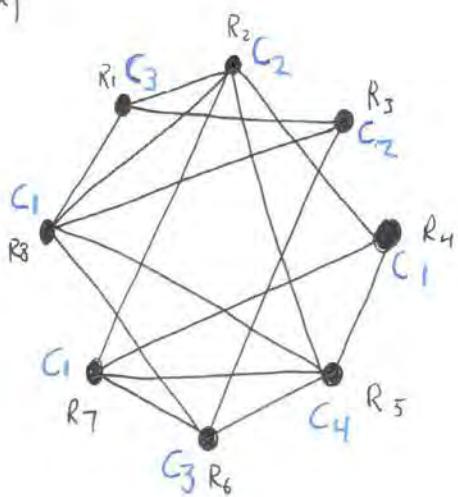
16/20

Brando Miranda

Collab: Erika

Kate Fan

a)



- \* colors represent timeslots, which we want to minimize
- \* Verticies represent different recitations
- \* Edges connect recitations that share a staff member

Question?  
W-Q  
oh

b) See graph above for coloring using 6 colors.

The colorings imply:

Timeslot 1 :	R8, R4, R7
" 2 :	R2, R3
3 :	R1, R6
4 :	R5

why 4 colors  
necessary?  
-1

## Problem 4

20/20

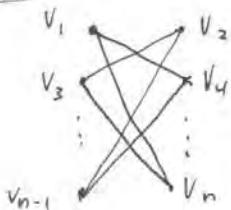
a)  $\chi(G) = 2$

because the chromatic number of a bipartite graph is 2.

b)  $\{v_1, v_3, v_5 \dots v_{n-1}, v_2, v_4, v_6 \dots v_n\}$

c)  $\{v_1, v_2, v_3, v_4, v_5 \dots v_n\}$

d) <sup>Thm</sup> For all even integers  $n$ , an ordering of the vertices such that the greedy algorithm uses exactly  $\frac{n}{2}$  colors is: All the odd vertices on the left subset  $V_L$  and all the even on the right  $V_R$ .

ExampleProof

(By induction)

P(n). All rows have different colors from each other, but same color within the row.

Base case $n=2$ : The graph has 2 nodes and 0 edges. $n/2 = 2/2$ : Graph is one-colorable!

Induction Step: Assume  $P(n)$  for the purposes of induction to show  $P(n+2)$

IF we add a pair of nodes that are directly across from each other, the greedy algorithm will assign the first <sup>in the right side</sup> one a new color (since it is connected to every other node <sup>in the right side</sup> and thus every color).

The greedy algorithm will then assign the second <sup>left side</sup> node the same color (since it is connected to every node on the right side except its pair, so that every pair of nodes  $\Rightarrow P(n+2)$  will use one more color than  $P(n)$ )

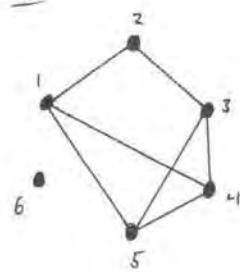
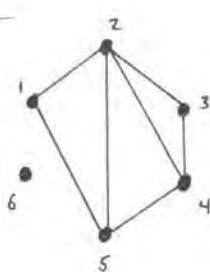
shares a color. Directly across from each other.

Since  $P(n)$  uses  $\frac{n}{2}$  colors and  $P(n+2)$  uses  $\frac{(n+2)}{2} = \frac{n+1}{2}$  colors

□

15/15 DN

## Problem 5

a)  $\underline{G_1}$  $\underline{G_2}$ 

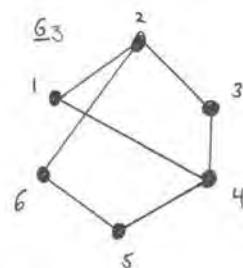
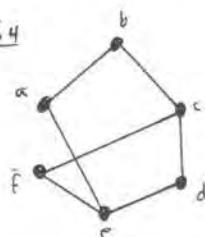
No isomorphism exists because the degree is not consistent with any mapping.

Ex: In  $G_2$ ,  $v(2)$  has degree = 4.

No vertex in  $G_1$  has degree = 4

□

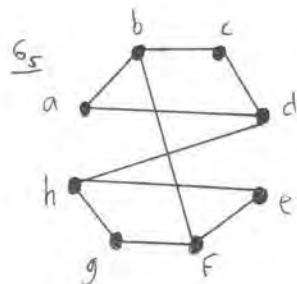
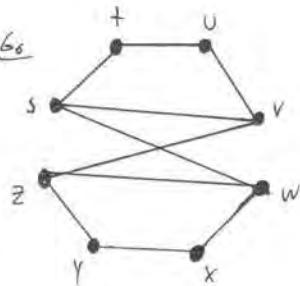
b)

 $\underline{G_4}$ 

Isomorphism:  $f(v)$  maps:

$$\begin{aligned} e &\rightarrow 2 \\ c &\rightarrow 4 \\ d &\rightarrow 3 \\ a &\rightarrow 6 \\ f &\rightarrow 1 \\ b &\rightarrow 5 \end{aligned}$$

c)

 $\underline{G_6}$ 

In  $G_5$ , each node with degree 2 is only adjacent to nodes with degree 3. This property does not hold in  $G_6$ .

Ex: In  $G_6$ ,  $v(\deg 2)$  is adjacent to  $v(\deg 3)$  and  $t(\deg 2)$

$\Rightarrow$  No isomorphism

(15)

## Problem 6

Ihm Let  $G = (V, E)$  be a graph. A matching in  $G$  is a set  $M \subseteq E$  s.t. no two edges in  $M$  are incident on a common vertex. Let  $M_1, M_2$  be two matchings of  $G$ . Then  $G' = (V, M_1 \cup M_2)$  is bipartite.

Proof

WLOG  $G'$  is connected because all  $G$ 's can be divided into components that by definition, are connected. If all components are bipartite  $\rightarrow G'$  is bipartite.

$\Rightarrow$  nodes in  $G'$  can have max deg = 2 b/c each node is connected to at most 2 (1 from each matching  $M_1, M_2$ )

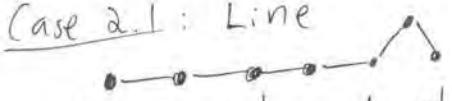
Lemma 1. Nodes of  $G'$  can be connected only in two ways.  
A path/line or loop.

Case 1: Path

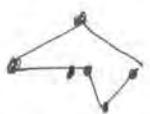
Assume  $V_i$  has deg = 1.  $\Rightarrow$  Nodes can only form a path since "branching out" would require at least 2 degrees.

Case 2: Loop

Assume  $V_i$  has deg = 2. There are two cases:

Case 2.1: Line

Connections terminate at a vertex with degree = 1  $\Rightarrow$  Line

Case 2.2 Loop

Connections terminate at the same point  $\Rightarrow$  Loop  $\square$

Proof of Theorem

Let's analyze both cases allowed by Lemma 1. that is line and loop.

Case 1 Path

Paths are always 2 colorable since we can simply assign the same color to every other vertex. This is Bipartite because we can put all the nodes of each color in a subset and no nodes would be connected to another in their subset. (since they have the same color)

stack

## Problem 1

$$\begin{array}{c} a^2 \\ \diagup \quad \diagdown \\ 107 \end{array} \quad -15$$

$$\frac{16}{16} \text{ LS}$$

- a) If the two vertices are distance of  $\leq \frac{n}{2}$  apart go left and right on the cycle in the direction of the vertex.  
 If they are greater than  $\frac{n}{2}$  distance apart, take a crossing edge and then go left or right on the cycle in the direction of the vertex.
- b) If  $n$  is odd the diameter of  $G$  is  $\frac{n+1}{2}$   
 If  $n$  is even, the diameter of  $G$  is  $\frac{n}{2}$
- c) Every node has to have a degree of at least  $k$  for  $G$  to be  $k$ -edge connected. In  $G$  all of the nodes have degree 3 and so the graph is not 4-edge connected since we could simply remove three edges incident to the same node, effectively disconnecting that node from the rest of the graph.
- d) Every vertex has degree three since it is connected to two other vertices in the cycle and one crossing edge.

Case 1: We remove  $a$  edges in the cycle.

Case 1.1: The removed edges are incident to the same node:

In this case the affected node is still connected to the rest of the graph by the crossing edge

Case 1.2: The removed edges are not incident to the same node

In this case, the graph remains connected since the crossing edges connect the rest of the graph and every node has degree at least 2.

Case 2: We remove a crossing <sup>edge</sup> and an edge in the cycle

In this case, we can still traverse from one end of the <sup>(now broken)</sup> cycle to the other end. Regardless of the removed crossing edge. <sup>new path that connects all nodes</sup>

Case 3: We remove two crossing edges.

The graph is still connected by the cycle,

14/14 DN

## Problem 2

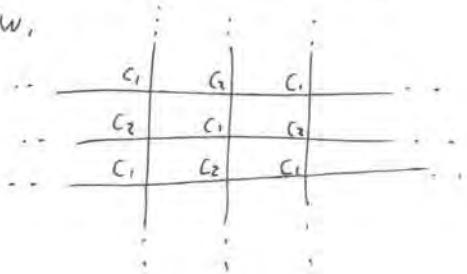
a)

Thm: IF both  $N$  and  $M$  are odd, then the  $N \times M$  grid is not Hamiltonian

Proof:

Lemma: Any  $N \times M$  2-dimensional undirected grid is bipartite.

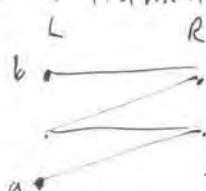
Proof: Any  $N \times M$  2-dimensional undirected grid is 2-colorable by alternating colors 1 and 2 down each row and column as pictured below.



Since the grid is 2-colorable it is also bipartite  $\square$

Proof of theorem:

If  $N \times M$  is odd there are an odd number of vertices. By Lemma an  $N \times M$  grid is bipartite and since the total number of vertices is odd we will end up on the same side of the bipartite graph that we started on. By definition of a bipartite graph this node cannot be connected to the start node and so a Hamiltonian cycle is not possible.



If we start at point b and end at a, there is no edge incident to both b and a and so a Hamiltonian cycle is not possible  $\square$

+5

b)

i) Thm: IF either  $N$  is even and  $M > 1$  or  $M$  is even and  $N > 1$ , then  $N \times M$  grid is Hamiltonian

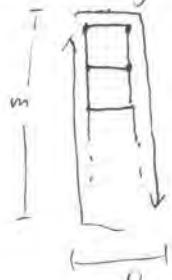
Proof: Without loss of generality, we can only consider the first case since the other would just be a rotated grid.

Proof (By induction)

$P(n)$ : IF  $n$  is even and  $m > 1$ , then  $n \times m$  is hamiltonian

Base case:

Consider a grid where  $m=2$  and  $n$  is any even number.

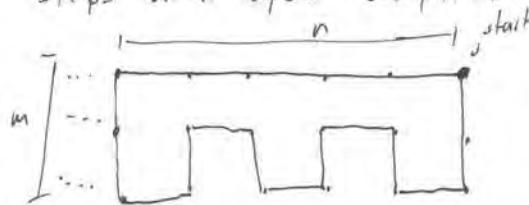


A hamiltonian cycle would transverse the first row, down the last column, on the last row move one vertex towards the beginning and back up to the start.

Inductive step:

$P(m)$ : Start at one of the corners. Travel along the entire first col and go down the entire last row. Then move in the last row move one vertex towards the beginning and back up that row till the second column. Then travel along one vertex towards the beginning and move down to the end of the row, repeat these steps until cycle completed. The cycle looks like this

+3



Assume  $P(m)$  for the purposes of induction to show  $P(m+1)$

Increasing  $m$  by one,  $P(m+1)$  has a similar shape for a cycle with the longer "grooves" on the cycle. In this way, there still exists a hamiltonian cycle if  $P(m)$  is true. so  $P(m) \Rightarrow P(m+1)$

□

2) Proof breaks down when  $N$  is odd since the base case would immediately fail. Additionally, if  $n$  is odd, we would not end up with our assumed shape for  $P(n)$  and our inductive step would break down. +3

3) Bruce and Sam would survive because there always exists a hamiltonian cycle if either  $N$  is even and  $M \geq 1$  or  $M$  is even and  $N \geq 1$ . Since they are guaranteed to visit every intersect once (by definition of a hamiltonian path) the location of the bomb does not make a difference in the outcome.

+3

(16/16)

## Problem 3

a) Thm: A simple connected graph with  $n$  nodes and  $n-1$  edges is a tree.Proof: A tree is defined to be connected and acyclic. Proving that the graph is acyclic would prove it is a tree (since it is also connected).

- You need at least  $n$  edges to create a cycle, since there are only  $n-1$  edges, there is no cycle.

Therefore, it is an acyclic connected graph  $\Rightarrow$  The graph is a tree  $\square$ b) Thm: Any connected graph has a spanning tree.Proof (by induction on the number of edges)Base case

A connected graph with zero edges must have only one vertex. This is a tree.

Inductive StepP(n): Any connected graph with  $n$  edges has a spanning tree.

Assume P(n) to prove P(n+1)

There are two cases:

Case 1:  $G$  is acyclic.This means that  $G$  is connected and acyclic so we are done.Case 2:  $G$  has a cycleWe can remove an edge in the cycle to get  $G'$ . Since  $G'$  is still connected and has  $n$  edges  $\Rightarrow G'$  has a spanning tree.Since all of the edges in  $G'$  are also in  $G$ , the spanning tree for  $G'$  is also a spanning tree for  $G$ .  $\Rightarrow P(n+1)$  $\square$

## Problem 4

+0/13

- a) Thm: Any minimum-weight spanning tree (MST) for  $G'$  is an MST for  $G$

Proof: Since  $G'$  is a subgraph of  $G$ .

We could remove any edge from  $G$  to create a tree, a minimum spanning tree would not make use of the maximum-weight edge in  $G$  so an MST for  $G'$  is also an MST for  $G$   $\square$

+0/8

- b) This procedure terminates with  $n-1$  edges since anything more than that would create a cycle.

To create a tree from a graph with 1 or more cycles, we could remove any edge from each cycle. Since we are removing the most costly edge for each we are guaranteed to terminate with an MST

why?

+95

DC



## Problem 5

As seen in class, a butterfly with a power of two can adjust each individual bit at each of the  $k$  levels. Since each level equates to a bit wise correction we know that after  $L$  levels,  $L$  bits have been corrected. This means  $2^L$  inputs can get to a switch at that level (at most) and, assuming the total number of levels to be  $k$ , there are  $2^{k-L}$  reachable outputs from that switch because the first  $L$  bits have already been shifted.

We define the congestion for each switch to equal the minimum of the  $2^L$  congestion in and the  $2^{k-L}$  congestion out.

Therefore the max congestion occurs at the switch with the maximum individual congestion. This occurs at  $L/2$  and

$$2^{L/2} = \sqrt{N}$$

## Problem 6

(20/20) 4

- a) Ihm: If  $2^m \equiv 1 \pmod{N-1}$ , then  $m$  perfect shuffles will return a deck of  $N$  cards to its original order.

Proof

Lemma 1: For a deck of  $N$  cards indexed at 0, a card at index  $i$  will move to index  $2i \pmod{N-1}$  after a shuffle

Proof: There are two cases

Case 1  $i < N/2$

A perfect shuffle will place  $2i$  cards in front of the card.  
 since  $2i = 2i \pmod{N-1}$  the lemma holds  
 (because there are  $i$  cards before  $i$ )

Case 2  $i \geq N/2$

If the deck is divided perfectly in half, card at index  $i$  has  $i - N/2$  cards in its half of the deck. Therefore after a shuffle it will have the  $i - N/2$  cards from its half in addition to the  $i - N/2 + 1$  cards from the other deck. in front.

In this case the card will be at index  $2(i - N/2) + 1$  after a shuffle.

$$\text{Since: } 2(i - N/2) + 1 = 2i - N + 1$$

$$\equiv 2i - 1 + 1 \pmod{N-1}$$

$$\equiv 2i \pmod{N-1} \text{ the lemma holds}$$

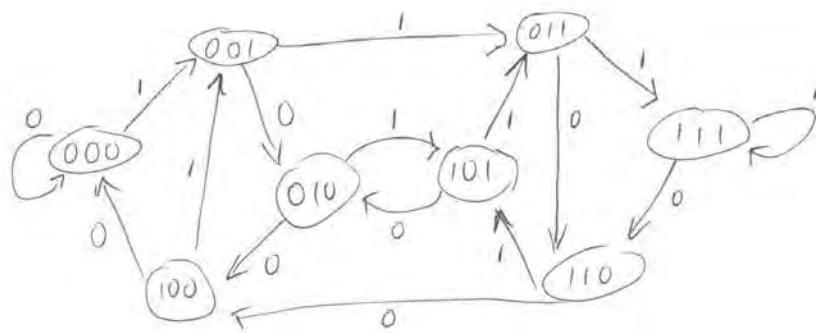
Proof of theorem

By lemma 1, a card at index  $i$  will be at index  $2i \pmod{N-1}$  after one shuffle. After  $m$  shuffles the card will be at index  $2^m i \pmod{N-1}$  because  $2^m \pmod{N-1} \equiv 1$   $2^m i \pmod{N-1} \equiv 1 \cdot i \pmod{N-1}$   $\equiv i \pmod{N-1} \equiv i$

This means the a card  $i$  will return to its original index after  $m$  perfect shuffles. This applies to any all of the cards so the deck returns to its original position  $\square$

## Problem 8

10/10



start at 000  $\circlearrowleft$  take path

001  
 011  
 111  
 111  
 110  
 101  
 011  
 110  
 100  
 001  
 010  
 101  
 010  
 100  
 000

This is a Eulerian Tour

16 bit de Bruijn Sequence

0 1 1 1 0 1 1 0 0 1 0 1 0 0 0 0
---------------------------------

## Problem 7

Diameter . 15

Since input 1 is the Farthest away from an edge that connects it to the other side, we pick it as our sample for the diameter. If we connect input 1 to output 6, the length of paths taken is 15. This is the longest path from an input to an output and so it is the diameter.

Congestion: At least 2

Consider the mapping shown below

$$\begin{aligned} I_0 &\rightarrow O_1 \\ I_1 &\rightarrow O_4 \\ I_2 &\rightarrow O_6 \\ I_3 &\rightarrow O_2 \\ I_4 &\rightarrow O_5 \\ I_5 &\rightarrow O_3 \end{aligned}$$

As with any mapping all six inputs need to get to the other half of the graph using only 3 switches. This means that at least one switch will have a congestion of at least 2. *it's 2!* -2

In the mapping described above, it is possible to get a congestion of 5 if we consider the minimum distance path for each mapping. X

## Problem 1

(20%/  
20)

- a) The distance between any two verticies is at most k.
- b) There is no path from i to j
- c) There is no path from i to i meaning there are no self loops. All disconnected or cycles.
- d) The graph is disconnected and has at least 2 connected components.

105  
105

missed Problem 3?

## Problem 2

15/15 μ

- a) Thm: A strongly connected graph has at most one set of PageRank values that are stationary

Proof

Lemma 1: For two sets of values of PageRank  $d_1$  and  $d_2$ , let  $\gamma \equiv \max_{x \in V} \frac{d_1(x)}{d_2(x)}$ , there exists a directed edge from  $y$  to  $z$  such that  $\frac{d_1(y)}{d_2(y)} < \gamma$  and  $\frac{d_1(z)}{d_2(z)} = \gamma$

Proof

Because the graph is strongly connected, every node must have at least one in-edge and at least out-edge. This means that there must be a directed edge from  $y$  to  $z$ .

We pick node  $z$  such that  $\frac{d_1(z)}{d_2(z)} = \gamma$ . We then look at a node with an in-edge from  $z$ , call it  $y$ .

Case 1:  $\frac{d_1(y)}{d_2(y)} = \gamma$

In this case, we make  $y$  the new  $z$  and repeat this process.

Case 2:  $\frac{d_1(y)}{d_2(y)} < \gamma$

The Lemma holds. ✓

Note that  $\frac{d_1(y)}{d_2(y)} > \gamma$  cannot occur since  $\gamma$  is defined to be the max ratio. Also note that case 2 will eventually happen since  $d_1$  and  $d_2$  are different (and PageRank values must add up to 1). (The only way to have  $\frac{d_1(y)}{d_2(y)} = \gamma$  for every node is if  $d_1$  was a scaled version of the other which cannot happen because PageRank values add up to 1.)

□

b)

Proof of Theorem

(by contradiction)

Assume (for the purposes of contradiction) that  $d_1$  and  $d_2$  are distinct stationary sets of values of PageRank.

Consider a node  $z$  where the max ratio  $\left(\frac{d_1(z)}{d_2(z)}\right)$  occurs.

Consider all of the nodes,  $x_1, x_2 \dots x_n$  that link to  $z$ .

Then, the total incoming PageRank to  $z$  is:

$$d_2(z) = \sum_{i=1}^n \frac{d_2(x_i)}{\deg(x_i)}$$

Similarly,  $d_1(z)$  can be expressed as:  $d_1(z) = \sum_{i=1}^n \frac{d_1(x_i)}{\deg(x_i)}$

$$\text{Then } d_1(z) = \nu d_2(z) = \sum_{i=1}^n \frac{\nu d_2(x_i)}{\deg(x_i)}$$

Since  $\nu d_2(x_i) \geq d_1(x_i)$  and  $\nu d_2(y) > d_1(y)$

$$d_1(z) = \nu(d_2(z)) = \sum_{i=1}^n \frac{\nu d_2(x_i)}{\deg(x_i)} > \sum_{i=1}^n \frac{d_1(x_i)}{\deg(x_i)} = d_1(z)$$

#

□

## Problem 4

a)  $R_n := \{(x, y) \in \mathbb{Z} \times \mathbb{Z} \mid x \equiv y \pmod{n}\}$

Symmetry

$$\forall x, y \in \mathbb{Z} \quad x \equiv y \pmod{n} \Rightarrow y \equiv x \pmod{n} \quad \checkmark$$

Reflexive

$$\forall x \in \mathbb{Z} \quad x \equiv x \pmod{n} \quad \checkmark$$

Transitive

$$\forall x, y, z \in \mathbb{Z} \quad x \equiv y \pmod{n} \wedge y \equiv z \pmod{n} \Rightarrow x \equiv z \pmod{n} \quad \checkmark$$

Equivalence Class:  $[x] = \{y \mid x \equiv y \pmod{n}\}$ 

b)  $R := \{(x, y) \in P \times P \mid x \text{ is taller than } y\}$   $P$  is the set of all the people in the world today.

$R$  is not an equivalence relation because, for example, it violates the reflexive relation since  $x$  cannot be taller than him/herself.

c)  $R := \{(x, y) \in \mathbb{Z} \times \mathbb{Z} \mid \underbrace{\gcd(x, y)}_{\text{Relatively prime.}} = 1\}$

$R$  is not an equivalence relation because, for example, it violates the reflexive relations since the  $\gcd(x, x) = x \neq 1$  if  $x \neq 1$ .

d)  $P_G := \text{the set of } (x, y) \in V \times V \mid V \text{ is the set of vertices of a graph } G, \text{ and there is a path } x, v_1, \dots, v_k, y \text{ from } x \text{ to } y \text{ along the edge of } G.$

-2 IF  $G$  is a directed graph, the symmetry relation will be violated since just because there is a path between  $x \rightarrow y$  does not necessarily mean there will be a path from  $y \rightarrow x$ .

Unless stated otherwise, assume  $G$  is not directed.

## Problem 5

a)  $R_1 \cap R_2 = R$  is an equivalence relationship.

10  
10 71

Reflexive

$R_1$  and  $R_2$  are (by definition) reflexive because they are both an equivalence relation. Since a relation in  $R$  must also be in  $R_1$  and  $R_2$ , it will maintain its reflexive properties.

$$(x, x) \in R_1 \wedge (x, x) \in R_2 \wedge x \in X \Rightarrow (x, x) \in (R_1 \cap R_2)$$

Symmetric

This implies

Suppose  $(x, y)$  is a relation in  $R_1 \cap R_2$ .  $\Rightarrow (x, y) \in R_1 \wedge (x, y) \in R_2$

Since both  $R_1$  and  $R_2$  are an equivalence relation, they satisfy the symmetric relation.

$$\begin{aligned} \text{If } (x, y) \in R_1 \text{ and } (x, y) \in R_2 &\Rightarrow (y, x) \in R_1 \wedge (y, x) \in R_2 \\ &\Rightarrow (y, x) \in R_1 \cap R_2 \quad \therefore R_1 \cap R_2 \text{ is symmetric} \end{aligned}$$

Transitive

$(y, p)$

$\wedge (y, p) \in R_1$

$\wedge (y, p) \in R_2$

Suppose  $(x, y)$  is a relation in  $R_1 \cap R_2$ .  $\Rightarrow (x, y) \in R_1 \wedge (x, y) \in R_2$   
Since both  $R_1$  and  $R_2$  are an equivalence relation, they satisfy the transitive property.

$$\begin{aligned} \text{If } (x, p) \in R_1 \text{ and } (x, p) \in R_2 &\Rightarrow (x, p) \in R_1 \wedge (x, p) \in R_2 \\ &\therefore R_1 \cap R_2 \text{ is transitive.} \end{aligned}$$

□

b)  $R_1 \cup R_2$  is not an equivalence relationship.

Proof (By counter example)

suppose:

$$R_1 = \{(y, y), (y, z), (x, x), (z, y)\}$$

$$R_2 = \{(x, x), (z, z), (x, z), (z, y)\}$$

$R_1$  and  $R_2$  satisfy the equivalence relation since they are  $\begin{cases} \text{transitive} \\ \text{symmetric} \\ \text{reflexive} \end{cases}$

We know  $(x, y), (y, z) \in R_1 \cup R_2$

However  $(x, z) \notin R_1 \cup R_2$

$\Rightarrow R_1 \cup R_2$  is not transitive

□

## Problem 3 (Credo - original was lost)

a)  $\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{3} & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{3} & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{3} & \frac{1}{2} & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \\ r_5 \\ r_6 \\ r_7 \\ r_8 \end{bmatrix}$

$$\begin{aligned} r = & \begin{bmatrix} \frac{1}{8} \\ \frac{1}{8} \end{bmatrix} \xrightarrow{\text{1st iteration}} \begin{bmatrix} .1250 \\ .0417 \\ .0625 \\ .1250 \\ .2292 \\ .2293 \\ .0625 \\ .1250 \end{bmatrix} \xrightarrow{\text{2nd iteration}} \begin{bmatrix} .0417 \\ .0417 \\ .0625 \\ .0938 \\ .3021 \\ .2812 \\ .1146 \\ .0625 \end{bmatrix} \\ & \text{Starting Rank} \end{aligned}$$

b) Thm: Let  $G$  be a graph consisting of two strongly connected components  $C_1$  and  $C_2$ . If there are edges from  $C_1$  to  $C_2$  (but not from  $C_2$  to  $C_1$ ). Then, the stationary PageRank values of  $G$  are entirely concentrated in  $C_2$ , that is, the PageRank values are all 0 in  $C_1$ .

Proof (By contradiction). Assume, for the purposes of contradiction, that in the stationary set of PageRank values, there are some that are not zero in  $C_1$ . This means that some non-zero value of PageRank is being circulated only among  $C_1$  nodes. This is not possible since there is an edge from  $C_1$  to  $C_2$ . Since  $C_1$  is strongly connected, there is a path from any two vertices in  $C_1$ . If there is a nonzero value of PageRank, it would contribute some non-zero value that would eventually arrive at a node in  $C_1$  that has an edge from it to  $C_2$ . Thus, a non-zero PageRank value would be passed into  $C_2$  and since there are no edges from  $C_2$  to  $C_1$ , the total PageRank in  $C_1$  decreases. This means that for some node in  $C_1$ , the total PageRank entering it is less than the total PageRank leaving it, so it is not stationary.  $\# \square$

## Problem 4

(15/15) M

Invariant: At the beginning of each round, there is exactly one pint in glass 1

Let  $w$  be the amount of wine at the beginning of a round.

$$\text{In round 0: } w_0 = 0 \text{ Pint}$$

$\frac{1}{3}$  pmt  
 $\frac{2}{3}$  total

$$\text{In round 1: } w_1 = \underbrace{\left(\frac{2}{3}\right)w_0}_{\frac{1}{4} \text{ pmt}} + \underbrace{\left(\frac{1}{4}\right)\left(1 - \left(\frac{2}{3}\right)w_0\right)}_{\text{Amount of water in mixture}}$$

$$= \frac{1}{4} + \frac{w_0}{2}$$

$$\text{In round 2: } w_2 = \left(\frac{1}{4}\right) + \left(\frac{1}{4} + \frac{w_1}{2}\right) = \frac{1}{4} + \frac{1}{8} + \frac{w_1}{4}$$

Generalizing thus for  $n$  rounds we get

$$\sum_{i=1}^n \left(\frac{1}{2}\right)^{i+1} + \frac{w}{2^n}$$

$\downarrow$

$$\frac{1}{2} \sum_{i=1}^n \left(\frac{1}{2}\right)^i$$

$\downarrow$

$$\frac{1 - \left(\frac{1}{2}\right)^{n+1}}{1 - \frac{1}{2}} - 1$$

since we start from  $i=1$  and is from  $i=0$

$\downarrow$

Putting all together

$$\Rightarrow \left(\frac{1}{2}\right) \left(-1 + \frac{1 - \left(\frac{1}{2}\right)^{n+1}}{1 - \frac{1}{2}}\right) + \frac{w}{2^n} = \underbrace{\frac{w}{2^n}}_{W \text{ is originally so...}} + \left(\frac{1}{2}\right) \cdot -\left(\frac{1}{2}\right)^{n+1}$$

Amount of wine in first glass after  $n$  rounds:

$$\boxed{\left(\frac{1}{2}\right) - \left(\frac{1}{2}\right)^{n+1}}$$

b)  $\lim_{n \rightarrow \infty} \left(\frac{1}{2} - \left(\frac{1}{2}\right)^{n+1}\right) = \frac{1}{2}$  in glass 1  $\xrightarrow{\text{invariant.}}$   $\frac{1}{2}$  in glass 2 because of the

Glass 1: $\frac{1}{2}$
Glass 2: $\frac{1}{2}$

Fernando Troyano

6.0412 Pset #7

19  
Fernando Miranda  
(collab: Erik Kal)

Problem 5

a)  $f(n) = \log_2 n$        $g(n) = \log_{10} n$

$$\boxed{f = \Theta(g), F = \Omega(g), F = O(g)}$$

b)  $f(n) = 2^n$        $g(n) = 10^n$

$$\boxed{f = O(g), F = o(g)}$$

$\frac{20}{20}$  DN

c)  $f(n) = 0$        $g(n) = 17$

$$\boxed{f = O(g), F = o(g)}$$

d)  $f(n) = 1 + \cos\left(\frac{\pi n}{2}\right)$        $g(n) = 1 + \sin\left(\frac{\pi n}{2}\right)$



None

e)  $f(n) = 1.000006001^n$        $g(n) = n^{100000000000}$

$$\boxed{f = \Omega(g), F = \omega(g)}$$

Exponential Always grows faster than polynomial.

Problem 6

 $+20/20$ 

a)  $n! = O((n+1)!) \Rightarrow \lim_{x \rightarrow \infty} \left| \frac{f(x)}{g(x)} \right| < \infty \Rightarrow \lim_{n \rightarrow \infty} \frac{n!}{(n+1)!}$

$$= \lim_{n \rightarrow \infty} \frac{n!}{(n+1)n!} = 0 < \infty$$

(TRUE)

$n! = \Omega((n+1)!) \Rightarrow \lim_{x \rightarrow \infty} \left| \frac{f(x)}{g(x)} \right| > \infty \Rightarrow \lim_{n \rightarrow \infty} \frac{n!}{(n+1)!}$

$$= \lim_{n \rightarrow \infty} \frac{n!}{(n+1)n!} = 0 \neq \infty$$

(FALSE)

$n! = \Theta((n+1)!) \Rightarrow \lim_{x \rightarrow \infty} \left| \frac{f(x)}{g(x)} \right| \neq \text{non-zero constant}$

(FALSE)

$n! = \omega((n+1)!) \Rightarrow \lim_{x \rightarrow \infty} \left| \frac{f(x)}{g(x)} \right| = \infty \quad 0 \neq \infty$

(FALSE)

$n! = o((n+1)!) \Rightarrow \lim_{x \rightarrow \infty} \left| \frac{f(x)}{g(x)} \right| = 0$

 $0 = 0$   
(TRUE)

b)  $n! = \omega\left(\left(\frac{n}{3}\right)^{n+e}\right) \Rightarrow \lim_{x \rightarrow \infty} \left| \frac{f(x)}{g(x)} \right| = \infty \Rightarrow \lim_{n \rightarrow \infty} \frac{n!}{\left(\frac{n}{3}\right)^{n+e}} = \lim_{n \rightarrow \infty} \frac{\left(\frac{n}{e}\right)^n \sqrt{2\pi n}}{\left(\frac{n}{3}\right)^{n+e}}$

$$\Rightarrow \lim_{n \rightarrow \infty} \frac{\sqrt{2\pi n} 3^n e^{e(n)}}{e^n n^e} = \infty \quad b/c \quad 3^n > e^n$$

 $\therefore f = \Omega(g)$   $\square$

$$c) n! = \mathcal{L}(2^n) \Rightarrow \lim_{x \rightarrow \infty} \left| \frac{f(x)}{g(x)} \right| > 0$$

$$\Rightarrow \lim_{n \rightarrow \infty} \frac{n!}{2^n} = \lim_{n \rightarrow \infty} \frac{\left(\frac{n}{e}\right)^n \sqrt{2\pi n} e^{E(n)}}{2^n} > 0$$

Since  $n^n \gg 2^n$

□

$$d) \sum_{k=1}^n k^6 = \Theta(n^7)$$

↓

$$S = 1 + 2^6 + 3^6 + 4^6 + \dots + (n-2)^6 + (n-1)^6 + n^6$$

$$\sum_{m=1}^n m^0 = n$$

$$\sum_{m=1}^n m = \frac{n(n+1)}{2} = \Theta(n^2)$$

$$\sum_{m=1}^n m^2 = \frac{n(n+1)(2n+1)}{6} = \Theta(n^3)$$

$$\sum_{m=1}^n m^p = \text{some polynomial with leading term being } cn^{p+1} = \Theta(n^{p+1})$$

$$\text{So } \sum_{k=1}^n k^6 \text{ is some polynomial with leading term being } cn^7 \\ = \Theta(n^7)$$

□

(10/10) JK

## Problem 3

$$\text{a)} \sum_{i=1}^{\infty} \frac{1}{(2i+1)^2} = \frac{1}{(2+1)^2} + \frac{1}{(2(2)+1)^2} + \frac{1}{(2(3)+1)^2} + \dots \\ = \frac{1}{9} + \frac{1}{25} + \frac{1}{49} + \frac{1}{81} + \frac{1}{121} + \dots$$

Positive and decreasing:

$$\frac{1}{9} + \int_2^{\infty} f(x) dx + F(n) \leq \sum_{i=2}^{\infty} \frac{1}{(2i+1)^2} \leq \int_2^{\infty} f(x) dx + F(2) + \frac{1}{9}$$

$$\int_2^{\infty} \frac{dx}{(2x+1)^2} + 0 \stackrel{f(\infty)}{\leq} \sum_{i=2}^{\infty} \frac{1}{(2i+1)^2} \leq \int_2^{\infty} \frac{dx}{(2x+1)^2} + \frac{1}{25}$$

$$= -\frac{1}{2} (2x+1)^{-1} \Big|_2^{\infty}$$

$$= 0 - \left[ -\frac{1}{2} (4+1)^{-1} \right] = \frac{1}{10}$$

$$\boxed{\frac{1}{10} + \frac{1}{9} \leq \sum_{i=2}^{\infty} \frac{1}{(2i+1)^2} \leq \frac{1}{9} + \frac{1}{10} + \frac{1}{25}}$$

$$\text{b)} 1.1 \sum_{i=1}^n \ln(i+1) \leq \int_0^n \ln(x+2) dx$$

$$2.1 \sum_{i=1}^n \ln(i+1) \leq \ln 2 + \int_1^n \ln(x+1) dx$$

1.1 Is True TRUE2.1. Since we only add  $\ln 2$  once the inequality does not hold.FALSE

## Problem 1

~~155~~  
~~165~~

- a) Let's define two mutually recursive procedures:

$P_1(n)$ : moves a stack of  $n$  disks 1 pole forward. It uses  $P_{n-1}$  (described below) to move the top  $n-1$  disks two poles forward. We then move the big disk on pole one over to the empty pole 2. We can then use  $P_{n-1}$  again to move the entire stack on the third pole over to the second pole (on top of the big disk).

$P_2(n)$ : We use  $P_2(n-1)$  to move the top  $n-1$  disks two poles forward. Again, we move the big disk onto the second pole. We can then use  $P_1(n-1)$  to move the stack on the third pole over to the first pole. We again move the big disk from the second pole to the third. Now we can use  $P_2(n-1)$  to move the  $n-1$  stack onto the third pole, on top of the big disk.

\* To move  $n-1$  rings two poles over we start by moving the top  $n-2$  rings two poles over, after moving  $n-1^{\text{th}}$  ring from the First to the second pole, then moving the entire stack on the third pole over to the First. We can now move the  $n-1^{\text{th}}$  ring from pole 2 to pole 3 and we can move the  $n-2$  stack on pole one to pole 2. Finally we move the  $n-1$  stack from pole 3 to pole 1.

- b) We let  $S_n$  be the # of moves required to solve the  $n$  disk problem.  
 Furthermore, let  $T_n$  be the # of moves required to move  $n$  disks forward two posts. We have the following recurrences

$$S_n = 2T_{n-1} + 1$$

$$T_n = 2T_{n-1} + S_{n-1} + 2$$

so,

$$T_{n-1} = \frac{S_{n-1}}{2} \Rightarrow T_n = 2\left(\frac{S_{n-1}}{2}\right) + S_{n-1} + 2$$

$$\Rightarrow T_n = S_{n-1} + S_{n-1} + 2 \Rightarrow T_{n-1} = S_{n-1} - 1 + \overline{S_{n-2}} + 2 \Rightarrow S_n = 2(S_{n-1} - 1 + \overline{S_{n-2}} + 2) + 1$$

$$\boxed{S_n = 2S_{n-1} + 2S_{n-2} + 3}$$

c) Closed form expression for  $S_n$ :

$$S_n = 2S_{n-1} + 2S_{n-2} + 3$$

(characteristic equation):

$$x^n = 2x^{n-1} + 2x^{n-2}$$

Roots:  $x = 1 - \sqrt{3}$

$$x = 1 + \sqrt{3}$$

Homogeneous solution

$$c_1 (1 - \sqrt{3})^n + c_2 (1 + \sqrt{3})^n$$

Particular solution:

guess:  $c$

Boundary conditions

$$\boxed{S_n = \frac{(1 + \sqrt{3})^n}{3 - \sqrt{3}} + \frac{(1 - \sqrt{3})^n}{3 + \sqrt{3}} - 1}$$



## Problem 2

$$f(n) = C \alpha^n + D \beta^n$$

General solution

$$x_n = C \alpha^n + D \beta^n$$

Roots  $\alpha$  and  $\beta$ 

$$(x - \alpha)(x - \beta) = x^2 - x\beta - x\alpha + \alpha\beta = 0$$

$$x^{n-2}(x^2 - x\beta - x\alpha + \alpha\beta) = 0$$

$$x^n - x^{n-1}\beta - x^{n-1}\alpha + x^{n-2}\alpha\beta = 0$$

$$x^n = x^{n-1}\beta + x^{n-1}\alpha - x^{n-2}\alpha\beta$$

$$x_n = \beta x_{n-1} + \alpha x_{n-1} - \alpha\beta x_{n-2}$$

$$x_0 = C + D$$

$$x_1 = C\alpha + D\beta$$

Fernando Trujano

### Problem 3

$$a) T(n) = 8T(\lfloor n/2 \rfloor) + n \quad \text{Plug and chug}$$

$$\begin{aligned}
 k=1 &= 8 \left( 8T_{n/4} + \frac{n}{2} \right) + n \\
 &= 8^2 T_{n/4} + 4n + n \\
 k=2 &= 8^2 \left( 8T_{n/8} + \frac{n}{4} \right) + 4n + n \\
 &= 8^3 T_{n/8} + 16n + 4n + n \\
 &= 8^3 T_{n/8} + 4^2 n + 4^1 n + 4^0 n \\
 &\vdots \\
 k=\log_2 n &= 8^{\log_2 n} \cdot T(1) + \underbrace{4^{\log_2 n - 1} \frac{n}{n}}_{\approx cn} + 4^{\log_2 n - 2} + \dots + \underbrace{4^0 n}_{1} \\
 &= (2^3)^{\log_2 n} + \sum_{i=1}^{\log_2 n} 4^{\log_2 n - i} \frac{n}{n} \\
 &\qquad\qquad\qquad \underbrace{\qquad\qquad\qquad}_{\approx cn \text{ For some constant } c} \\
 &= n^3 + cn \\
 &= \Theta(n^3)
 \end{aligned}$$

$$b) T(n) = 2T\left(\lfloor n/8 \rfloor + 1/n\right) + n \quad \text{Akra-Bazzi: since } \forall n \ O\left(\frac{n}{\log n}\right)$$

$$a = 2, \quad b = 1/8 \quad z(1/8)^p = 1$$

$$\Rightarrow p = 1/3$$

$$\begin{aligned}
 T(n) &= \Theta\left(n^{1/3} \left(1 + \int_1^n \frac{n}{v^{4/3}} dv\right)\right) \\
 &= \Theta\left(n^{1/3} \left[1 + \frac{3}{2} v^{2/3}\Big|_1^n\right]\right) \\
 &= \Theta\left(n^{1/3} \left[1 + \frac{3}{2} n^{2/3} - \frac{3}{2}\right]\right) \\
 &= \Theta\left(n^{1/3} \left(1 + \frac{3}{2} n^{2/3} - \frac{3}{2}\right)\right) \\
 &= \Theta\left(\frac{3}{2}n - \cancel{\frac{3}{2}n^{1/3}}\right)
 \end{aligned}$$

$$= \Theta(n)$$

$$c) T(n) = 7T(\lfloor n/20 \rfloor) + 2T(\lfloor n/8 \rfloor) + n$$

Akra Buzz:

$$\begin{aligned} &= \Theta\left(n^p \left(1 + \int_1^n \frac{u}{u^{p+1}} du\right)\right) \\ &= \Theta\left(n^p \left(1 + \int_1^n u^{-p} du\right)\right) \\ &= \Theta\left(n^p \left(1 + \frac{1}{-p+1} u^{-p+1}\Big|_1^n\right)\right) \\ &= \Theta\left(n^p + \frac{1}{-p+1} n\right) \\ &\sum_{i=1}^k a_i b_i^p = 1 \Rightarrow 7\left(\frac{1}{20}\right)^p + 2\left(\frac{1}{8}\right)^p = 1 \\ &\Rightarrow p = .8 \\ &\Rightarrow = \boxed{\Theta(n)} \end{aligned}$$

$$d) T(n) = 2T(\lfloor n/4 \rfloor + 1) + n^{1/2}$$

Akra Buzz: b/c  $1 = O\left(\frac{n}{\log^2 n}\right)$

$$\begin{aligned} q &= 2 \\ b &= 1/4 \quad 2\left(\frac{1}{4}\right)^p = 1 \Rightarrow p = 1/2 \end{aligned}$$

$$\begin{aligned} \bar{T}(n) &= \Theta\left(n^p \left(1 + \int_1^n \frac{u^{1/2}}{u^{p+1}} du\right)\right) \\ &= \Theta\left(n^{1/2} \left(1 + \int_1^n u^{-p-1} du\right)\right) \\ &= \Theta\left(n^{1/2} \left(1 + \ln n - 0\right)\right) \\ &= \Theta\left(n^{1/2} + n^{1/2} \ln n\right) \\ &= \boxed{\Theta(n^{1/2} \ln n)} \end{aligned}$$

$$e) T(n) = 3T(\lfloor n^{1/9} + n^{1/9} \rfloor) + 1$$

$$\begin{aligned} a &= 3 \quad b = 1/a \quad 3\left(\frac{1}{9}\right)^p = 1 \\ &\Rightarrow p = 1/2 \end{aligned}$$

$$\begin{aligned} T(n) &= \Theta\left(n^{1/2} \left(1 + \int_1^n \frac{1}{u^{3/2}} du\right)\right) \\ &= \Theta\left(n^{1/2} \left(1 - 2(n^{-1/2} - 1)\right)\right) \\ &= \Theta\left(n^{1/2} (3 - 2n^{-1/2})\right) \\ &= \boxed{\Theta(n^{1/2})} \end{aligned}$$

Problem 4

(30/30) HLC

a)  $x_n = 4x_{n-1} - x_{n-2} - 6x_{n-3}$

$$\begin{aligned}x_0 &= 3 \\x_1 &= 4 \\x_2 &= 14\end{aligned}$$

1) (Characteristic Equation):

$$\frac{r^n - 4r^{n-1} + r^{n-2} - 6r^{n-3}}{r^n}$$

$$1 = 4r^{-1} - r^{-2} - 6r^{-3}$$

$$0 = \frac{4}{r} - \frac{1}{r^2} - \frac{6}{r^3} - 1$$

$$\begin{aligned}r &= -1 \\r &= 2 \\r &= 3\end{aligned}$$

2) Homogeneous Solution

$$c_1(-1)^n + c_2 2^n + c_3 3^n$$

3) Particular solution

via inhomogeneous part

4) General Solution

$$x_n = c_1(-1)^n + c_2 2^n + c_3 3^n$$

5) Boundary Conditions

$$\begin{aligned}3 &= x_0 = c_1(-1)^0 + c_2 2^0 + c_3 3^0 & 4 &= x_1 = c_1(-1) + c_2 2 + 3c_3 \\&= c_1 + c_2 + c_3 = 3 & &= 2c_2 - c_1 + 3c_3 = 4\end{aligned}$$

$$\begin{aligned}14 &= x_2 = c_1(-1)^2 + c_2 2^2 + c_3 3^2 \\&= c_1 + 4c_2 + 9c_3 = 14\end{aligned}$$

$$\begin{aligned}&\Rightarrow c_1 = 1 \\&c_2 = 1 \\&c_3 = 1\end{aligned}$$

$$x_n = (-1)^n + 2^n + 3^n$$

✓

$$b) x_n = -x_{n-1} + 2x_{n-2} + n \quad x_0 = 5 \\ x_1 = -4/9$$

1) Characteristic Equation

$$x^n = -1 x^{n-1} + 2x^{n-2} \\ x^2 + x - 2 = 0 \\ \text{Roots: } x = -2 \\ x = 1$$

2) Homogeneous Solution

$$c_1(-2)^n + c_2 1^n$$

3) Particular Solution

$$\text{Guess: } f(n) = bn + c$$

$$bn + d = -(b(n-1) + d) + 2(b(n-2) + d) + n \\ = -(bn - b + d) + 2bn - 4b + 2d + n \\ = -bn + b - d + 2bn - 4b + 2d + n$$

But

$$bn + d \neq bn - 3b + d + n$$

$\Rightarrow$  Incorrect Guess!

$$\text{New guess: } f(n) = an^2 + bn + d$$

$$an^2 + bn + d = -(a(n-1)^2 + b(n-1) + d) + 2(a(n-2)^2 + b(n-2) + d) + n \\ = -an^2 + 2an - a - bn + b - d + 2an^2 - 2a - 8n + 8a + 2bn + 2d - 4b + n \\ = an^2 - 6an + 7a + bn - 3b + d + n$$

$$6an - 7a + 3b = n \quad -7a + 3b = 0$$

$$\Rightarrow a = 1/6 \quad b = 7/18 \quad d = 0$$

4) General Solution

$$f(n) = c_1(-2)^n + c_2 + \frac{1}{6}n^2 + \frac{7}{18}n$$

$$5) \text{ Boundary Conditions} \quad x_0 = 5 \\ x_1 = -4/9$$

$$x_0 = f(0) = 5 = c_1 + c_2$$

$$f(1) = -4/9 = c_1(-2) + c_2 + 1/6 + 7/18$$

$$0 = -4/9 + 3(c_2 + 1) = 0$$

$$\Rightarrow c_2 = 3 \\ 5 = c_1 + 3 \\ \Rightarrow c_1 = 2$$



$$f(n) = 2(-2)^n + 3 + \frac{1}{6}n^2 + \frac{7}{18}n$$

50

## Problem 5

a) choose Rank of pair      2 suits  
 $\binom{13}{1} \cdot \binom{4}{2} \cdot \binom{12}{3} \cdot \binom{4}{1} \cdot \binom{4}{1} \cdot \binom{4}{1}$

$$= \frac{13 \cdot 4!}{2! 2!} \cdot \frac{12!}{3! 9!} \cdot 4^3$$

$$= \boxed{1,098,240}$$

b)  $\binom{48}{5} + \binom{4}{1} \cdot \binom{48}{4}$  =  $\underbrace{2490624}_{\text{0 or 1 king}} - \binom{52}{5} = 2490624 = \boxed{108,336}$   
 all cards except kings  
 suit

c)  $\binom{50}{4} + \binom{50}{4} + \binom{50}{3}$  =  $\boxed{480,200}$   
 A spades      A clubs      A spade A clubs

d)  $\sum_{i=0}^k x_i = n$  Similar to the doughnut example in class, we write  $x_i$  0's for each  $x_i$  and use a | as a delimiter.

Our final bit sequence contains  $n$  0's and  $k$  1's  $\Rightarrow \boxed{\binom{n+k}{k}}$

e)  $\sum_{i=0}^k x_i \leq n$  Same as  $\sum_{i=0}^{k+1} x_i = n$  since subtracting  $x_{k+1}$  would give

a summation that is less than or equal to  $n$ .  $\Rightarrow \boxed{\binom{n+k+1}{k}}$

f)  $3n$  students into  $n$  groups of 3

$\frac{1}{n!} \cdot \binom{3n}{3} \cdot \binom{3n-3}{3} \cdot \binom{3n-6}{3} \cdot \binom{3n-9}{3} \cdots \binom{3}{3}$   
 First group      Second  
 ordering does not matter

$$= \frac{1(3n)! \cdot (3n-3)! \cdots 3!}{n! 3! (3n-3)! 3! (3n-6)! \cdots 3! 0!} = \boxed{\frac{(3n)!}{(3!)^n n!}}$$

## Problem 7

a) Since we are sorting  $n$  distinct numbers, their range won't matter (in terms of comparisons made)

b)  $n!$  permutations

c)

Everytime we ask a question, we can eliminate half of the available permutations. This means we can narrow everything down <sup>with</sup>  $\log_2(n!)$  questions. This means that  $2^k \leq n!$ .

$$\begin{aligned} \log_2(n!) &= k \\ 2^k &= n! \end{aligned}$$

*use induction - 2*

d)  $2^k \leq n!$

$$k = c \cdot n \log n \Rightarrow 2^{cn \log n} = 2^{\log n^{cn}} \leq n!$$

$$\Rightarrow 2^{cn} \leq n! \Rightarrow 2^{cn} \leq \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$$

*c could be anything that makes this equation hold*

e)  $\Theta(n \log n)$  implies that  $n \log n$  will always be an upperbound for the comparisons. That is, not equal to.

As seen in d), there is a constant  $c > 0$  such that  $k = c n \log n$  questions cannot guarantee a win. This means that the number of questions is not always upperbounded by  $n \log n$ , but could be equal to  $n \log n$

*use contradiction - 1*

Problem 6

$$\frac{1}{10}$$

Let  $X$  be all the sets of 7 cards that the audience might select

Let  $Y$  be the sequence of 4 cards that the Assistant might reveal.

$$|X| = \binom{52}{7} = 133784560 \quad (\text{By the subset rule})$$

$$|Y| = 52 \cdot 51 \cdot 50 \cdot 49 = 6497400 \quad (\text{By Generalized Product Rule})$$

$$\left\lceil \frac{133784560}{6497400} \right\rceil = 21 \quad (\text{By Pigeonhole Principle})$$

This means the Assistant must reveal the same sequence of 4 cards four at least 21 different 7 card hands. So there are 21 possibilities for the last cards.

However, if we can consider the temporal order, that is the order in which the cards are placed on the table by the assistant, we essentially have two orderings:

The ordering on the table

The order in which the cards were placed.

Using this method we can multiply  $|Y|$  by  $4!$  and get

$$|Y|4! = 155937600$$

So

$$\left\lceil \frac{133784560}{8 \cdot 155937600} \right\rceil = .8579 = 1$$

So the magician will be able to figure out the card.

YES

10110 EW

## Problem 1

- a) ✓ We know that 2 numbers are relatively prime if they are consecutive that is  $n$  is for sure relatively prime to  $p = n+1$ . This is because  $\gcd(n, n+1) = 1$ .

If we choose  $n+1$  distinct numbers from a set of size  $2n$ , at least one of these numbers must be consecutive to another number by the Pigeon Hole Principle.

- b) ✓ We know that a node in a <sup>connected</sup> finite <sup>undirected</sup> graph with  $n \geq 2$  nodes must have a minimum degree of  $1$ . A node can also have a max degree of  $n-1$ , one edge for each of the other nodes.

By the Pigeon Hole principle, there cannot be a unique ~~degree~~ number for all of the  $n$  nodes  $\Rightarrow$  2 vertices must have the same degree.

IJ

$$\begin{array}{r} 150 \\ \hline 150 \end{array}$$

## Problem 2

Let  $P_{18062}$  be the set of all permutations in which 18062 appears.

We define  $P_{6042}$  and  $P_{35876}$  similarly.

$$|P_{18062}| = 6!$$

$$|P_{6042}| = 7!$$

$$|P_{35876}| = 6!$$

$$|P_{6042} \cap P_{35876}| = 3!$$

$$\Rightarrow \text{Invalid 10 digit passwords} = |P_{18062}| + |P_{6042}| + |P_{35876}| - |P_{6042} \cap P_{35876}|$$

$$= 6! + 7! + 6! - 3!$$

Since pa  
have sha  
di

$$\begin{aligned}
 & - |P_{6042} \cap P_{18062}| \xrightarrow{0} \\
 & - |P_{18062} \cap P_{35876}| \xrightarrow{0} \\
 & + |P_{18062} \cap P_{35876} \cap P_{6042}| \xrightarrow{0}
 \end{aligned}$$

Total permutations Invalid passwords  
 Valid passwords =  $10! - (6! + 7! + 6! - 3!)$   
 = 3622326

## Problem 3

15/15

$$\text{Thm } n2^{n-1} = \sum_{k=1}^n k \binom{n}{k}$$

$$\text{Proof } \binom{n}{k} = \frac{n!}{(k-1)!(n-k)!}$$

The number of all length- $n$  sequences can be represented in two different ways.

First: <sup>left hand side</sup> We choose where to put the star  $\binom{n}{1}$ , all of the other  $(n-1)$  spots can be a 0 or a 1. Multiplying these two we get  $n2^{n-1}$

Second: The number of ways can also be represented by the

$$\text{multinomial } \sum_{k=0}^n \binom{n}{k, n-k-1, 1} = \frac{n!}{k! (n-k-1)!}$$

$$\text{This is equivalent to } \sum_{k=1}^n \frac{n!}{(k-1)!(n-k)!} = \sum_{k=1}^n k \binom{n}{k}$$

Since both of these numbers,  $n2^{n-1}$  and  $\sum_{k=1}^n k \binom{n}{k}$ , count the same thing. We can conclude that they are equal.

□

Brando Mira  
Collab: Erika

**20/20**

## Problem 4

$$\text{a) } \underbrace{\left(\begin{array}{c} 10 \\ 1 \end{array}\right)\left(\begin{array}{c} 9 \\ 1 \end{array}\right)\left(\begin{array}{c} 9 \\ 1 \end{array}\right)\dots\left(\begin{array}{c} 9 \\ 1 \end{array}\right)}_{n \text{ times}} \Rightarrow \boxed{\overline{10 \cdot 9^{n-1}}} \quad \text{m}$$

↑      ↗  
 First digit      2nd, 3rd ... nth  
 can be anything      digit can't be the previous  
 digits so  $\left(\begin{array}{c} 9 \\ 1 \end{array}\right)$ .

- b) Let  $P_2$  be the set of all #'s in the range  $[1\dots 700]$  that are divisible by 2. We define  $P_5$  and  $P_7$  similarly.

$$\begin{aligned} |P_2| &= 350 && \leftarrow \text{since half of the #'s are even and thus divisible by 2} \\ |P_5| &= 140 && \leftarrow \text{since for every 100, there are 20 #'s divisible by 5. } 20 \times 7 = 140 \\ |P_7| &= 100 && \leftarrow \text{since } 700/7 = 100 \end{aligned}$$

$$\begin{aligned} P_2 \cap P_5 &= P_{10} & P_2 \cap P_7 &= P_{14} & P_5 \cap P_7 &= P_{35} \\ |P_{10}| &= 70 & |P_{14}| &= 50 & |P_{35}| &= 20 \end{aligned}$$

$$P_2 \cap P_5 \cap P_7 = P_{70} \quad |P_{70}| = 10$$

# of numbers in range  $[1\dots 700]$  divisible by 2, 5 or 7 =

$$\begin{aligned} |P_2| + |P_5| + |P_7| - |P_2 \cap P_5| - |P_2 \cap P_7| - |P_5 \cap P_7| + |P_2 \cap P_5 \cap P_7| \\ 350 + 140 + 100 - 70 - 50 - 20 + 10 \end{aligned}$$

$$\boxed{= 460}$$

c) Assume q is not in the first two digits:

$$\text{Total}_1 = \binom{8}{5} \cdot q^4 \cdot 8$$

+-----+  
 |      |  
 | 9 \times 8 ways to choose first two digits  
 | ways to choose where q's will go.  
 | 9^3 \times 8 ways to choose last 3 digits

Now assume q is in one of the first two digits

$$\text{Total}_2 = \binom{8}{4} \cdot q^5 \cdot 2$$

+-----+  
 |      |  
 | ways to choose where the rest of the q's will go.  
 | Since q can be on first or second digit  
 | 9^5 ways to choose the other 4 digits

Adding both totals.

$$\text{TOTAL} = \text{Total}_1 + \text{Total}_2$$

$$= \binom{8}{5} q^4 \cdot 8 + \binom{8}{4} \cdot q^5 \cdot 2$$

$$= 11206188$$

## Problem 5

- a) We can represent the total # of ways to distribute  $n$  dollars among  $k$  people with a bijection. Let's construct a bit string where a 0 represents \$1 and a 1 delimits each person. A bit string of 0010010001 means one person has \$2, one person has \$2 and another \$3. For  $n$  dollars and  $k$  people, we will need a bit string with  $n$  zeroes and  $k-1$  ones for a total length of  $n+k-1$ . The number of valid bitstrings of this form is simply:

$$\boxed{\binom{n+k-1}{k-1}}$$

- b) This problem can be solved similar to a) except we begin by taking  $k$  dollars and giving every person a dollar. We proceed as in a), starting with  $k$  fewer dollars.

$$= \binom{(n-k)+k-1}{k-1} = \boxed{\binom{n-1}{k-1}}$$

- c) Consider a similar bit string as described in a) using 0's for books and 1's for book shelf as dividers.

There are  $(n+k-1)!$  ways to order all of these sequences. We then need to divide by  $(k-1)!$  since order matters only for the books.

Assuming shelves are distinguishable.

$$= \boxed{\frac{(n+k-1)!}{(k-1)!}}$$

d) This problem can be thought in a similar way to b) except order matters.

Begin by choosing  $k$  books:  $\binom{n}{k}$ , we multiply this by  $n^{k!}$  to find all of different orderings for these books (one per shelf), now we proceed to assign the remaining books to the book shelves.

$$\boxed{k! \binom{n}{k} \frac{(n-1)!}{(k-1)!}}$$

This is again, assuming the shelves are distinguishable.

(25/25)

HL

## Problem 6

- a) Only one way to make  $n$  cents using only pennies.

$$\langle 1, 1, 1, 1, \dots \rangle \longleftrightarrow P(x) = \frac{1}{1-x}$$

- b)  $\langle 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, \dots \rangle = 1 + x^5 + x^{10} \dots$

$$y = x^5 \Rightarrow 1 + x + x^2 + x^3 = \frac{1}{1-x} \Rightarrow N(x) = \frac{1}{1-x^5}$$

- c) Since  $P(x)$  and  $N(x)$  are disjoint, the generating function for  $P \cup N = P(x)N(x)$

↳ The product of the generating functions,

$$\left( \frac{1}{1-x} \right) \left( \frac{1}{1-x^5} \right)$$

- d) For Dimes:

$$D(x) : \langle 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, \dots \rangle = 1 + x^{10} + x^{20} \dots$$

$$D(x) = \frac{1}{1-x^{10}}$$

For Quarters:

$$Q(x) : \langle 1, \underbrace{0 \dots 0}_{24 \text{ zeros}}, 1, \dots \rangle$$

$$Q(x) = \frac{1}{1-x^{25}}$$

For half dollars

$$H(x) = \frac{1}{1-x^{50}}$$

so total # of ways to make  $n$  cents

$$= P(x)N(x)D(x)Q(x)H(x) = \frac{1}{(1-x)(1-x^5)(1-x^{10})(1-x^{25})(1-x^{50})}$$

- e) The # of ways to change 50 cents is the coefficient of  $x^{50}$  in the polynomial of  $T(x)$



## Problem 7

Thm:  $\forall k \in \mathbb{N}$  the generating function for the nonnegative integer  $k^{\text{th}}$  power is a quotient of polynomials in  $x$ ,

$$\text{So } \forall k \in \mathbb{N} \exists R_k(x) S_k(x) \mid [x^n] \left( \frac{R_k(x)}{S_k(x)} \right) = n^k$$

$\hookrightarrow$  coefficient of  $x^n = n^k$

Proof: (By induction)

$$P(k) : \forall k \in \mathbb{N} \exists R_k(x) S_k(x) \mid [x^n] \left( \frac{R_k(x)}{S_k(x)} \right) = n^k$$

Base Case:

$$k=0 \quad 1+x+x^2+x^3+x^4+\dots = \frac{1}{1-x} \Rightarrow \text{Quotient } \checkmark$$

Inductive Step:

Assume  $P(k)$  for the purposes of induction, to show  $P(k) \Rightarrow P(k+1)$

Recall,

The generating function for the non negative integer  $k^{\text{th}}$  power is given by

$$0 + x + 2^k x^2 + 3^k x^3 + \dots + n^k x^n = \frac{R_k(x)}{S_k(x)}$$

Taking the derivative of both sides yields

$$0 + 1 + 4^k x + 9^k x^2 + \dots + (n^2)^k x^{n-1} = \frac{R'_k(x)}{S'_k(x)}$$

Multiplying both sides by  $x$  yields

$$0 + x + 4^k x^2 + 9^k x^3 + \dots + (n^{k+1}) x^n = \frac{R'_k(x)}{S'_k(x)} x$$

Since the derivative of a polynomial is also a polynomial

$$P(k) \Rightarrow P(k+1)$$

The proof follows by induction  $\rightarrow$  the statement holds... by induction.

□

The proof is already complete!

b)

If  $f(n)$  is a function on the nonnegative integers defined recursively in the form

$$f(n) = af(n-1) + bf(n-2) + cf(n-3) + p(n) \alpha^n$$

where the  $a, b, c, \alpha \in \mathbb{C}$  and  $p$  is a polynomial with complex coefficient

We know that  $\sum_{n \in \mathbb{N}} f(n)x^n$  is the generating function for the sequence

$$f(0), f(1), f(2), \dots$$

Also,  $f(n) = af(n-1) + bf(n-2) + cf(n-3) + p(n) \alpha^n$

$$\left( \sum_{n \in \mathbb{N}} f(n)x^n \right) (1 - ax - bx^2 - cx^3) = \sum_{n \in \mathbb{N}} p(n) \alpha^n x^n$$

← each term will be in the form!      ↑ each term will be in the form  $p^k \alpha^k x^k$

$$(f(k) - af(k-1) - bf(k-2) - cf(k-3))x^k$$

$$\Rightarrow \sum_{n \in \mathbb{N}} f(n)x^n = \frac{\sum_{n \in \mathbb{N}} p(n) \alpha^n x^n}{1 - ax - bx^2 - cx^3}$$

Now, to show the left hand side is a quotient of polynomials, we need to show

that  $\sum_{n \in \mathbb{N}} p(n) \alpha^n x^n$  is a quotient of polynomials

From I we know that any generating function in the form  $\sum_{n \in \mathbb{N}} n^k x^n$  is a quotient of polynomials.  $p(n)$  is a polynomial so  $p(n) \alpha^n x^n$  will be in the form of  $c n^k x^n$  for a constant  $c$ .  $\Rightarrow \sum_{n \in \mathbb{N}} p(n) x^n$  is a quotient of polynomials. Let  $x = \alpha x \Rightarrow \sum_{n \in \mathbb{N}} p(n) \alpha^n x^n$  also quotient of poly.

Since

$$\sum_{n \in \mathbb{N}} f(n)x^n = \frac{\sum_{n \in \mathbb{N}} p(n) \alpha^n x^n}{1 - ax - bx^2 - cx^3} \quad ; \text{ it is also a quotient of polynomials}$$

(by above) □

## Problem 8

$$\text{Thm } \binom{i+j}{k} = \sum_{\ell=0}^k \binom{i}{\ell} \binom{j}{k-\ell}$$

Proof (using generating function)

The left hand side  $\binom{i+j}{k}$  is the  $k^{\text{th}}$  coefficient of the polynomial  $(1+x)^{i+j}$  by the binomial Theorem.

The right hand side  $\sum_{\ell=0}^k \binom{i}{\ell} \binom{j}{k-\ell}$  is the  $k^{\text{th}}$  coefficient of the polynomial  $(1+x)^i (1+x)^j = (1+x)^{i+j}$

This is because

$$\text{convolution } \left[ \binom{i}{0} (x^0) \binom{j}{k} x^k + \binom{i}{1} (x^1) \binom{j}{k-1} x^{k-1} + \dots \right] = \sum_{\ell=0}^k \binom{i}{\ell} \binom{j}{k-\ell}$$

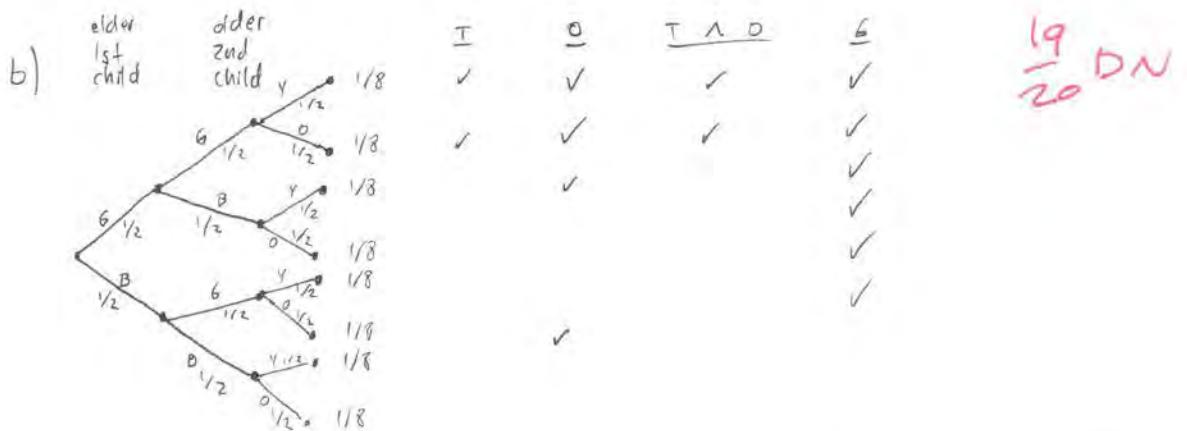
Since both the left hand side and the right hand side are counting the same thing, they must be equal.

$\frac{15}{15}$  DN

□

## Problem 1

a)  $(G, S, Y), (G, G, \emptyset) \in \text{E}$  -1



$$\Pr(T|O) = \frac{\Pr(T \cap O)}{\Pr(O)} = \frac{\frac{2}{8}}{\frac{4}{8}} = \frac{\frac{1}{4}}{\frac{1}{2}} = \boxed{\frac{1}{2}}$$

c)

The probability that a girl answers the door is  $P(O) = 1/2$

The probability that there is at least one girl in the household is  $P(G) = 5/8$

The solution claims that

$$P(T|O) = \frac{P(T|G)}{P(G)}$$

which implies that  $P(O) = P(G)$   
but  $1/2 \neq 5/8$

~~89/168~~  $\frac{84}{95}$

## Problem 2

$$\text{a) } P(A) = \underbrace{P(\text{all } m \text{ birthdays are different})}_{e^{-\frac{m(m-1)}{2N}}} \text{ and } P\left(\begin{array}{l} \text{none of other } k-m \text{ people have} \\ \text{the same birthday as one of the } m \end{array}\right) \underbrace{\left(\frac{N-m}{N}\right)^{k-M}}$$

These two events are independent - from the birthday assumption  
so

$$P(A) \sim e^{-\frac{m(m-1)}{2N}} \left(\frac{N-m}{N}\right)^{k-M}$$

$$P(A) \sim e^{\frac{m(m-2k)}{2N}}$$

$$\text{b) } e^{\frac{m(m-2k)}{2N}} = \frac{1}{2}$$

Solve for  $m$

$$-\ln 2 = \frac{m(m-2k)}{2N}$$

$$2N \ln 2 = m(m-2k)$$

$$2N \ln 2 = m^2 - 2km$$

$$0 = m^2 - 2km + N \ln 2$$

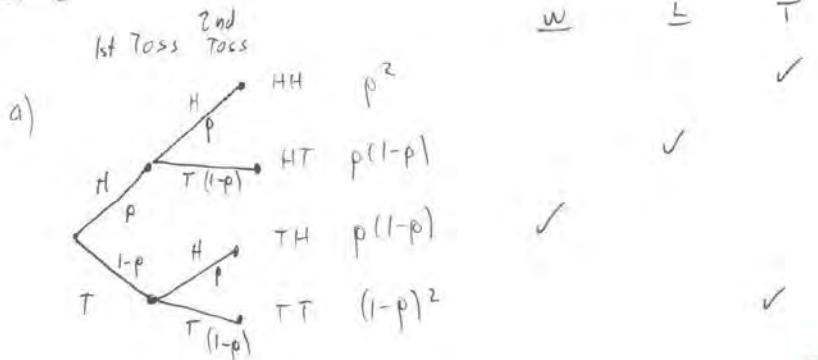
$$m = \frac{-2k \pm \sqrt{(2k)^2 - 4(1)(N)}}{2} = k \pm \sqrt{k^2 - 2N \ln 2}$$

$$\boxed{m \approx \frac{N \ln 2}{k}}$$

We want the negative solution since it is the smaller positive number.

5x5

## Problem 3



*Explain*  
-3

$$P(W) = p(1-p) = P(L) \approx \chi_2$$

Since HH and TT are discarded each new round of two tosses has the same probability

b)

$$P(T) = p^2 + (1-p)^2 = (p+1-p)^2$$

$$\begin{aligned} P(\text{TIE}^{\text{TOTAL}}) &= P(T)^N = \left(p^2 + (1-p)^2\right)^N \\ &= \left(p^2 + 1 - 2p + p^2\right)^N \\ &= \underbrace{\left(2p^2 - 2p + 1\right)}_{\text{Less than } 1 \text{ since } p < 1}^N \end{aligned}$$

*Explain*  
-2

$$p^2 < p$$

$$\Rightarrow \lim_{N \rightarrow \infty} (2p^2 - 2p + 1)^N = 0$$

## Problem 4

(2c)  
/2c

a) Disjoint  $\Rightarrow P(A \cap B) = 0$

Independent  $\Rightarrow P(A \cap B) = P(A)P(B)$

$P(A \cap B) = P(A)P(B) = 0$   
IFF  $P(A) = 0 \vee P(B) = 0$

b)  $P(A \cap \bar{B}) = P(A) - P(A \cap B)$

$P(A \cap \bar{B}) = P(A) - P(A \cap B)$

$P(A \cap \bar{B}) = P(A) - P(A)P(B)$

$= P(A)(1 - P(B))$

$P(A \cap \bar{B}) = P(A)P(\bar{B})$

 $\Rightarrow A, \bar{B}$  independent

## c) Example

We define

 $X = 1, 2, 3, \text{ or } 4$  with equal probability of  $1/4$ 

Let:

$$\begin{array}{ll}
 \text{not independent} & \\
 \begin{cases} 
 A: X = 1 \text{ or } 2 \\ 
 B: X = 1 \text{ or } 3 \\ 
 C: X = 1 \text{ or } 4 
 \end{cases} & \rightarrow B \cup C: X = 1 \text{ or } 3 \text{ or } 4
 \end{array}
 \quad \Pr(A \cap B \cup C) = 1/4 \neq 3/8$$

## d) We know:

$P(C \cap A) = P(C)P(A)$

$P(C \cap B) = P(C)P(B)$

$P(C \cap (A \cap B)) = P(C)P(A \cap B)$

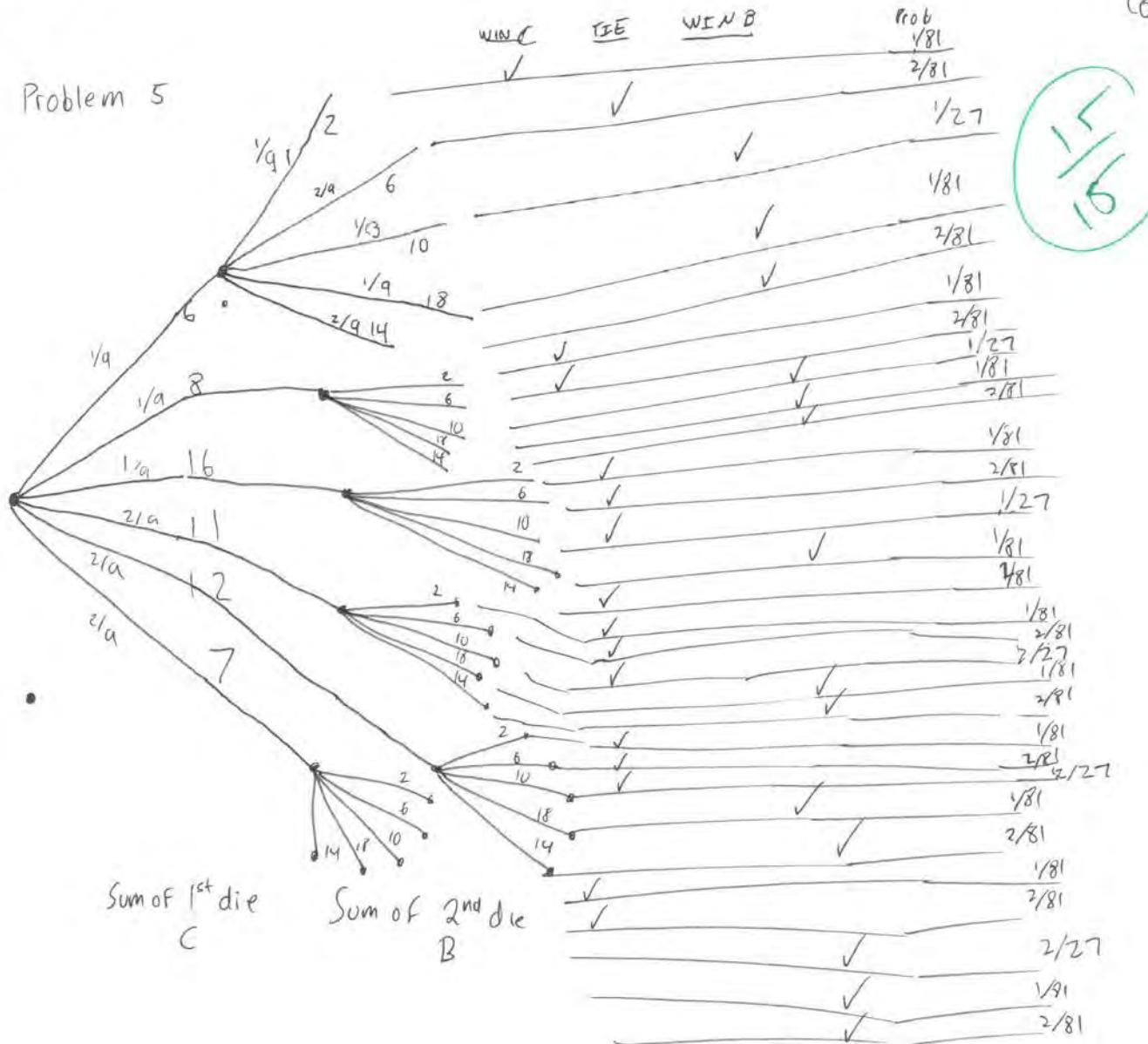
$P(A \cup B | C) = P(A) + P(B) - P(A \cap B) \quad \leftarrow \begin{matrix} \text{derived} \\ \text{in class!} \end{matrix}$

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$P(A \cup B) = P(A \cup B | C)$

 $\Rightarrow C$  is independent of  $A \cup B$

### Problem 5



$$\Pr(\text{C}_\text{win}) = 42/81$$

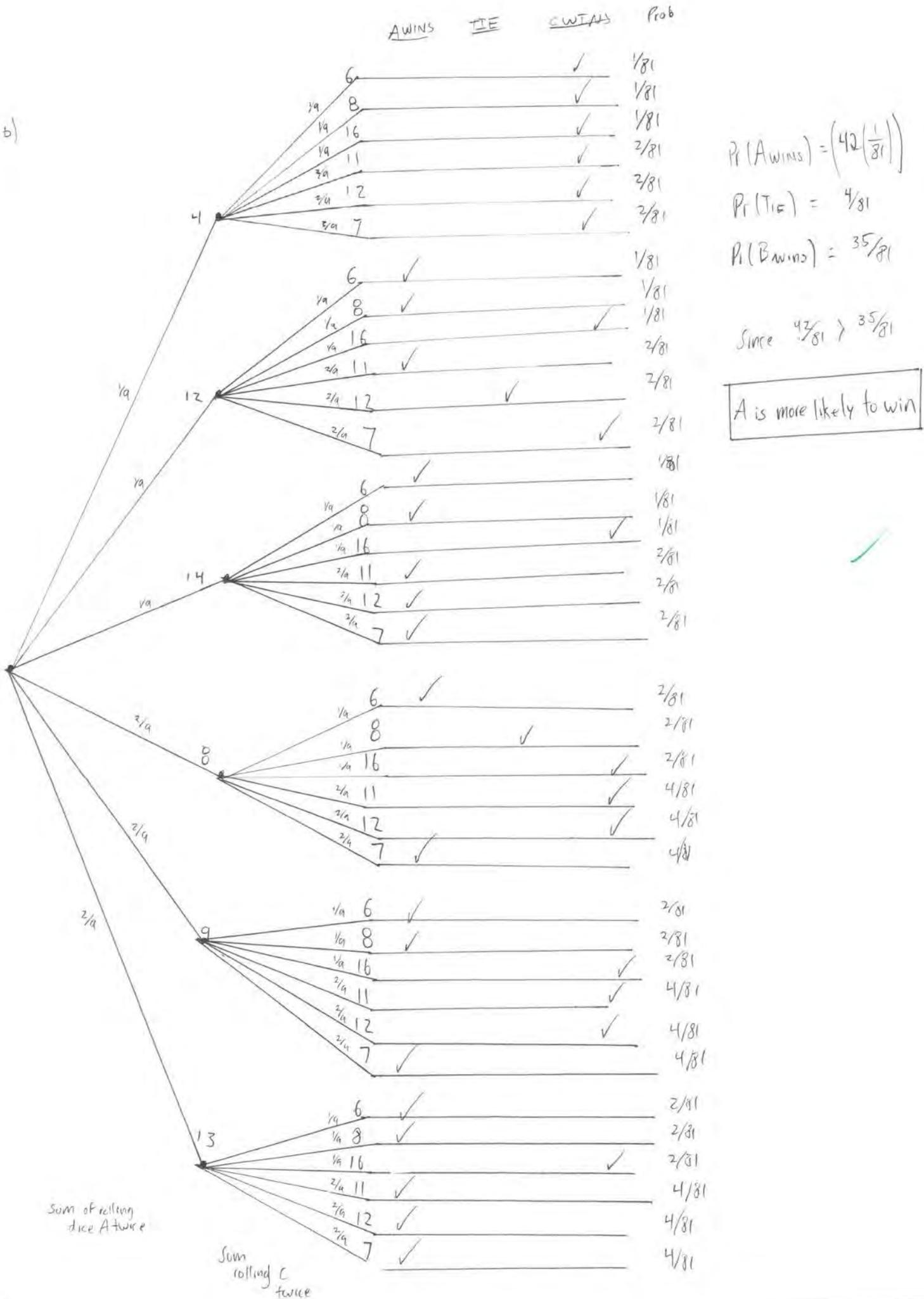
$$\Pr(\text{Tie}) = 2/81$$

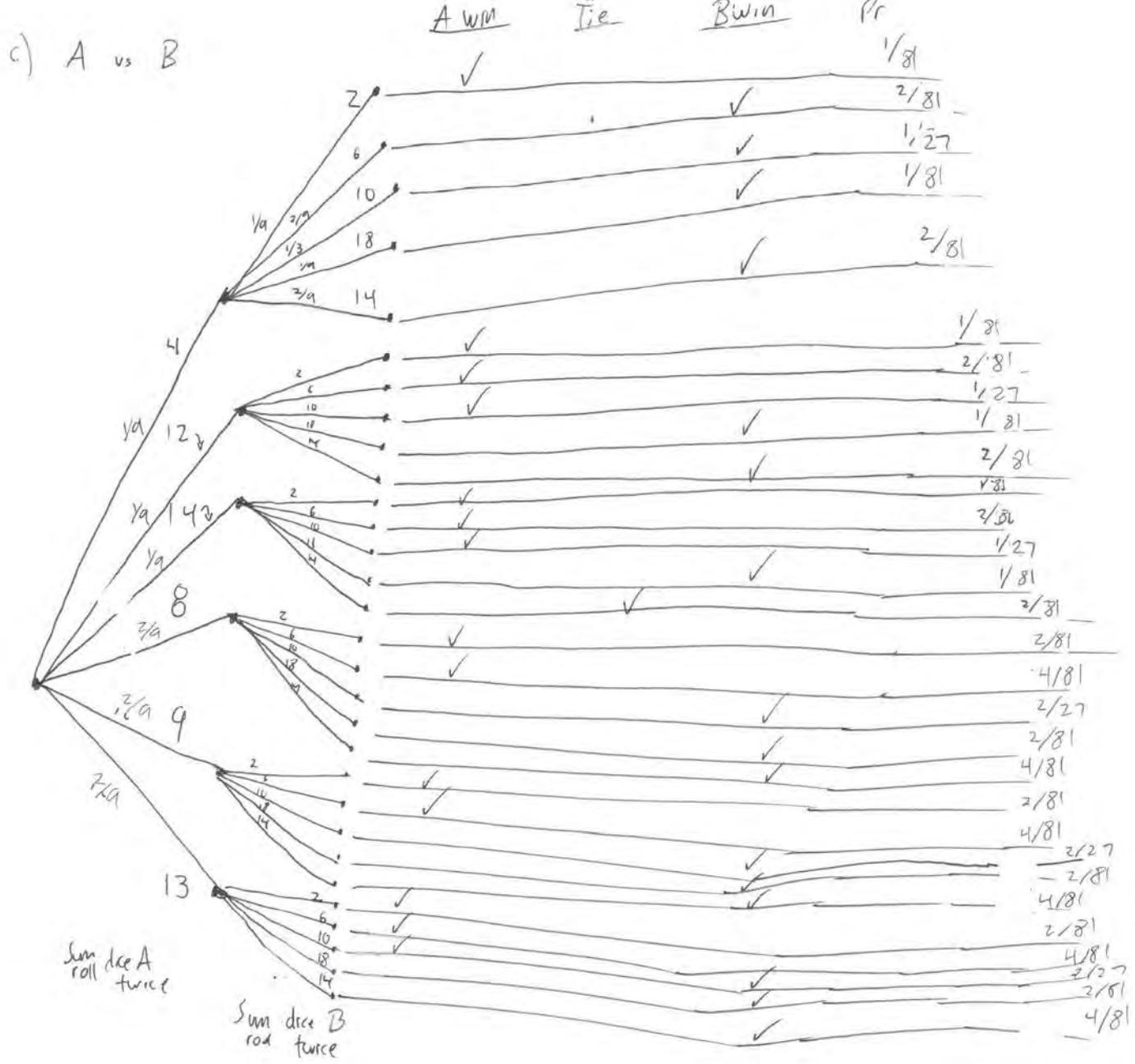
$$\Pr(B_{wm}) = 37/81$$

$$\text{Since } \left(\frac{42}{81}\right) > \left(\frac{37}{81}\right)$$

C is more likely to win

b)





$$\Pr(A_{\text{win}}) = \left( 25 \left( \frac{1}{81} \right) + 4 \left( \frac{1}{27} \right) \right) = .45 \quad P(\text{tie}) = \frac{2}{81}$$

$$\Pr(B_{\text{win}}) = .51$$

Since  $.51 > .45$

B is more likely to win

(15/15) JK

## Problem 6

- a) We consider any sequence of flips that leads to the sequence HHT. The inverse of this, replacing heads with tails and vice versa, yields TTH. There exists a bijection between these two sequences so their two probabilities must be equal. Since either sequence must come before the other one, the probabilities must sum to one.

By this argument:  $\Pr(\text{HHT}_{\text{first}}) = 1/2$

- b) Let  $A = \text{event where you get the sequence HHT before you see TTH}$

$$\Pr(A) = \Pr(A|H)\Pr(H) + \Pr(A|T)\Pr(T) \quad \text{Total probability}$$

$$\Pr(A) = \Pr(A|H)\frac{1}{2} + \underbrace{\Pr(A|T)}_{(\rightarrow \Pr(A) \sim \text{"reset"})} \frac{1}{2}$$

$$\Pr(A|T) = \Pr(A|H)\left(\frac{1}{2}\right) + \Pr(A|HT)\left(\frac{1}{2}\right)$$

$$\left(\frac{1}{2}\right)\Pr(A|T) = \Pr(A|H)\left(\frac{1}{2}\right)$$

Therefore:  $\Pr(A|T) = \Pr(A|H)$

$$\Pr(A|H) = \underbrace{\Pr(A|HH)\Pr(H)}_{=1, \text{ can either get H and repeat or T and win.}} + \Pr(A|HT)\Pr(T) \quad \text{Total probability}$$

$$(1) \quad \Pr(A|H) = \frac{1}{2} + \underbrace{\left(\frac{1}{2}\right)\Pr(A|HT)}_{\Pr(A|H)}$$

$$\Pr(A|HT) = \Pr(A|HTH)\Pr(H) + \Pr(A|HTT)\Pr(T) \quad \text{Total probability}$$

$$\Pr(A|HT) = \left(\frac{1}{2}\right) \underbrace{\Pr(A|HTH)}_{\Pr(A|H)}$$

$$\Pr(A|HT) = \left(\frac{1}{2}\right) \Pr(A|H)$$

$$\text{so from (1)} \quad \Pr(A|H) = \left(\frac{1}{2}\right) + \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\Pr(A|H)$$

$$\frac{3}{4}\Pr(A|H) = \frac{1}{2}$$

$$\Pr(A|H) = \frac{1}{2} \cancel{\left(\frac{1}{2}\right)} = \frac{1}{2} = \frac{2}{3}$$

$$\Pr(A|H) = \Pr(A|T) = \Pr(A)$$

$$\boxed{\cancel{\frac{2}{3}}}$$

## Problem 1

a) Expected net gain = 0

$$\text{Expected } \$ = \underbrace{.5(2N)}_{\text{double}} + \underbrace{.5(0)}_{\text{nothing}} = \boxed{N}$$

$$\text{Var} = E((X-N)^2)$$

$$\frac{1}{2}(2N-N)^2 + \frac{1}{2}(-N)^2$$

$$\frac{1}{2}(4N^2 - 4N^2 + N^2) + \frac{1}{2}N^2 = \boxed{N^2}$$

b) Let  $M_i = \$\$ in stock; . \quad \text{Expected } \$ = \sum_{i=1}^N \underbrace{E(M_i)}_i = \boxed{N}$

$$2 \cdot \frac{1}{2} + \frac{1}{2} \cdot 0 = 1$$

$$\text{Var}(\sum M_i) = \sum \text{Var}(M_i)$$

<sup>all independent</sup>

$$\text{Var}(M_i) \frac{1}{2}(2-1)^2 + \frac{1}{2}(-1)^2$$

$$\frac{1}{2} + \frac{1}{2} = 1$$

$$\sum_{i=1}^N \underbrace{\text{Var}(M_i)}_i = \boxed{N}$$

c) Option b) has less variance and so less risk

d) Expected  $\$ = \frac{1}{6}(1) + \frac{1}{6}(2) + \frac{1}{6}(3) + \frac{1}{6}(4) + \frac{1}{6}(5) + \frac{1}{6}(6) = \boxed{3.5}$

$$\text{Var}(R) = E((R-E(R))^2)$$

$$= \frac{2.5^2 + 1.5^2 + .5^2 + .5^2 + 1.5^2 + 2.5^2}{6} = \boxed{2.916666}$$

e)  $\frac{1}{6}(1+8+27+64+125+216) = 441/6$

$$\text{Since } \text{Var}(X) = E(X^2) - E^2[X]$$

$$= \frac{1}{6}(1+64+27^2+64^2+125^2+216^2) - \boxed{(441/6)^2} = \boxed{5792.9}$$

## Problem 2

a) Proposition only False if all terms False so

$$1 - \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{7}{8}$$

So # True propositions is  $\sum_{i=1}^7 T_i$

$$\begin{aligned} E(T_1 + \dots + T_7) &= E(T_1) + \dots + E(T_7) \\ &= \frac{7}{8} + \frac{7}{8} + \dots + \frac{7}{8} \\ &= \boxed{\frac{49}{8}} \end{aligned}$$

b) since  $r, v T$  can't be less than  $E(T)$

$$\exists T \mid T_1 + \dots + T_7 \geq 6\frac{1}{8}$$

$\Rightarrow T_1 + \dots + T_7 = 7$  w/ some assignment  
to variables that makes all propositions true.

### Problem 3

goal: calculate  $n$

$$tolerance \\ t = 5/8$$

$H$  = # heads in first  $n$  flips

Fair coin

$$\Pr(H > \frac{5}{8}n) \leq 0.05$$

so we can be 95% sure

$$\Rightarrow \text{CDF}_H(\frac{5}{8}n) \geq .95$$

↑  
cumulative  
distribution function

Biased coin

$$\Pr(H \leq (\frac{5}{8})n) \leq .05$$

$$\Rightarrow \text{CDF}_H(\frac{5}{8}n) \leq .95$$



Now we would have to find the minimum  $n$  that satisfies both of these equations using methods explained in the book to calculate the binomial cumulative distribution function.

## Problem 4

a) Not independent since the value of previous outcomes can tell me about the # of balls in other boxes.

Say for example that I tell you that  $n-1$  boxes are empty, now I know that the other box has  $n$  balls.

In order for box  $i$  to be empty, all other  $n$  balls must land in another box.

$$\Pr(\text{1 ball does not land in box } i) = \frac{(n-1)}{n} = 1 - \frac{1}{n}$$

Since balls are thrown independently,

$$\Pr(X_i = 1) = \left(1 - \frac{1}{n}\right)^n$$

b)  $E(\#\text{empty boxes}) = nE(X_i) = n\left(1 - \frac{1}{n}\right)^n$

wolfram  $\sim n \cdot \frac{1}{e}$

so  $\boxed{c = \frac{1}{e}}$

c)

$$\Pr(\text{k balls Fall in First box}) = \left(\frac{1}{n}\right)^k$$

So the event that at least  $k$  balls land in the first box can be easily calculated using the Union Bound

$$\Pr(\text{at least } k \text{ balls Fall in First box}) \leq \binom{n}{k} \cdot \left(\frac{1}{n}\right)^k$$

d)  $\Pr(R \geq k) \leq \frac{n}{k!}$

$$\Pr(R \geq k) \leq n \Pr(\text{at least } k \text{ balls fall in First box})$$

$$= n \left( \frac{n(n-1) \dots (n-k+1)}{k! n^k} \right)$$

$$= \frac{n}{k!} \left( \dots \right)$$

$$\leq \frac{n}{k!} \quad \text{so} \quad \boxed{\Pr(R \geq k) \leq \frac{n}{k!}}$$

$$e) \lim_{n \rightarrow \infty} P_e \{ R \geq n^\epsilon \} = 0 \quad \forall \epsilon > 0$$

From last problem

$$P_e \{ R \geq k \} \leq \frac{n}{k!}$$

By stirling's formula

$$\begin{aligned} &\approx \frac{n}{\sqrt{2\pi k} \left(\frac{k}{e}\right)^k} \leq \left(\frac{n}{ke}\right)^k = \frac{n e^k}{k^k} \\ &= \frac{e^{k+\ln n}}{e^{k \ln k}} \end{aligned}$$

$$\text{Let } k = n^\epsilon$$

$$= \frac{e^{n^\epsilon + \ln n}}{e^{k \ln n^\epsilon}}$$

$$n^\epsilon + \ln n = o(n^\epsilon \ln n^\epsilon)$$

So

$$\lim_{n \rightarrow \infty} \frac{e^{n^\epsilon + \ln n}}{e^{n^\epsilon \ln n^\epsilon}} = 0$$

## Problem 5

a) Thm

$$\mathbb{E}\left(\frac{1}{X}\right) \geq \frac{1}{\mathbb{E}(X)}$$

Proof (By induction on # of outcomes of  $X$ )

Base Case  $X$  has one outcome  $x$ ,

$$\mathbb{E}\left(\frac{1}{X}\right) = \frac{1}{x} = \frac{1}{\mathbb{E}(x)} \quad \checkmark$$

Inductive Step

Let  $Y$  be a r.v. of first  $n$  outcomes with probabilities

$$\frac{p_1}{1-p_{n+1}}, \dots, \frac{p_n}{1-p_{n+1}}$$

(normalized)

$$\frac{1}{\sum_{i=1}^n \frac{p_i}{1-p_{n+1}} x_i} = \frac{1}{\mathbb{E}(Y)} \leq \mathbb{E}\left(\frac{1}{Y}\right)$$

$$= \sum_{i=1}^n \frac{p_i}{(1-p_{n+1}) x_i}$$

$$\begin{aligned} \sum_{i=1}^{n+1} \frac{p_i x_i}{1-p_{n+1}} &= p_{n+1} x_{n+1} + \sum_{i=1}^n \frac{p_i x_i}{1-p_{n+1}} \\ &= p_{n+1} x_{n+1} + (1-p_{n+1}) \sum_{i=1}^n \frac{p_i}{1-p_{n+1}} x_i \\ &\leq \frac{p_{n+1}}{x_{n+1}} + (1-p_{n+1}) \sum_{i=1}^n \frac{p_i}{1-p_{n+1}} x_i \\ &\leq \frac{p_{n+1}}{x_{n+1}} + (1-p_{n+1}) \sum_{i=1}^n \frac{p_i}{(1-p_{n+1}) x_i} \\ &= \sum_{i=1}^{n+1} \frac{p_i}{x_i}. \quad \square \end{aligned}$$

$$b) \mathbb{E}\left(\frac{R}{T}\right) = \mathbb{E}\left(R \cdot \frac{1}{T}\right) = \mathbb{E}(R) \mathbb{E}\left(\frac{1}{T}\right) = \left(\frac{1}{\mathbb{E}(T)}\right) \mathbb{E}(R) \quad \square$$

## Problem 6

We will use linearity of Expectation

We define  $T$  as the # of rolls to get all 6 numbers.

$t_i = \# \text{ of rolls until we see the } i^{\text{th}} \text{ number.}$

From these definitions

$$T = \sum_{i=0}^6 t_i$$

For each  $t_i$ , we will roll the die until we see one of the  $7-i$  remaining numbers.

So on each roll

$$\Pr(\text{New \#}) = \frac{7-i}{6}$$

$$E(t_i) = \frac{6}{7-i}$$

$$\begin{aligned} E[T] &= \sum_{i=0}^6 E[t_i] \\ &= \boxed{14.7} \end{aligned}$$

## Problem 7

Let  $X_i$  be the indicator variable for the event that the  $i$ th element in the vector is larger than all the previous elements.

We will update the current max

$$\sum_{i=1}^N X_i$$

We can use linearity of expectation to find

$$E[X_1 + \dots + X_n]$$

$$\Pr(X_i = 1) = \frac{1}{i}$$

$\approx \text{largest}$

$$E(x) = \sum_i \Pr(X_i = 1)$$

$$\boxed{\sum_{i=1}^n \frac{1}{i}}$$

$$\approx \ln n$$

## Problem 8

We define  $\sigma^2 = \text{Var}(\lambda)$

$$\sigma_e^2 = \frac{\sum y_i^2}{n} - \left( \frac{\sum y_i}{n} \right)^2$$

$$E(\sigma_e^2) = E \left[ \frac{\sum y_i^2}{n} - \left( \frac{\sum y_i}{n} \right)^2 \right]$$

$$= \frac{\sum E(y_i^2)}{n} - \frac{E(\sum y_i)^2}{n^2}$$

$$= \underbrace{\sum}_{n} (\sigma^2 + E(x)^2) - \text{Var}(\sum y_i) + E^2(\sum y_i)$$

$$= \frac{n(\sigma^2 + E(x)^2)}{n} - \frac{n\sigma^2 + n^2 E(x)^2}{n^2}$$

$$= \sigma^2 \left( 1 - \frac{1}{n} \right)$$

Biased

To make unbiased, we multiply it by  $\frac{n}{n-1}$

$$\underbrace{| \overline{E \left( \frac{n\sigma^2}{n-1} \right)} = \sigma^2 |}_{\text{Biased}}$$