Final Exam

YOUR NAME: _____

- You may use **two** $8.5 \times 11''$ sheets with notes in your own handwriting on both sides, but no other reference materials. Calculators are not allowed.
- You may assume all results presented in class.
- 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.

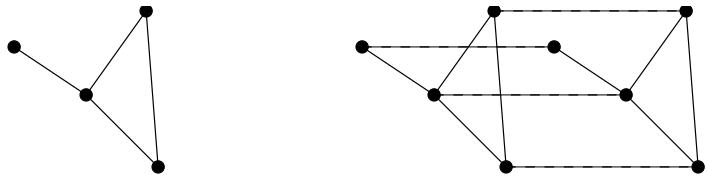
GOOD LUCK!

Problem	Points	Grade	Grader
1	13		
2	15		
3	12		
4	12		
5	12		
6	12		
7	12		
8	12		
Total	100		

Problem 1. [13 points] Give an inductive proof that the Fibonacci numbers F_n and F_{n+1} are relatively prime for all $n \geq 0$. The Fibonacci numbers are defined as follows:

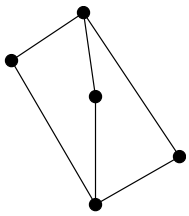
$$F_0 = 0 \qquad F_1 = 1 \qquad F_n = F_{n-1} + F_{n-2} \quad \text{(for } n \geq 2\text{)}$$

Problem 2. [15 points] The *double* of a graph G consists of two copies of G with edges joining corresponding vertices. For example, a graph appears below on the left and its double appears on the right.



Some edges in the graph on the right are dashed to clarify its structure.

(a) Draw the double of the graph shown below.



(b) Suppose that G_1 is a bipartite graph, G_2 is the double of G_1, G_3 is the double of G_2, and so forth. Use induction on n to prove that G_n is bipartite for all $n \geq 1$.

Problem 3. [12 points] *Finalphobia* is a rare disease in which the victim has the delusion that he or she is being subjected to an intense mathematical examination.

- A person selected uniformly at random has finalphobia with probability $1/100$.
- A person with finalphobia has shaky hands with probability $9/10$.
- A person without finalphobia has shaky hands with probability $1/20$.

What is the probablility that a person selected uniformly at random has finalphobia, given that he or she has shaky hands?

Problem 4. [12 points] Suppose that you roll five 6-sided dice that are fair and mutually independent. For the problems below, answers alone are sufficient, but we can award partial credit only if you show your work. Also, you do not need to simplify your answers; you may leave factorials, binomial coefficients, and arithmetic expressions unevaluated.

(a) What is the probability that all five dice show different values?

Example: $(1, 2, 3, 4, 5)$ is a roll of this type, but $(1, 1, 2, 3, 4)$ is not.

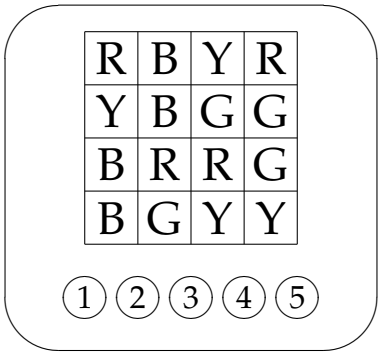
(b) What is the probability that two dice show the same value and the remaining three dice all show different values?

Example: $(6, 1, 6, 2, 3)$ is a roll of this type, but $(1, 1, 2, 2, 3)$ and $(4, 4, 4, 5, 6)$ are not.

(c) What is the probability that two dice show one value, two different dice show a second value, and the remaining die shows a third value?

Example: $(6, 1, 2, 1, 2)$ is a roll of this type, but $(4, 4, 4, 4, 5)$ and $(5, 5, 5, 6, 6)$ are not.

Problem 5. [12 points] An electronic toy displays a 4×4 grid of colored squares. At all times, four are red, four are green, four are blue, and four are yellow. For example, here is one possible configuration:



For parts (a) and (b) below, you need not simplify your answers.

(a) How many such configurations are possible?

(b) Below the display, there are five buttons numbered 1, 2, 3, 4, and 5. The player may press a sequence of buttons; however, the same button can not be pressed twice in a row. How many different sequences of n button-presses are possible?

(c) Each button press scrambles the colored squares in a complicated, but nonrandom way. Prove that there exist two *different* sequences of 32 button presses that both produce the *same* configuration, if the puzzle is initially in the state shown above. (Hint: $4^{32} = 16^{16} > 16!$)

.

Problem 6. [12 points] MIT students sometimes delay laundry for a few days. Assume all random values described below are mutually independent.

(a) A ***busy*** student must complete 3 problem sets before doing laundry. Each problem set requires 1 day with probability $2/3$ and 2 days with probability $1/3$. Let B be the number of days a busy student delays laundry. What is $\mathrm{Ex}(B)$?

Example: If the first problem set requires 1 day and the second and third problem sets each require 2 days, then the student delays for $B = 5$ days.

(b) A ***relaxed*** student rolls a fair, 6-sided die in the morning. If he rolls a 1, then he does his laundry immediately (with zero days of delay). Otherwise, he delays for one day and repeats the experiment the following morning. Let R be the number of days a relaxed student delays laundry. What is $\mathrm{Ex}(R)$?

Example: If the student rolls a 2 the first morning, a 5 the second morning, and a 1 the third morning, then he delays for $R = 2$ days.

(c) Before doing laundry, an ***unlucky*** student must recover from illness for a number of days equal to the product of the numbers rolled on two fair, 6-sided dice. Let U be the expected number of days an unlucky student delays laundry. What is $\mathrm{Ex}(U)$?

Example: If the rolls are 5 and 3, then the student delays for $U = 15$ days.

(d) A student is ***busy*** with probability $1/2$, ***relaxed*** with probability $1/3$, and ***unlucky*** with probability $1/6$. Let D be the number of days the student delays laundry. What is $\mathrm{Ex}(D)$? *Leave your answer in terms of* $\mathrm{Ex}(B)$, $\mathrm{Ex}(R)$, *and* $\mathrm{Ex}(U)$.

Problem 7. [12 points] I have twelve cards:

1	1	2	2	3	3	4	4	5	5	6	6

I shuffle them and deal them in a row. For example, I might get:

1	2	3	3	4	6	1	4	5	5	2	6

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.

We can award partial credit only if you show your work.

Problem 8. [12 points] Each time a baseball player bats, he hits the ball with some probability. The table below gives the hit probability and number of chances to bat next season for five players.

player	prob. of hit	# chances to bat
Player A	$1/3$	300
Player B	$1/4$	200
Player C	$1/4$	400
Player D	$1/5$	250
Player E	$2/5$	500

(a) Let X be the total number times these five players hit the ball next season. What is $\mathrm{Ex}(X)$?

(b) Give a nontrivial upper bound on $\Pr(X \geq 1500)$ and justify your answer. *Do not* assume that hits happen mutually independently.

(c) Using a Chernoff inequality, give a nontrivial upper bound on $\Pr(X \leq 400)$. For this part, you *may* assume that all hits happen mutually independently.

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Final Exam

YOUR NAME: _____

:

- You may use **two** $8.5 \times 11''$ sheets with notes in your own handwriting on both sides, but no other reference materials. Calculators are not allowed.
- You may assume all results presented in class.
- 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.

GOOD LUCK!

Problem	Points	Grade	Grader
1	13		
2	15		
3	12		
4	12		
5	12		
6	12		
7	12		
8	12		
Total	100		

Problem 1. [13 points] Give an inductive proof that the Fibonacci numbers F_n and F_{n+1} are relatively prime for all $n \geq 0$. The Fibonacci numbers are defined as follows:

$$F_0 = 0 \qquad F_1 = 1 \qquad F_n = F_{n-1} + F_{n-2} \quad \text{(for } n \geq 2\text{)}$$

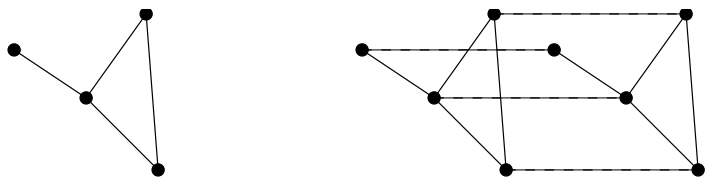
Solution. We use induction on n. Let $P(n)$ be the proposition that F_n and F_{n+1} are relatively prime.

Base case: $P(0)$ is true because $F_0 = 0$ and $F_1 = 1$ are relatively prime.

Inductive step: Assume that $P(n)$ is true where $n \geq 0$; that is, F_n and F_{n+1} are relatively prime. We must show that F_{n+1} and F_{n+2} are relatively prime as well. If F_{n+1} and F_{n+2} had a common divisor $d > 1$, then d would also divide the linear combination $F_{n+2} - F_{n+1} = F_n$, contradicting the assumption that F_n and F_{n+1} are relatively prime. So F_{n+1} and F_{n+2} are relatively prime.

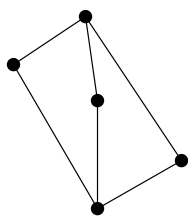
The theorem follows by induction.

Problem 2. [15 points] The *double* of a graph G consists of two copies of G with edges joining corresponding vertices. For example, a graph appears below on the left and its double appears on the right.

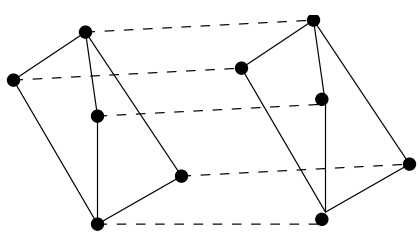


Some edges in the graph on the right are dashed to clarify its structure.

(a) Draw the double of the graph shown below.



Solution.



(b) Suppose that G_1 is a bipartite graph, G_2 is the double of G_1, G_3 is the double of G_2, and so forth. Use induction on n to prove that G_n is bipartite for all $n \geq 1$.

Solution. We use induction. Let $P(n)$ be the proposition that G_n is bipartite.

Base case: $P(1)$ is true because G_1 is bipartite by assumption.

Inductive step: For $n \geq 1$, we assume $P(n)$ in order to prove $P(n+1)$. The graph G_{n+1} consists of two subgraphs isomorphic to G_n with edges joining corresponding vertices. Remove these extra edges. By the assumption $P(n)$, we can color each vertex of one subgraph black or white so that adjacent vertices get different colors. If we color the corresponding vertices in the other subgraph oppositely, then adjacent vertices get different colors within that subgraph as well. And now if we add back the extra edges, each of these joins a white vertex and a black vertex. Therefore, G_{n+1} is bipartite.

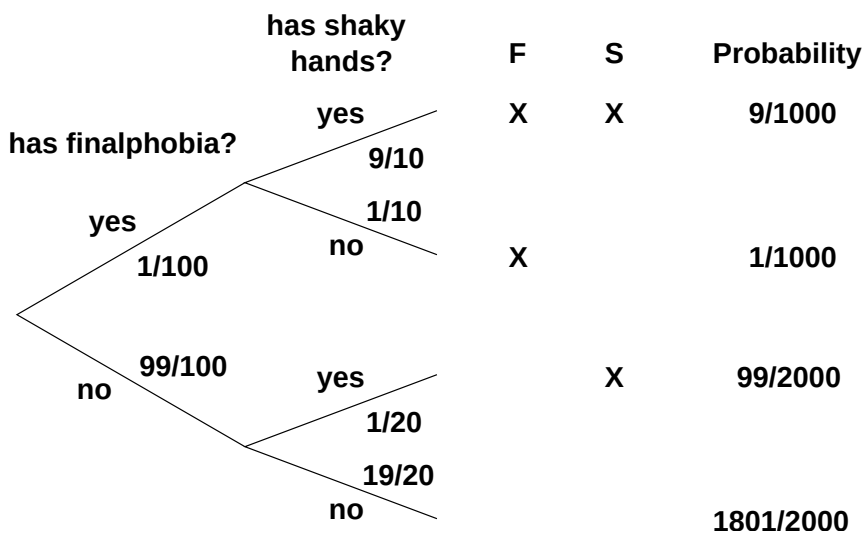
The theorem follows from the principle of induction.

Problem 3. [12 points] *Finalphobia* is a rare disease in which the victim has the delusion that he or she is being subjected to an intense mathematical examination.

- A person selected uniformly at random has finalphobia with probability $1/100$.
- A person with finalphobia has shaky hands with probability $9/10$.
- A person without finalphobia has shaky hands with probability $1/20$.

What is the probablility that a person selected uniformly at random has finalphobia, given that he or she has shaky hands?

Solution. Let F be the event that the randomly-selected person has finalphobia, and let S be the event that he or she has shaky hands. A tree diagram is worked out below:



The probability that a person has finalphobia, given that he or she has shaky hands is:

$$\begin{aligned}\Pr(F \mid S) &= \frac{\Pr(F \cap S)}{\Pr(S)} \\ &= \frac{9/1000}{9/1000 + 99/2000} \\ &= \frac{18}{18 + 99} \\ &= \frac{18}{117}\end{aligned}$$

Problem 4. [12 points] Suppose that you roll five 6-sided dice that are fair and mutually independent. For the problems below, answers alone are sufficient, but we can award partial credit only if you show your work. Also, you do not need to simplify your answers; you may leave factorials, binomial coefficients, and arithmetic expressions unevaluated.

(a) What is the probability that all five dice show different values?

Example: $(1, 2, 3, 4, 5)$ is a roll of this type, but $(1, 1, 2, 3, 4)$ is not.

Solution. The probability space is the uniform distribution on the 6^5 possible numbers rolled on the five (distinguishable) dice. So the probability that all dice are different is the number of outcomes in which the dice have distinct values divided by 6^5. There are $(6)_5 = 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2$ such outcomes so

$$\Pr(\text{all rolls distinct}) = \frac{6 \cdot 5 \cdot 4 \cdot 3 \cdot 2}{6^5} = \frac{5!}{6^4}$$

An alternative approach uses the observation that the conditional probability that the $i+1$st die value differs from the preceding rolls, given that the first i values differ, is $(6-i)/6$ for $1 \leq i \leq 4$, and the probability that all five values are different is the product of these conditional probabilities, namely

$$\Pr(\text{all rolls distinct}) = \frac{5}{6} \cdot \frac{4}{6} \cdot \frac{3}{6} \cdot \frac{2}{6} = \frac{5!}{6^4}$$

(b) What is the probability that two dice show the same value and the remaining three dice all show different values?

Example: $(6, 1, 6, 2, 3)$ is a roll of this type, but $(1, 1, 2, 2, 3)$ and $(4, 4, 4, 5, 6)$ are not.

Solution. There are $\binom{5}{2}$ possible pairs of rolls that might have the same value and 6 possibilities for what this value is. There $5 \cdot 4 \cdot 3$ possible distinct values for the remaining three rolls. So

$$\Pr(\text{exactly two values the same}) = \frac{\binom{5}{2} \cdot 6 \cdot 5 \cdot 4 \cdot 3}{6^5} = \frac{100}{6^3}$$

An alternative way to count is: there are $\binom{6}{4}$ sets of four values among the five dice, 4 choices for which of these values is repeated, and by the Bookkeeper rule, $\binom{5}{2,1,1,1} = 5!/2$ permutations of a sequence consisting of five values, one of which appears twice. So,

$$\Pr(\text{exactly two values the same}) = \frac{\binom{6}{4} \cdot 4 \cdot 5!/2}{6^5} = \frac{100}{6^3}$$

(c) What is the probability that two dice show one value, two different dice show a second value, and the remaining die shows a third value?

Example: $(6, 1, 2, 1, 2)$ is a roll of this type, but $(4, 4, 4, 4, 5)$ and $(5, 5, 5, 6, 6)$ are not.

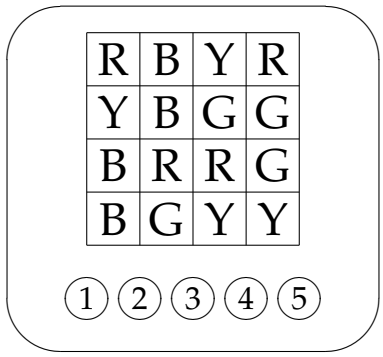
Solution. There are $\binom{6}{2}$ sets of two values that might be duplicated. There are $\binom{5}{2}$ rolls where larger duplicated value may come up and $\binom{3}{2}$ remaining rolls where the smaller duplicated value may come up. There is only 1 remaining roll where the nonduplicated value may then come up, and 4 remaining values it could take. So,

$$\Pr\left(\text{exactly two pairs of same values}\right) = \frac{\binom{6}{2} \cdot \binom{5}{2} \cdot \binom{3}{2} \cdot 4}{6^5} = \frac{50}{6^3}$$

Alternatively, there are $\binom{6}{3}$ sets of three values among the five dice, 3 choices for which of these values is not repeated, and by the Bookkeeper rule, $\binom{5}{2,2,1}$ permutations of a sequence consisting of five values, two of which appear twice. So,

$$\Pr\left(\text{exactly two pairs of same values}\right) = \frac{\binom{6}{3} \cdot 3 \cdot \binom{5}{2,2,1}}{6^5} = \frac{50}{6^3}$$

Problem 5. [12 points] An electronic toy displays a 4×4 grid of colored squares. At all times, four are red, four are green, four are blue, and four are yellow. For example, here is one possible configuration:



For parts (a) and (b) below, you need not simplify your answers.

(a) How many such configurations are possible?

Solution. This is equal to the number of sequences containing 4 R's, 4 G's, 4 B's, and 4 Y's, which is:

$$\frac{16!}{(4!)^4}$$

(b) Below the display, there are five buttons numbered 1, 2, 3, 4, and 5. The player may press a sequence of buttons; however, the same button can not be pressed twice in a row. How many different sequences of n button-presses are possible?

Solution. There are 5 choices for the first button press and 4 for each subsequence press. Therefore, the number of different sequences of n button presses is $5 \cdot 4^{n-1}$.

(c) Each button press scrambles the colored squares in a complicated, but nonrandom way. Prove that there exist two *different* sequences of 32 button presses that both produce the *same* configuration, if the puzzle is initially in the state shown above. (Hint: $4^{32} = 16^{16} > 16!$)

Solution. We use the Pigeonhole Principle. Let A be the set of all sequences of 32 button presses, let B be the set of all configurations, and let $f : A \to B$ map each sequence of button presses to the configuration that results. Now:

$$|A| > 4^{32} > 16! > |B|$$

Thus, by the Pigeonhole Principle, f is not injective; that is, there exist distinct elements $a_1, a_2 \in A$ such taht $f(a_1) = f(a_2)$. In other words, there are two different sequences of button presses that produce the same configuration.

Problem 6. [12 points] MIT students sometimes delay laundry for a few days. Assume all random values described below are mutually independent.

(a) A ***busy*** student must complete 3 problem sets before doing laundry. Each problem set requires 1 day with probability $2/3$ and 2 days with probability $1/3$. Let B be the number of days a busy student delays laundry. What is $\text{Ex}(B)$?

Example: If the first problem set requires 1 day and the second and third problem sets each require 2 days, then the student delays for $B = 5$ days.

Solution. The expected time to complete a problem set is:

$$1 \cdot \frac{2}{3} + 2 \cdot \frac{1}{3} = \frac{4}{3}$$

Therefore, the expected time to complete all three problem sets is:

$$\begin{aligned}
\text{Ex}(B) &= \text{Ex}(\text{pset1}) + \text{Ex}(\text{pset2}) + \text{Ex}(\text{pset3}) \\
&= \frac{4}{3} + \frac{4}{3} + \frac{4}{3} \\
&= 4
\end{aligned}$$

(b) A ***relaxed*** student rolls a fair, 6-sided die in the morning. If he rolls a 1, then he does his laundry immediately (with zero days of delay). Otherwise, he delays for one day and repeats the experiment the following morning. Let R be the number of days a relaxed student delays laundry. What is $\text{Ex}(R)$?

Example: If the student rolls a 2 the first morning, a 5 the second morning, and a 1 the third morning, then he delays for $R = 2$ days.

Solution. If we regard doing laundry as a failure, then the mean time to failure is $1/(1/6) = 6$. However, this counts the day laundry is done, so the number of days delay is $6 - 1 = 5$. Alternatively, we could derive the answer as follows:

$$\begin{aligned}
\text{Ex}(R) &= \sum_{k=0}^{\infty} \Pr(R > k) \\
&= \frac{5}{6} + \left(\frac{5}{6}\right)^2 + \left(\frac{5}{6}\right)^3 + \ldots \\
&= \frac{5}{6} \cdot \left(1 + \frac{5}{6} + \left(\frac{5}{6}\right)^2 + \ldots\right) \\
&= \frac{5}{6} \cdot \frac{1}{1 - 5/6} \\
&= 5
\end{aligned}$$

(c) Before doing laundry, an ***unlucky*** student must recover from illness for a number of days equal to the product of the numbers rolled on two fair, 6-sided dice. Let U be the expected number of days an unlucky student delays laundry. What is $\text{Ex}(U)$?

Example: If the rolls are 5 and 3, then the student delays for $U = 15$ days.

Solution. Let D_1 and D_2 be the two die rolls. Recall that a die roll has expectation $7/2$. Thus:

$$\begin{aligned}\text{Ex}(U) &= \text{Ex}(D_1 \cdot D_2)\\ &= \text{Ex}(D_1) \cdot \text{Ex}(D_2)\\ &= \frac{7}{2} \cdot \frac{7}{2}\\ &= \frac{49}{4}\end{aligned}$$

(d) A student is ***busy*** with probability $1/2$, ***relaxed*** with probability $1/3$, and ***unlucky*** with probability $1/6$. Let D be the number of days the student delays laundry. What is $\text{Ex}(D)$? *Leave your answer in terms of* $\text{Ex}(B)$, $\text{Ex}(R)$, *and* $\text{Ex}(U)$.

Solution.

$$\text{Ex}(D) = \frac{1}{2}\text{Ex}(B) + \frac{1}{3}\text{Ex}(R) + \frac{1}{6}\text{Ex}(U)$$

Problem 7. [12 points] I have twelve cards:

1	1	2	2	3	3	4	4	5	5	6	6

I shuffle them and deal them in a row. For example, I might get:

1	2	3	3	4	6	1	4	5	5	2	6

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.

We can award partial credit only if you show your work.

Solution. Consider an adjacent pair. The left card matches only one of the other 11 cards, which is equally likely to be in any of the 11 other positions. Therefore, the probability that an adjacent pair matches is $1/11$. Since there are 11 adjacent pairs, the expected number of matches is $11 \cdot 1/11 = 1$ by linearity of expectation.

Problem 8. [12 points] Each time a baseball player bats, he hits the ball with some probability. The table below gives the hit probability and number of chances to bat next season for five players.

player	prob. of hit	# chances to bat
Player A	$1/3$	300
Player B	$1/4$	200
Player C	$1/4$	400
Player D	$1/5$	250
Player E	$2/5$	500

(a) Let X be the total number times these five players hit the ball next season. What is $\mathrm{Ex}(X)$?

Solution. The expected number of events that happen is equal to the sum of the event probabilities. So we have:

$$\begin{aligned}\mathrm{Ex}(X) &= 300 \cdot \frac{1}{3} + 200 \cdot \frac{1}{4} + 400 \cdot \frac{1}{4} + 250 \cdot \frac{1}{5} + 500 \cdot \frac{2}{5} \\ &= 100 + 50 + 100 + 50 + 200 \\ &= 500\end{aligned}$$

(b) Give a nontrivial upper bound on $\Pr(X \geq 1500)$ and justify your answer. *Do not* assume that hits happen mutually independently.

Solution. We use the Markov Inequality:

$$\Pr(X \geq c) \leq \frac{\mathrm{Ex}(X)}{c}$$

This gives:

$$\begin{aligned}\Pr(X \geq 1500) &\leq \frac{\mathrm{Ex}(X)}{1500} \\ &= \frac{1}{3}\end{aligned}$$

(c) Using a Chernoff inequality, give a nontrivial upper bound on $\Pr(X \leq 400)$. For this part, you *may* assume that all hits happen mutually independently.

Solution. We use the Chernoff Inequality:

$$\Pr(X \leq (1-\delta)\,\mathrm{Ex}(X)) \leq e^{-\delta^2\,\mathrm{Ex}(X)/2}$$

This gives:

$$\begin{aligned}\Pr(X \leq 400) &= \Pr(X \leq (1-1/5)\,\mathrm{Ex}(X)) \\ &\leq e^{-(1/5)^2 500/2} \\ &= e^{-10}\end{aligned}$$

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Final

- This final is **closed book**, but you may have three $8.5 \times 11''$ sheet with notes in your own handwriting on both sides.
- Calculators are not allowed.
- You may assume all of the results presented in class.
- 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.
- For this final, $\mathbb{N}$ is the set of nonnegative integers (including 0): $\mathbb{N} = \{0, 1, \ldots, \}$.
- GOOD LUCK!
- **Important:** If you show your reasoning, even if your answer is wrong, you could earn partial credit.

Problem	Points	Grade	Grader
1	8		
2	20		
3	8		
4	10		
5	10		
6	28		
7	16		
Total	100		

Problem 1. [8 points] Prove that for all $n \in \mathbb{N}$, the following identity holds

$$\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}.$$

Problem 2. [20 points] Coin-Flip is a 2 player game. Each player wins with probability exactly 0.5. There are no ties.

n people are playing a Coin-Flip tournament. Every person plays a Coin-Flip game with every other person exactly once. Thus everybody plays $n-1$ games. The outcomes of all the games are mutually independent of one another.

We say that the tournament is a *success* if for every $i \in \{0, 1, \ldots, n-1\}$, there is exactly one player, which we will refer to as p_i, with exactly i wins.

(a) [10 points] Prove that if the tournament is a success, then for any integers j, k with $0 \leq k < j \leq n-1$, p_j defeats p_k.

(b) [6 points] What is the probability that the tournament will be a success?

(c) [4 points] Show that your answer to part (b) is $o(1)$.

Problem 3. [8 points] A person is passing time by advancing a token on the set of natural numbers. In the beginning, a token is placed on 0.

The person keeps playing *moves* forever. Each move proceeds as follows:

1. First the person tosses a fair coin (with heads/tails equally likely).
2. Suppose the token is currently placed on n. If heads came up, then the person moves the token to $n+3$, otherwise he moves the token to $n+4$.

For each $n \in \mathbb{N}$, let E_n be the event "There was a move on which the token landed on n". Let $p_n = \Pr[E_n]$.

Find a recurrence relation for p_n. *You do not need to solve the recurrence, but you should specify the boundary conditions that would be necessary to find a solution to the recurrence.*

Problem 4. [10 points] Exactly $1/5$th of the people in a town have Beaver Fever©.

There are two tests for Beaver Fever, TEST1 and TEST2. When a person goes to a doctor to test for Beaver Fever, with probability $2/3$ the doctor conducts TEST1 on him and with probability $1/3$ the doctor conducts TEST2 on him.

When TEST1 is done on a person, the outcome is as follows:

- If the person has the disease, the result is positive with probability $3/4$.
- If the person does not have the disease, the result is positive with probability $1/4$.

When TEST2 is done on a person, the outcome is as follows:

- If the person has the disease, the result is positive with probability 1.
- If the person does not have the disease, the result is positive with probability $1/2$.

A person is picked uniformly at random from the town and is sent to a doctor to test for Beaver Fever. The result comes out positive. What is the probability that the person has the disease?

Problem 5. [10 points] Two identical complete decks of cards, each with 52 cards, have been mixed together. A hand of 5 cards is picked uniformly at random from amongst all subsets of exactly 5 cards.

(a) [5 points] What is the probability that the hand has no identical cards (i.e., cards with the same suit and value. For example, the hand $\langle Q\heartsuit, 5\spadesuit, 6\spadesuit, 8\clubsuit, Q\heartsuit\rangle$ *has* identical cards.)?

(b) [5 points] What is the probability that the hand has exactly one pair of identical cards?

Problem 6. [28 points] Scores for a final exam are given by picking an integer uniformly at random from the set $\{50, 51, \ldots, 97, 98\}$. The scores of all 128 students in the class are assigned in this manner. For parts (a), (b), (c) and (d) you may NOT assume that these scores are assigned independently. For parts (e), (f), (g) and (h) you MAY assume that these scores are assigned independently.

Let $S_1, \ldots, S_{128}$ be their scores. Let $S = \frac{1}{128}(\sum_{i=1}^{128} S_i)$ be the average score of the class.

(a) [3 points] For $i \in \{1, \ldots, 128\}$, what is $\mathbb{E}[S_i]$?

(b) [2 points] Show that $\mathbb{E}[S] = 74$. Make no independence assumptions.

(c) [4 points] Prove that

$$\Pr[S \geq 88] \leq \frac{37}{44}.$$

Make no independence assumptions.

(d) [5 points] Improve your previous bound by using the fact that the minimum possible score is 50. Prove that

$$\Pr[S \geq 88] \leq \frac{12}{19}.$$

Make no independence assumptions.

(e) [4 points] For the remaining problems, assume that all the scores are assigned mutually independently. Use Problem 1 of this final to find $Var[S_i]$.

(f) [3 points] What is $Var[S]$?

(g) [2 points] What is the standard deviation of S?

(h) [5 points] Prove, using the Chebyshev Inequality, that

$$\Pr[S \leq 69] \leq \frac{1}{16}.$$

Problem 7. [16 points] 1000 files $F_1, F_2, \ldots, F_{1000}$ have just reached a disk manager for writing onto disk. Each file's size is between $0MB$ and $1MB$. The sum of all files' sizes is $400MB$.

The disk manager has 4 disks under its control. For each file F_i, the disk manager chooses a disk uniformly at random from amongst the 4 disks, and F_i is written to that disk. The choices of disk for the different files are mutually independent.

(a) [2 points] What is the expected number of files that will be written to the first disk?

We can use indicator variables. For each file, $P_i = 1$ if F_i is written to the first disk. The chance of an individual file being written to the first disk is 1/4. By linearity of expectation, the expected number of files written to the first disk is the sum of the expected values of P_i's. The expected value of each indicator variable is $1/4$, and $\sum_{i=1}^{(} 1000)1/4 = 250$, so the expected number of files to be written to the first disk is 250.

(b) [2 points] What is the expected number of bytes written on the first disk?

We can say that each file F_i has bit size S_i. Each file has a $1/4$ chance of being written do the first disk. Therefore, by linearity of expectation, the expected number of bytes written to the first disk is the sum of the expected number of bytes per file written to the first disk, which is:

$$\sum_{i=1}^{1000} 1/4 \cdot S_i = 1/4 \sum_{i=1}^{1000} S_i = 1/4 \cdot 400 = 100$$

(c) [8 points] Find the best upper bound you can on the probability that $200MB$ or more are written on the first disk?

For this we can use the first Chernoff bound, which is:

$$\Pr\left(X \geq c\,\mathrm{Ex}\left(X\right)\right) \leq e^{-(c\ln c - c + 1)\,\mathrm{Ex}\left(X\right)}$$

The Chernoff bound only works if X is the sum of random variables that each take on a value between 0 and 1. The file size of each file in the first disk is between 0 and 1Mb . So we can define X to be the total number of bytes in disk 1. The expected value of X is 100, so we take c to be 2. We get:

$$\Pr\left(X \geq 2 \cdot 100\right) \leq e^{-(2\ln 2 - 2 + 1)100}$$

(d) [4 points] Find the best upper bound you can on the probability that there is some disk with $200MB$ or more written on it?

For this we can use the Union Bound along with our result from above. The probability of this event happening in one or more disks is upper bounded by the sum of the probabilities of the event happening in each disk. This gives us an upper bound of

$$4 \cdot e^{-(2\ln 2 - 1)100}$$

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Final

Problem 1. [8 points] Prove that for all $n \in \mathbb{N}$, the following identity holds

$$\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}.$$

Solution. By induction on $n \geq 1$, with induction hypothesis

$$P(n) : \sum_{i=0}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$$

for all $n \in \mathbb{N}$

Base case ($n = 1$):

$$\frac{1(1+1)(2+1)}{6} = \frac{6}{6} = 1 = 1^2$$

Inductive step: Assume $P(n)$, we need to show that $P(n+1)$ holds.

$$\begin{aligned}
\sum_{i=0}^{n+1} i^2 &= (\sum_{i=0}^{n} i^2) + (n+1)^2 \\
&= \frac{n(n+1)(2n+1)}{6} + (n+1)^2 \\
&= \frac{n(n+1)(2n+1) + 6(n+1)^2}{6} \\
&= \frac{(n+1)(2n^2 + n + 6n + 6)}{6} \\
&= \frac{(n+1)(2n^2 + 7n + 6)}{6} \\
&= \frac{(n+1)(n+2)(2n+3)}{6} \\
&= \frac{(n+1)((n+1)+1)(2(n+1)+1)}{6} \\
&\Rightarrow P(n+1)
\end{aligned}$$

as required.

We have shown that $P(n) \Rightarrow P(n+1)$. Thus, $P(n)$ is true for all $n \in \mathbb{N}$.

Problem 2. [20 points] Coin-Flip is a 2 player game. Each player wins with probability exactly 0.5. There are no ties.

n people are playing a Coin-Flip tournament. Every person plays a Coin-Flip game with every other person exactly once. Thus everybody plays $n-1$ games. The outcomes of all the games are mutually independent of one another.

We say that the tournament is a *success* if for every $i \in \{0, 1, \ldots, n-1\}$, there is exactly one player, which we will refer to as p_i, with exactly i wins.

(a) [10 points] Prove that if the tournament is a success, then for any integers j, k with $0 \leq k < j \leq n-1$, p_j defeats p_k.

Solution. We prove it by induction on k. The inductive hypothesis $P(k)$ is that for all $0 \leq \ell \leq k$ and all $\ell < j < n$, p_j defeats p_ℓ.

The base case is $k = 0$. Now p_0 loses all $n-1$ games. Thus, for every $j > 0$, p_j defeats p_0. Suppose $P(k)$ and let us show $P(k+1)$. Since the tournament is a sucess, p_{k+1} wins exactly $k+1$ games. Because $P(k)$ holds, p_{k+1} defeats the $k+1$ players $p_0, p_1, \ldots, p_k$. Thus, for all $k+1 < j < n$, p_j defeats p_{k+1}. So $P(k+1)$ holds.

(b) [6 points] What is the probability that the tournament will be a success?

Solution. Let π be a permutation of $\{0, 2, \ldots, n-1\}$ and define the event $\mathcal{E}_\pi$ to be that the tournament is a success with players p_i winning exactly i games, where p_i is the $\pi(i)$th player. Then the events $\mathcal{E}_\pi$ and $\mathcal{E}_\sigma$ are disjoint for $\pi \neq \sigma$. Moreover, by symmetry, $\Pr[\mathcal{E}_\pi] = \Pr[\mathcal{E}_\sigma]$ for all permutations σ, π. Let $\mathcal{E}$ be the event the tournament is a success. Since the events are disjoint,

$$\Pr[\mathcal{E}] = \sum_\pi \Pr[\mathcal{E}_\pi] = n! \Pr[\mathcal{E}_\sigma],$$

where σ is some fixed permutation. Now $\mathcal{E}_\sigma$ determines all the outcomes of the games. As there are $\binom{n}{2}$ games, we have

$$\Pr[\mathcal{E}_\sigma] = \left(\frac{1}{2}\right)^{\binom{n}{2}}.$$

In total,

$$\Pr[\mathcal{E}] = n! \left(\frac{1}{2}\right)^{\binom{n}{2}}.$$

(c) [4 points] Show that your answer to part (b) is $o(1)$. **Solution.** We have,

$$\Pr[\mathcal{E}] \leq n^n \left(\frac{1}{2}\right)^{\binom{n}{2}} = 2^{n \log n} \left(\frac{1}{2}\right)^{\binom{n}{2}} = \left(\frac{1}{2}\right)^{\binom{n}{2} - n \log n}.$$

Now, $\binom{n}{2} - n \log n = \Omega(n^2)$, so there is a positive constant c such that for sufficiently large n, this probability is at most

$$\left(\frac{1}{2}\right)^{cn^2},$$

which for sufficiently large n, is clearly less than any positive constant, and thus is $o(1)$.

Problem 3. [8 points] A person is passing time by advancing a token on the set of natural numbers. In the beginning, a token is placed on 0.

The person keeps playing *moves* forever. Each move proceeds as follows:

1. First the person tosses a fair coin (with heads/tails equally likely).
2. Suppose the token is currently placed on n. If heads came up, then the person moves the token to $n+3$, otherwise he moves the token to $n+4$.

For each $n \in \mathbb{N}$, let E_n be the event "There was a move on which the token landed on n". Let $p_n = \Pr[E_n]$.

Find a recurrence relation for p_n. *You do not need to solve the recurrence, but you should specify the boundary conditions that would be necessary to find a solution to the recurrence.*

Solution. For all $n \geq 4$,

$$p_n = \frac{1}{2}p_{n-3} + \frac{1}{2}p_{n-4},$$

with boundary conditions $p_0 = 1$, $p_1 = 0$, $p_2 = 0$, $p_3 = 1/2$.

Problem 4. [10 points] Exactly $1/5$th of the people in a town have Beaver Fever©.

There are two tests for Beaver Fever, TEST1 and TEST2. When a person goes to a doctor to test for Beaver Fever, with probability $2/3$ the doctor conducts TEST1 on him and with probability $1/3$ the doctor conducts TEST2 on him.

When TEST1 is done on a person, the outcome is as follows:

- If the person has the disease, the result is positive with probability $3/4$.
- If the person does not have the disease, the result is positive with probability $1/4$.

When TEST2 is done on a person, the outcome is as follows:

- If the person has the disease, the result is positive with probability 1.
- If the person does not have the disease, the result is positive with probability $1/2$.

A person is picked uniformly at random from the town and is sent to a doctor to test for Beaver Fever. The result comes out positive. What is the probability that the person has the disease?

Solution. Let B be the event that the person has BLAH. Let $T1$ be the event that the person is tested with test1. Let $T2$ be the event that the person is tested with test2. Let P be the event that the test comes out positive.

A tree diagram is worked out below with the given information:

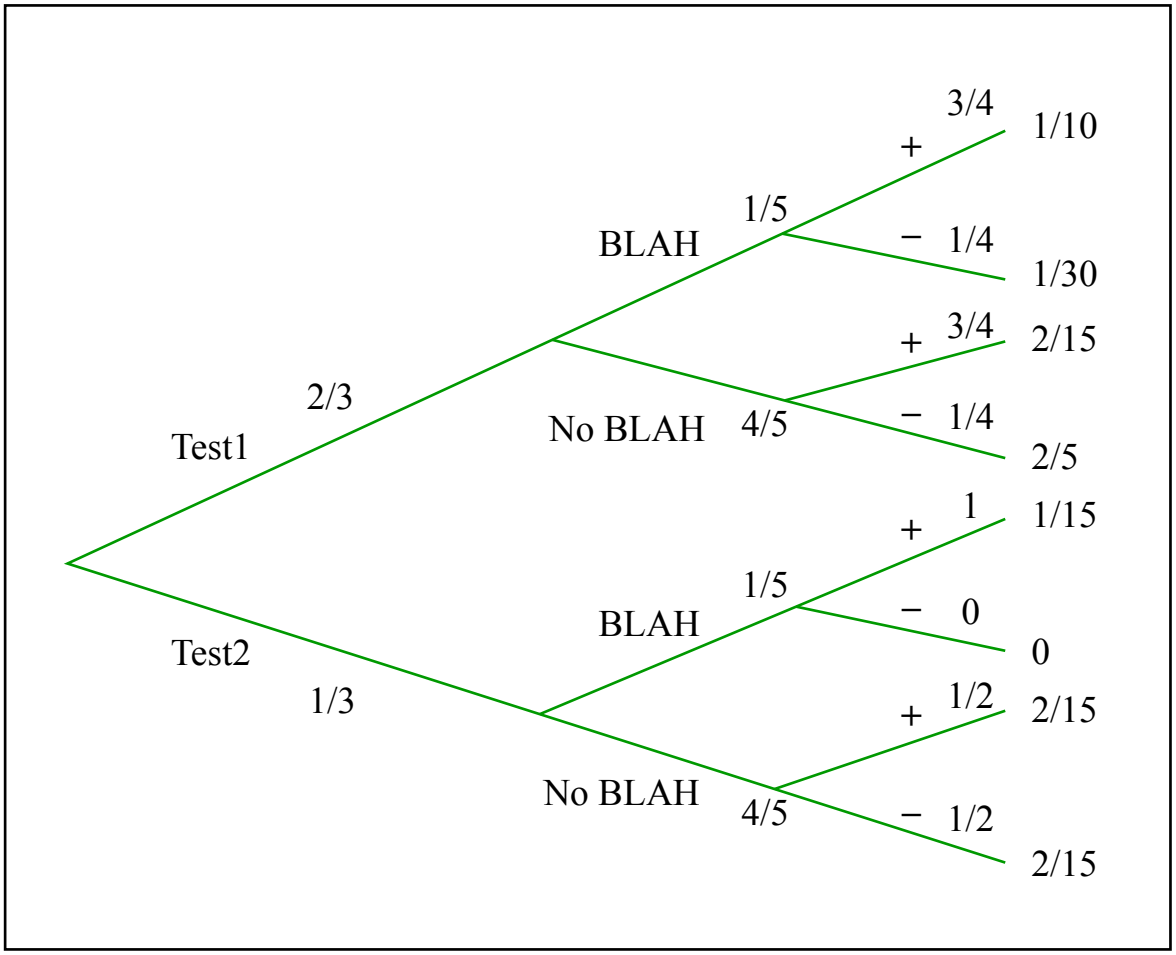


Image by MIT OpenCourseWare.

The probability that a person has BLAH, given that the test comes out positive is:

$$
\begin{aligned}
\Pr\{B \mid S\} = & \quad \Pr\{B \mid T1 \cap P\} \cdot \Pr(T1) + \Pr\{B \mid T2 \cap P\} \cdot \Pr(T2) \\
= & \quad \frac{\Pr(B \cap T1 \cap P)}{\Pr(T1 \cap P)} \cdot \Pr(T1) + \frac{\Pr(B \cap T2 \cap P)}{\Pr(T2 \cap P)} \cdot \Pr(T2) \\
= & \quad \frac{\Pr(D \cap T1 \cap P)}{\Pr(D \cap T1 \cap P) + \Pr\left(\bar{D} \cap T1 \cap P\right)} \cdot \Pr(T1) + \\
& \quad \frac{\Pr(D \cap T2 \cap P)}{\Pr(D \cap T2 \cap P) + \Pr\left(\bar{D} \cap T2 \cap P\right)} \cdot \Pr(T2) \\
= & \quad \frac{\frac{1}{10}}{\frac{1}{10} + \frac{2}{15}} \cdot \frac{2}{3} + \frac{\frac{1}{15}}{\frac{1}{15} + \frac{2}{15}} \cdot \frac{1}{3} \\
= & \quad \frac{5}{13}
\end{aligned}
$$

Problem 5. [10 points] Two identical complete decks of cards, each with 52 cards, have been mixed together. A hand of 5 cards is picked uniformly at random from amongst all subsets of exactly 5 cards.

(a) [5 points] What is the probability that the hand has no identical cards (i.e., cards with the same suit and value. For example, the hand $\langle Q\heartsuit, 5\spadesuit, 6\spadesuit, 8\clubsuit, Q\heartsuit\rangle$ *has* identical cards.)? We can calculate this probability by computing

$$\frac{\text{hands with no identical cards}}{\text{total possible hands}}$$

There are 104 cards. There are 5 cards in a hand. Order does not matter. The total number of possible hands is:

$\binom{104}{5}$

There are 52 possible card faces, and we can choose 5 of them if no identical cards are allowed. Additionally, each card can be from either deck 1 or deck 2. Therefore the number of hands with no identical cards, chosen from 2 decks is:

$$\binom{52}{5} \cdot 2^5$$

Therefore the probability of drawing a hand with no identical cards is:

$$\frac{\binom{52}{5} \cdot 2^5}{\binom{104}{5}}$$

(b) [5 points] What is the probability that the hand has exactly one pair of identical cards? This can be solved by a similar approach. A hand of this type can be distinguished by the face (suit and value) of the repeated card, and by the faces of the 3 non-repeated cards. There are 52 possible values for the face of the repeated card. There are $\binom{51}{3}$ possible faces for the non-repeated cards, since none of these can be repeated. Each of these could come from either the 1st deck or the 2nd deck. There are $\binom{104}{5}$ possible hands, as before. So the probability of getting a hand with exactly one pair of identical cards is:

$$\frac{52 \cdot \binom{51}{3} \cdot 2^3}{\binom{104}{5}}$$

Problem 6. [28 points] Scores for a final exam are given by picking an integer uniformly at random from the set $\{50, 51, \ldots, 97, 98\}$. The scores of all 128 students in the class are assigned in this manner. For parts (a), (b), (c) and (d) you may NOT assume that these scores are assigned independently. For parts (e), (f), (g) and (h) you MAY assume that these scores are assigned independently.

Let $S_1, \ldots, S_{128}$ be their scores. Let $S = \frac{1}{128}(\sum_{i=1}^{128} S_i)$ be the average score of the class.

(a) [3 points] For $i \in \{1, \ldots, 128\}$, what is $\mathbb{E}[S_i]$?

(b) [2 points] Show that $\mathbb{E}[S] = 74$. Make no independence assumptions.

(c) [4 points] Prove that

$$\Pr[S \geq 88] \leq \frac{37}{44}.$$

Make no independence assumptions.

(d) [5 points] Improve your previous bound by using the fact that the minimum possible score is 50. Prove that

$$\Pr[S \geq 88] \leq \frac{12}{19}.$$

Make no independence assumptions.

(e) [4 points] For the remaining problems, assume that all the scores are assigned mutually independently. Use Problem 1 of this final to find $Var[S_i]$.

(f) [3 points] What is $Var[S]$?

(g) [2 points] What is the standard deviation of S?

(h) [5 points] Prove, using the Chebyshev Inequality, that

$$\Pr[S \leq 69] \leq \frac{1}{16}.$$

Solution.

(a) We simply take the average of the numbers from 50 to 98. Thus, $\mathbb{E}[S_i] = \frac{50+98}{2} = 74$.

(b) By linearity of expectation,

$$\mathbb{E}[S] = \mathbb{E}[\frac{1}{128}(\sum_{i=1}^{128} S_i)] = \frac{1}{128}(128 * \mathbb{E}[S_1]) = \mathbb{E}[S_1] = 74$$

(c) By Markov's inequality,

$$\Pr[S \geq 88] \leq \frac{\mathbb{E}[S]}{88} = \frac{74}{88} = \frac{37}{44}.$$

(d) We define a random variable $T = S - 50$, and thus $\mathbb{E}[T] = \mathbb{E}[S] - 50 = 24$. Now we just apply Markov's inequality:

$$\Pr[S \geq 88] = \Pr[T \geq 38] \leq \frac{\mathbb{E}[T]}{38} = \frac{24}{38} = \frac{12}{19}.$$

(e) We define $T_i = S_i - 50$.

$$Var[S_i] = Var[T_i] = \mathbb{E}[T_i^2] - \mathbb{E}^2[T_i] = (\frac{1}{49}\sum_{i=0}^{48} i^2) - \mathbb{E}^2[T_i] = \frac{1}{49}\frac{(48)(49)(97)}{6} - (24)^2 = 776 - 576 = 200.$$

(f)

$$Var[S] = Var[\frac{1}{128}(\sum_{i=1}^{128} S_i)] = (\frac{1}{128})^2(128 * Var[S_1]) = \frac{Var[S_1]}{128} = \frac{200}{128} = \frac{25}{16}.$$

(g) The standard deviation of S is simply the square root of the variance of S:

$$\sigma_S = \sqrt{\frac{25}{16}} = \frac{5}{4}.$$

(h) Using Chebyshev's inequality,

$$\Pr[S \leq 69] \leq \Pr[|S - 74| \leq 5] = \Pr[|S - \mathbb{E}[S]| \leq 4 * \sigma_S] \leq \frac{1}{4^2} = \frac{1}{16}.$$

Problem 7. [16 points] 1000 files $F_1, F_2, \ldots, F_{1000}$ have just reached a disk manager for writing onto disk. Each file's size is between $0MB$ and $1MB$. The sum of all files' sizes is $400MB$.

The disk manager has 4 disks under its control. For each file F_i, the disk manager chooses a disk uniformly at random from amongst the 4 disks, and F_i is written to that disk. The choices of disk for the different files are mutually independent.

(a) [2 points] What is the expected number of files that will be written to the first disk?

We can use indicator variables. For each file, $P_i = 1$ if F_i is written to the first disk. The chance of an individual file being written to the first disk is 1/4. By linearity of expectation, the expected number of files written to the first disk is the sum of the expected values of P_i's. The expected value of each indicator variable is $1/4$, and $\sum_{i=1}^{(} 1000)1/4 = 250$, so the expected number of files to be written to the first disk is 250.

(b) [2 points] What is the expected number of bytes written on the first disk?

We can say that each file F_i has bit size S_i. Each file has a $1/4$ chance of being written do the first disk. Therefore, by linearity of expectation, the expected number of bytes written to the first disk is the sum of the expected number of bytes per file written to the first disk, which is:

$$\sum_{i=1}^{1000} 1/4 \cdot S_i = 1/4 \sum_{i=1}^{1000} S_i = 1/4 \cdot 400 = 100$$

(c) [8 points] Find the best upper bound you can on the probability that $200MB$ or more are written on the first disk?

For this we can use the first Chernoff bound, which is:

$$\Pr\left(X \geq c\,\mathrm{Ex}\left(X\right)\right) \leq e^{-(c\ln c - c + 1)\,\mathrm{Ex}\left(X\right)}$$

The Chernoff bound only works if X is the sum of random variables that each take on a value between 0 and 1. The file size of each file in the first disk is between 0 and 1Mb . So we can define X to be the total number of bytes in disk 1. The expected value of X is 100, so we take c to be 2. We get:

$$\Pr\left(X \geq 2 \cdot 100\right) \leq e^{-(2\ln 2 - 2 + 1)100}$$

(d) [4 points] Find the best upper bound you can on the probability that there is some disk with $200MB$ or more written on it?

For this we can use the Union Bound along with our result from above. The probability of this event happening in one or more disks is upper bounded by the sum of the probabilities of the event happening in each disk. This gives us an upper bound of

$$4 \cdot e^{-(2\ln 2 - 1)100}$$

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Final Exam

Name: _____

- This final is **closed book**, but you may have three 8.5″ × 11″ sheets with notes in your own handwriting on both sides.
- You may not use a calculator or a Python interpreter, and while exercising skill with a PostScript interpreter would highly impress at least one member of the course staff, you are not allowed to use that either. You *may* work with any of the following: a slide rule, an abacus, a Curta, Napier's bones, any original version of the Antikythera mechanism, an Enigma machine, and/or the difference engine. You are also permitted to use a perpetual motion machine as a source of energy, and an antigravity device to elevate yourself above the rest of the class.
- You may assume all of the results presented in class.
- 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.
- Please resist the urge to roll on the floor laughing out loud.
- You have three hours to complete the exam.

Problem	1	2	3	4	5	6	7	8	9	10	Total
Points	25	25	25	15	15	15	25	15	20	20	200
Score											
Grader											

Problem 1. [25 points] The Final Breakdown

Suppose the 6.042 final consists of:

- 36 true/false questions worth 1 point each.
- 1 induction problem worth 15 points.
- 1 giant problem that combines everything from the semester, worth 49 points.

Grading goes as follows:

- The TAs choose to grade the easy true/false questions. For each individual point, they flip a fair coin. If it comes up heads, the student gets the point.
- Marten and Brooke split the task of grading the induction problem.
  - With 1/3 probability, Marten grades the problem. His grading policy is as follows: Either he gets exasperated by the improper use of math symbols and gives 0 points (which happens with 2/5 probability), or he finds the answer satisfactory and gives 15 points (which happens with 3/5 probability).
  - With 2/3 probability, Brooke grades the problem. Her grading policy is as follows: She selects a random integer point value from the range from 0 to 15, inclusive, with uniform probability.
- Finally, Tom grades the giant problem. He rolls two fair **seven**-sided dice (which have values from 1 to 7, inclusive), takes their product, and subtracts it from 49 to determine the score. (Example: Tom rolls a 3 and a 4. The score is then $49 - 3 \cdot 4 = 37$.)

Assume all random choices during the grading process are mutually independent.

The problem parts start on the next page. Show your work to receive partial credit.

(a) [7 pts] What is the expected score on the exam?

(b) [5 pts] What is the variance on the 36 true/false questions?

(c) [5 pts] What is the variance on the induction score, given that Marten graded the problem?

(d) [3 pts] Argue why the Markov bound can be used to determine an upper bound on the probability that the score on the exam is ≥ 80. You do not need to compute the actual bound.

(e) [5 pts] Use the Chebyshev bound to determine an upper bound on the probability that the score on the true/false questions is ≥ 24.

Problem 2. [25 points] Woodchucks Chucking Wood

All woodchucks can chuck wood, but only some can do it well.

- 1/3 of all woodchucks like to chuck wood.
- 2/3 of all woodchucks can chuck wood well.
- 1/2 of those that like chucking wood can do it well.
- The expected amount of wood chucked by a woodchuck (randomly chosen with uniform probability) is 7 kg/day.
- The expected amount of wood chucked by a woodchuck that likes chucking wood but can't do it well is 1 kg/day.
- A woodchuck that does not like chucking wood does not chuck any wood at all, regardless of its wood-chucking skillz or lack thereof.

(a) [10 pts] What is the probability that a woodchuck (randomly chosen with uniform probability) likes chucking wood, given that it can do it well?

(b) [15 pts] On average, how much wood would a woodchuck chuck if the woodchuck could chuck wood well?

Problem 3. [25 points] Cardsharing☆Revolution

Three 6.042 students—Kirari, Noelle, and Cobeni—are playing a game of Tan Tan Taan!. During each round of Tan Tan Taan!, each player is dealt 4 cards of their own, and one additional card is shared among all players, so that each player has 5 cards that they can use (the 4 cards of their own along with the single shared card). Cards are uniformly distributed from a 52-card deck. If you get four of a kind (for example, four aces or four 2's), you can continue playing in the next round. If you don't get four of a kind, you must quit and return to doing your 6.042 homework. Cards from round to round are mutually independent. This game is so fun that even if two of the three players must quit and return to their 6.042 homework, the third player will continue playing alone as long as they are able to.

(a) [5 pts] What is the probability that Kirari has four aces in the first round?

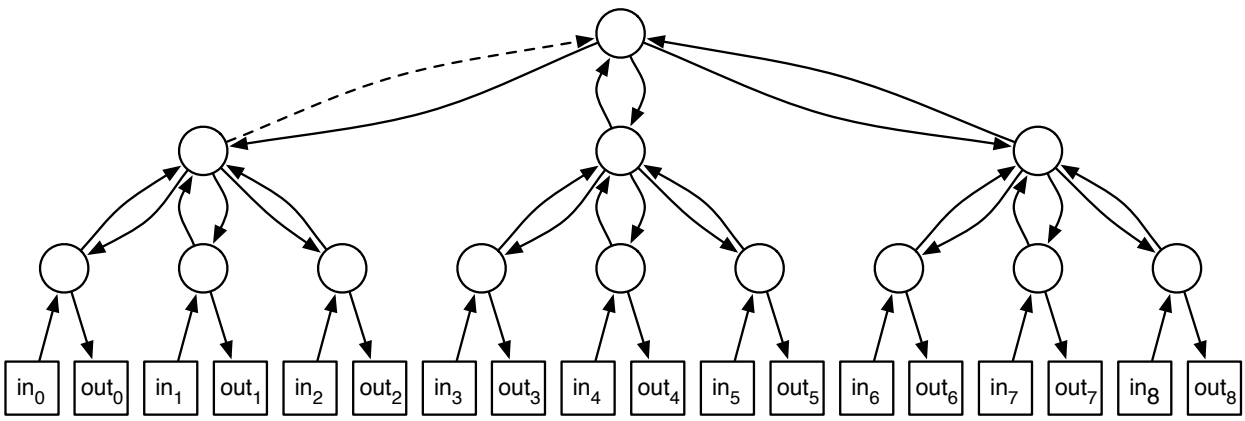
(b) [5 pts] What is the probability that Kirari doesn't get four of a kind in the first round (and must quit playing)?

(c) [5 pts] What is the expected number of rounds that Kirari will play?

(d) [10 pts] What is the probability that all three can play a second round?

Problem 4. [15 points] Packet Racket!

Consider the complete ternary-tree network with 9 inputs and 9 outputs shown below where packets are routed randomly. The route each packet takes is the shortest path between input and output. Let I_0, I_1, and I_2 be indicator random variables for the events that a packet originating at in_0, in_1, and in_2, respectively, crosses the dashed edge in the figure. Let $T = I_0 + I_1 + I_2$ be a random variable for the number of packets passing through the dashed edge.



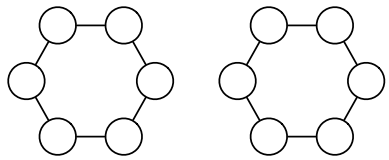
(a) [10 pts] Suppose that each input sends a single packet to an output selected uniformly at random; the packet destinations are mutually independent. (Note that outputs may receive packets from multiple inputs including their corresponding input.)

What are the expectation and variance of T?

(b) [5 pts] Now consider the situation where a permutation of inputs to outputs is chosen uniformly at random; each input sends a packet to a distinct output. What is the expected value of T? Briefly justify your answer.

Problem 5. [15 points] Connected or Not? That Is the Question

Suppose we have a simple, undirected graph G with $2n$ vertices and $2n$ edges, where $n \geq 3$. The graph consists of two disjoint cycles with n edges each. For example, if $n = 6$, the graph would look like this:



(a) [5 pts] A pair of vertices u and v from G is selected uniformly at random from the pairs of distinct vertices with no edge between them. A new graph G' is constructed to be the same as G, except that there is an edge between u and v. What is the probability that G' is connected?

(b) [10 pts] k pairs of vertices from G are selected uniformly at random from the pairs of distinct vertices with no edge between them. Repetition is allowed; it is possible, for example, that the same pair appears multiple times in the set of k pairs. A new graph G'' is constructed to be the same as G, except that there are k new edges: the edges that correspond to the k selected pairs. What is the probability that G'' is **not** connected?

(*Hint: For $k = 1$, the sum of your answers to part (a) and part (b) should equal* 1.)

Problem 6. [15 points] 6.042: The Ultimate Showdown

There are 100 homework problems in 6.042 throughout the term. Let T_i, $1 \leq i \leq 100$, be the random variable indicating the fraction of a day that is needed by a student to solve the ith problem of 6.042.

The distribution for each T_i is different and unknown. We only know that the T_i are mutually independent and that for all i, $0 \leq T_i \leq 1$ and $\text{Ex}[T_i] = 0.3$.

Let T be the sum of all T_i's; T represents the total number of days needed by a student to complete all homework problems for 6.042. Prove that the probability that T is greater than $30e$ is exceedingly small by deriving the best bound you can on this probability. (*Hint: We do not consider* $1/e$ *to be exceedingly small.*)

Problem 7. [25 points] Gotta Count 'Em All!

An unusual species inhabits the forest surrounding Functional City. Each member of the species can take one of three possible forms, called *Schemander*, *Haskeleon*, and *Camlizard*.

In January of every year, each individual undergoes "evolution"—a process by which the individual splits into two individuals, whose forms depend on the form of the original:

- A Schemander splits into a Schemander and a Haskeleon.
- A Haskeleon splits into a Schemander and a Camlizard.
- A Camlizard splits into a Schemander and a Haskeleon.

We are investigating the distribution of forms within a large population of this species over time. It is known that in June of year 0, the population consisted of a single Schemander. Assume that no individual ever dies and that all individuals successfully undergo evolution exactly once every January.

(a) [3 pts] Let S_n, H_n, and C_n be the number of Schemanders, Haskeleons, and Camlizards, respectively, in June of year n. Express S_n, H_n, and C_n in terms of S_{n-1}, H_{n-1}, and C_{n-1}, for $n > 0$.

(b) [5 pts] Let $T_n = S_n + H_n + C_n$ be the total number of individuals in June of year n. Use induction to prove that $T_n = 2^n$ for all $n \geq 0$.

(c) [2 pts] Show that $H_n = T_{n-1} - H_{n-1}$ for $n > 0$.

(d) [15 pts] Give a closed-form expression for H_n. You may use, without proof, the fact stated in part (b) and the recurrence given in part (c).

Problem 8. [15 points] Asymptotic Awesomeness

For each row in the following table, determine whether there exist functions f and g that satisfy all the properties marked **Yes** and do *not* satisfy the properties marked **No**. You do not have to provide examples.

	$f = \Theta(g)$	$f = O(g)$	$f = o(g)$	$f = \Omega(g)$	$f = \omega(g)$	Do f, g exist?
(a)	Yes	Yes	Yes	No	No	
(b)	No	No	No	Yes	Yes	
(c)	No	No	Yes	No	No	
(d)	Yes	Yes	No	Yes	No	
(e)	No	Yes	No	No	No	
(f)	No	No	No	No	No	

Problem 9. [20 points] Yet Another Graph Proof

Prove that in a finite directed graph, if every node has at least one outgoing edge, then the graph has a cycle.

(*Hint: Consider the longest path.*)

Problem 10. [20 points] Revenge of the Slipped Disc Puzzle™: The Curse of 6.042

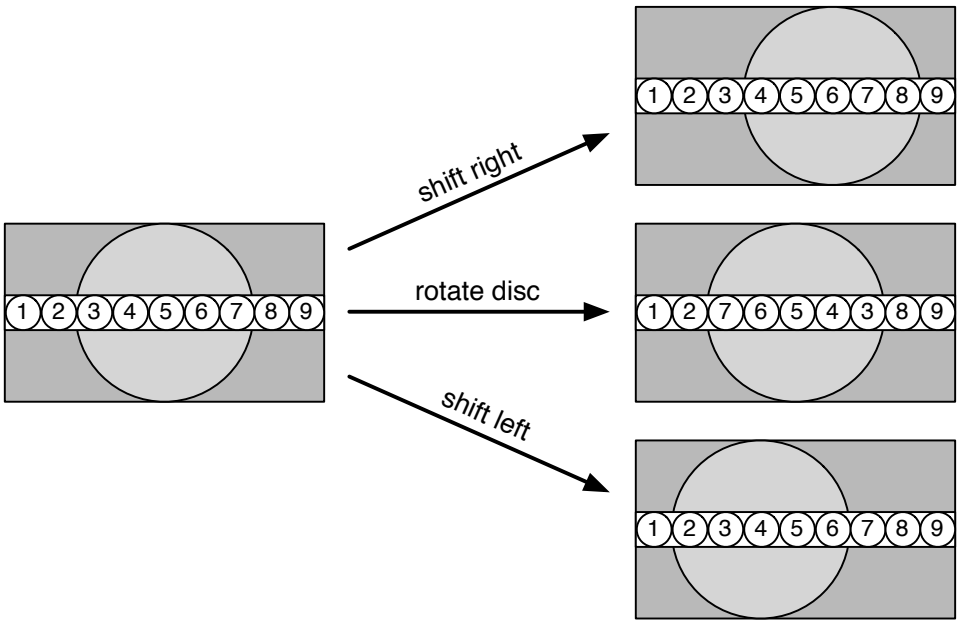
(This problem is similar to the Slipped Disc Puzzle™ of Quiz 1, but here we rotate 5 tiles instead of 4.)

The Super Awesome Extreme zomgroflolwut Spifftastic-to-the-Max Slipped Disc Puzzle™ consists of a track holding 9 circular tiles. In the middle is a disc that can slide left and right and rotate 180° to change the positions of *exactly* **five** tiles. As shown below, there are three ways to manipulate the puzzle:

Shift Right: The center disc is moved one unit to the right (if there is space).

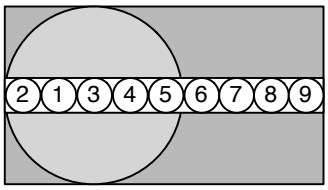
Rotate Disc: The **five** tiles in the center disc are reversed.

Shift Left: The center disc is moved one unit to the left (if there is space).

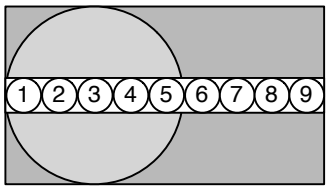


Prove that if the puzzle starts in an initial state with all but tiles 1 and 2 in their natural order, then it is impossible to reach a goal state where all the tiles are in their natural order. The initial and goal states are shown below:

Initial State



Goal State



Write your proof on the next page...

Room for Problem 10...

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Final Exam

Problem 1. [25 points] The Final Breakdown

Suppose the 6.042 final consists of:

- 36 true/false questions worth 1 point each.
- 1 induction problem worth 15 points.
- 1 giant problem that combines everything from the semester, worth 49 points.

Grading goes as follows:

- The TAs choose to grade the easy true/false questions. For each individual point, they flip a fair coin. If it comes up heads, the student gets the point.
- Marten and Brooke split the task of grading the induction problem.
  - With 1/3 probability, Marten grades the problem. His grading policy is as follows: Either he gets exasperated by the improper use of math symbols and gives 0 points (which happens with 2/5 probability), or he finds the answer satisfactory and gives 15 points (which happens with 3/5 probability).
  - With 2/3 probability, Brooke grades the problem. Her grading policy is as follows: She selects a random integer point value from the range from 0 to 15, inclusive, with uniform probability.
- Finally, Tom grades the giant problem. He rolls two fair **seven**-sided dice (which have values from 1 to 7, inclusive), takes their product, and subtracts it from 49 to determine the score. (Example: Tom rolls a 3 and a 4. The score is then $49 - 3 \cdot 4 = 37$.)

Assume all random choices during the grading process are mutually independent.

The problem parts start on the next page. Show your work to receive partial credit.

(a) [7 pts] What is the expected score on the exam?

Solution.

$$36/2 + (1/3)15(3/5) + (2/3)15/2 + (49 - (4 * 4)) = 18 + 3 + 5 + 33 = 59$$

The expected score on the exam is the sum of the expected scores on the individual problems.

$$\text{Ex}\,(\text{test score}) = \text{Ex}\,(\text{mc score}) + \text{Ex}\,(\text{ind score}) + \text{Ex}\,(\text{giant score})$$

- The expected multiple choice score is just the sum of the expectations on 36 coin tosses. Since the coin is fair, the expected number of heads on each flip is 1/2. Therefore:
$$\text{Ex}\,(\text{mc score}) = 36 \cdot 1/2 = 18$$
- The expected induction score is a weighted sum of the expectations on problems graded by Marten and Brooke. Let M be the event that Marten graded the problem, and B be the event that Brooke graded the problem. Therefore:
$$\text{Ex}\,(\text{ind score} \mid M) = 0 \cdot 2/5 + 15 \cdot 3/5 = 9$$
$$\text{Ex}\,(\text{ind score} \mid B) = \sum_{i=0}^{15} (i \cdot 1/16) = 15/2$$
$$\begin{aligned}\text{Ex}\,(\text{ind score}) &= \text{Ex}\,(\text{ind score} \mid M)\,\Pr\{M\} + \text{Ex}\,(\text{ind score} \mid B)\,\Pr\{B\} \\ &= 9 \cdot 1/3 + 15/2 \cdot 2/3 = 8\end{aligned}$$
- The expected giant problem score is the expectation on 49 minus the product of rolls of two fair seven-sided dice. We can pull out the constant 49, and (since the dice are independent) the expectation on the product of the rolls becomes the product of the expectations on the rolls, which is 4 in this case (average of numbers 1 to 7). Therefore, if R is the random variable representing the roll of a single die:
$$\text{Ex}\,(\text{giant score}) = \text{Ex}\left(49 - R^2\right) = 49 - \text{Ex}\,(R)^2 = 49 - (4)^2 = 49 - 16 = 33$$

Therefore, the overall expectation is:

$$\text{Ex}\,(\text{test score}) = 18 + 8 + 33 = 59$$

■

(b) [5 pts] What is the variance on the 36 true/false questions?

Solution. Since the coin flips are independent, we can sum the variances of each flip.

$$\text{Var}\,[\text{mc score}] = 36 \cdot 1/2 \cdot 1/2 = 9$$

■

(c) [5 pts] What is the variance on the induction score, given that Marten graded the problem?

Solution. Using the equation $\text{Var}[X] = \text{Ex}\left(X^2\right) - \text{Ex}(X)^2$:

$$\text{Var}[\text{ind score}] = (15^2 \cdot 3/5) - (9)^2 = 135 - 81 = 54$$

■

(d) [3 pts] Argue why the Markov bound can be used to determine an upper bound on the probability that the score on the exam is ≥ 80. You do not need to compute the actual bound.

Solution. The Markov bound can be used if there is a lower bound on the possible values of the random variable. In this case, all test scores are ≥ 0. ■

(e) [5 pts] Use the Chebyshev bound to determine an upper bound on the probability that the score on the true/false questions is ≥ 24.

Solution. If C be the random variable representing the multiple choice score:

$$\begin{aligned}
\Pr\{C \geq 24\} &\leq \Pr\{\mid C - 18 \mid \geq 6\} \\
&= \Pr\{\mid C - \text{Ex}(C) \mid \geq 6\} \\
&\leq \frac{\text{Var}[C]}{6^2} \\
&= \frac{9}{36} = \frac{1}{4}
\end{aligned}$$

■

Problem 2. [25 points] Woodchucks Chucking Wood

All woodchucks can chuck wood, but only some can do it well.

- 1/3 of all woodchucks like to chuck wood.
- 2/3 of all woodchucks can chuck wood well.
- 1/2 of those that like chucking wood can do it well.
- The expected amount of wood chucked by a woodchuck (randomly chosen with uniform probability) is 7 kg/day.
- The expected amount of wood chucked by a woodchuck that likes chucking wood but can't do it well is 1 kg/day.
- A woodchuck that does not like chucking wood does not chuck any wood at all, regardless of its wood-chucking skillz or lack thereof.

(a) [10 pts] What is the probability that a woodchuck (randomly chosen with uniform probability) likes chucking wood, given that it can do it well?

Solution. Let L be the event that a woodchuck likes chucking wood and C be the event that a woodchuck can chuck wood well. We are given:

$$\begin{aligned} Pr(L) &= 1/3 \\ Pr(C) &= 2/3 \\ Pr(C|L) &= 1/2 \end{aligned}$$

We wish to find $Pr(L|C)$. Using Bayes' rule, we have

$$\begin{aligned} Pr(L|C) &= \frac{Pr(C|L)Pr(L)}{Pr(C)} \\ &= \frac{(1/2)(1/3)}{(2/3)} \\ &= 1/4 \end{aligned}$$

■

(b) [15 pts] On average, how much wood would a woodchuck chuck if the woodchuck could chuck wood well?

Solution. Let W be a random variable representing the amount of wood the woodchuck chucks. We are given (in units of kg/day):

$$\begin{aligned} E(W) &= 7 \\ E(W|(L \cap \bar{U})) &= 1 \end{aligned}$$

Using the law of total expectation, we can partition the sample space into three events: (1) the woodchuck can chuck wood well; (2) the woodchuck can't chuck wood well but likes chucking wood; (3) the woodchuck can't chuck wood well and doesn't like chucking wood.

$$\begin{aligned} E(W) &= E(W|C)Pr(C) + E(W|(\bar{C} \cap L))Pr(\bar{C} \cap L) + E(W|(\bar{C} \cap \bar{L}))Pr(\bar{C} \cap \bar{L}) \\ 7 &= E(W|C)(2/3) + 1 \cdot Pr(\bar{C} \cap L) + E(W|(\bar{C} \cap \bar{L}))Pr(\bar{C} \cap \bar{L}) \end{aligned}$$

Since a woodchuck that doesn't like chucking wood does not chuck any wood, whether or not it can do so, $E(W|\bar{C} \cap \bar{L}) = 0$, so the last product in the sum vanishes:

$$7 = E(W|C)(2/3) + Pr(\bar{C} \cap L)$$

Finally, from the definition of conditional expection, we have

$$\begin{aligned} Pr(\bar{C} \cap L) &= P(\bar{C}|L)P(L) \\ &= (1 - P(C|L))P(L) \\ &= (1 - 1/2)(1/3) \\ &= 1/6 \end{aligned}$$

We wish to find $E(W|C)$, the expected amount of wood chucked by a woodchuck that can chuck wood well:

$$\begin{aligned} 7 &= E(W|C)(2/3) + 1/6 \\ 41/6 &= E(W|C)(2/3) \\ 41/4 &= E(W|C) \end{aligned}$$

So $E(W|C) = 41/4 = 10.25$.

■

Problem 3. [25 points] Cardsharing☆Revolution

Three 6.042 students—Kirari, Noelle, and Cobeni—are playing a game of Tan Tan Taan!. During each round of Tan Tan Taan!, each player is dealt 4 cards of their own, and one additional card is shared among all players, so that each player has 5 cards that they can use (the 4 cards of their own along with the single shared card). Cards are uniformly distributed from a 52-card deck. If you get four of a kind (for example, four aces or four 2's), you can continue playing in the next round. If you don't get four of a kind, you must quit and return to doing your 6.042 homework. Cards from round to round are mutually independent. This game is so fun that even if two of the three players must quit and return to their 6.042 homework, the third player will continue playing alone as long as they are able to.

(a) [5 pts] What is the probability that Kirari has four aces in the first round?

Solution. The total number of hands that Kirari can possibly get is $\binom{52}{5}$. Now we count how many ways they can make quad aces. There is only one way to get all four aces, and $52 - 4 = 48$ choices for the last card in the hand. So there are 48 hands that correspond to quad aces, and the probability of making quad aces is

$$\frac{48}{\binom{52}{5}}$$

■

(b) [5 pts] What is the probability that Kirari doesn't get four of a kind in the first round (and must quit playing)?

Solution. The different four-of-a-card hands do not overlap at all (no one hand is both four-of-a-kind for two different numbers), so the probability of getting any four-of-a-kind hand is

$$Pr[\text{ any four-card hand }] = \sum_{i=\text{Ace}}^{\text{King}} Pr[\text{ four-of-a-kind of card i}]$$

By symmetry, all these probabilities are the same, so the probability of getting four-of-a-kind is $13 * \frac{48}{\binom{52}{5}}$. The event of not getting a four-of-a-kind is the complement of this set, and so has probability

$$1 - \frac{624}{\binom{52}{5}}$$

■

(c) [5 pts] What is the expected number of rounds that Kirari will play?

Solution. This is a mean-time-to-failure problem. Imagine flipping a coin that has "heads" with probability

$$p = 1 - \frac{624}{\binom{52}{5}}$$

. We flip the coin until we get heads. From results in our class (or through doing the summation ourselves), we know that the result is $1/p$. ■

(d) [10 pts] What is the probability that all three can play a second round?

Solution. There are two problems in play here. First, we have to figure out the total number of hands that can be assigned to everyone: second, we have to figure out the number of ways that everyone can get four-of-a-kind.

First, consider the problem of finding the total number of hands that can be given out. Note that we can represent each dealing of hands as an ordered list of 13 cards chosen from the 52 cards in the deck, where the first 4 cards belong to Kirari, the second 4 belong to Noelle, the third four belong to Cobeni, and the final card is the communal card. There are $52!/(52-13)!$ ways to do this. However, we must remember that the ordering of cards in Kirari, Noelle, and Cobeni's hands do not matter - so we must "remove" the ordering in each of those four-card hands. This is done by dividing out by 4! three times. So the total number of ways we can deal cards is:

$$52!/39!4!4!4!$$

Now, we must count the number of ways to have all three people make four-of-a-kind hands. There are two cases: when the communal card is part of a four-of-a-kind, and when it is not.

- First consider when it is not part of a four-of-a-kind. Choose first the four-of-a-kinds that could happen: there are $13 * 12 * 11$ ways to assign numbers to the three people. After these numbers have been assigned, there are $52 - 12 = 40$ choices for the communal card. So there are $13 * 12 * 11 * 40$ ways to make a four-of-a-kind this way.
- Now say the communal card *is* part of a four-of-a-kind. Say it was part of Kirari's four-of-a-kind (by symmetry, the counting is the same if it were part of the other two's four-of-a-kinds). Once again, there are 13 choices for Kirari, 12 choices for

Noelle, and 11 choices for Cobeni. And once again, there are 40 choices for the extra card in Kirari's hand. So in total, there are $13 * 12 * 11 * 40 * 3$ ways for four-of-a-kind to happen in this case.

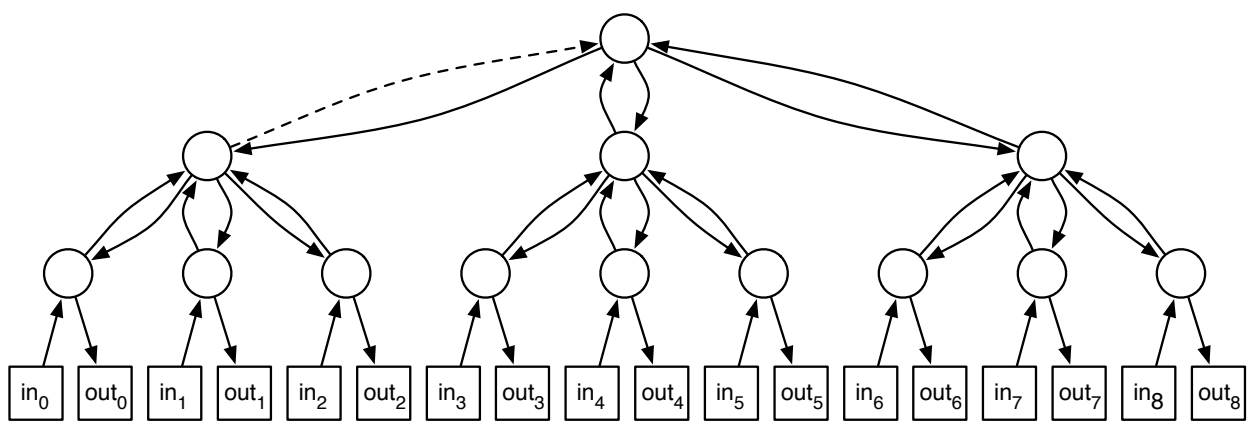
In total, there are $4(13 * 12 * 11 * 40)$ ways to make four-of-a-kind hands in this format. So the probability of everyone getting a four-of-a-kind is

$$\frac{4 * 13 * 12 * 11 * 40 * (39!4!4!4!)}{52!}$$

■

Problem 4. [15 points] Packet Racket!

Consider the complete ternary-tree network with 9 inputs and 9 outputs shown below where packets are routed randomly. The route each packet takes is the shortest path between input and output. Let I_0, I_1, and I_2 be indicator random variables for the events that a packet originating at in_0, in_1, and in_2, respectively, crosses the dashed edge in the figure. Let $T = I_0 + I_1 + I_2$ be a random variable for the number of packets passing through the dashed edge.



(a) [10 pts] Suppose that each input sends a single packet to an output selected uniformly at random; the packet destinations are mutually independent. (Note that outputs may receive packets from multiple inputs including their corresponding input.)

What are the expectation and variance of T?

Solution. A packet will pass through the dashed edge if it originates in inputs 0–2 and is destined for outputs 3–8. For $j \in \{0, 1, 2\}$ Let I_j be an indicator random variable for the event that a packet leaving input j passes through the dashed edge. The probability of this event is $\frac{2}{3}$. It follows that:

$$\begin{aligned} T &= I_1 + I_2 + I_3 \\ Ex[T] &= Ex[I_1 + I_2 + I_3] \\ &= Ex[I_1] + Ex[I_2] + Ex[I_3] \\ &= 3 \cdot \frac{2}{3} \\ &= 2 \end{aligned}$$

Similarity, but the linearity of variance for independent random variables:

$$\begin{aligned} T &= I_1 + I_2 + I_3 \\ Var[T] &= Var[I_1 + I_2 + I_3] \\ &= Var[I_1] + Var[I_2] + Var[I_3] \\ &= 3 \cdot \frac{2}{3} \cdot \left(1 - \frac{2}{3}\right) \\ &= \frac{2}{3} \end{aligned}$$

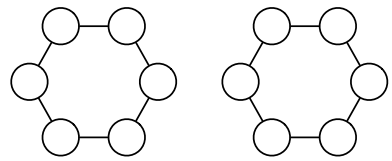
■

(b) [5 pts] Now consider the situation where a permutation of inputs to outputs is chosen uniformly at random; each input sends a packet to a distinct output. What is the expected value of T? Briefly justify your answer.

Solution. Once again, $T = I_1 + I_2 + I_3$. By linearity of expectation, $\mathrm{E}[T] = \mathrm{E}[I_1] + \mathrm{E}[I_2] + \mathrm{E}[I_3]$. We know that $\mathrm{E}[I_i] = 2/3$ still, because each input is equally likely routed to any of the outputs (even when we only restrict ourselves to permutations). Thus, $\mathrm{E}[T] = 3 * 2/3 = 2$. ■

Problem 5. [15 points] Connected or Not? That Is the Question

Suppose we have a simple, undirected graph G with $2n$ vertices and $2n$ edges, where $n \geq 3$. The graph consists of two disjoint cycles with n edges each. For example, if $n = 6$, the graph would look like this:



(a) [5 pts] A pair of vertices u and v from G is selected uniformly at random from the pairs of distinct vertices with no edge between them. A new graph G' is constructed to be the same as G, except that there is an edge between u and v. What is the probability that G' is connected?

Solution. G' is connected if and only if u and v come from different cycles. There are n^2 pairs of vertices consisting of vertices in different cycles. In all, there are $\binom{2n}{2} - 2n$ pairs of vertices with no edge between them, since there are $\binom{2n}{2}$ pairs of vertices and $2n$ of these pairs have an edge between them. The desired probability p can be computed as follows:

$$\begin{aligned} p &= \frac{n^2}{\binom{2n}{2} - 2n} \\ &= \frac{n^2}{\frac{2n(2n-1)}{2} - 2n} \\ &= \frac{n^2}{2n^2 - n - 2n} \\ &= \frac{n}{2n-3} \end{aligned}$$

■

(b) [10 pts] k pairs of vertices from G are selected uniformly at random from the pairs of distinct vertices with no edge between them. Repetition is allowed; it is possible, for example, that the same pair appears multiple times in the set of k pairs. A new graph G'' is constructed to be the same as G, except that there are k new edges: the edges that correspond to the k selected pairs. What is the probability that G'' is **not** connected?

(*Hint: For $k = 1$, the sum of your answers to part (a) and part (b) should equal* 1.)

Solution. Note that the probability of not connecting the graph in one sampling of two nonadjacent vertices is

$$p = 1 - \frac{n}{2n-3}$$

Because we are able to choose the same pair many times, we are simply taking independent samples of two nonadjacent vertices. Furthermore, in k samples, the graph is not connected if and only if none of the pairs chosen have connected the graph. The probability of this happening is

$$p^k = (1 - \frac{n}{2n-3})^k$$

■

Problem 6. [15 points] 6.042: The Ultimate Showdown

There are 100 homework problems in 6.042 throughout the term. Let T_i, $1 \leq i \leq 100$, be the random variable indicating the fraction of a day that is needed by a student to solve the ith problem of 6.042.

The distribution for each T_i is different and unknown. We only know that the T_i are mutually independent and that for all i, $0 \leq T_i \leq 1$ and $\text{Ex}[T_i] = 0.3$.

Let T be the sum of all T_i's; T represents the total number of days needed by a student to complete all homework problems for 6.042. Prove that the probability that T is greater than $30e$ is exceedingly small by deriving the best bound you can on this probability. (*Hint: We do not consider* $1/e$ *to be exceedingly small.*)

Solution. We know that $T = \sum_i T_i$. Thus, we know that $\text{Ex}[T] = \text{Ex}[\sum_i T_i] = \sum_i \text{Ex}[T_i] = 100(.3) = 30$

Now, from the Chernoff bound (which we can use because the T_i are mutually independent), we have that

$$Pr\{T \geq 30e\} \leq e^{-(e-e+1)\cdot 30} = e^{-30}$$

Which is quite a small bound.

■

Problem 7. [25 points] Gotta Count 'Em All!

An unusual species inhabits the forest surrounding Functional City. Each member of the species can take one of three possible forms, called *Schemander*, *Haskeleon*, and *Camlizard*.

In January of every year, each individual undergoes "evolution"—a process by which the individual splits into two individuals, whose forms depend on the form of the original:

- A Schemander splits into a Schemander and a Haskeleon.
- A Haskeleon splits into a Schemander and a Camlizard.
- A Camlizard splits into a Schemander and a Haskeleon.

We are investigating the distribution of forms within a large population of this species over time. It is known that in June of year 0, the population consisted of a single Schemander. Assume that no individual ever dies and that all individuals successfully undergo evolution exactly once every January.

(a) [3 pts] Let S_n, H_n, and C_n be the number of Schemanders, Haskeleons, and Camlizards, respectively, in June of year n. Express S_n, H_n, and C_n in terms of S_{n-1}, H_{n-1}, and C_{n-1}, for $n > 0$.

Solution. For each form, we look at what forms can be possible parents for it. A Schemander can have any of the three forms as a parent, so the number of Schemanders at time n is the number of Schemanders, Haskeleons, and Camlizards at time $n-1$. Likewise, a Haskeleon can have either a Schemander or a Camlizard as a parent, and a Camlizard can have only a Haskeleon as a parent.

Therefore, we can the recurrence equations as

$$\begin{aligned} S_n &= S_{n-1} + H_{n-1} + C_{n-1} \\ H_n &= S_{n-1} + C_{n-1} \\ C_n &= H_{n-1} \end{aligned}$$

■

(b) [5 pts] Let $T_n = S_n + H_n + C_n$ be the total number of individuals in June of year n. Use induction to prove that $T_n = 2^n$ for all $n \geq 0$.

Solution. We expand each of the three terms using the recurrence:

$$\begin{aligned} T_n &= S_n + H_n + C_n \\ &= (S_{n-1} + H_{n-1} + C_{n-1}) + (S_{n-1} + C_{n-1}) + (H_{n-1}) \\ &= 2(S_{n-1} + H_{n-1} + C_{n-1}) \\ &= 2(T_{n-1}) \end{aligned}$$

We can see that the population doubles every year. Since $T_0 = 1$ (just a single individual), $T_n = 2^n$. ■

(c) [2 pts] Show that $H_n = T_{n-1} - H_{n-1}$ for $n > 0$.

Solution. We use the expression for H_n from the recurrence:

$$\begin{aligned} H_n &= S_{n-1} + C_{n-1} \\ &= (S_{n-1} + H_{n-1} + C_{n-1}) - H_{n-1} \\ &= T_{n-1} - H_{n-1} \end{aligned}$$

■

(d) [15 pts] Give a closed-form expression for H_n. You may use, without proof, the fact stated in part (b) and the recurrence given in part (c).

Solution. From parts (b) and (c), we have

$$\begin{aligned} H_n &= T_{n-1} - H_{n-1} \\ &= 2^{n-1} - H_{n-1} \\ &= (1/2)2^n - H_{n-1} \\ H_n + H_{n-1} &= (1/2)2^n \end{aligned}$$

This is a linear recurrence. We first solve for the particular solution. Since the $(1/2)2^n$ term is an exponential, we try $f(n) = a2^n$. Plugging this into the recurrence gives us

$$\begin{aligned} a2^n + a2^{n-1} &= (1/2)2^n \\ a2^n + (1/2)a2^n &= (1/2)2^n \\ (3/2)a &= 1/2 \\ a &= 1/3 \end{aligned}$$

Next we solve for the homogeneous solution:

$$\begin{aligned} H_n + H_{n-1} &= 0 \\ r + 1 &= 0 \\ r &= -1 \end{aligned}$$

So our expression for H_n is of the form

$$H_n = A(-1)^n + (1/3)2^n$$

We solve for A by using the initial condition $H_0 = 0$, since there are no Haskeleons in year 0.

$$\begin{aligned} 0 &= A(-1)^0 + (1/3)2^0 \\ &= A + 1/3 \\ A &= -1/3 \end{aligned}$$

The final expression for H_n is thus

$$H_n = (-1/3)(-1)^n + (1/3)2^n$$

■

Problem 8. [15 points] Asymptotic Awesomeness

For each row in the following table, determine whether there exist functions f and g that satisfy all the properties marked **Yes** and do *not* satisfy the properties marked **No**. You do not have to provide examples.

Solution. **(a)** No. $f = \Theta(g)$ implies that $f = \Omega(g)$.

(b) Yes. Example: $f(n) = n^2; g(n) = n$.

(c) No. $f = o(g)$ implies that $f = O(g)$.

(d) Yes. Example: $f(n) = n; g(n) = 2n$.

(e) Yes. Example: $f(n) = n^2; g(n) = n^{(-1)^n+1}$.

(f) Yes. Example: $f(n) = n^{(-1)^{n+1}+1}; g(n) = n^{(-1)^n+1}$.

■

Problem 9. [20 points] Yet Another Graph Proof

Prove that in a finite directed graph, if every node has at least one outgoing edge, then the graph has a cycle.

(*Hint: Consider the longest path.*)

Solution. Suppose that every node has at least one outgoing edge. Since the digraph is finite, there exists a longest path $v_1 \to v_2 \to \ldots \to v_h$. Node v_h has an outgoing edge $v_h \to v$. If $v \notin \{v_1, v_2, \ldots, v_h\}$, then $v_1 \to v_2 \to \ldots \to v_h \to v$ is a longer path of length $h+1$. Therefore, $v \in \{v_1, v_2, \ldots, v_h\}$, that is, $v = v_i$ for some $1 \leq i \leq h$. This means that the graph has a cycle $v_i \to \ldots \to v_h \to v_i$. ■

Problem 10. [20 points] Revenge of the Slipped Disc Puzzle™: The Curse of 6.042

(This problem is similar to the Slipped Disc Puzzle™ of Quiz 1, but here we rotate 5 tiles instead of 4.)

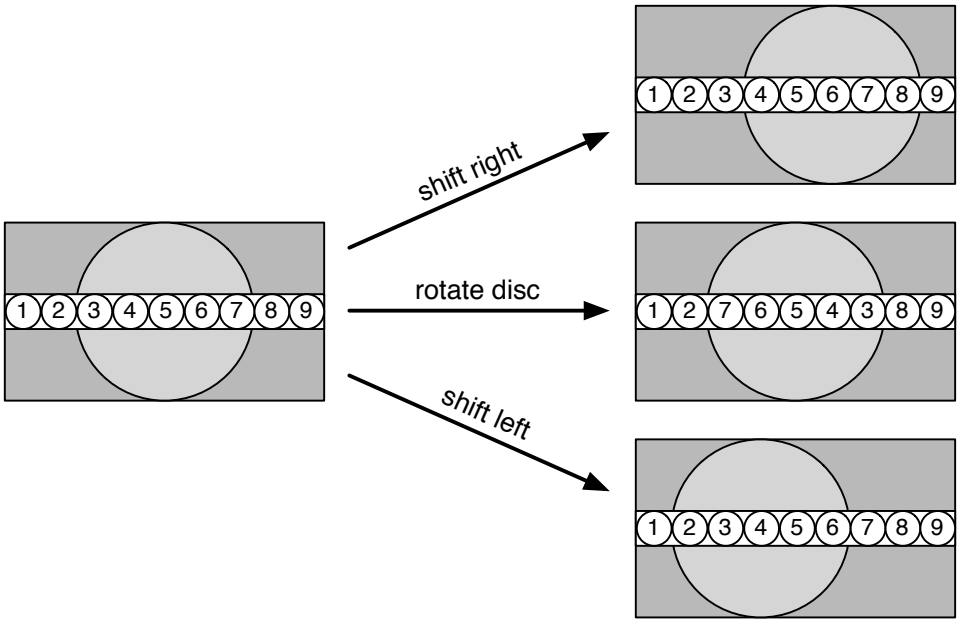
	$f = \Theta(g)$	$f = O(g)$	$f = o(g)$	$f = \Omega(g)$	$f = \omega(g)$	Do f, g exist?
(a)	Yes	Yes	Yes	No	No	**No**
(b)	No	No	No	Yes	Yes	**Yes**
(c)	No	No	Yes	No	No	**No**
(d)	Yes	Yes	No	Yes	No	**Yes**
(e)	No	Yes	No	No	No	**Yes**
(f)	No	No	No	No	No	**Yes**

The Super Awesome Extreme zomgroflolwut Spifftastic-to-the-Max Slipped Disc Puzzle™ consists of a track holding 9 circular tiles. In the middle is a disc that can slide left and right and rotate 180° to change the positions of *exactly* **five** tiles. As shown below, there are three ways to manipulate the puzzle:

Shift Right: The center disc is moved one unit to the right (if there is space).

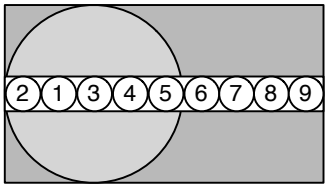
Rotate Disc: The **five** tiles in the center disc are reversed.

Shift Left: The center disc is moved one unit to the left (if there is space).

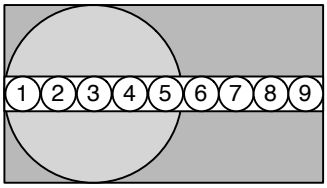


Prove that if the puzzle starts in an initial state with all but tiles 1 and 2 in their natural order, then it is impossible to reach a goal state where all the tiles are in their natural order. The initial and goal states are shown below:

Initial State



Goal State



Write your proof on the next page...

Solution. Order the tiles from left to right in the puzzle. Define an *inversion* to be a pair of tiles that is out of their natural order (e.g. 4 appearing to the left of 3).

Lemma. *Starting from the initial state there is an odd number of inversions after any number of transitions.*

Proof. The proof is by induction. Let $P(n)$ be the proposition that starting from the initial state there is an odd number of inversions after n transitions.

Base case: After 0 transitions, there is one inversion, so $P(0)$ holds.

Inductive step: Assume $P(n)$ is true. Say we have a configuration that is reachable after $n+1$ transitions.

1. Case 1: The last transition was a shift left or shift right

   In this case, the left-to-right order of the discs does not change and thus the number of inversions remains the same as in

2. The last transition was a rotate disc.

   In this case, six pairs of disks switch order. If there were x inversions among these pairs after n transitions, there will be $6-x$ inversions after the reversal. If x is odd, $6-x$ is odd, so after $n+1$ transitions the number of inversions is odd.

□

Conclusion: Since all reachable states have an odd number of inversions and the goal state has an even number of inversions (specifically 0), the goal state cannot be reached. ■

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Final

Name: _____

- This quiz is **closed book**, but you may have two 8.5″ × 11″ sheets with notes in your own handwriting on both sides.
- Calculators are not allowed.
- You may assume all of the results presented in class.
- 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.
- The exam ends at 4:30 PM.

Problem	Points	Grade	Grader	Problem	Points	Grade	Grader
1	20			8	20		
2	12			9	11		
3	10			10	10		
4	10			11	10		
5	15			12	32		
6	10			13	10		
7	10			total	180		

Problem 1. [20 points]

We define the following recurrence for $n \geq 0$:

$$T_{n+2} = T_{n+1} + 2T_n$$

where $T_0 = T_1 = 1$.

(a) [8 pts] Prove by induction that T_n is odd for $n \geq 0$. You do not need to solve the recurrence for this.

(b) [12 pts] Prove by induction that $\gcd(T_{n+1}, T_n) = 1$ for $n \geq 0$. You may assume that T_n is odd for all n. You do not need to solve the recurrence for this.

Problem 2. [12 points]

Find a closed-form solution to the following recurrence:

$$\begin{aligned} x_0 &= 4 \\ x_1 &= 23 \\ x_n &= 11x_{n-1} - 30x_{n-2} \text{ for } n \geq 2. \end{aligned}$$

Problem 3. [10 points] Note: in this question, you may use "choose" notation or factorials in your answers for both (a) and (b). In the card game of bridge, you are dealt a hand of 13 cards from the standard 52-card deck.

(a) [5 pts] A balanced hand is one in which a player has roughly the same number of cards in each suit. How many different hands are there where the player has 4 cards in one suit and 3 cards in each of the other suits?

(b) [5 pts] Not surprisingly, a non-balanced hand is one in which a player has more cards in some suits than others. Hands that are very desired are ones where over half the cards are in one suit. How many different hands are there where there are exactly 7 cards in one suit?

Problem 4. [10 points] Three pairs of twins are sharing a bucket of 24 pieces of chicken from Olive Oyl's. Through the magic of twinship, each person always eats exactly as many pieces of chicken as his or her twin. If each person eats at least one piece of chicken, how many ways can the chicken pieces be distributed amongst this family? For the purposes of this problem, chicken pieces are indistinguishable, and every piece is eaten.

Problem 5. [15 points] A family of seven folk eat a meal at Kansas Flightless Chicken that comes with unlimited free biscuits (the restaurant owner has deep pockets). Being a caring family, they **only** eat biscuits by sharing them with other family members. However, the family obeys the following two rules:

- A biscuit must be shared between **exactly two** people
- If two people share a biscuit, they cannot share another biscuit with each other.

Prove that two members of this family must have shared the same number of biscuits over the course of dinner.

Problem 6. [10 points] Prove that for any m, n, k such that $0 \leq k < m < n - k$, the following identity holds:

$$\binom{n}{k} = \sum_{i=0}^{k} \binom{m}{i} \cdot \binom{n-m}{k-i}$$

Problem 7. [10 points] Let G be a bipartite graph with n nodes and k connected components. You (mutually) independently color each of the nodes of G red or black with equal probabilities. What is the probability that your coloring is a valid 2-coloring of G? (Hint: the answer does not depend on the number of edges.)

Problem 8. [20 points] In the far-off land of Spain, two soccer teams, Barcelona and Madrid have been battling each other for centuries to see who is more awesome. They play two games a year: one in the Spring and one in the Fall. The Madrid team has a tendency to fire their coaches very often to try to improve their results. The outcomes of the games are as follows:

1. If Madrid has not fired their coach since the previous game:
   - Barcelona wins with probability $\frac{2}{5}$
   - Madrid wins with probability $\frac{2}{5}$
   - They tie with probability $\frac{1}{5}$
2. If Madrid has fired their coach since the previous game (they haven't yet realized that it's a bad idea):
   - Barcelona wins with probability $\frac{3}{5}$
   - Madrid wins with probability $\frac{1}{5}$
   - They tie with probability $\frac{1}{5}$

Now, Madrid does not fire their coach if they win or tie, but if they lose, they will fire their coach with 90% probability following the loss. For the following two questions, assume that Madrid did not fire their coach before the Spring 2010 game. Here, everything is mutually independent unless otherwise specified. You can use the following timeline to help you visualize the situation.

(a) [10 pts] Given that Madrid lost the Fall 2010 game, what is the probability that they fired their coach between the Spring and Fall games of 2010?

(b) [10 pts] What is the probability that Madrid fired their coach between the Spring and Fall games in 2010 given that they lost BOTH of the games in 2010?

Problem 9. [11 points] Harvard's loan officer is evaluating her portfolio of student loans. There are 4 especially suspicious loans. Loan A is given to an average Harvard student who does average Harvard stuff. It has a probability of 1/10 of being paid back. Loan B was given to Bill Gates and has a 1/5 probability of being paid back. Loan C was given to David the evil TA and has a probability of 1/3 of being paid back. Loan D was given to a 6.042 student and has 1/2 probability of being paid back. In each of these parts, express any numbers in your answer as ratios of integers.

(a) [3 pts] If the probabilities of being paid back are mutually independent, what is the probability that all four loans are paid back?

(b) [4 pts] If the probabilities are pairwise independent, what is the most that you can say about the probability that all four loans get paid back?

(c) [4 pts] If you cannot assume any independence, what is the most you can say about the probability that all four loans are paid back?

Problem 10. [10 points]

Consider a length n vector of integers, $x = (x_1, x_2, \ldots, x_n)$ where the entries of the vector are integers in the set $\{1, 2, \ldots, n\}$, and the x_i are selected uniformly at random. Note that this means numbers can repeat; x could, for example, be a vector of all 1's. Let A be the number of entries x_i in the vector for which $x_i \leq i$. Calculate $\mathrm{Ex}[A]$ and give it in closed form.

Problem 11. [10 points] Let R be a positive random variable with $\text{Ex}[R^3] = k$. Prove that $Pr(R \geq x) \leq k/x^3$ for any $x > 0$.

Problem 12. [32 points]

In all parts of this problem, assume that we are using fair, regular dice (six-sided with values $1, 2, 3, 4, 5, 6$ appearing equally likely). Furthermore, assume that all dice rolls are mutually independent events.

(a) [4 pts] You roll two dice and look at the sum of the faces that come up. What is the expected value of this sum? Express your answer as a real number.

(b) [7 pts] Assuming that the two dice are independent, calculate the variance of their sum. Express your answer as a real number.

(c) [7 pts] You repeatedly roll two fair dice and look at the sum. What is the probability that you will roll a sum of 4 before you roll a sum of 7? Express your answer as a real number.

(d) [7 pts] What is the expected number of rolls until you get a sum of 4 or a sum of 7? (For example, if you get 7 on the first roll, the number of rolls is 1.) Express your answer as a real number.

(e) [7 pts] You roll 10 dice. Using the Chernoff Bound, give an upper bound for the probability that 8 or more of them rolled a 1 or a 2? You don't need to calculate the value with a calculator (since you do not have one), but please write it in simplest terms.

Problem 13. [10 points]

An MIT student is walking along Mass Ave and is torn between making a trip to the pub and going to her dorm to finish her 6.042 homework. The student is n steps from the pub and $T - n$ steps from the dorm where $T >= 3$ is the distance between the pub and the dorm, both located on Mass Ave (which we will assume to be a straight line). The student flips a coin at each step to decide which way to go. With probability p, the student takes 2 steps toward the pub and with probability $1 - p$, she takes 1 step toward her dorm. If the student is one or two steps away from the pub and takes two steps toward the pub, then she reaches the pub. If she is one step away from her dorm and takes a step toward her dorm, then she reaches the dorm.. Let X_n be the probability that the student reaches the pub before she reaches the dorm. Assume that all coin tosses are mutually independent. Find a recurrence for X_n. (You do not need to solve the recurrence, but your answer should contain sufficient information such that a 6.042 could solve it given enough time.)

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Midterm Practice Problems

Name: _____

- This quiz is **closed book**, but you may have one 8.5×11" sheet with notes in your own handwriting on both sides.
- Calculators are not allowed.
- You may assume all of the results presented in class.
- 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.

Problem 1. [10 points] In problem set 1 you showed that the `nand` operator by itself can be used to write equivalent expressions for all other Boolean logical operators. We call such an operator *universal*. Another universal operator is `nor`, defined such that $P\texttt{nor}Q \Leftrightarrow \neg(P \vee Q)$.

Show how to express $P \wedge Q$ in terms of: `nor`, P, Q, and grouping parentheses.

Problem 2. [15 points] We define the sequence of numbers

$$a_n = \begin{cases} 1 & \text{if } 0 \leq n \leq 3, \\ a_{n-1} + a_{n-2} + a_{n-3} + a_{n-4} & \text{if } n \geq 4. \end{cases}$$

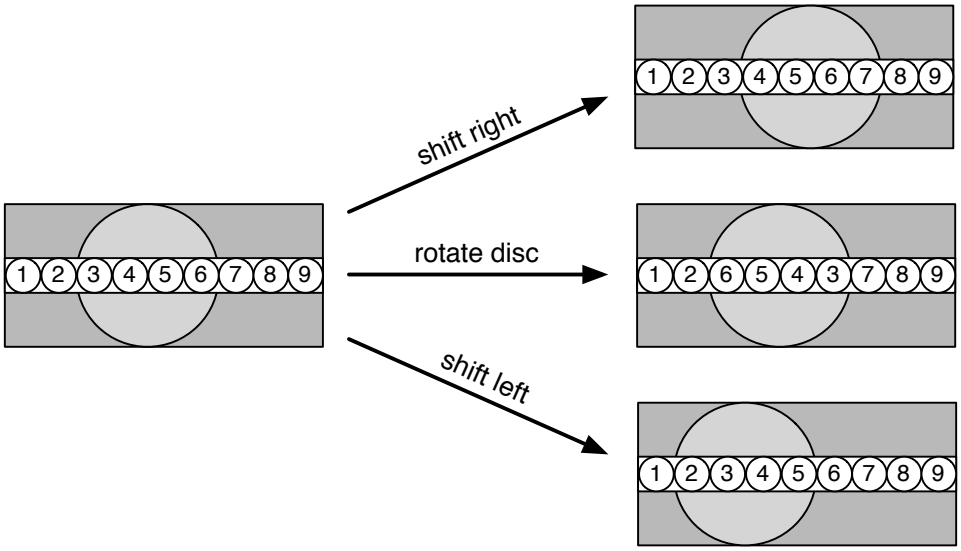
Prove that $a_n \equiv 1 \pmod 3$ for all $n \geq 0$.

Problem 3. [20 points] The Slipped Disc Puzzle™ consists of a track holding 9 circular tiles. In the middle is a disc that can slide left and right and rotate 180° to change the positions of *exactly* four tiles. As shown below, there are three ways to manipulate the puzzle:

Shift Right: The center disc is moved one unit to the right (if there is space)

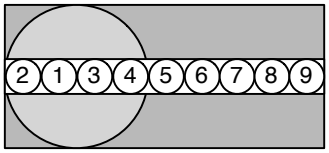
Rotate Disc: The four tiles in the center disc are reversed

Shift Left: The center disc is moved one unit to the left (if there is space)

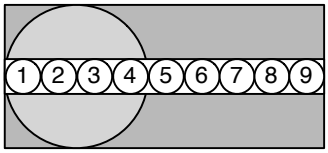


Prove that if the puzzle starts in an initial state with all but tiles 1 and 2 in their natural order, then it is impossible to reach a goal state where all the tiles are in their natural order. The initial and goal states are shown below:

Initial State



Goal State

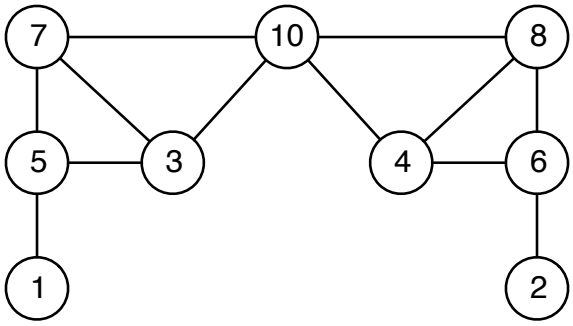


Write your proof on the next page...

room for problem 3...

Problem 4. [10 points] Find the multiplicative inverse of 17 modulo 72 in the range $\{0, 1, \ldots, 71\}$.

Problem 5. [15 points] Consider a graph representing the main campus buildings at MIT.

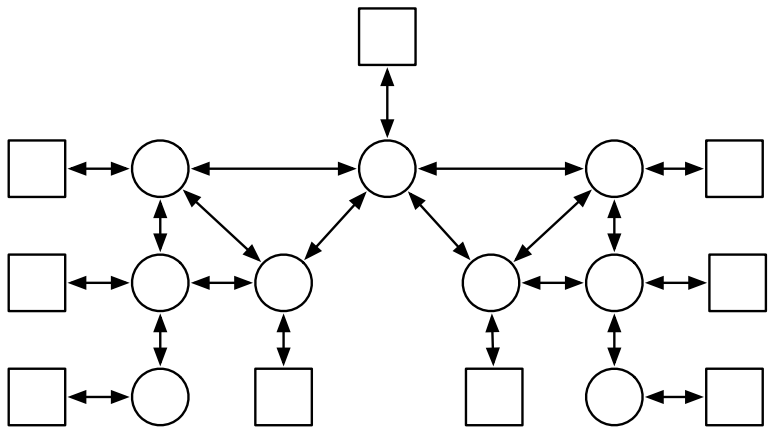


(a) [3 pts] Is this graph bipartite? Provide a brief argument for your answer.

(b) [4 pts] Does this graph have an Euler circuit? Provide a brief argument for your answer.

Problem 5 continued...

Now suppose each building has separate mail collection and drop-off boxes and each collection box has a single package destined for a unique drop-off box (i.e. a permutation). We can model this as a permutation routing problem by treating the buildings as switches, attaching an input and output terminal to each of the nine buildings, and treating the existing edges as bidirectional as in the graph below:



(c) [4 pts] Give the diameter of this graph:

(d) [4 pts] What is the max congestion of this graph? That is, in the worst case permutation, how many packages would need to pass through a single building? Provide a brief argument for your answer.

Problem 6. [10 points]

A tournament graph $G = (V, E)$ is a directed graph such that there is either an edge from u to v or an edge from v to u for *every* distinct pair of nodes u and v. (The nodes represent players and an edge $u \to v$ indicates that player u beats player v.)

Consider the "beats" relation implied by a tournament graph. Indicate whether or not each of the following relational properties hold *for all* tournament graphs and briefly explain your reasoning. You may assume that a player never plays herself.

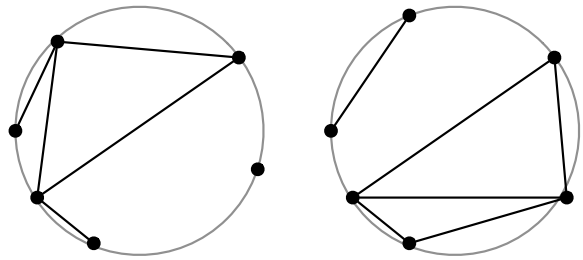
1. **transitive**

2. **symmetric**

3. **antisymmetric**

4. **reflexive**

Problem 7. [20 points] An outerplanar graph is an undirected graph for which the vertices *can be* placed on a circle in such a way that no edges (drawn as straight lines) cross each other. For example, the complete graph on 4 vertices, K_4, is not outerplanar but any proper subgraph of K_4 with strictly fewer edges is outerplanar. Some examples are provided below:



Prove that any outerplanar graph is 3-colorable. A fact you may use without proof is that any outerplanar graph has a vertex of degree at most 2.

Problem 8. [10 points] Give upper and lower bounds for the following expression which differ by at most 1.

$$\sum_{i=1}^{n} \frac{1}{i^3}$$

Problem 9. [15 points] Circle every symbol on the left that could correctly appear in the box to its right. For each of the six parts you may need to circle any number of symbols.

(a) $O \quad \Omega \quad \Theta \quad o \quad \omega \quad \sim$ $\qquad 6n^2 + 7n - 10 \quad = \boxed{} \, (n^2)$

(b) $O \quad \Omega \quad \Theta \quad o \quad \omega \quad \sim$ $\qquad 6^n \quad = \boxed{} \, (n^6)$

(c) $O \quad \Omega \quad \Theta \quad o \quad \omega \quad \sim$ $\qquad n! \quad = \boxed{} \, (n^n)$

(d) $O \quad \Omega \quad \Theta \quad o \quad \omega \quad \sim$ $\qquad \sum_{j=1}^{n} \frac{1}{j} \quad = \boxed{} \, (\ln n)$

(e) $O \quad \Omega \quad \Theta \quad o \quad \omega \quad \sim$ $\qquad \ln(n^3) \quad = \boxed{} \, (\ln n)$

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Midterm Practice Problems

Problem 1. [10 points] In problem set 1 you showed that the `nand` operator by itself can be used to write equivalent expressions for all other Boolean logical operators. We call such an operator *universal*. Another universal operator is `nor`, defined such that $P \texttt{nor} Q \Leftrightarrow \neg(P \vee Q)$.

Show how to express $P \wedge Q$ in terms of: `nor`, P, Q, and grouping parentheses.

Solution. $(\neg P)\ \texttt{nor}\ (\neg Q) = (P\ \texttt{nor}\ P)\ \texttt{nor}\ (Q\ \texttt{nor}\ Q)$. ■

Problem 2. [15 points] We define the sequence of numbers

$$a_n = \begin{cases} 1 & \text{if } 0 \le n \le 3, \\ a_{n-1} + a_{n-2} + a_{n-3} + a_{n-4} & \text{if } n \ge 4. \end{cases}$$

Prove that $a_n \equiv 1 \pmod 3$ for all $n \ge 0$.

Solution. Proof by strong induction. Let $P(n)$ be the predicate that $a_n \equiv 1 \pmod 3$.

Base case: For $0 \le n \le 3$, $a_n = 1$ and is therefore $\equiv 1 \pmod 3$.

Inductive step: For $n \ge 4$, assume $P(k)$ for $0 \le k \le n$ in order to prove $P(n+1)$.

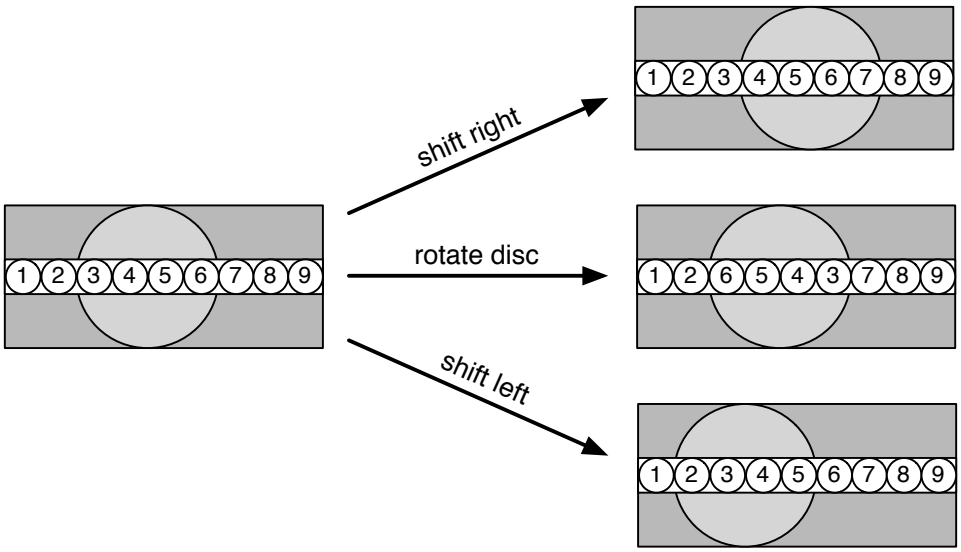
In particular, since each of a_{n-4}, a_{n-3}, a_{n-2} and a_{n-1} is $\equiv 1 \pmod 3$, their sum must be $\equiv 4 \equiv 1 \pmod 3$. Therefore, $a_n \equiv 1 \pmod 3$ and $P(n+1)$ holds. ■

Problem 3. [20 points] The Slipped Disc Puzzle™ consists of a track holding 9 circular tiles. In the middle is a disc that can slide left and right and rotate 180° to change the positions of *exactly* four tiles. As shown below, there are three ways to manipulate the puzzle:

Shift Right: The center disc is moved one unit to the right (if there is space)

Rotate Disc: The four tiles in the center disc are reversed

Shift Left: The center disc is moved one unit to the left (if there is space)



Prove that if the puzzle starts in an initial state with all but tiles 1 and 2 in their natural order, then it is impossible to reach a goal state where all the tiles are in their natural order. The initial and goal states are shown below:

Write your proof on the next page...

Solution. Order the tiles from left to right in the puzzle. Define an *inversion* to be a pair of tiles that is out of their natural order (e.g. 4 appearing to the left of 3).

Lemma. *Starting from the initial state there is an odd number of inversions after any number of transitions.*

Proof. The proof is by induction. Let $P(n)$ be the proposition that starting from the initial state there is an odd number of inversions after n transitions.

Base case: After 0 transitions, there is one inversion, so $P(0)$ holds.

Inductive step: Assume $P(n)$ is true. Say we have a configuration that is reachable after $n+1$ transitions.

1. Case 1: The last transition was a shift left or shift right

   In this case, the left-to-right order of the discs does not change and thus the number of inversions remains the same as in

2. The last transition was a rotate disc.

   In this case, six pairs of disks switch order. If there were x inversions among these pairs after n transitions, there will be $6 - x$ inversions after the reversal. If x is odd, $6 - x$ is odd, so after $n + 1$ transitions the number of inversions is odd.

$\square$

Conclusion: Since all reachable states have an odd number of inversions and the goal state has an even number of inversions (specifically 0), the goal state cannot be reached. $\blacksquare$

Problem 4. [10 points] Find the multiplicative inverse of 17 modulo 72 in the range $\{0, 1, \ldots, 71\}$.

Solution. Since 17 and $72 = 2^3 3^2$ are relatively prime, an inverse exists and can be found by either Euler's theorem or the Pulverizer.

Solution 1: Euler's Theorem

$$\begin{aligned}
\phi(72) &= \phi(2^3 \cdot 3^2) \\
&= \phi(2^3) \cdot \phi(3^2) && \text{(since } 2^3 \text{ and } 3^2 \text{ are rel. prime)} \\
&= (2^3 - 2^2)(3^2 - 3^1) && \text{(since 2 and 3 are prime)} \\
&= 4 \cdot 6 = 24
\end{aligned}$$

Therefore, $17^{\phi(72)-1} = 17^{23}$ is *an* inverse of 17. To find *the* inverse in the range $\{0, 1, \ldots, 71\}$ we take the remainder using the method of repeated squaring:

$$\begin{aligned}
17 &= 17 \\
17^2 &= 289 \\
&\equiv 1 && \text{(since } 289 = 4 \cdot 72 + 1) \\
17^4 &\equiv 1^2 = 1 \\
17^8 &\equiv 1 \\
&\ldots \textit{etc.}
\end{aligned}$$

Therefore the inverse of 17 in the range $\{0, 1, \ldots, 71\}$ is given by,

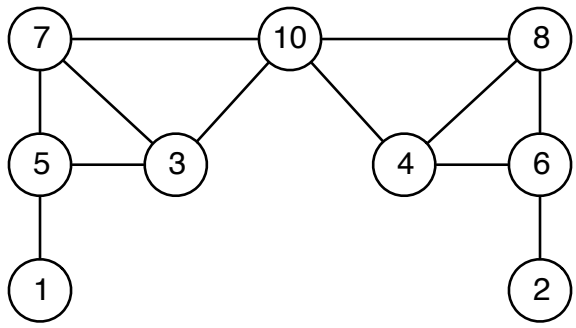
$$\begin{aligned}
17^{23} &= 17^{16} 17^4 17^2 17^1 \\
&\equiv 1 \cdot 1 \cdot 1 \cdot 17 \\
&= 17
\end{aligned}$$

Solution 2: The Pulverizer

x	y	rem xy	$=$	$x - q \cdot y$
72	17	4	$=$	$72 - 4 \cdot 17$
17	4	1	$=$	$17 - 4 \cdot 4$
			$=$	$17 - 4 \cdot (72 - 4 \cdot 17)$
			$=$	$17 \cdot 17 - 4 \cdot 72$
4	1	0		

Since $17^2 - 4 \cdot 72 = 1$, $17^2 \equiv 1 \pmod{72}$ and so 17 is self inverse. ■

Problem 5. [15 points] Consider a graph representing the main campus buildings at MIT.



(a) [3 pts] Is this graph bipartite? Provide a brief argument for your answer.

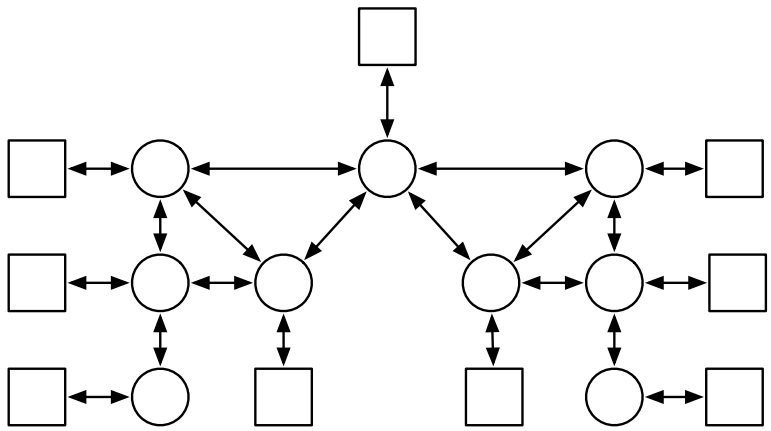
Solution. No, there is an odd-length cycle ■

(b) [4 pts] Does this graph have an Euler circuit? Provide a brief argument for your answer.

Solution. This graph does not have an Euler circuit because there are vertices with odd degree ■

Problem 5 continued...

Now suppose each building has separate mail collection and drop-off boxes and each collection box has a single package destined for a unique drop-off box (i.e. a permutation). We can model this as a permutation routing problem by treating the buildings as switches, attaching an input and output terminal to each of the nine buildings, and treating the existing edges as bidirectional as in the graph below:



(c) [4 pts] Give the diameter of this graph:

Solution. The diameter is 8, the length of a shortest path between the terminals for buildings 1 and 2. ■

(d) [4 pts] What is the max congestion of this graph? That is, in the worst case permutation, how many packages would need to pass through a single building? Provide a brief argument for your answer.

Solution. The maximum congestion is 9. Consider a permutation where all the packages on the left side are destined for drop-offs on the right side and vice versa with the building 10 package destined for building 10. In this case, all 9 packages must pass through building 10, giving a congestion of 9, which is the maximum for a graph with nine input-output pairs. ■

Problem 6. [10 points]

A tournament graph $G = (V, E)$ is a directed graph such that there is either an edge from u to v or an edge from v to u for *every* distinct pair of nodes u and v. (The nodes represent players and an edge $u \to v$ indicates that player u beats player v.)

Consider the "beats" relation implied by a tournament graph. Indicate whether or not each of the following relational properties hold *for all* tournament graphs and briefly explain your reasoning. You may assume that a player never plays herself.

1. **transitive**

   Solution. The "beats" relation is not transitive because there could exist a cycle of length 3 where x beats y, y beats z and z beats x. By the definition of a tournament, x cannot then beat y in such a situation. ■

2. **symmetric**

   Solution. The "beats" relation is not symmetric by the definition of a tournament: if x beats y then y does not beat x. ■

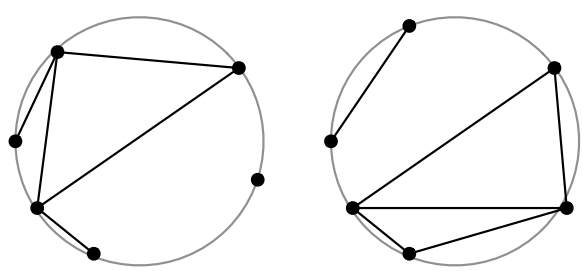
3. **antisymmetric**

   Solution. The "beats" relation is antisymmetric since for any distinct players x and y, if x beats y then y does not beat x. ■

4. **reflexive**

   Solution. The "beats" relation is not reflexive since a tournament graph has no self-loops. ■

Problem 7. [20 points] An outerplanar graph is an undirected graph for which the vertices *can be* placed on a circle in such a way that no edges (drawn as straight lines) cross each other. For example, the complete graph on 4 vertices, K_4, is not outerplanar but any proper subgraph of K_4 with strictly fewer edges is outerplanar. Some examples are provided below:



Prove that any outerplanar graph is 3-colorable. A fact you may use without proof is that any outerplanar graph has a vertex of degree at most 2.

Solution. *Proof.* Proof by induction on the number of nodes n with the induction hypothesis $P(n) =$ "every outerplanar graph with n vertices is 3-colorable."

Base case: For $n = 1$ the single node graph with no edges is trivially outerplanar and 3-colorable.

Inductive step: Assume $P(n)$ holds and let G_{n+1} be an outerplanar graph with $n+1$ vertices. There must exist a vertex v in G_{n+1} with degree at most 2. Removing v and all its incident edges leaves a subgraph G_n with n vertices.

Since G_{n+1} could be drawn with its vertices on a circle and its edges drawn as straight lines without intersections, any subgraph can also be drawn in such a way and so G_n is also an outerplanar graph. $P(n)$ implies G_n is 3-colorable. Therefore we can color all the vertices in G_{n+1} other than v using only 3 colors and since $\deg(v) \leq 2$ we may color it a color different than the vertices adjacent to it using only 3 colors. Therefore, G_{n+1} is 3-colorable and $P(n+1)$ holds.

□

■

Problem 8. [10 points] Give upper and lower bounds for the following expression which differ by at most 1.

$$\sum_{i=1}^{n} \frac{1}{i^3}$$

Solution. To find upper and lower bounds, we use the integral method:

$$\begin{aligned}
\sum_{i=1}^{n} \frac{1}{i^3} &\leq 1 + \int_1^n \frac{1}{x^3} dx \\
&= 1 - \frac{1}{2} x^{-2} \Big|_1^n \\
&= 1 - \frac{1}{2}\left(\frac{1}{n^2} - 1\right) = \frac{3}{2} - \frac{1}{2n^2} \\
\sum_{i=1}^{n} \frac{1}{i^3} &\geq \frac{1}{n^3} + \int_1^n \frac{1}{x^3} dx \\
&= \frac{1}{n^3} - \frac{1}{2} x^{-2} \Big|_1^n \\
&= \frac{1}{n^3} - \frac{1}{2}\left(\frac{1}{n^2} - 1\right) = \frac{1}{2} + \frac{1}{n^3} - \frac{1}{2n^2}
\end{aligned}$$

■

Problem 9. [15 points] Circle every symbol on the left that could correctly appear in the box to its right. For each of the six parts you may need to circle any number of symbols.

(a) $O \quad \Omega \quad \Theta \quad o \quad \omega \quad \sim$ $\qquad 6n^2 + 7n - 10 \quad = \boxed{O, \Omega, \Theta} \left(n^2\right)$

(b) $O \quad \Omega \quad \Theta \quad o \quad \omega \quad \sim$ $\qquad 6^n \quad = \boxed{\Omega, \omega} \left(n^6\right)$

(c) $O \quad \Omega \quad \Theta \quad o \quad \omega \quad \sim$ $\qquad n! \quad = \boxed{O, o} \left(n^n\right)$

(d) $O \quad \Omega \quad \Theta \quad o \quad \omega \quad \sim$ $\qquad \sum_{j=1}^{n} \frac{1}{j} \quad = \boxed{O, \Omega, \Theta, \sim} (\ln n)$

(e) $O \quad \Omega \quad \Theta \quad o \quad \omega \quad \sim$ $\qquad \ln(n^3) \quad = \boxed{O, \Omega, \Theta} (\ln n)$

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Midterm

Name: _____

- This quiz is **closed book**, but you may have one $8.5 \times 11''$ sheet with notes in your own handwriting on both sides.
- Calculators are not allowed.
- You may assume all of the results presented in class.
- 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.
- The exam ends at 9:30 PM.

Problem	Points	Grade	Grader
1	10		
2	10		
3	20		
4	15		
5	20		
6	25		
7	10		
8	10		
Total	120		

Problem 1. [10 points]

Consider these two propositions:

$$\texttt{P}: \ (A \vee B) \Rightarrow C$$

$$\texttt{Q}: \ (\neg C \Rightarrow \neg A) \vee (\neg C \Rightarrow \neg B)$$

Which of the following best describes the relationship between P and Q? Please circle exactly one answer.

1. P and Q are equivalent
2. $P \Rightarrow Q$
3. $Q \Rightarrow P$
4. All of the above
5. None of the above

Draw a truth table to illustrate your reasoning. You can use as many columns as you need.

Problem 2. [10 points]

Let $G_0 = 1$, $G_1 = 3$, $G_2 = 9$, and define

$$G_n = G_{n-1} + 3G_{n-2} + 3G_{n-3} \tag{1}$$

for $n \geq 3$. Show by induction that $G_n \leq 3^n$ for all $n \geq 0$.

Problem 3. [20 points]

In the game of Squares and Circles, the players (you and your computer) start with a shared finite collection of shapes: some circles and some squares. Players take turns making moves. On each move, a player chooses any two shapes from the collection. These two are replaced with a single one according to the following rule:

A pair of identical shapes is replaced with a square. A pair of different shapes is replaced with a circle.

At the end of the game, when only one shape remains, you are a winner if the remaining shape is a circle. Otherwise, your computer wins.

(a) [5 pts] Prove that the game will end.

(b) [15 pts] Prove that you will win if and only if the number of circles initially is odd.

Problem 4. [15 points]

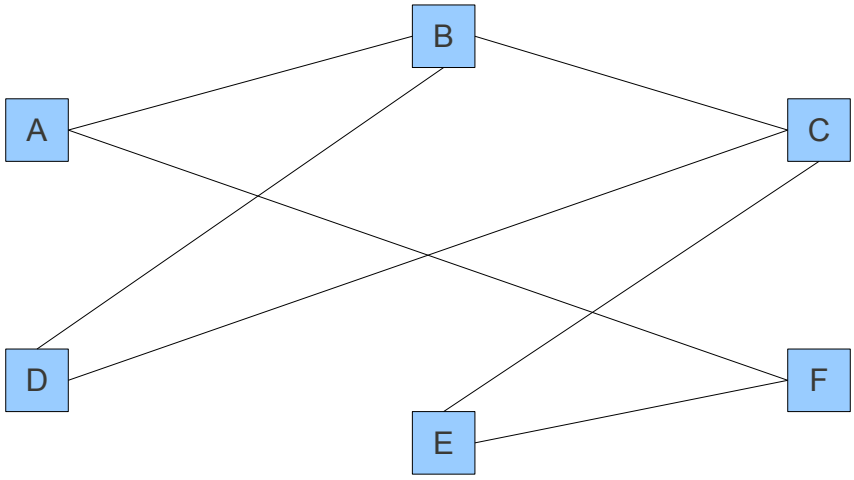
(a) [8 pts] Find a number $x \in \{0, 1, \ldots, 112\}$ such that $11x \equiv 1 \pmod{113}$.

(b) [7 pts] Find a number $y \in \{0, 1, \ldots, 112\}$ such that $11^{112111} \equiv y \pmod{113}$ (*Hint: Note that* 113 *is a prime.*)

Problem 5. [20 points]

Consider the simple graph G given in figure 1.

Figure 1: Simple graph G



(a) [4 pts] Give the diameter of G.

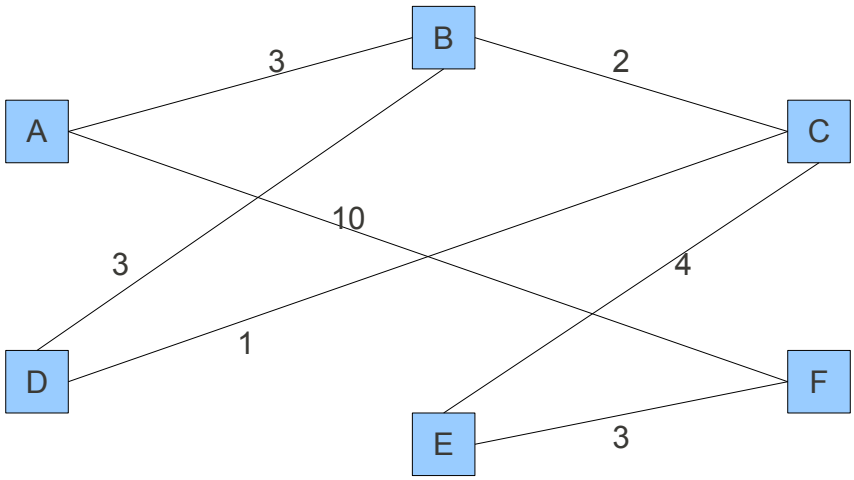
(b) [4 pts] Give a Hamiltonian Cycle on G.

(c) [4 pts] Give a coloring on G and show that it uses the smallest possible number of colors.

(d) [4 pts] Does G have an Eulerian cycle? Justify your answer.

Now consider graph H, which is like G but with weighted edges, in figure 2:

Figure 2: Weighted graph H



(e) [4 pts] Give a list of edges reflecting the order in which one of the greedy algorithms presented in class (i.e. in lecture, recitation, or the course text) would choose edges when finding an MST on H.

Problem 6. [25 points] Let G be a graph with m edges, n vertices, and k components. Prove that G contains at least $m - n + k$ cycles. (Hint: Prove this by induction on the number of edges, m)

Problem 7. [10 points] For the following sum, find an upper and a lower bound that differ by at most 1.

$$\sum_{i=1}^{\infty} \frac{1}{\sqrt{i^3}}$$

Problem 8. [10 points] State whether each of the following claims is True or False and prove your answer.

(a) [2 pts] $x \ln x$ is $O(x)$

(b) [2 pts] $x/100$ is $o(x)$

(c) [2 pts] x^{n+1} is $\Omega\left(x^n\right)$

(d) [4 pts] $n!$ is $\Theta\left(n^n\right)$.

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Midterm

Problem 1. [10 points]

Consider these two propositions:

$$\texttt{P}: \ (A \vee B) \Rightarrow C$$

$$\texttt{Q}: \ (\neg C \Rightarrow \neg A) \vee (\neg C \Rightarrow \neg B)$$

Which of the following best describes the relationship between P and Q? Please circle exactly one answer.

1. P and Q are equivalent
2. $P \Rightarrow Q$
3. $Q \Rightarrow P$
4. All of the above
5. None of the above

Draw a truth table to illustrate your reasoning. You can use as many columns as you need.

Solution.

A	B	C	P	Q
T	T	T	T	T
T	T	F	F	F
T	F	T	T	T
T	F	F	F	T
F	T	T	T	T
F	T	F	F	T
F	F	T	T	T
F	F	F	T	T

Observe from the last two columns of the table that $P \Rightarrow Q$ is always true, but $Q \Rightarrow P$ is not always true (e.g. line 4). Thus P and Q are not equivalent but $P \Rightarrow Q$.

■

Problem 2. [10 points]

Let $G_0 = 1$, $G_1 = 3$, $G_2 = 9$, and define

$$G_n = G_{n-1} + 3G_{n-2} + 3G_{n-3} \tag{1}$$

for $n \geq 3$. Show by induction that $G_n \leq 3^n$ for all $n \geq 0$.

Solution. The proof is by strong induction with hypothesis $P(n) := G_n \leq 3^n$.

Proof. **Base Cases**

$n = 0$**:** $G_0 = 1 = 3^0$.

$n = 1$**:** $G_1 = 3 \leq 3^1$.

$n = 2$**:** $G_2 = 9 \leq 3^2$.

Inductive Step: Assume $n \geq 3$ and $P(k)$ for all k such that $0 \leq k \leq n$.

$$\begin{aligned}
G_n &= G_{n-1} + 3G_{n-2} + 3G_{n-3} && \text{by (1)} \\
&\leq 3^{n-1} + (3)3^{n-2} + (3)3^{n-3} && \text{by induction hypothesis} \\
&= 3^{n-2}[3 + 3 + 1] \\
&= (7)3^{n-2} \\
&< (9)3^{n-2} \\
&= 3^n
\end{aligned}$$

□

■

Problem 3. [20 points]

In the game of Squares and Circles, the players (you and your computer) start with a shared finite collection of shapes: some circles and some squares. Players take turns making moves. On each move, a player chooses any two shapes from the collection. These two are replaced with a single one according to the following rule:

A pair of identical shapes is replaced with a square. A pair of different shapes is replaced with a circle.

At the end of the game, when only one shape remains, you are a winner if the remaining shape is a circle. Otherwise, your computer wins.

(a) [5 pts] Prove that the game will end.

Solution. *Proof.* We use induction on the number of turns to prove the statement. Let n be the number of shapes originally, and let P(k) be the proposition that if $0 \leq k \leq n-1$ then after k turns, the number of remaining shapes is $n-k$. Thus the game ends after $n-1$ steps.

Base case: $P(0)$ is true by definition; the number of reamaining shapes after 0 turns is $n-0=n$, the original number of shapes.

Inductive step: Now we must show that $P(k)$ implies $P(k+1)$ for all $k \geq 0$. If $k >= n-1$, $P(k)$ implies $P(k+1)$ would always be true, since $P(k+1)$ would be trivially true. So we only need to prrove this for $k < n-1$. So assume for your inductive hypothesis that $P(k)$ is true, where $0 \leq k < n-1$; that is, after k turns the number of remaining shapes in $n-k$. Since $k < n-1$, the number of remaining shapes is $n-k > 1$. Hence there are at least 2 shapes to choose from and the game has not ended yet. In the k+1st turn either the computer will choose 2 shapes, or you will choose two shapes. In either case the two shapes chosen, will be replaced by exactly once. Hence the number of shapes remaining will be $n-k-2+1 = n-k-1 = n-(k+1)$ as desired. This proves that $P(k)$ implies $P(k+1)$ for all $k \geq 0$.

By the principle of induction, $P(k)$ is true for all $k \geq 0$.

Hence, by our inductve hypothes after n-1 turns, 1 shape remains, which by the problem definition implies the game ends. □

■

(b) [15 pts] Prove that you will win if and only if the number of circles initially is odd.

Solution. *Proof.* We use induction on the number of turns to prove the statement. Let a be the number of circles initially, and let P(k) be the proposition that if $0 \leq k \leq n-1$ then after k turns, the number of remaining ciircles is $a - 2i$, for some nonnegative integer i. Thus if a is odd initially, at turn $n-1$, when the game ends, $a - 2i$ circles - an odd number - remain, and since there is only one shape remaining, there must be exactly 1 circle left, and you win.

Base case: $P(0)$ is true by definition; the number of reamaining circles after 0 turns is $a - 2 * 0 = a$, the original number of shapes.

Inductive step: Now we must show that $P(k)$ implies $P(k+1)$ for all $k \geq 0$. If $k >= n-1$, $P(k)$ implies $P(k+1)$ would always be true, since $P(k+1)$ would be trivially true. So we only need to prrove this for $k < n-1$. So assume for your inductive hypothesis that $P(k)$ is true, where $0 \leq k < n-1$; that is, after k turns the number of remaining circles is $a - 2i_1$. for some nonnegative integer i_1. Since $k < n-1$, the number of remaining shapes is $n - k > 1$ (from part a), hence there are at least 2 shapes to choose from and the game has not ended yet. In the k+1st turn either the computer will choose 2 shapes, or you will choose two shapes. If the two shapes chosen are both squares, then they are replaced by a square, and the number of circles doe not change, and hence is still $a - 2i_1$. If the two shapes chosen are both circles, then they are replaced by a square, and the number of circles gets decreased by 2, and is $a - 2i - 2 = a - 2(i+1)$. If one of the shapes chosen was a circle and the other was a square, they get replaced by a circle, and again the number of circles does not change and remains $a - 2i$. Hence in all three transitions P(k+1) holds.

By the principle of induction, $P(k)$ is true for all $k \geq 0$.

Hence, by induction when the game ends the parity of the number of circles is the same as the original parity of the number of circles. So you will win only if the number of circles to begin with was odd.

□

■

Problem 4. [15 points]

(a) [8 pts] Find a number $x \in \{0, 1, \ldots, 112\}$ such that $11x \equiv 1 \pmod{113}$.

Solution. We can do this using the pulverizer. Specifically, if we find a pair (s, t) such that $11s + 113t = 1$, then we know that $11s \equiv 1 \pmod{113}$.

x	y	$\text{rem}(x, y)$	$=$	$x - q \cdot y$
113	11	3	$=$	$113 - 10 \cdot 11$
11	3	2	$=$	$11 - 3 \cdot 3$
			$=$	$11 - 3 \cdot (113 - 10 \cdot 11)$
			$=$	$-3 \cdot 113 + 31 \cdot 11$
3	2	1	$=$	$3 - 2$
			$=$	$(113 - 10 \cdot 11) - (-3 \cdot 113 + 31 \cdot 18)$
			$=$	$(4 \cdot 113 - 41 \cdot 11)$

From the above work we see that $1 = 4 \cdot 113 - 41 \cdot 11$, and so $-41 \cdot 11 \equiv 1 \pmod{113}$. Therefore -41 is *an* inverse of 113. However, it is not the *unique* inverse of 113 in the range $\{1, \ldots, 113\}$, which is given by $\text{rem}(-41, 113) = 72$. So the multiplicative inverse of 11 modulo 113 is 72. ■

(b) [7 pts] Find a number $y \in \{0, 1, \ldots, 112\}$ such that $11^{112111} \equiv y \pmod{113}$ (*Hint: Note that* 113 *is a prime.*)

Solution. By Fermat's Theorem, since 113 is prime and 113 and 11 are relatively prime, it must be that

$$11 \cdot 11^{111} \equiv 11^{113-1} \equiv 1 \pmod{113},$$

so $x \equiv 111 \pmod{113}$ (where x is defined as in part a). As a result,

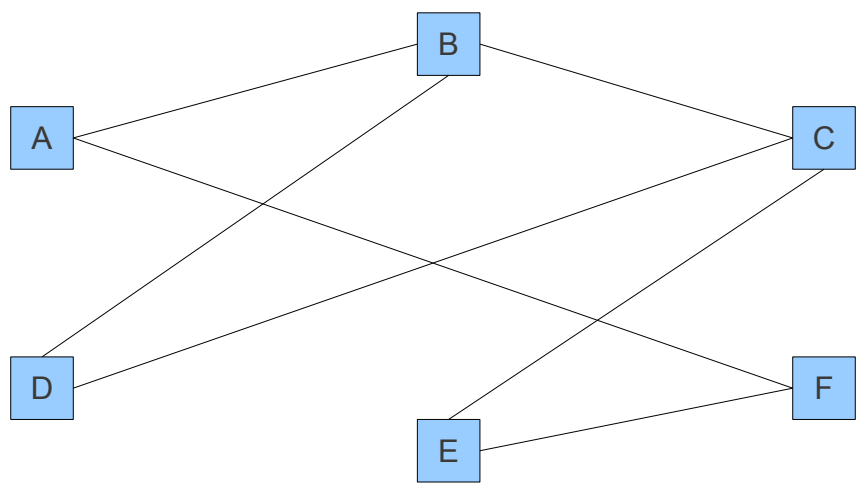
$$11^{112111} \equiv 11^{112 \cdot 1000 + 111} \equiv 11^{112^{1000}} \cdot 11^{111} \equiv 1^{1000} \cdot x \equiv x \equiv 72 \pmod{113},$$

so the answer is 72. ■

Problem 5. [20 points]

Consider the simple graph G given in figure 1.

Figure 1: Simple graph G



(a) [4 pts] Give the diameter of G.

Solution. Recall that the diameter is the maximum of all shortest path lengths between pairs of vertices. Note that the shortest path length between D and F is 3, and all other pairs of non-adjacent vertices share a neighbor. ■

(b) [4 pts] Give a Hamiltonian Cycle on G.

Solution. One possible solution is (A, F, E, C, D, B, A). This cycle and its reverse should constitute all possible solutions. ■

(c) [4 pts] Give a coloring on G and show that it uses the smallest possible number of colors.

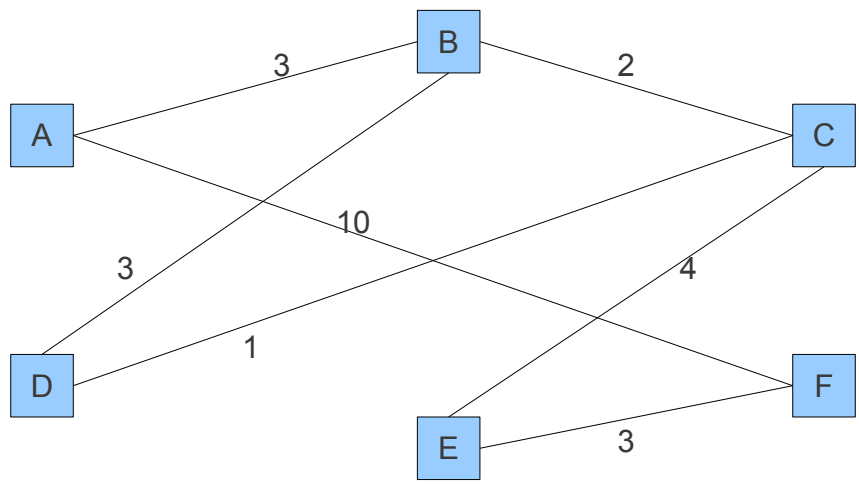
Solution. One possible 3-coloring is: $\{A, D, E\}$ red; $\{B, F\}$ green; C blue. Because there exists an odd-length cycle (e.g. (B, D, C)), no 2-coloring exists. Therefore the given coloring uses the least possible number of colors. ■

(d) [4 pts] Does G have an Eulerian cycle? Justify your answer.

Solution. No. This follows from the fact that there exist vertices with odd degree; e.g. B. ■

Now consider graph H, which is like G but with weighted edges, in figure 2:

Figure 2: Weighted graph H

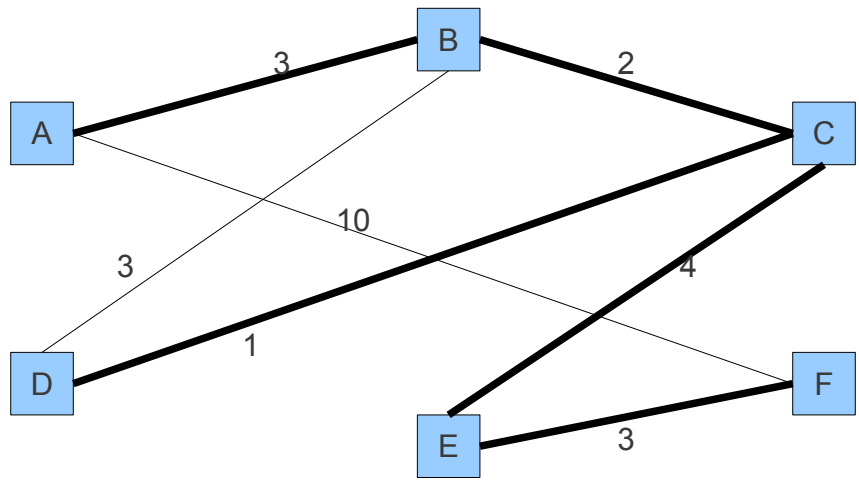


(e) [4 pts] Give a list of edges reflecting the order in which one of the greedy algorithms presented in class (i.e. in lecture, recitation, or the course text) would choose edges when finding an MST on H.

Solution. Kruskal's alg (building up an acyclic subgraph) gives two possible orders: $((C,D),(B,C),(A,B),(E,F),(C,E))$ and $((C,D),(B,C),(E,F),(A,B),(C,E))$. Prim's algorithm (building up a connected, acyclic subgraph) gives one possible order: $((C,D),(B,C),(A,B),(C,E),(E,F))$. Figure 3 below gives the MST generated in any case.

■

Figure 3: MST on graph H



Problem 6. [25 points] Let G be a graph with m edges, n vertices, and k components. Prove that G contains at least $m - n + k$ cycles. (Hint: Prove this by induction on the number of edges, m)

Solution. The proof is by induction on m with hypothesis P(n):= If G is a graph with n vertices, m edges and k components, then G contains at least $m + k - n = c$ cycles

Proof. **Base Case** $m = 0$: Let G be any graph with 0 edges and n vertices. Then since there are no edges, each vertex is its own connected component, hence there are $k = n$ connected components. Since there are no edges there are also no cycles. Lastly we note that $m + k - n = n + 0 - n = 0$, and hence our base case.

Inductive Step Assume that P(m) holds, that is any graph with m edges, n vertices, and k components has at least $m - n + k$ cycles. We must show that P(m+1) holds.

Consider an arbitrary graph G with $m + 1$ edges, n vertices, and k components. Suppose we remove an arbitrary edge, e, from G to obtain G'. This edge, e, was either in a cycle in G or not:

Case 1: e is part of a cycle in G

If e is in a cycle in G then removing it removes at least one cycle. Furthermore, removing e does not disconnect the graph so the number of components remains the same. So G' has m edges, n vertices, and k components, which by the inductive hypothesis tells us it has at least $m + k - n$ cycles. But G has at least one more cycle than G' (since e is part of a cycle). So G has at least $m + k - n + 1$, or $(m + 1) + k - n$ cycles, as desired.

Case 2: e is not part of a cycle in G If e is not part of a cycle removing it disconnects the graph of G, so the number of components in G' is $k + 1$. So, G' contains m edges, n

vertices, and $k+1$ components, so by the inductive hypothesis it contains at least $m + (k+1) - n$ cycles. Now since e was not part of a cycle, G and G' have the same number of cycles. So G also has at least$m + (k+1) - n$ cycles. Rearranging we get that G has at least $(m+1) + k - n$ cycles as desired.

$\square$

$\blacksquare$

Problem 7. [10 points] For the following sum, find an upper and a lower bound that differ by at most 1.

$$\sum_{i=1}^{\infty} \frac{1}{\sqrt{i^3}}$$

Solution. To find the upper bound, we use the integral method, where $f(i) = \frac{1}{\sqrt{i^3}}$:

$$\begin{aligned}\sum_{i=1}^{\infty} \frac{1}{\sqrt{i^3}} &\leq f(1) + \int_1^{\infty} \frac{1}{\sqrt{i^3}} dx \\ &= 1 - 2\frac{1}{\sqrt{i}}\Big|_1^{\infty} \\ &= 1 - 2\,(0 - 1) = 3\end{aligned}$$

To find the lower bound, we use also use the intergral method:

$$\begin{aligned}\sum_{i=1}^{\infty} \frac{1}{\sqrt{i^3}} &\geq \lim_{x\to\infty} \frac{1}{\sqrt{i^3}} + \int_1^{\infty} \frac{1}{\sqrt{i^3}} dx \\ &= 0 + 2 = 2\end{aligned}$$

The two bounds differ by exactly 1. We conclude that

$$2 \leq \sum_{i=1}^{\infty} \frac{1}{\sqrt{i^3}} \leq 3.$$

■

Problem 8. [10 points] State whether each of the following claims is True or False and prove your answer.

(a) [2 pts] $x \ln x$ is $O(x)$

Solution. False. $\lim_{x\to\infty} x \ln x / x = \lim_{x\to\infty} \ln x = \infty > 0$ ■

(b) [2 pts] $x/100$ is $o(x)$

Solution. False. In this case we have $\dfrac{x/100}{x} = \dfrac{1}{100} \to 1/100 > 0$ as $x \to \infty$ ■

(c) [2 pts] x^{n+1} is $\Omega\left(x^n\right)$

Solution. True. Taking the quotient we arrive to $\dfrac{x^{n+1}}{x^n} = x \to \infty > 0$ ■

(d) [4 pts] $n!$ is $\Theta\left(n^n\right)$.

Solution. False. Stirling's formula asserts $n! \sim \dfrac{n^n\sqrt{2\pi n}}{e^n}$ so $\lim_{n\to\infty} n!/n^n = \lim_{n\to\infty} \dfrac{n^n\sqrt{2\pi n}}{n^n e^n} = \lim_{n\to\infty} \dfrac{\sqrt{2\pi n}}{e^n} = 0$. Hence $n!$ is not $\Theta\left(n^n\right)$. ■

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Problem Set 1

Problem 1. [24 points]

Translate the following sentences from English to predicate logic. The domain that you are working over is X, the set of people. You may use the functions $S(x)$, meaning that "x has been a student of 6.042," $A(x)$, meaning that "x has gotten an 'A' in 6.042," $T(x)$, meaning that "x is a TA of 6.042," and $E(x, y)$, meaning that "x and y are the same person."

(a) [6 pts] There are people who have taken 6.042 and have gotten A's in 6.042

(b) [6 pts] All people who are 6.042 TA's and have taken 6.042 got A's in 6.042

(c) [6 pts] There are no people who are 6.042 TA's who did not get A's in 6.042.

(d) [6 pts] There are at least three people who are TA's in 6.042 and have not taken 6.042

Problem 2. [24 points]

Use a truth table to prove or disprove the following statements:

(a) [12 pts]

$$\neg(P \vee (Q \wedge R)) = (\neg P) \wedge (\neg Q \vee \neg R)$$

(b) [12 pts]

$$\neg(P \wedge (Q \vee R)) = \neg P \vee (\neg Q \vee \neg R)$$

Problem 3. [24 points]

The binary logical connectives $\wedge$ (*and*), $\vee$ (*or*), and $\Rightarrow$ (*implies*) appear often in not only computer programs, but also everyday speech. In computer chip designs, however, it is considerably easier to construct these out of another operation, `nand`, which is simpler to represent in a circuit. Here is the truth table for `nand`:

P	Q	P `nand` Q
`true`	`true`	`false`
`true`	`false`	`true`
`false`	`true`	`true`
`false`	`false`	`true`

(a) [12 pts] For each of the following expressions, find an equivalent expression using only `nand` and $\neg$ (*not*), as well as grouping parentheses to specify the order in which the operations apply. You may use A, B, and the operators any number of times.

(i) $A \wedge B$

(ii) $A \vee B$

(iii) $A \Rightarrow B$

(b) [4 pts] It is actually possible to express each of the above using only `nand`, without needing to use $\neg$. Find an equivalent expression for $(\neg A)$ using only `nand` and grouping parentheses.

(c) [8 pts] The constants `true` and `false` themselves may be expressed using only `nand`. Construct an expression using an arbitrary statement A and `nand` that evaluates to `true` regardless of whether A is `true` or `false`. Construct a second expression that always evaluates to `false`. Do not use the constants `true` and `false` themselves in your statements.

Problem 4. [10 points] You have 12 coins and a balance scale, one of which is fake. All the real coins weigh the same, but the fake coin weighs less than the rest. All the coins visually appear the same, and the difference in weight is imperceptible to your senses. In at most 3 weighings, give a strategy that detects the fake coin. (Note: the scale in this problem is a scale with two dishes, which tips toward the side that is heavier. For clarification, do an image search for "balance scale").

Problem 5. [6 points] Prove the following statement by proving its contrapositive: if r is irrational, then $r^{1/5}$ is irrational. (Be sure to *state* the contrapositive explicitly.)

Problem 6. [12 points] Suppose that $w^2 + x^2 + y^2 = z^2$, where w, x, y, and z always denote positive integers. (Hint: It may be helpful to represent even integers as $2i$ and odd integers as $2j + 1$, where i and j are integers)

Prove the proposition: z is even if and only if w, x, and y are even. Do this by considering all the cases of w, x, y being odd or even.

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Problem Set 2

Problem 1. [12 points] Define a *3-chain* to be a (not necessarily contiguous) subsequence of three integers, which is either monotonically increasing or monotonically decreasing. We will show here that any sequence of five distinct integers will contain a *3-chain*. Write the sequence as a_1, a_2, a_3, a_4, a_5. Note that a monotonically increasing sequences is one in which each term is greater than or equal to the previous term. Similarly, a monotonically decreasing sequence is one in which each term is less than or equal to the previous term. Lastly, a subsequence is a sequence derived from the original sequence by deleting some elements without changing the location of the remaining elements.

(a) [4 pts] Assume that $a_1 < a_2$. Show that if there is no *3-chain* in our sequence, then a_3 must be less than a_1. (Hint: consider a_4!)

(b) [2 pts] Using the previous part, show that if $a_1 < a_2$ and there is no *3-chain* in our sequence, then $a_3 < a_4 < a_2$.

(c) [2 pts] Assuming that $a_1 < a_2$ and $a_3 < a_4 < a_2$, show that any value of a_5 must result in a *3-chain*.

(d) [4 pts] Using the previous parts, prove by contradiction that any sequence of five distinct integers must contain a *3-chain*.

Problem 2. [8 points]

Prove by either the Well Ordering Principle or induction that for all nonnegative integers, n:

$$\sum_{i=0}^{n} i^3 = \left(\frac{n(n+1)}{2}\right)^2. \tag{1}$$

Problem 3. [25 points] The following problem is fairly tough until you hear a certain one-word clue. The solution is elegant but is slightly tricky, so don't hesitate to ask for hints!

During 6.042, the students are sitting in an $n \times n$ grid. A sudden outbreak of beaver flu (a rare variant of bird flu that lasts forever; symptoms include yearning for problem sets and craving for ice cream study sessions) causes some students to get infected. Here is an example where $n = 6$ and infected students are marked $\times$.

×				×	
	×				
		×	×		
		×			
			×		×

Now the infection begins to spread every minute (in discrete time-steps). Two students are considered *adjacent* if they share an edge (i.e., front, back, left or right, but NOT diagonal); thus, each student is adjacent to 2, 3 or 4 others. A student is infected in the next time step if either

- the student was previously infected (since beaver flu lasts forever), or
- the student is adjacent to *at least two* already-infected students.

In the example, the infection spreads as shown below.

×				×	
	×				
		×	×		
		×			
			×		×

⇒

×	×			×	
×	×	×			
	×	×	×		
		×			
		×	×		
		×	×	×	×

⇒

×	×	×		×	
×	×	×	×		
×	×	×	×		
	×	×	×		
		×	×	×	
		×	×	×	×

In this example, over the next few time-steps, all the students in class become infected.

Theorem. *If fewer than n students in class are initially infected, the whole class will never be completely infected.*

Prove this theorem.

Hint: When one wants to understand how a system such as the above "evolves" over time, it is usually a good strategy to (1) identify an appropriate property of the system at the initial stage, and (2) prove, by induction on the number of time-steps, that the property is preserved at every time-step. So look for a property (of the set of infected students) that remains invariant as time proceeds.

If you are stuck, ask your recitation instructor for the one-word clue and even more hints!

Problem 4. [10 points] Find the flaw in the following *bogus* proof that $a^n = 1$ for all nonnegative integers n, whenever a is a nonzero real number.

Proof. The *bogus* proof is by induction on n, with hypothesis

$$P(n) ::= \forall k \leq n.\, a^k = 1,$$

where k is a nonnegative integer valued variable.

Base Case: $P(0)$ is equivalent to $a^0 = 1$, which is true by definition of a^0. (By convention, this holds even if $a = 0$.)

Inductive Step: By induction hypothesis, $a^k = 1$ for all $k \in \mathbb{N}$ such that $k \leq n$. But then

$$a^{n+1} = \frac{a^n \cdot a^n}{a^{n-1}} = \frac{1 \cdot 1}{1} = 1,$$

which implies that $P(n+1)$ holds. It follows by induction that $P(n)$ holds for all $n \in \mathbb{N}$, and in particular, $a^n = 1$ holds for all $n \in \mathbb{N}$. $\square$

Problem 5. [10 points] Let the sequence $G_0, G_1, G_2, \ldots$ be defined recursively as follows: $G_0 = 0$, $G_1 = 1$, and $G_n = 5G_{n-1} - 6G_{n-2}$, for every $n \in \mathbb{N}, n \geq 2$.

Prove that for all $n \in \mathbb{N}$, $G_n = 3^n - 2^n$.

Problem 6. [20 points]

In the 15-puzzle, there are 15 lettered tiles and a blank square arranged in a 4×4 grid. Any lettered tile adjacent to the blank square can be slid into the blank. For example, a sequence of two moves is illustrated below:

A	*B*	*C*	*D*
E	*F*	*G*	*H*
I	*J*	*K*	*L*
M	*O*	*N*	

$\rightarrow$

A	*B*	*C*	*D*
E	*F*	*G*	*H*
I	*J*	*K*	*L*
M	*O*		**N**

$\rightarrow$

A	*B*	*C*	*D*
E	*F*	*G*	*H*
I	*J*		*L*
M	*O*	**K**	*N*

In the leftmost configuration shown above, the O and N tiles are out of order. Using only legal moves, is it possible to swap the N and the O, while leaving all the other tiles in their original position and the blank in the bottom right corner? In this problem, you will prove the answer is "no".

Theorem. *No sequence of moves transforms the board below on the left into the board below on the right.*

A	*B*	*C*	*D*
E	*F*	*G*	*H*
I	*J*	*K*	*L*
M	**O**	**N**	

A	*B*	*C*	*D*
E	*F*	*G*	*H*
I	*J*	*K*	*L*
M	**N**	**O**	

(a) [2 pts] We define the "order" of the tiles in a board to be the sequence of tiles on the board reading from the top row to the bottom row and from left to right within a row. For example, in the right board depicted in the above theorem, the order of the tiles is A, B, C, D, E, etc.

Can a row move change the order of the tiles? Prove your answer.

(b) [2 pts] How many pairs of tiles will have their relative order changed by a column move? More formally, for how many pairs of letters L_1 and L_2 will L_1 appear earlier in the order of the tiles than L_2 before the column move and later in the order after the column move? Prove your answer correct.

(c) [2 pts] We define an *inversion* to be a pair of letters L_1 and L_2 for which L_1 precedes L_2 in the alphabet, but L_1 appears after L_2 in the order of the tiles. For example, consider the following configuration:

A	B	C	E
D	H	G	F
I	J	K	L
M	N	O	

There are exactly four inversions in the above configuration: E and D, H and G, H and F, and G and F.

What effect does a row move have on the parity of the number of inversions? Prove your answer.

(d) [4 pts] What effect does a column move have on the parity of the number of inversions? Prove your answer.

(e) [8 pts] The previous problem part implies that we must make an *odd* number of column moves in order to exchange just one pair of tiles (N and O, say). But this is problematic, because each column move also knocks the blank square up or down one row. So after an *odd* number of column moves, the blank can not possibly be back in the last row, where it belongs! Now we can bundle up all these observations and state an *invariant*, a property of the puzzle that never changes, no matter how you slide the tiles around.

Lemma. *In every configuration reachable from the position shown below, the parity of the number of inversions is different from the parity of the row containing the blank square.*

row 1	A	B	C	D
row 2	E	F	G	H
row 3	I	J	K	L
row 4	M	O	N	

Prove this lemma.

(f) [2 pts] Prove the theorem that we originally set out to prove.

Problem 7. [15 points] There are two types of creature on planet Char, Z-lings and B-lings. Furthermore, every creature belongs to a particular generation. The creatures in each generation reproduce according to certain rules and then die off. The subsequent generation consists entirely of their offspring.

The creatures of Char pair with a mate in order to reproduce. First, as many Z-B pairs as possible are formed. The remaining creatures form Z-Z pairs or B-B pairs, depending on whether there is an excess of Z-lings or of B-lings. If there are an odd number of creatures, then one in the majority species dies without reproducing. The number and type of offspring is determined by the types of the parents

- If both parents are Z-lings, then they have three Z-ling offspring.
- If both parents are B-lings, then they have two B-ling offspring and one Z-ling offspring.
- If there is one parent of each type, then they have one offspring of each type.

There are 200 Z-lings and 800 B-lings in the first generation. Use induction to prove that the number of Z-lings will always be at most twice the number of B-lings.

Hint: You may want to use a stronger hypothesis for the induction.

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Problem Set 3

Problem 1. [16 points] Warmup Exercises

For the following parts, a correct numerical answer will only earn credit if accompanied by it's derivation. Show your work.

(a) [4 pts] Use the Pulverizer to find integers s and t such that $135s + 59t = \gcd(135, 59)$.

(b) [4 pts] Use the previous part to find the inverse of 59 modulo 135 in the range $\{1, \ldots, 134\}$.

(c) [4 pts] Use Euler's theorem to find the inverse of 17 modulo 31 in the range $\{1, \ldots, 30\}$.

(d) [4 pts] Find the remainder of 34^{82248} divided by 83. (*Hint: Euler's theorem.*)

Problem 2. [16 points]

Prove the following statements, assuming all numbers are positive integers.

(a) [4 pts] If $a \mid b$, then $\forall c$, $a \mid bc$

(b) [4 pts] If $a \mid b$ and $a \mid c$, then $a \mid sb + tc$.

(c) [4 pts] $\forall c$, $a \mid b \Leftrightarrow ca \mid cb$

(d) [4 pts] $\gcd(ka, kb) = k \gcd(a, b)$

Problem 3. [20 points] In this problem, we will investigate numbers which are squares modulo a prime number p.

(a) [5 pts] An integer n is a square modulo p if there exists another integer x such that $n \equiv x^2 \pmod{p}$. Prove that $x^2 \equiv y^2 \pmod{p}$ if and only if $x \equiv y \pmod{p}$ or $x \equiv -y \pmod{p}$. (*Hint:* $x^2 - y^2 = (x + y)(x - y)$)

(b) [5 pts] There is a simple test we can perform to see if a number n is a square modulo p. It states that

Theorem 1 (Euler's Criterion)**.** *:*

1. *If n is a square modulo p then $n^{\frac{p-1}{2}} \equiv 1 \pmod{p}$.*
2. *If n is not a square modulo p then $n^{\frac{p-1}{2}} \equiv -1 \pmod{p}$.*

Prove the first part of Euler's Criterion. (*Hint: Use Fermat's theorem.*)

(c) [10 pts] Assume that $p \equiv 3 \pmod 4$ and $n \equiv x^2 \pmod p$. Given n and p, find one possible value of x. (*Hint: Write p as $p = 4k+3$ and use Euler's Criterion. You might have to multiply two sides of an equation by n at one point.*)

Problem 4. [10 points] Prove that for any prime, p, and integer, $k \geq 1$,

$$\phi(p^k) = p^k - p^{k-1},$$

where ϕ is Euler's function. (*Hint: Which numbers between 0 and $p^k - 1$ are divisible by p? How many are there?*)

Problem 5. [18 points] Here is a *very, very fun* game. We start with two distinct, positive integers written on a blackboard. Call them x and y. You and I now take turns. (I'll let you decide who goes first.) On each player's turn, he or she must write a new positive integer on the board that is a common divisor of two numbers that are already there. If a player can not play, then he or she loses.

For example, suppose that 12 and 15 are on the board initially. Your first play can be 3 or 1. Then I play 3 or 1, whichever one you did not play. Then you can not play, so you lose.

(a) [6 pts] Show that every number on the board at the end of the game is either x, y, or a positive divisor of $\gcd(x, y)$.

(b) [6 pts] Show that every positive divisor of $\gcd(x, y)$ is on the board at the end of the game.

(c) [6 pts] Describe a strategy that lets you win this game every time.

Problem 6. [20 points] In one of the previous problems, you calculated square roots of numbers modulo primes equivalent to 3 modulo 4. In this problem you will prove that there are an infinite number of such primes!

(a) [6 pts] As a warm-up, prove that there are an infinite number of prime numbers. (*Hint: Suppose that the set F of all prime numbers is finite, that is $F = \{p_1, p_2, \ldots, p_k\}$ and define $n = p_1 p_2 \ldots p_k + 1$.*)

(b) [2 pts] Prove that if p is an odd prime, then $p \equiv 1 \pmod 4$ or $p \equiv 3 \pmod 4$.

(c) [6 pts] Prove that if $n \equiv 3 \pmod 4$, then n has a prime factor $p \equiv 3 \pmod 4$.

(d) [8 pts] Let F be the set of all primes p such that $p \equiv 3 \pmod 4$. Prove by contradiction that F has an infinte number of primes.

(*Hint: Suppose that F is finite, that is $F = \{p_1, p_2, \ldots, p_k\}$ and define $n = 4p_1p_2 \ldots p_k - 1$. Prove that there exists a prime $p_i \in F$ such that $p_i | n$.*)

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Problem Set 4

Problem 1. [15 points] Let $G = (V, E)$ be a graph. A *matching* in G is a set $M \subset E$ such that no two edges in M are incident on a common vertex.

Let M_1, M_2 be two matchings of G. Consider the new graph $G' = (V, M_1 \cup M_2)$ (i.e. on the same vertex set, whose edges consist of all the edges that appear in either M_1 or M_2). Show that G' is bipartite.

Helpful definition: A *connected component* is a subgraph of a graph consisting of some vertex and every node and edge that is connected to that vertex.

Problem 2. [20 points] Let $G = (V, E)$ be a graph. Recall that the *degree* of a vertex $v \in V$, denoted d_v, is the number of vertices w such that there is an edge between v and w.

(a) [10 pts] Prove that

$$2|E| = \sum_{v \in V} d_v.$$

(b) [5 pts] At a 6.042 ice cream study session (where the ice cream is plentiful and it helps you study too) 111 students showed up. During the session, some students shook hands with each other (everybody being happy and content with the ice-cream and all). Turns out that the University of Chicago did another spectacular study here, and counted that each student shook hands with exactly 17 other students. Can you debunk this too?

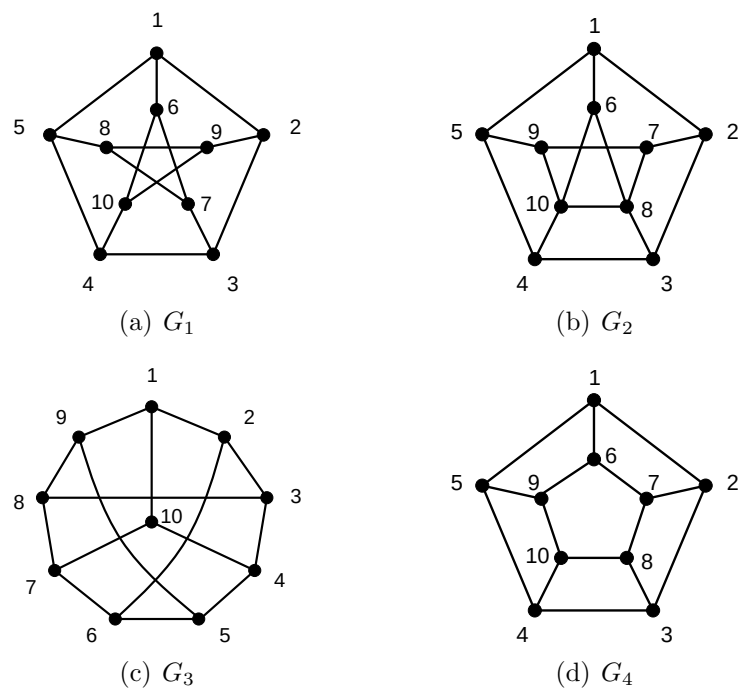
(c) [5 pts] And on a more dull note, how many edges does K_n, the complete graph on n vertices, have?

Problem 3. [15 points] Two graphs are isomorphic if they are the same up to a relabeling of their vertices (see Definition 5.1.3 in the book). A property of a graph is said to be *preserved under isomorphism* if whenever G has that property, every graph isomorphic to G also has that property. For example, the property of having five vertices is preserved under isomorphism: if G has five vertices then every graph isomorphic to G also has five vertices.

(a) [5 pts] Some properties of a simple graph, G, are described below. Which of these properties is *preserved under isomorphism*?

1. G has an even number of vertices.
2. None of the vertices of G is an even integer.
3. G has a vertex of degree 3.
4. G has exactly one vertex of degree 3.

(b) [10 pts] Determine which among the four graphs pictured in the Figures are isomorphic. If two of these graphs are isomorphic, describe an isomorphism between them. If they are not, give a property that is preserved under isomorphism such that one graph has the property, but the other does not. For at least one of the properties you choose, *prove* that it is indeed preserved under isomorphism (you only need prove one of them).



(a) G_1 (b) G_2 (c) G_3 (d) G_4

Figure 1: Which graphs are isomorphic?

Problem 4. [15 points] Recall that a ***coloring*** of a simple graph is an assignment of a color to each vertex such that no two adjacent vertices have the same color. A ***k-coloring*** is a coloring that uses at most k colors.

False Claim. *Let G be a (simple) graph with maximum degree at most k. If G also has a vertex of degree less than k, then G is k-colorable.*

(a) [5 pts] Give a counterexample to the False Claim when $k = 2$.

(b) [10 pts] Consider the following proof of the False Claim:

Proof. Proof by induction on the number n of vertices:

Induction hypothesis: $P(n)$ is defined to be: Let G be a graph with n vertices and maximum degree a most k. If G also has a vertex of degree less than k, then G is k-colorable.

Base case: (n=1) G has only one vertex and so is 1-colorable. So $P(1)$ holds.

Inductive step:

We may assume $P(n)$. To prove $P(n+1)$, let G_{n+1} be a graph with $n+1$ vertices and maximum degree at most k. Also, suppose G_{n+1} has a vertex, v, of degree less than k. We need only 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. Removing v reduces the degree of all vertices adjacent to v by 1. So in G_n, each of these vertices has degree less than k. Also the maximum degree of G_n remains at most k. So G_n satisfies the conditions of the induction hypothesis $P(n)$. 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}. □

Identify the exact sentence where the proof goes wrong.

Problem 5. [15 points] Prove or disprove the following claim: for some $n \geq 3$ (n boys and n girls, for a total of $2n$ people), there exists a set of boys' and girls' preferences such that every dating arrangement is stable.

Problem 6. [20 points]

Let $(s_1, s_2, ..., s_n)$ be an arbitrarily distributed sequence of the number $1, 2, ..., n-1, n$. For instance, for $n = 5$, one arbitrary sequence could be $(5, 3, 4, 2, 1)$.

Define the graph G=(V,E) as follows:

1. $V = \{v_1, v_2, ..., v_n\}$
2. $e = (v_i, v_j) \in E$ if either:
   (a) $j = i + 1$, for $1 \leq i \leq n-1$
   (b) $i = s_k$, and $j = s_{k+1}$ for $1 <= k <= n-1$

(a) [10 pts] Prove that this graph is 4-colorable for any $(s_1, s_2, ..., s_n)$.

Hint: First show that a line graph is 2-colorable. Note that a line graph is defined as follows: The n-node graph containing $n-1$ edges in sequence is known as the line graph L_n.

(b) [10 pts] Suppose $(s_1, s_2, ..., s_n) = (1, a_1, 3, a_2, 5, a_3, ...)$ where $a_1, a2....$ is an arbitrary distributed sequence of the even numbers in $1, ..., n-1$. Prove that the resulting graph is 2-colorable.

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Problem Set 5

Readings: Section 5.4 to 5.7 and 6.1-6.2.

Problem 1. [20 points] Recall that a tree is a connected acyclic graph. In particular, a single vertex is a tree. We define a *Splitting Binary Tree*, or *SBTree* for short, as either the lone vertex, or a tree with the following properties:

1. exactly one node of degree 2 (called the root).
2. every other node is of degree 3 or 1 (called internal nodes and leaves, respectively).

For the case of one single vertex (see above), that vertex is considered to be a leaf. It is easier to understand the definition visually, so an example is shown in Figure 1. An example of a tree which is not an SBTree is shown in Figure 2.

(a) [10 pts] Show if an SBTree has more than one vertex, then the induced subgraph obtained by removing the unique root consists of two disconnected SBTrees. You may assume that by removing the root you obtain two separate connected componenents, so all you need to prove is that those two components are SBTrees.

(b) [10 pts] Prove that two SBTrees with the same number of leaves must also have the same total number of nodes. *Hint: As a conjecture, guess an expression for the total number of nodes in terms of the number of leaves* $N(l)$. *Then use induction to prove that it holds for all trees with the same* l

Problem 2. [20 points]

In "Die Hard: The Afterlife", the ghosts of Bruce and Sam have been sent by the evil Simon on another mission to save midtown Manhattan. They have been told that there is a bomb on a street corner that lies in Midtown Manhattan, which Simon defines as extending from 41st Street to 59th Street and from 3rd Avenue to 9th Avenue. Additionally, the code that they need to defuse the bomb is on another street corner. Simon, in a good mood, also tosses them two carrots:

- He will have a helicopter initially lower them to the street corner where the bomb is.

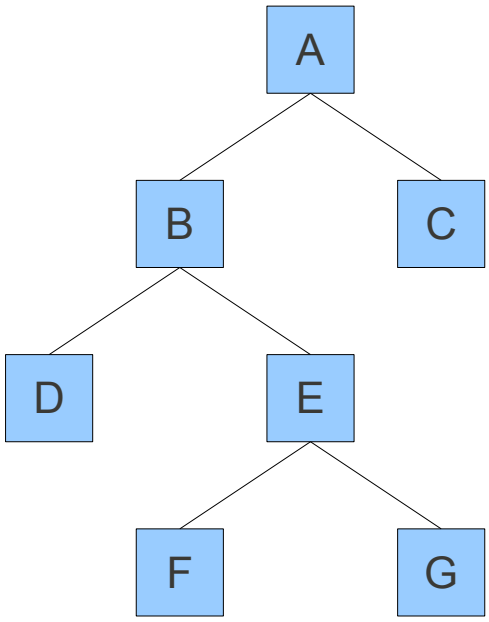


Figure 1: Splitting Binary Tree: Node A is the root, B and E are internal nodes, and C, D, F, and G are leaves. Notice how all internal nodes have degree 3.

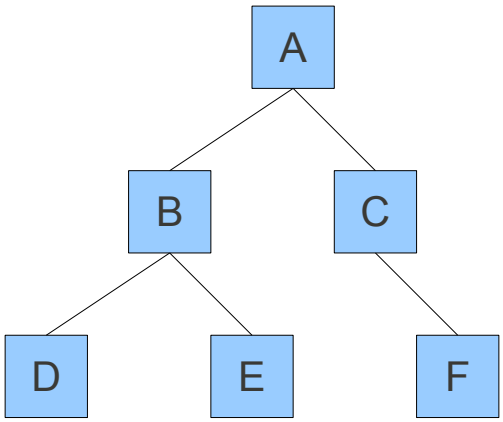


Figure 2: This is an example of a tree which is NOT a Splitting Binary Tree. Notice how both A and C have degree 2, when a BSTree can only have one such node.

- He promises that the code is placed only on a corner of a numbered street and a numbered avenue, so they don't have to search Broadway.

The map of midtown Manhattan is an example of an $N \times M$ (undirected) grid. In particular, midtown Manhattan is a 19×7 grid.

Bruce and Sam need to check all $19 \cdot 7 = 133$ street corners for the code. Once they are at a corner, they don't need any additional time to verify if the code is there. Once they find the code and return to the bomb, they can disarm it in 2 minutes (even, or especially, as the timer ticks down to 0). Also, they can run one block (in any of the four directions) in exactly 1 minute. They are given 135 minutes total in which to find the code and disarm the bomb, which means that they need to return to the bomb, code in hand, in 133 minutes.

Sam realizes that the map of NYC is actually a graph, and that they need to use a cool new 6.042 concept: A *Hamiltonian cycle* is a path that visits each vertex in a graph exactly once and ends at its starting point (so it is a cycle). A graph is *Hamiltonian* if it has a Hamiltonian cycle.

Hamiltonian graphs are really useful because you can visit each node and return to the starting point by taking only n steps, where n is the number of nodes – if a graph is not Hamiltonian, you would need at least $n + 1$ steps to visit each of the n nodes and return to the starting point.

In general, we don't know how to efficiently determine whether a general graph is Hamiltonian or not. However, Sam is very excited because he thinks that he can show that Midtown Manhattan is Hamiltonian. If it is, Bruce and Sam can save the day! Will they make it?

(a) [10 pts] Show that they cannot do it – that is, more generally, show that if both N and M are odd, then the $N \times M$ grid is *not* Hamiltonian. *Hint: First show that any $N \times M$ 2-dimensional undirected grid is bipartite.*

(b) [10 pts] Suppose Simon defined Midtown in the more standard way as extending from 40th Street to 59th Street and from 3rd Avenue to 9th Avenue (that is suppose Midtown Manhattan was a 20×7 grid), and gave them another 7 minutes,

1. Show that if either N is even and $M > 1$ or M is even and $N > 1$, then the $N \times M$ grid is Hamiltonian.
2. Explain why your proof breaks down when N and M are odd.
3. Would they survive? Does it depend on where the bomb is placed?

Problem 3. [20 points]

An n-node graph is said to be *tangled* if there is an edge leaving every set of $\lceil \frac{n}{3} \rceil$ or fewer vertices. As a special case, the graph consisting of a single node is considered tangled. (Recall that the notation $\lceil x \rceil$ refers to the smallest integer greater than or equal to x.)

(a) [7 pts] Find the error in the proof of the following claim.

Claim. Every non-empty, tangled graph is connected.

Proof. The proof is by strong induction on the number of vertices in the graph. Let $P(n)$ be the proposition that if an n-node graph is tangled, then it is connected. In the base case, $P(1)$ is true because the graph consisting of a single node is defined to be tangled and is trivially connected.

In the inductive step, for $n \geq 1$ assume $P(1), \ldots, P(n)$ to prove $P(n+1)$. That is, we want to prove that if an $(n+1)$-node graph is tangled, then it is connected. Let G be a tangled, $(n+1)$-node graph. Arbitrarily partition G into two pieces so that the first piece contains exactly $\lceil \frac{n}{3} \rceil$ vertices, and the second piece contains all remaining vertices. Note that since $n \geq 1$, the graph G has a least two vertices, and so both pieces contain at least one vertex. By induction, each of these two pieces is connected. Since the graph G is tangled, there is an edge leaving the first piece, joining it to the second piece. Therefore, the entire graph is connected. This shows that $P(1), \ldots, P(n)$ imply $P(n+1)$, and the claim is proved by strong induction. $\square$

(b) [5 pts] Draw a tangled graph that is not connected.

(c) [8 pts] An n-node graph is said to be *mangled* if there is an edge leaving every set of $\lceil \frac{n}{2} \rceil$ or fewer vertices. Again, as a special case, the graph consisting of a single node is considered mangled. Prove the following claim. *Hint: Prove by contradiction.*

Claim. Every non-empty, mangled graph is connected.

Problem 4. [15 points]

(a) [5 pts] Suppose that G is a simple, connected graph on n nodes. Show that G has exactly $n-1$ edges *iff* G is a tree.

(b) [10 pts] Prove by induction that any connected graph has a spanning tree.

Problem 5. [15 points] The adjacency matrix of a graph is given below (Section 5.1.6 in the book defines adjacency matrices)

$$\begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 1 & 0 \end{bmatrix}$$

(a) [4 pts] Draw the graph defined by this adjacency matrix. Label the vertices of your graph $1, 2, \ldots, 6$ so that vertex i corresponds to row and column i of the matrix.

(b) [4 pts] In a graph, we define the *distance* between to vertices to be the length of the shortest path between them. We define the *diameter* of a graph to be the largest distance between any two nodes. What is the diameter of this graph? Explain why.

(c) [3 pts] Find a cycle in this graph of maximum length and explain why it has maximum length.

(d) [4 pts] Give a coloring of the vertices that uses the minimum number of colors. Prove that this a minimum coloring.

Problem 6. [10 points] Let G be a graph. In this problem we show every vertex of odd degree is connected to at least one other vertex of odd degree in G.

(a) [6 pts] Let v be an odd degree node. Consider the longest walk starting at v that does not repeat any edges (though it may omit some). Let w be the final node of that walk. Show that $w \neq v$.

(b) [4 pts] Show that w must also have odd degree.

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Problem Set 6

Problem 1. [20 points] [15] For each of the following, either prove that it is an equivalence relation and state its equivalence classes, or give an example of why it is not an equivalence relation.

(a) [5 pts] $R_n := \{(x, y) \in \mathbb{Z} \times \mathbb{Z} \text{ s.t. } x \equiv y \pmod{n}\}$

(b) [5 pts] $R := \{(x, y) \in P \times P \text{ s.t. } x \text{ is taller than } y\}$ where P is the set of all people in the world today.

(c) [5 pts] $R := \{(x, y) \in \mathbb{Z} \times \mathbb{Z} \text{ s.t. } gcd(x, y) = 1\}$

(d) [5 pts] $R_G :=$ the set of $(x, y) \in V \times V$ such that V is the set of vertices of a graph G, and there is a path $x, v_1, \ldots, v_k, y$ from x to y along the edges of G.

Problem 2. [20 points] Every function has some subset of these properties:

injective surjective bijective

Determine the properties of the functions below, and briefly explain your reasoning.

(a) [5 pts] The function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x \sin(x)$.

(b) [5 pts] The function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 99x^{99}$.

(c) [5 pts] The function $f : \mathbb{R} \to \mathbb{R}$ defined by $\tan^{-1}(x)$.

(d) [5 pts] The function $f : \mathbb{N} \to \mathbb{N}$ defined by $f(x) =$ the number of numbers that divide x. For example, $f(6) = 4$ because $1, 2, 3, 6$ all divide 6. *Note: We define here the set* $\mathbb{N}$ *to be the set of all positive integers* $(1, 2, \ldots)$.

Problem 3. [20 points] In this problem we study partial orders (posets). Recall that a weak partial order $\preceq$ on a set X is reflexive ($x \preceq x$), anti-symmetric ($x \preceq y \wedge y \preceq x \to x = y$), and transitive ($x \preceq y \wedge y \preceq z \to x \preceq z$). Note that it may be the case that neither $x \preceq y$ nor $y \preceq x$. A chain is a list of *distinct* elements $x_1, \ldots, x_i$ in X for which $x_1 \preceq x_2 \preceq \cdots \preceq x_i$. An antichain is a subset S of X such that for all distinct $x, y \in S$, neither $x \preceq y$ nor $y \preceq x$.

The aim of this problem is to show that any sequence of $(n-1)(m-1)+1$ integers either contains a non-decreasing subsequence of length n or a decreasing subsequence of length m. Note that the given sequence may be out of order, so, for instance, it may have the form $1, 5, 3, 2, 4$ if $n = m = 3$. In this case the longest non-decreasing and longest decreasing subsequences have length 3 (for instance, consider $1, 2, 4$ and $5, 3, 2$).

(a) [7 pts] Label the given sequence of $(n-1)(m-1)+1$ integers $a_1, a_2, \ldots, a_{(n-1)(m-1)+1}$. Show the following relation $\preceq$ on $\{1, 2, 3, \ldots, (n-1)(m-1)+1\}$ is a weak poset: $i \preceq j$ if and only if $i \leq j$ and $a_i \leq a_j$ (as integers).

For the next part, we will need to use Dilworth's theorem, as covered in lecture. Recall that Dilworth's theorem states that if $(X, \preceq)$ is any poset whose longest chain has length n, then X can be partitioned into n disjoint antichains.

(b) [7 pts] Show that in any sequence of $(n-1)(m-1)+1$ integers, either there is a non-decreasing subsequence of length n or a decreasing subsequence of length m.

(c) [6 pts] Construct a sequence of $(n-1)(m-1)$ integers, for arbitrary n and m, that has no non-decreasing subsequence of length n and no decreasing subsequence of length m. Thus in general, the result you obtained in the previous part is best-possible.

Problem 4. [20 points] Louis Reasoner figures that, wonderful as the Beneš network may be, the butterfly network has a few advantages, namely: fewer switches, smaller diameter, and an easy way to route packets through it. So Louis designs an N-input/output network he modestly calls a *Reasoner-net* with the aim of combining the best features of both the butterfly and Beneš nets:

> The ith input switch in a Reasoner-net connects to two switches, a_i and b_i, and likewise, the jth output switch has two switches, y_j and z_j, connected to it. Then the Reasoner-net has an N-input Beneš network connected using the a_i switches as input switches and the y_j switches as its output switches. The Reasoner-net also has an N-input butterfly net connected using the b_i switches as inputs and¡ the z_j switches as outputs.

In the Reasoner-net the minimum latency routing does not have minimum congestion. The *latency for min-congestion* (LMC) of a net is the best bound on latency achievable using routings that minimize congestion. Likewise, the *congestion for min-latency* (CML) is the best bound on congestion achievable using routings that minimize latency.

Fill in the following chart for the Reasoner-net and briefly explain your answers.

diameter	**switch size(s)**	**# switches**	**congestion**	**LMC**	**CML**

Problem 5. [20 points] Let B_n denote the butterfly network with $N = 2^n$ inputs and N outputs, as defined in Notes 6.3.8. We will show that the congestion of B_n is exactly $\sqrt{N}$ when n is even.

Hints:

- For the butterfly network, there is a unique path from each input to each output, so the congestion is the maximum number of messages passing through a vertex for any matching of inputs to outputs.
- If v is a vertex at level i of the butterfly network, there is a path from exactly 2^i input vertices to v and a path from v to exactly 2^{n-i} output vertices.
- At which level of the butterfly network must the congestion be worst? What is the congestion at the node whose binary representation is all 0s at that level of the network?

(a) [10 pts] Show that the congestion of B_n is at most $\sqrt{N}$ when n is even.

(b) [10 pts] Show that the congestion achieves $\sqrt{N}$ somewhere in the network and conclude that the congestion of B_n is exactly $\sqrt{N}$ when n is even.

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Problem Set 7

Problem 1. [15 points] Express

$$\sum_{i=0}^{n} i^2 x^i$$

as a closed-form function of n.

Problem 2. [20 points]

(a) [5 pts] What is the product of the first n odd powers of two: $\prod_{k=1}^{n} 2^{2k-1}$?

(b) [5 pts] Find a closed expression for

$$\sum_{i=0}^{n} \sum_{j=0}^{m} 3^{i+j}$$

(c) [5 pts] Find a closed expression for

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

(d) [5 pts] Find a closed expression for

$$\prod_{i=1}^{n} \prod_{j=1}^{n} 2^i \cdot 3^j$$

Problem 3. [10 points]

(a) [6 pts] Use integration to find upper and lower bounds that differ by at most 0.1 for the following sum. (You may need to add the first few terms explicitly and then use integrals to bound the sum of the remaining terms.)

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

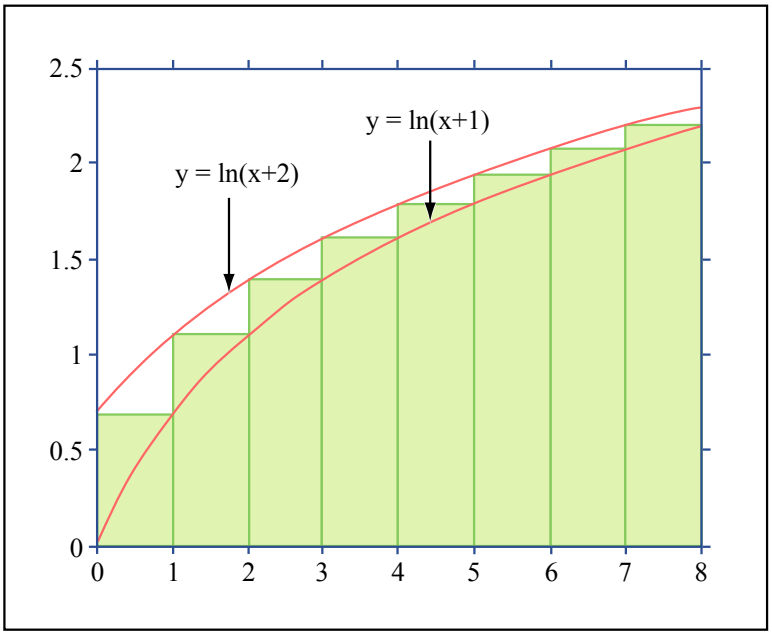


Image by MIT OpenCourseWare.

(b) [4 pts] Assume n is an integer larger than 1. Which of the following inequalities, if any, hold. You may find the graph helpful.

1. $\displaystyle\sum_{i=1}^{n} \ln(i+1) \leq \int_0^n \ln(x+2)dx$
2. $\displaystyle\sum_{i=1}^{n} \ln(i+1) \leq \ln 2 + \int_1^n \ln(x+1)dx$

Problem 4. [15 points] There is a bug on the edge of a 1-meter rug. The bug wants to cross to the other side of the rug. It crawls at 1 cm per second. However, at the end of each second, a malicious first-grader named Mildred Anderson *stretches* the rug by 1 meter. Assume that her action is instantaneous and the rug stretches uniformly. Thus, here's what happens in the first few seconds:

- The bug walks 1 cm in the first second, so 99 cm remain ahead.
- Mildred stretches the rug by 1 meter, which doubles its length. So now there are 2 cm behind the bug and 198 cm ahead.
- The bug walks another 1 cm in the next second, leaving 3 cm behind and 197 cm ahead.
- Then Mildred strikes, stretching the rug from 2 meters to 3 meters. So there are now $3 \cdot (3/2) = 4.5$ cm behind the bug and $197 \cdot (3/2) = 295.5$ cm ahead.
- The bug walks another 1 cm in the third second, and so on.

Your job is to determine this poor bug's fate.

(a) [5 pts] During second i, what *fraction* of the rug does the bug cross?

(b) [5 pts] Over the first n seconds, what fraction of the rug does the bug cross altogether? Express your answer in terms of the Harmonic number H_n.

(c) [5 pts] Approximately how many seconds does the bug need to cross the entire rug?

Problem 5. [20 points] For each of the following six pairs of functions f and g (parts (a) through (f)), state which of these order-of-growth relations hold (more than one may hold, or none may hold):

$$f = o(g) \qquad f = O(g) \qquad f = \omega(g) \qquad f = \Omega(g) \qquad f = \Theta(g) \qquad f \sim g$$

(a)	$f(n) = \log_2 n$	$g(n) = \log_{10} n$
(b)	$f(n) = 2^n$	$g(n) = 10^n$
(c)	$f(n) = 0$	$g(n) = 17$
(d)	$f(n) = 1 + \cos\left(\frac{\pi n}{2}\right)$	$g(n) = 1 + \sin\left(\frac{\pi n}{2}\right)$
(e)	$f(n) = 1.0000000001^n$	$g(n) = n^{10000000000}$

Problem 6. [15 points] This problem continues the study of the asymptotics of factorials.

(a) [5 pts]

Either prove or disprove each of the following statements.

- $n! = O((n+1)!)$
- $n! = \Omega((n+1)!)$
- $n! = \Theta((n+1)!)$
- $n! = \omega((n+1)!)$
- $n! = o((n+1)!)$

(b) [5 pts] Show that $n! = \omega\left(\left(\frac{n}{3}\right)^{n+e}\right)$.

(c) [5 pts] Show that $n! = \Omega(2^n)$

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Problem Set 8

Problem 1. [25 points] Find Θ bounds for the following divide-and-conquer recurrences. Assume $T(1) = 1$ in all cases. Show your work.

(a) [5 pts] $T(n) = 8T(\lfloor n/2 \rfloor) + n$

(b) [5 pts] $T(n) = 2T(\lfloor n/8 \rfloor + 1/n) + n$

(c) [5 pts] $T(n) = 7T(\lfloor n/20 \rfloor) + 2T(\lfloor n/8 \rfloor) + n$

(d) [5 pts] $T(n) = 2T(\lfloor n/4 \rfloor + 1) + n^{1/2}$

(e) [5 pts] $T(n) = 3T(\lfloor n/9 + n^{1/9} \rfloor) + 1$

Problem 2. [30 points] It is easy to misuse induction when working with asymptotic notation.

False Claim If

$$T(1) = 1 \text{ and}$$
$$T(n) = 4T(n/2) + n$$

Then T(n) = O(n).

False Proof We show this by induction. Let $P(n)$ be the proposition that $T(n) = O(n)$.

Base Case: $P(1)$ is true because $T(1) = 1 = O(1)$.

Inductive Case: For $n \geq 1$, assume that $P(n-1), \ldots, P(1)$ are true. We then have that

$$T(n) = 4T(n/2) + n = 4O(n/2) + n = O(n)$$

And we are done.

(a) [5 pts] Identify the flaw in the above proof.

(b) [10 pts] A simple attempt to prove $T(n) \neq O(n)$ via induction ultimately fails. We assume for sake of contradiction that $T(n) = O(n)$. Then there exists positive integer n_0 and positive real number c such that for all $n \geq n_0$, $T(n) \leq cn$. We then define $P(n)$ as the proposition that $T(n) \leq cn$.

We then proceed with strong induction.

Base Case, $n = n_0$: $P(n_0)$ is true, by assumption.

Inductive Step: We assume $P(n_0)$, $P(n_0 + 1)$, $\ldots$, $P(n - 1)$ true.

Fill in the rest of this proof attempt, and explain why it doesn't work.

Note: As this problem was updated so late, the graders will be instructed to be exceedingly lenient when grading this.

(c) [5 pts] Using Akra-Bazzi theorem, find the correct asymptotic behavior of this recurrence.

(d) [10 pts] We have now seen several recurrences of the form $T(n) = aT(\lfloor n/b \rfloor) + n$. Some of them give a runtime that is $O(n)$, and some don't. Find the relationship between a and b that yields $T(n) = O(n)$, and prove that this is sufficient.

Problem 3. [15 points] Define the sequence of numbers A_i by

$A_0 = 2$

$A_{n+1} = A_n/2 + 1/A_n$ (for $n \geq 1$)

Prove that $A_n \leq \sqrt{2} + 1/2^n$ for all $n \geq 0$.

Problem 4. [30 points] Find closed-form solutions to the following linear recurrences.

(a) [15 pts] $x_n = 4x_{n-1} - x_{n-2} - 6x_{n-3} \quad (x_0 = 3, x_1 = 4, x_2 = 14)$

(b) [15 pts] $x_n = -x_{n-1} + 2x_{n-2} + n \quad (x_0 = 5, x_1 = -4/9)$

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Problem Set 9

Problem 1. [10 points]

(a) [5 pts] Show that of any $n+1$ distinct numbers chosen from the set $\{1, 2, \ldots, 2n\}$, at least 2 must be relatively prime. (*Hint:* $\gcd(k, k+1) = 1$.)

(b) [5 pts] Show that any finite connected undirected graph with $n \geq 2$ vertices must have 2 vertices with the same degree.

Problem 2. [10 points] Under Siege!

Fearing retribution for the many long hours his students spent completing problem sets, Prof. Leighton decides to convert his office into a reinforced bunker. His only remaining task is to set the 10-digit numeric password on his door. Knowing the students are a clever bunch, he is not going to pick any passwords containing the forbidden consecutive sequences "18062", "6042" or "35876" (his MIT extension).

How many 10-digit passwords can he pick that don't contain forbidden sequences if each number $0, 1, \ldots, 9$ can only be chosen once (i.e. without replacement)?

Problem 3. [50 points] Be sure to show your work to receive full credit. In this problem we assume a standard card deck of 52 cards.

(a) [4 pts] How many 5-card hands have a single pair and no 3-of-a-kind or 4-of-a-kind?

(b) [4 pts] For fixed positive integers n and k, how many nonnegative integer solutions $x_0, x_1, \ldots, x_k$ are there to the following equation?

$$\sum_{i=0}^{k} x_i = n$$

(c) [4 pts] For fixed positive integers n and k, how many nonnegative integer solutions $x_0, x_1, \ldots, x_k$ are there to the following equation?

$$\sum_{i=0}^{k} x_i \leq n$$

(d) [4 pts] How many simple undirected graphs are there with n vertices?

(e) [4 pts] How many directed graphs are there with n vertices (self loops allowed)?

(f) [4 pts] How many tournament graphs are there with n vertices?

(g) [4 pts] How many acyclic tournament graphs are there with n vertices?

(h) [4 pts] How many numbers are there that are in the range [1..700] which are divisible by $2, 5$ or 7?

(i) [9 pts] In how many ways can you arrange n books on k bookshelf (assuming the order of books on a shelf matters?)

(j) [9 pts] How about if there has to be at least 1 book at each bookshelf?

Problem 4. [15 points] Give a combinatorial proof of the following theorem:

$$n2^{n-1} = \sum_{k=1}^{n} k \binom{n}{k}$$

(*Hint: Consider the set of all length-n sequences of 0's, 1's and a single *.*)

Problem 5. [15 points] At a congressional hearing, there are $2n$ members present. Exactly n of them are Democrats and n of them are Republicans. The members want to select a smaller subcommittee of size n from within those present at the hearing. However, since the Democrats currently hold majority, they want there to be more Democrats then Republicans in the committee. In how many ways can you select such a committee? (*Hint: Consider two cases: n odd and n even.*)

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Problem Set 10

Problem 1. [15 points] Suppose $\Pr\{\} : \mathcal{S} \to [0,1]$ is a *probability function* on a sample space, $\mathcal{S}$, and let B be an event such that $\Pr\{B\} > 0$. Define a function $\Pr_B\{\cdot\}$ on outcomes $w \in \mathcal{S}$ by the rule:

$$\Pr_B\{w\} = \begin{cases} \Pr\{w\} / \Pr\{B\} & \text{if } w \in B, \\ 0 & \text{if } w \notin B. \end{cases} \tag{1}$$

(a) [7 pts] Prove that $\Pr_B\{\cdot\}$ is also a probability function on $\mathcal{S}$ according to Definition 14.4.2.

(b) [8 pts] Prove that

$$\Pr_B\{A\} = \frac{\Pr\{A \cap B\}}{\Pr\{B\}}$$

for all $A \subseteq \mathcal{S}$.

Problem 2. [20 points]

(a) [10 pts] Here are some handy rules for reasoning about probabilities that all follow directly from the Disjoint Sum Rule. Use Venn Diagrams, or another method, to prove them.

$$\Pr\{A - B\} = \Pr\{A\} - \Pr\{A \cap B\} \qquad \text{(Difference Rule)}$$

$$\Pr\left\{\bar{A}\right\} = 1 - \Pr\{A\} \qquad \text{(Complement Rule)}$$

$$\Pr\{A \cup B\} = \Pr\{A\} + \Pr\{B\} - \Pr\{A \cap B\} \qquad \text{(Inclusion-Exclusion)}$$

$$\Pr\{A \cup B\} \leq \Pr\{A\} + \Pr\{B\}. \qquad \text{(2-event Union Bound)}$$

$$\text{If } A \subseteq B, \text{ then } \Pr\{A\} \leq \Pr\{B\}. \qquad \text{(Monotonicity)}$$

(b) [10 pts] Prove the following probabilistic identity, referred to as the **Union Bound**. You may assume the theorem that the probability of a union of *disjoint* sets is the sum of their probabilities.

Theorem. *Let $A_1, \ldots, A_n$ be a collection of events on some sample space. Then*

$$\Pr(A_1 \cup A_2 \cup \cdots \cup A_n) \leq \sum_{i=1}^{n} \Pr(A_i).$$

(*Hint: Induction*)

Problem 3. [15 points] Recall the strange dice from lecture:

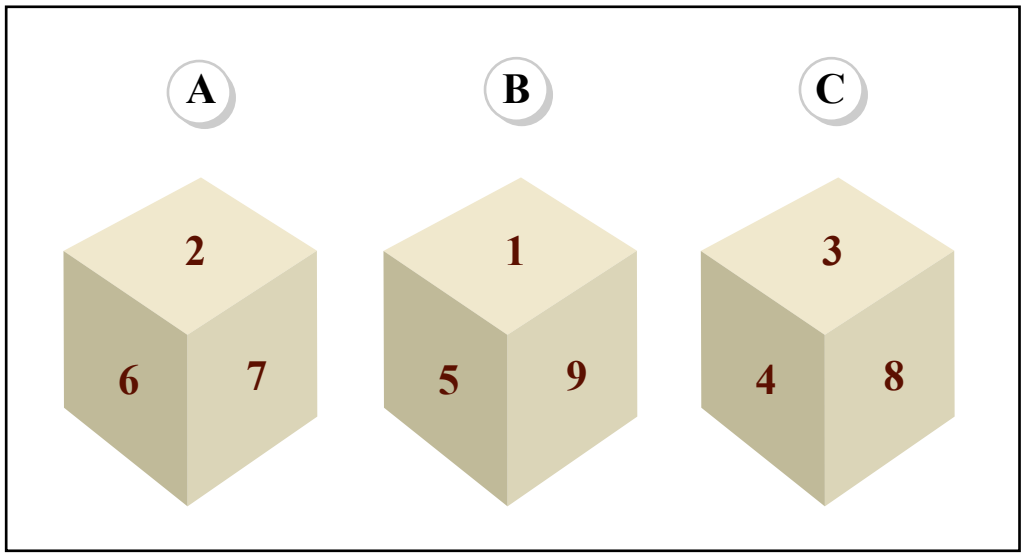


Image by MIT OpenCourseWare.

In the book we proved that if we roll each die once, then die A beats B more often, die B beats die C more often, and die C beats die A more often. Thus, contrary to our intuition, the "beats" relation $>$ is not transitive. That is, we have $A > B > C > A$.

We then looked at what happens if we roll each die twice, and add the result. In this problem we will show that the "beats" relation reverses in this game, that is, $A < B < C < A$, which is very counterintuitive!

(a) [5 pts] Show that rolling die C twice is more likely to win than rolling die B twice.

(b) [5 pts] Show that rolling die A twice is more likely to win that rolling die C twice.

(c) [5 pts] Show that rolling die B twice is more likely to win than rolling die A twice.

Problem 4. [14 points] Let's play a game! We repeatedly flip a fair coin. You have the sequence HHT, and I have the sequence HTT. If your sequence comes up first, then you win. If my sequence comes up first, then I win. For example, if the sequence of tosses is:

$$TTHTHT\underline{HHT}$$

then you win. What is the probability that you win? It may come as a surprise that the answer is very different from 1/2.

This problem is tricky, because the game could go on for an arbitrarily long time. Draw enough of the tree diagram to see a pattern, and then sum up the probabilities of the (infinitely many) outcomes in which you win.

It turns out that for any sequence of three flips, there is another sequence that is likely to come up before it. So there is no sequence of three flips which turns up earliest! ... and given any sequence of three flips, knowing how to pick another sequence that comes up sooner more than half the time gives you a nice chance to fool people gambling at a bar :-)

Problem 5. [15 points] We're interested in the probability that a randomly chosen poker hand (5 cards from a standard 52-card deck) contains cards from at most two suits.

(a) [7 pts] What is an appropriate sample space to use for this problem? What are the outcomes in the event, $\mathcal{E}$, we are interested in? What are the probabilities of the individual outcomes in this sample space?

(b) [8 pts] What is $\Pr(\mathcal{E})$?

Problem 6. [21 points]

I have a deck of 52 regular playing cards, 26 red, 26 black, randomly shuffled. They all lie face down in the deck so that you can't see them. I will draw a card off the top of the deck and turn it face up so that you can see it and then put it aside. I will continue to turn up cards like this but at some point while there are still cards left in the deck, you have to declare that you want the next card in the deck to be turned up. If that next card turns up black you win and otherwise you lose. Either way, the game is then over.

(a) [4 pts] Show that if you take the first card before you have seen any cards, you then have probability 1/2 of winning the game.

(b) [4 pts] Suppose you don't take the first card and it turns up red. Show that you have then have a probability of winning the game that is greater than 1/2.

(c) [4 pts] If there are r red cards left in the deck and b black cards, show that the probability of winning in you take the next card is $b/(r+b)$.

(d) [9 pts] Either,

1. come up with a strategy for this game that gives you a probability of winning strictly greater than 1/2 and prove that the strategy works, or,
2. come up with a proof that no such strategy can exist.

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Problem Set 11

Problem 1. [20 points] You are organizing a neighborhood census and instruct your census takers to knock on doors and note the sex of any child that answers the knock. Assume that there are two children in a household and that girls and boys are equally likely to be children and to open the door.

A sample space for this experiment has outcomes that are triples whose first element is either `B` or `G` for the sex of the elder child, likewise for the second element and the sex of the younger child, and whose third coordinate is `E` or `Y` indicating whether the *e*lder child or *y*ounger child opened the door. For example, $(\mathtt{B}, \mathtt{G}, \mathtt{Y})$ is the outcome that the elder child is a boy, the younger child is a girl, and the girl opened the door.

(a) [5 pts] Let T be the event that the household has two girls, and O be the event that a girl opened the door. List the outcomes in T and O.

(b) [5 pts] What is the probability $\Pr(T \mid O)$, that both children are girls, given that a girl opened the door?

(c) [10 pts] Where is the mistake in the following argument for computing $\Pr(T \mid O)$?

> If a girl opens the door, then we know that there is at least one girl in the household. The probability that there is at least one girl is
>
> $$1 - \Pr(\text{both children are boys}) = 1 - (1/2 \times 1/2) = 3/4.$$
>
> So,
>
> $$\begin{aligned} &\Pr(T \mid \text{there is at least one girl in the household}) \\ &= \frac{\Pr(T \cap \text{there is at least one girl in the household})}{\Pr\{\text{there is at least one girl in the household}\}} \\ &= \frac{\Pr(T)}{\Pr\{\text{there is at least one girl in the household}\}} \\ &= (1/4)/(3/4) = 1/3. \end{aligned}$$
>
> Therefore, given that a girl opened the door, the probability that there are two girls in the household is 1/3.

Problem 2. [20 points] Professor Plum, Mr. Green, and Miss Scarlet are all plotting to shoot Colonel Mustard. If one of these three has both an *opportunity* and the *revolver*, then that person shoots Colonel Mustard. Otherwise, Colonel Mustard escapes. Exactly one of the three has an *opportunity* with the following probabilities:

$$\begin{aligned}\Pr\{\text{Plum has opportunity}\} &= 1/6\\ \Pr\{\text{Green has opportunity}\} &= 2/6\\ \Pr\{\text{Scarlet has opportunity}\} &= 3/6\end{aligned}$$

Exactly one has the *revolver* with the following probabilities, regardless of who has an opportuntity:

$$\begin{aligned}\Pr\{\text{Plum has revolver}\} &= 4/8\\ \Pr\{\text{Green has revolver}\} &= 3/8\\ \Pr\{\text{Scarlet has revolver}\} &= 1/8\end{aligned}$$

(a) [5 pts] Draw a tree diagram for this problem. Indicate edge and outcome probabilities.

(b) [5 pts] What is the probability that Colonel Mustard is shot?

(c) [5 pts] What is the probability that Colonel Mustard is shot, given that Miss Scarlet does not have the revolver?

(d) [5 pts] What is the probability that Mr. Green had an opportunity, given that Colonel Mustard was shot?

Problem 3. [15 points] In lecture we discussed the Birthday Paradox. Namely, we found that in a group of m people with N possible birthdays, if $m \ll N$, then:

$$\Pr\{\text{all } m \text{ birthdays are different}\} \sim e^{-\frac{m(m-1)}{2N}}$$

To find the number of people, m, necessary for a half chance of a match, we set the probability to 1/2 to get:

$$m \sim \sqrt{(2\ln 2)N} \approx 1.18\sqrt{N}$$

For $N = 365$ days we found m to be 23.

We could also run a different experiment. As we put on the board the birthdays of the poeple surveyed, we could ask the class if anyone has the same birthday. In this case, before we reached a match amongst the surveyed people, we would already have found other people in the rest of the class who have the same birthday as someone already surveyed. Let's investigate why this is.

(a) [5 pts] Consider a group of m people with N possible birthdays amongst a larger class of k people, such that $m \leq k$. Define $\Pr\{A\}$ to be the probability that m people all have

different birthdays *and* none of the other $k-m$ people have the same birthday as one of the m.

Show that, if $m \ll N$, then $\Pr\{A\} \sim e^{\frac{m(m-2k)}{2N}}$. (Notice that the probability of no match is $e^{-\frac{m^2}{2N}}$ when k is m, and it gets smaller as k gets larger.)

Hints: For $m \ll N$: $\frac{N!}{(N-m)!N^m} \sim e^{-\frac{m^2}{2N}}$, and $(1-\frac{m}{N}) \sim e^{-\frac{m}{N}}$.

(b) [10 pts] Find the approximate number of people in the group, m, necessary for a half chance of a match (your answer will be in the form of a quadratic). Then simplify your answer to show that, as k gets large (such that $\sqrt{N} \ll k$), then $m \sim \frac{N \ln 2}{k}$.

Hint: For $x \ll 1$: $\sqrt{1-x} \sim (1-\frac{x}{2})$.

Problem 4. [10 points] We're covering probability in 6.042 lecture one day, and you volunteer for one of Professor Leighton's demonstrations. He shows you a coin and says he'll bet you \$1 that the coin will come up heads. Now, you've been to lecture before and therefore suspect the coin is biased, such that the probability of a flip coming up heads, $\Pr\{H\}$, is p for $1/2 < p \leq 1$.

You call him out on this, and Professor Leighton offers you a deal. He'll allow you to come up with an algorithm using the biased coin to *simulate* a fair coin, such that the probability you win and he loses, $\Pr\{W\}$, is equal to the probability that he wins and you lose, $\Pr\{L\}$. You come up with the following algorithm:

1. Flip the coin twice.
2. Based on the results:
   - $TH \Rightarrow$ you win $[W]$, and the game terminates.
   - $HT \Rightarrow$ Professor Leighton wins $[L]$, and the game terminates.
   - $(HH \vee TT) \Rightarrow$ discard the result and flip again.
3. If at the end of N rounds nobody has won, declare a tie.

As an example, for $N = 3$, an outcome of HT would mean the game ends early and you lose, $HHTH$ would mean the game ends early and you win, and $HHTTTT$ would mean you play the full N rounds and result in a tie.

(a) [5 pts] Assume the flips are mutually independent. Show that $\Pr\{W\} = \Pr\{L\}$.

(b) [5 pts] Show that, if $p < 1$, the probability of a tie goes to 0 as N goes to infinity.

Problem 5. [20 points]

(a) [5 pts] Suppose A and B are *disjoint* events. Prove that A and B are *not independent*, unless $\Pr(A)$ or $\Pr(B)$ is zero.

(b) [5 pts] If A and B are independent, prove that A and $\bar{B}$ are also independent.

Hint: $\Pr\left(A \cap \bar{B}\right) = \Pr(A) - \Pr(A \cap B)$.

(c) [5 pts] Give an example of events A,B, C such that A is independent of B, A is independent of C, but A is not independent of $B \cup C$.

(d) [5 pts] Prove that if C is independent of A, and C is independent of B, and C is independent of $A \cap B$, then C is independent of $A \cup B$.

Hint: Calculate $\Pr(A \cup B \mid C)$.

Problem 6. [15 points] Three very rare DNA markers were found in the DNA collected at a crime scene. Only one in every $1,000$ people has marker A, one in every $3,000$ people has marker B, and one in every $5,000$ people has marker C. Joe the plumber was arrested and accused of committing the crime, because he had all those markers present in his DNA. The prosecutor argues that the chances of any person having all three DNA markers is

$$\frac{1}{1000} \cdot \frac{1}{3000} \cdot \frac{1}{5000} = \frac{1}{15,000,000,000}$$

, which is more than 1 over the number of people in the world. This, plus the fact that Joe the plumber lives only 100 miles away from the crime scene must clearly mean that he is guilty. Having taken 6.042, you should be suspicious of this reasoning.

(a) [2 pts] What assumption has the prosecutor made (even though he hasn't realized it) about the presence of the 3 markers in human DNA?

(b) [4 pts] What would be the probability of a person having all three markers assuming that the markers appear pairwise independently? Under this assumption, can it be stated with such certainty that Joe the plumber commited the crime?

(c) [4 pts] What can you say about the probability of a person having all three markers if there is no independence between the markers?

(d) [5 pts] In fact, it turns out that neither of the above assumptions is correct. A researcher from MIT (who was actually in your recitation section for 6.042 back in the day) has discovered that while markers B and C appear independently, the probability of having marker B if you have marker A is $\frac{1}{2}$ and the probability of having marker C if you have marker A is $\frac{1}{3}$. The defence attorney now argues that the probability of a randomly selected person having all three markers is

$$Pr[A \cap B \cap C] = Pr[A] \cdot Pr[B|A] \cdot Pr[C|A] = \frac{1}{1000} \cdot \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6,000}.$$

Called as a witness, the MIT researcher points out that this is not necessarily true and that in fact he himself does not know what the probability is. What is wrong with the defence attorney's reasoning? (We assume that the MIT researcher published correct information and that, since he took 6.042, he knows what he is talking about.)

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Problem Set 12

Problem 1. [15 points]

In this problem, we will (hopefully) be making tons of money! Use your knowledge of probability and statistics to keep from going broke!

Suppose the stock market contains N types of stocks, which can be modelled by independent random variables. Suppose furthermore that the behavior of these stocks is modelled by a double-or-nothing coin flip. That is, stock S_i has half probability of doubling its value and half probability of going to 0. The stocks all cost a dollar, and you have N dollars. Say you only keep these stocks for one time-step (that is, at the end of this timestep, all stocks would have doubled in value or gone to 0).

(a) [3 pts] What is your expected amount of money if you spend all your money on one stock? Your variance?

(b) [3 pts] Suppose instead you diversified your purchases and bought N shares of all different stocks. What is your expected amount of money then? Your variance?

(c) [3 pts] The money that you have invested came from your financially conservative mother. As a result, your goals are much aligned with hers. Given this, which investment strategy should you take?

(d) [3 pts] Now instead say that you make money on rolls of dice. Specifically, you play a game where you roll a standard six-sided dice, and get paid an amount (in dollars) equal to the number that comes up. What is your expected payoff? What is the variance?

(e) [3 pts] We change the rules of the game so that your payoff is the cube of the number that comes up. In that case, what is your expected payoff? What is its variance?

Problem 2. [10 points] Here are seven propositions:

$$\begin{array}{rcrcr}
x_1 & \vee & x_3 & \vee & \neg x_7 \\
\neg x_5 & \vee & x_6 & \vee & x_7 \\
x_2 & \vee & \neg x_4 & \vee & x_6 \\
\neg x_4 & \vee & x_5 & \vee & \neg x_7 \\
x_3 & \vee & \neg x_5 & \vee & \neg x_8 \\
x_9 & \vee & \neg x_8 & \vee & x_2 \\
\neg x_3 & \vee & x_9 & \vee & x_4
\end{array}$$

Note that:

1. Each proposition is the OR of three terms of the form x_i or the form $\neg x_i$.
2. The variables in the three terms in each proposition are all different.

Suppose that we assign true/false values to the variables $x_1, \ldots, x_9$ independently and with equal probability.

(a) [5 pts] What is the expected number of true propositions?

(b) [5 pts] Use your answer to prove that there exists an assignment to the variables that makes *all* of the propositions true.

Problem 3. [20 points] MIT students sometimes delay laundry for a few days (to the chagrin of their roommates). Assume all random variables described below are mutually independent.

(a) [5 pts] A ***busy*** student must complete 3 problem sets before doing laundry. Each problem set requires 1 day with probability 2/3 and 2 days with probability 1/3. Let B be the number of days a busy student delays laundry. What is $\mathrm{E}[B]$?

Example: If the first problem set requires 1 day and the second and third problem sets each require 2 days, then the student delays for $B = 5$ days.

(b) [5 pts] A ***relaxed*** student rolls a fair, 6-sided die in the morning. If he rolls a 1, then he does his laundry immediately (with zero days of delay). Otherwise, he delays for one day and repeats the experiment the following morning. Let R be the number of days a relaxed student delays laundry. What is $\mathrm{E}[R]$?

Example: If the student rolls a 2 the first morning, a 5 the second morning, and a 1 the third morning, then he delays for $R = 2$ days.

(c) [5 pts] Before doing laundry, an ***unlucky*** student must recover from illness for a number of days equal to the product of the numbers rolled on two fair, 6-sided dice. Let U be the expected number of days an unlucky student delays laundry. What is $\mathrm{E}[U]$?

Example: If the rolls are 5 and 3, then the student delays for $U = 15$ days.

(d) [5 pts] A student is ***busy*** with probability 1/2, ***relaxed*** with probability 1/3, and ***unlucky*** with probability 1/6. Let D be the number of days the student delays laundry. What is $\mathrm{E}[D]$?

Problem 4. [10 points] We have two coins: one is a fair coin and the other is a coin that produces heads with probability 3/4. One of the two coins is picked, and this coin is tossed n times. Explain how to calculate the number of tosses to make us 95% confident which coin was chosen. You do not have to calculate the minimum value of n, though we'd be pleased if you did.

Problem 5. [13 points] Each 6.042 final exam (out of 100 points) will be graded according to a rigorous procedure:

- With probability $\frac{4}{7}$ the exam is graded by a *TA*; with probability $\frac{2}{7}$ it is graded by a *lecturer*; and with probability $\frac{1}{7}$, it is accidentally dropped behind the radiator and arbitrarily given a score of 84.
- *TAs* score an exam by scoring each problem individually and then taking the sum.
  - There are ten true/false questions worth 2 points each. For each, full credit is given with probability 3/4, and no credit is given with probability 1/4.
  - There are four questions worth 15 points each. For each, the score is determined by rolling two fair dice, summing the results, and adding 3.
  - The single 20 point question is awarded either 12 or 18 points with equal probability.
- *Lecturers* score an exam by rolling a fair die twice, multiplying the results, and then adding a "general impression" score.
  - With probability $\frac{4}{10}$, the general impression score is 40.
  - With probability $\frac{3}{10}$, the general impression score is 50.
  - With probability $\frac{3}{10}$, the general impression score is 60.

Assume all random choices during the grading process are independent.

(a) [5 pts] What is the expected score on an exam graded by a TA?

(b) [5 pts] What is the expected score on an exam graded by a lecturer?

(c) [3 pts] What is the expected score on a 6.042 final exam?

Problem 6. [32 points]

Suppose n balls are thrown randomly into n boxes, so each ball lands in each box with uniform probability. Also, suppose the outcome of each throw is independent of all the other throws.

(a) [5 pts] Let X_i be an indicator random variable whose value is 1 if box i is empty and 0 otherwise. Write a simple closed form expression for the probability distribution of X_i. Are $X_1, X_2, \ldots, X_n$ independent random variables?

(b) [5 pts] Find a constant, c, such that the expected number of empty boxes is asymptotically equal ($\sim$) to cn.

(c) [5 pts] Show that

$$\Pr\left(\text{at least } k \text{ balls fall in the first box}\right) \leq \binom{n}{k}\left(\frac{1}{n}\right)^k.$$

(d) [7 pts] Let R be the maximum of the numbers of balls that land in each of the boxes. Conclude from the previous parts that

$$\Pr\{R \geq k\} \leq \frac{n}{k!}.$$

(e) [10 pts] Conclude that

$$\lim_{n\to\infty} \Pr\{R \geq n^\epsilon\} = 0$$

for all $\epsilon > 0$.

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Problems for Recitation 1

1 Team Problem: A Mystery

A certain cabal within the 6.042 course staff is plotting to make the final exam *ridiculously hard*. ("Problem 1. Prove that the axioms of mathematics are complete and consistent. Express your answer in Mayan hieroglyphics.") The only way to stop their evil plan is to determine exactly who is in the cabal. The course staff consists of nine people:

$$\{\text{Oscar}, \text{Stav}, \text{Darren}, \text{Patrice}, \text{David}, \text{Nick}, \text{Martyna}, \text{Marten}, \text{Tom}\}$$

The cabal is a subset of these nine. A membership roster has been found and appears below, but it is deviously encrypted in logic notation. The predicate *incabal* indicates who is in the cabal; that is, $incabal(x)$ is true if and only if x is a member. Translate each statement below into English and deduce who is in the cabal.

(i) $\exists x\ \exists y\ \exists z\ (x \neq y \wedge x \neq z \wedge y \neq z \wedge incabal(x) \wedge incabal(y) \wedge incabal(z))$

(ii) $\neg(incabal(\text{Stav}) \wedge incabal(\text{David}))$

(iii) $(incabal(\text{Martyna}) \vee incabal(\text{Patrice})) \rightarrow \forall x\ incabal(x)$

(iv) $incabal(\text{Stav}) \rightarrow incabal(\text{David})$

(v) $incabal(\text{Darren}) \rightarrow incabal(\text{Martyna})$

(vi) $(incabal(\text{Oscar}) \vee incabal(\text{Nick})) \rightarrow \neg incabal(\text{Tom})$

(vii) $(incabal(\text{Oscar}) \vee incabal(\text{David})) \rightarrow \neg incabal(\text{Marten})$

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Notes for Recitation 1

1 Logic

How can one discuss mathematics with logical precision, when the English language is itself riddled with ambiguities? For example, imagine that you ask a friend what kind of dessert was offered at the party you couldn't make it to last week, and your friend says,

You could have cake or ice cream.

Does this mean that you could have *both* cake and ice cream? Or does it mean you had to choose either one or the other?

To cope with such ambiguities, mathematicians have defined precise meanings for some key words and phrases. Furthermore, they have devised symbols to represent those words. For example, if P is a proposition, then "not P" is a new proposition that is true whenever P is false and vice versa. The symbolic representation for "not P" is $\neg P$ or $\overline{P}$.

Two propositions, P and Q, can be joined by "and", "or", "implies", or "if and only if" to form a new proposition. The truth of this new proposition is determined by the truth of P and Q according to the table below. Symbolic equivalents are given in parentheses.

P	Q	"P and Q" ($P \wedge Q$)	"P or Q" ($P \vee Q$)	"P implies Q" or "if P, then Q" ($P \Rightarrow Q$)	"P if and only if Q" or "P iff Q" ($P \Leftrightarrow Q$)
F	F	F	F	T	T
F	T	F	T	T	F
T	F	F	T	F	F
T	T	T	T	T	T

There are a couple notable features hidden in this table:

- The phrase "P or Q" is true if P is true, Q is true, or both. Thus, you can have your cake and ice cream too.
- The phrase "P implies Q" (equivalently, "if P, then Q") is true when P *is false or* Q *is true*. Thus, "if the moon is made of green cheese, then there will be no final in 6.042" is a true statement.

There are two more important phrases in mathematical writing: "for all" (symbolized by $\forall$) and "there exists" (symbolized by $\exists$). These are called ***quantifiers***. A quantifier is always followed by a variable (and perhaps an indication of the range of that variable) and then a predicate, which typically involves that variable. Here are two examples:

$$\begin{aligned} \forall\ x \in \mathbb{R}^+ \quad & e^x < (1+x)^{1+x} \\ \exists\ n \in \mathbb{N} \quad & 2^n > (100n)^{100} \end{aligned}$$

The first statement says that e^x is less than $(1+x)^{1+x}$ for every positive real number x. The second statement says that there exists a natural number n such that $2^n > (100n)^{100}$.

The special symbols such as $\forall$, $\exists$, $\neg$, and $\vee$ are useful to logicians trying to express mathematical ideas without resorting to English at all. And other mathematicians often use these symbols as a shorthand. We recommend using them sparingly, however, because decrypting statements written in this symbolic language can be challenging!

2 Proving an Implication

Let's try to prove the following theorem.

Theorem 1. *Let $P(a,b)$ be any predicate defined for all $a \in \mathcal{A}$ and $b \in \mathcal{B}$. Then:*

$$\big(\exists\ a \in \mathcal{A} \quad \forall\ b \in \mathcal{B} \quad P(a,b)\big) \quad \Rightarrow \quad \big(\forall\ b \in \mathcal{B} \quad \exists\ a \in \mathcal{A} \quad P(a,b)\big)$$

Yuck! Now you *know* you're in a math class! Let's impose a specific interpretation in order to give concrete meaning to this claim. Define:

$$\begin{aligned} \mathcal{A} &= \{\text{6.042 students}\} \\ \mathcal{B} &= \{\text{6.042 lectures}\} \\ P(a,b) &= \text{"student } a \text{ falls asleep during lecture } b\text{"} \end{aligned}$$

Interpreting the left side in these terms gives:

$$\exists\ a \in \mathcal{A} \quad \forall\ b \in \mathcal{B} \quad P(a,b) = \text{"there exists a student that falls asleep in every lecture"}$$

So this side asserts that some particular student — let's call him Snoozer — always falls asleep. Now on the right side, we have:

$$\forall\ b \in \mathcal{B} \quad \exists\ a \in \mathcal{A} \quad P(a,b) = \text{"in every lecture, some student falls asleep"}$$

This is a slightly different assertion, because there might be a different sleeper in each lecture. Intuitively, the left side should imply the right; if Snoozer sleeps in every lecture, then in every lecture some student is surely asleep.

The implication in Theorem 1 is actually true for *every* predicate P and choice of sets $\mathcal{A}$ and $\mathcal{B}$. A universally-true statement, like this one, is called a ***validity***. (Every tautology (*cf.* Lecture Notes 9/4, p.6) is a validity, but validities may also involve quantifiers.) The ***converse*** of an implication $P \Rightarrow Q$ is the reverse implication $Q \Rightarrow P$. In this case, the converse is:

$$\big(\forall\, b \in \mathcal{B} \quad \exists\, a \in \mathcal{A} \quad P(a,b)\big) \quad \Rightarrow \quad \big(\exists\, a \in \mathcal{A} \quad \forall\, b \in \mathcal{B} \quad P(a,b)\big)$$

Under our interpretation, this says, "If in every lecture some student falls asleep, then there is some student who falls asleep in every lecture." This is not *necessarily* true, although it might be true for certain choices of predicate and sets. But since the truth of this converse proposition depends on the particular choice of predicate and sets, it is *not* a validity.

Anyway, let's prove the theorem.

Proof. We consider two cases.

Case 1: Suppose that the left side of the implication is false. Then the claim as a whole is true by default.

Case 2: Suppose that the left side of the implication is true. Then there exists some element $a_0 \in \mathcal{A}$ such that $P(a_0, b)$ is true for all $b \in \mathcal{B}$. Thus, for all $b \in \mathcal{B}$ there exists an $a \in \mathcal{A}$ (namely, a_0) such that $P(a,b)$ is true. Therefore, the right side of the implication is also true.

In both cases, the left side implies the right side, and so the theorem holds. □

Broadly speaking, we just proved that $P \Rightarrow Q$ for some nasty-looking propositions P and Q. When P was false (case 1), the implication held trivially. When P was true (case 2), we had to do some work to show that Q was also true. Every implication proof has this same structure: all the substance is in case 2. Thus, ordinarily no one even bothers to write down case 1 or even to identify two cases! Instead, when proving an implication, you may dispense with everything except for the body of case 2; the boxed text alone is considered a valid proof of the theorem. In summary, *in order to prove that P implies Q, you should assume that P is true and prove that Q is also true subject to that assumption.*

3 Team Problem: A Mystery

A certain cabal within the 6.042 course staff is plotting to make the final exam *ridiculously hard*. ("Problem 1. Prove that the axioms of mathematics are complete and consistent. Express your answer in Mayan hieroglyphics.") The only way to stop their evil plan is to determine exactly who is in the cabal. The course staff consists of nine people:

$$\{\text{Oscar}, \text{Stav}, \text{Darren}, \text{Patrice}, \text{David}, \text{Nick}, \text{Martyna}, \text{Marten}, \text{Tom}\}$$

The cabal is a subset of these nine. A membership roster has been found and appears below, but it is deviously encrypted in logic notation. The predicate *incabal* indicates who is in the cabal; that is, $incabal(x)$ is true if and only if x is a member. Translate each statement below into English and deduce who is in the cabal.

(i) $\exists x\ \exists y\ \exists z\ (x \neq y \wedge x \neq z \wedge y \neq z \wedge incabal(x) \wedge incabal(y) \wedge incabal(z))$

Solution. A direct English paraphrase would be "There exist people we'll call x, y, and z, who are all different, such that x, y and z are each in the cabal." A better version would use the fact that there's no need in this case to give names to the people. Namely, a better paraphrase is, "There are 3 different people in the cabal." Perhaps a simpler way to say this is, "The cabal is of size at least 3." ■

(ii) $\neg(incabal(\text{Stav}) \wedge incabal(\text{David}))$

Solution. Stav and David are not both in the cabal. Equivalently: at least one of Stav and David is not in the cabal. ■

(iii) $((incabal\text{Martyna}) \vee incabal(\text{Patrice})) \rightarrow \forall x\ incabal(x)$

Solution. If either Martyna or Patrice is in the cabal, then everyone is. ■

(iv) $incabal(\text{Stav}) \rightarrow incabal(\text{David})$

Solution. If Stav is in the cabal, then David is also. ■

(v) $incabal(\text{Darren}) \rightarrow incabal(\text{Martyna})$

Solution. If Darren is in the cabal, then Martyna is also. ■

(vi) $(incabal(\text{Oscar}) \vee incabal(\text{Nick})) \rightarrow \neg incabal(\text{Tom})$

Solution. If either of Oscar or Nick is in the cabal, then Tom is not. Equivalently, if Tom *is* in the cabal, then neither Oscar nor Nick is. ■

(vii) $(incabal(\text{Oscar}) \vee incabal(\text{David})) \rightarrow \neg incabal(\text{Marten})$

Solution. If either of Oscar or David is in the cabal, then Marten is not. Equivalently, if Marten *is* in the cabal, then neither Oscar nor David is. ■

So much for the translations. We now argue that the only cabal satisfying all seven propositions above is one whose members are exactly Oscar, David, and Nick.

We first observe that by (ii), there must be someone — either Stav or David — who is not in the cabal. But if either Martyna or Patrice were in the cabal, then by (iii), everyone would be. So we conclude by contradiction that

$$\text{Martyna and Patrice are not in the cabal.} \tag{1}$$

Now consider that (v) implies its contrapositive: if Martyna is not in the cabal, then neither is Darren. Therefore, since Martyna is not in the cabal,

$$\text{Darren is not in the cabal.} \tag{2}$$

Next observe that if Stav were in the cabal, then by (iv), David would be too, contradicting (ii). So by again contradiction, we conclude that

$$\text{Stav is not in the cabal.} \tag{3}$$

Now suppose Tom is in the cabal. Then by (vi), Oscar and Nick are not. We already know Martyna, Patrice, Darren, and Stav are not in the cabal, leaving only three who could be — Tom, Marten, and David. 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 2. *If Tom is in the cabal, then Marten and David are in the cabal.*

But by (vii), if David is the cabal, then Marten is not. That is,

Lemma 3. *David and Marten cannot both be in the cabal.*

Now from Lemma 3 we conclude that the conclusion of Lemma 2 is false. So by contrapositive, the hypothesis of Lemma 2 must also be false, namely,

$$\text{Tom is not in the cabal.} \tag{4}$$

Finally, suppose Marten is in the cabal. Then by (vii), Oscar and David are not, and we already know Martyna, Patrice, Darren, Stav, and Tom are not. So the cabal must consist of at most two people (Marten and Nick). This contradicts (i), and we conclude by contradiction that

$$\text{Marten is not in the cabal.} \tag{5}$$

So the only remaining people who could be in the cabal are Oscar, David, and Nick. Since the cabal must have at least three members, we conclude that

Lemma 4. *The only possible cabal consists of Oscar, David, and Nick.*

But we're not done yet: we haven't shown that a cabal consisting of Oscar, David, and Nick actually does satisfy all seven conditions. So let $\mathcal{A} = \{\text{Oscar}, \text{David}, \text{Nick}\}$, and let's quickly check that $\mathcal{A}$ satisfies (i)–(vii):

- $|A| = 3$, so A satisfies (i).
- Stav is not in A, so A satisfies (ii) and (iv).
- Neither Martyna nor Patrice is in A, so the hypothesis of (iii) is false, which means that A satisfies (iii).
- Darren is not in A, so A satisfies (v).
- Finally, Tom and Marten are not in A, so the conclusions of both (vi) and (vii) are true, and so A satisfies (vi) and (vii).

So now we have proved

Proposition. $\{$*Oscar*, *David*, *Nick*$\}$ *is the* unique *cabal satisfying conditions* (i)–(vii).

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Problems for Recitation 2

1 Problem: A Geometric Sum

Perhaps you encountered this classic formula in school:

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

First use the well ordering principle, and then 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"!

2 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?

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.)

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Notes for Recitation 2

1 Induction

Recall the principle of induction:

Principle of Induction. Let $P(n)$ be a predicate. If

- $P(0)$ is true, and
- for all $n \in \mathbb{N}$, $P(n)$ implies $P(n+1)$,

then $P(n)$ is true for all $n \in \mathbb{N}$.

We'll use induction to prove the following conjecture:

Conjecture. *For all positive integers, n:*

$$1 \cdot 2 + 2 \cdot 3 + 3 \cdot 4 + \ldots + n(n+1) = \frac{n(n+1)(n+2)}{3}$$

Remember that an induction proof has five parts, though the last one is often omitted:

1. Say that the proof is by induction.
2. Define the induction hypothesis, a predicate P defined on the natural numbers.
3. Handle the base case: prove that $P(0)$ is true.
4. Handle the inductive step: prove that $P(n)$ implies $P(n+1)$ for all integers $n \geq 0$.
5. Conclude that $P(n)$ is true for all $n \in \mathbb{N}$ by the principle of induction.

We noted in Lecture that while the base case is usually $n = 0$, it could be any nonnegative integer, k, in which case the conclusion would simply be that $P(n)$ holds for all $n \geq k$.

Proof. We use induction. Let $P(n)$ be the proposition that:

$$1 \cdot 2 + 2 \cdot 3 + 3 \cdot 4 + \ldots + n(n+1) = \frac{n(n+1)(n+2)}{3} \tag{1}$$

Base case $n = 1$*:* $P(1)$ is true, because the left-hand side of (1) is $1 \cdot 2 = 2$, and the right-hand side is $(1 \cdot 2 \cdot 3)/3 = 2$.

Inductive step: We must show that $P(n)$ implies $P(n+1)$ for all $n \geq 1$. So assume that $P(n)$ is true, where n denotes a positive integer. Then we can reason as follows:

$$\begin{aligned}
&1 \cdot 2 + 2 \cdot 3 + \cdots + (n+1)(n+2) \\
&= [1 \cdot 2 + 2 \cdot 3 + \cdots + n(n+1)] + (n+1)(n+2) \\
&= \frac{n(n+1)(n+2)}{3} + (n+1)(n+2) && \text{by ind. hypothesis (1)} \\
&= \frac{n(n+1)(n+2) + 3(n+1)(n+2)}{3} \\
&= \frac{(n+1)(n+2)(n+3)}{3}
\end{aligned}$$

This shows that $P(n+1)$ is true, and so $P(n)$ implies $P(n+1)$ for all $n \geq 1$.

By the induction principle, $P(n)$ is true for all $n \geq 1$, which proves the claim. $\square$

2 Problem: A Geometric Sum

Perhaps you encountered this classic formula in school:

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

First use the well ordering principle, and then 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 Well Ordering Principle

Proof. The proof is by contradiction and use of the Well Ordering Principle. Assume that the theorem is false. Then, some nonnegative integers serve as counterexamples to it. Lets collect these counterexamples in a set: $C ::= \{n \in \mathbb{N} | 1 + r + r^2 + r^3 + \ldots + r^n \neq \frac{1-r^{n+1}}{1-r}\}$

By our assumption that the theorem admits counterexamples, C is a nonempty set of nonnegative integers. So, by the Well Ordering Principle, C has a minimum element, call it c. That is, c is the smallest counterexample to the theorem.

Since c is the smallest counterexample, we know that formula is false for $n = c$ but true for all nonnegative integers $n < c$. But the formula is true for $n = 0$ since $1 = \frac{1-r^{0+1}}{1-r}$. Hence $c > 0$. This means $c - 1$ is a nonnegative integer, and since it is less than c, the formula is true for $c - 1$. That is,

$$1 + r + r^2 + r^3 + \ldots + r^{c-1} = \frac{1 - r^c}{1 - r} \tag{2}$$

But then, adding r^c to both sides of equation (2) gives us

$$\begin{aligned} 1 + r + r^2 + r^3 + \ldots + r^{c-1} + r^c &= \frac{1 - r^c}{1 - r} + r^c \\ &= \frac{1 - r^c + (1 - r)r^c}{1 - r} \\ &= \frac{1 - r^c + r^c - r^{c+1}}{1 - r} \\ &= \frac{1 - r^{c+1}}{1 - r} \end{aligned}$$

which means that the formula does hold for c, after all! This is a contradiction, and we are done.

□

Proof by Induction

Proof. We use induction. Let $P(n)$ be the proposition that the following equation holds for all $r \neq 1$:

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

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:

$$\begin{aligned} 1 + r + r^2 + r^3 + \ldots + r^n + r^{n+1} &= \frac{1 - r^{n+1}}{1 - r} + r^{n+1} \\ &= \frac{1 - r^{n+1} + (1 - r) \cdot r^{n+1}}{1 - r} \\ &= \frac{1 - r^{n+2}}{1 - r} \end{aligned}$$

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 $n \in \mathbb{N}$. □

Note: You may have encountered a different proof of this formula. We'll write down a sequence of equations and then explain the reasoning.

$$\begin{aligned} S &= 1 + r + r^2 + r^3 + \ldots + r^n \\ rS &= r + r^2 + r^3 + \ldots + r^{n+1} \\ S - rS &= 1 - r^{n+1} \\ S &= \frac{1 - r^{n+1}}{1 - r} \end{aligned}$$

We define S on the first line, multiply by r to get the second equation, subtract the second equation from the first to get the third, and then solve for S. This gives the formula above!

This argument is great! It is a derivation of the formula rather than just a verification. But, at some level, we've only hidden the use of induction, since the operations we're doing on n-term sums are justified using — you guessed it — induction. ■

3 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 pth 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 1. *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.

2. If $p = n$, then all purple-eyed people leave the island.

3. If $p < n$, then all purple-eyed people are already gone.

Base case: We must verify that the three parts of $P(n)$ hold on day $n = 1$.

1. Suppose $p > 1$. Consider events on day 1 from the perspective of a purple-eyed islander. The producer says that someone has a purple eye, and she can indeed see at least one other person with a purple eye. Therefore, the facts available to her are consistent with her having either a purple or red eye. So she survives the day.

2. Suppose $p = 1$. The single purple-eyed islander sees no one else with a purple eye, concludes that he must have a purple eye, and leaves the island at the end of the show. No one else leaves because everyone else does see the purple-eyed islander, and they have no reason at this point to think they too are purple-eyed.

3. This statement is vacuously true, because the if-part ($p < 1$) is false; the problem statement says that $p \geq 1$.

Therefore, $P(1)$ is true.

Inductive step: Now suppose that $P(n)$ is true where $n \geq 0$. We must verify the three parts of $P(n + 1)$.

1. Suppose $p > n + 1$. Then $p > n$ so all the purple-eyed contestants survived the preceding day by part 1 of $P(n)$. Furthermore, each purple-eyed islander can see at least $n + 1 > n$ other purple-eyed people, so the observation that everyone survives is consistent with she herself having either a purple or a red eye by $P(n)$ as well. Thus, each purple-eyed islander survives the day.

2. Suppose $p = n + 1$. Then $p > n$, so all the purple-eyed contestants survived the preceding day by part 1 of $P(n)$. Thus, on day $n + 1$ each purple-eyed islander knows $p > n$, but sees only n other people with a purple eye. Thus, each purple-eyed islander realizes that she has a purple eye and leaves the island at the end of the show.

3. Suppose $p < n + 1$. Then either $p = n$ (in which case all the purple-eyed contestants left the island on day n by part 2 of $P(n)$) or else $p < n$ (in which case all the purple-eyed contestants were already gone on day n by part 3 of $P(n)$). In either case, all the purple-eyed people are already out of luck.

Therefore $P(n)$ implies $P(n + 1)$ for all $n \geq 0$.

By the principle of induction, $P(n)$ is true for all $n \geq 0$, and the theorem follows. $\square$

$\blacksquare$

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

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 2×2. The first split makes two pieces, both 2×1. 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.

2 Problem: The Temple of Forever

Each monk entering the Temple of Forever is given a bowl with 15 red beads and 12 green beads. Each time the Gong of Time rings, a monk must do one of two things:

1. *Exchange*: If he has at least 3 red beads in his bowl, then he may exchange 3 red beads for 2 green beads.
2. *Swap*: He may replace each green bead in his bowl with a red bead and replace each red bead in his bowl with a green bead. That is, if he starts with i red beads and j green beads, then after he performs this operation, he will have j red beads and i green beads.

A monk may leave the Temple of Forever only when he has exactly 5 red beads and 5 green beads in his bowl.

Let's look at how we can represent this problem as a state machine.

- What do the states of the machine look like?
- Use the notation you developed above to represent the allowable transitions in the state machine.
- Expand the state machine diagram to the first three or four levels. Label the transitions according to the operation type (E for *exchange* or S for *swap*).

Now we'll show that no monk can ever escape the Temple of Forever because the state $(5, 5)$ violates an invariant of the Temple of Forever machine.

Theorem 1. *No one ever leaves the Temple of Forever.*

Prove this theorem by induction. Begin by searching for an invariant that holds initially and is maintained by each operation, but would be violated by the condition required for departure.

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.)

Inside the Temple of Forever, the Gong of Time rings on. As you may well imagine, the monks begin to recognize that no matter how many ways they try to exchange or swap their beads, they always seem to end up in some state they've already been in before! For one or two monks, this realization is all they need to propel them instantly into a state of

enlightenment. For the overwhelming majority, however, this knowledge does nothing but weaken their resolve. They just get depressed. Taking note of the mental state of this second group, the Keeper of the Temple makes an unannounced appearance and proclaims to the group, "From now on, any monk who is able to visit 108 (108 being the mystical number that encompasses all of existence[1]) unique states will be allowed to leave the Temple of Forever."

Do the monks have any chance of leaving the Temple of Forever?

Theorem 3. *It is not possible to visit 108 unique states in the Temple of Forever machine.*

Prove this theorem. (Hint: Consider a proof by contradiction.)

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?

[1]See http://astrologyforthesoul.com/vp/mysticalnumber108.html. Also consider: $42 + 24 + 42 = 108$.

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Notes for Recitation 3

1 State Machines

Recall from Lecture 3 (9/16) that an *invariant* is a property of a system (in lecture, that system was the 8-puzzle) that does not change, regardless of the system's behavior. Now we'll take a look at how we can make use of an important modeling tool — a state machine — to analyze systems. We'll see that finding invariants of state machines can often help us prove propositions.

A *state machine* is an abstract model of a step-by-step process. The model consists of a collection of *states* and *transitions* between those states. More formally, we can define a state machine as

- a set of states: Q
- a designated start state: $q_0 \in Q$
- a set of allowed transitions between states: $\delta \subseteq Q \times Q$.

1.1 Example: Bounded Counter

Take, for example, a bounded counter that counts from 0 to 99 and overflows at 100. A diagram representing the corresponding state machine was shown in recitation. Note that in this example, there is a finite number of states. However, in general, the set of states might be infinite.

1.2 Example: Unbounded Counter

An unbounded counter is similar to a bounded counter except that there is an infinite number of states and no overflow. See Figure 2.

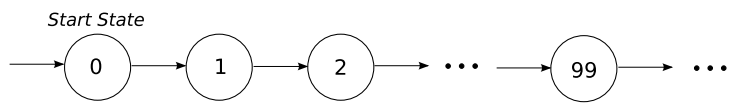


Figure 1: State machine for the unbounded counter

2 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 2×2. The first split makes two pieces, both 2×1. 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.

To do this, first break the chocolate bar of size $k + 1$ into two smaller pieces of size p and q where $p + q = k + 1$. This is certainly possible because the size of the bar is at least two. Now the pieces of sizes p and q are between one and k, so by strong induction, breaking these two pieces into single squares requires only $p - 1$ and $q - 1$ splits, respectively. The total number of splits required to break the bar of size $k + 1$ into single squares is therefore at most $1 + (p - 1) + (q - 1) = p + q - 1 = (k + 1) - 1 = k$.

This shows that $P(k)$ implies $P(k + 1)$, and the claim is proved by strong induction. $\square$

$\blacksquare$

3 Problem: The Temple of Forever

Each monk entering the Temple of Forever is given a bowl with 15 red beads and 12 green beads. Each time the Gong of Time rings, a monk must do one of two things:

1. *Exchange*: If he has at least 3 red beads in his bowl, then he may exchange 3 red beads for 2 green beads.
2. *Swap*: He may replace each green bead in his bowl with a red bead and replace each red bead in his bowl with a green bead. That is, if he starts with i red beads and j green beads, then after he performs this operation, he will have j red beads and i green beads.

A monk may leave the Temple of Forever only when he has exactly 5 red beads and 5 green beads in his bowl.

Let's look at how we can represent this problem as a state machine.

- What do the states of the machine look like?

  Solution. We can use variables such as r to represent the number of red beads and g to represent the number of green beads. Then the states can be represented as pairs (r, g) for $r \geq 0, g \geq 0$. ■

- Use the notation you developed above to represent the allowable transitions in the state machine.

  Solution. There are two possible transitions:

  1. trade 3 red for 2 green (*exchange*): $(r, g) \to (r - 3, g + 2), r \geq 3$
  2. swap red and green (*swap*): $(r, g) \to (g, r)$

  ■

- Expand the state machine diagram to the first three or four levels. Label the transitions according to the operation type (E for *exchange* or S for *swap*).

  Solution. You can see the diagram in Figure 3.

  ■

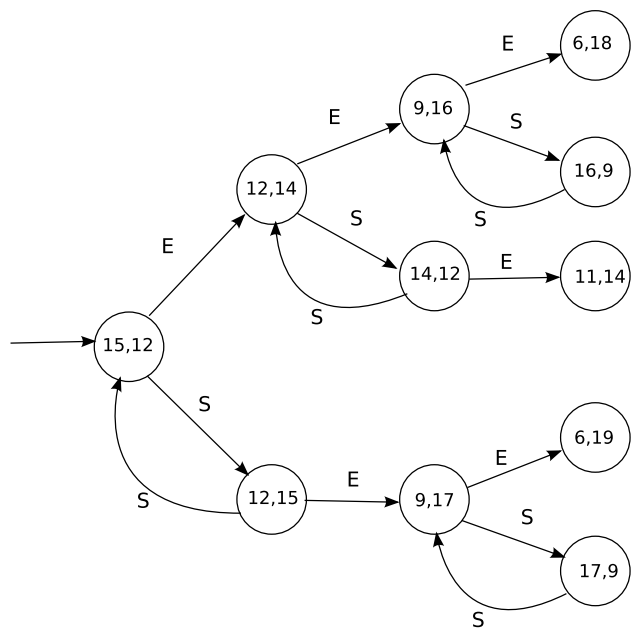


Figure 2: State machine diagram for first few levels

3.1 Invariants and State Machines

The Temple of Forever machine models every possible way of exchanging beads at each time step according to the rules. We would like to know whether the machine ever reaches the state $(5, 5)$. By starting at the start state and following the transition arrows through the state diagram, we can trace the possible paths, or *executions*, of the machine. So another way of stating our question is whether the state we are interested in appears in some execution of the machine. If so, then we can say that that state is *reachable*.

Definition 1. *A state is called* reachable *if there is a path to it starting from the start state, that is, if it appears in some execution.*

A useful technique for analyzing reachability is the identification of *invariants* of the machine. An invariant will hold for all reachable states of the machine. More formally,

Definition 2. *An* invariant *for a state machine is a predicate P on state machines such that $P(q_0)$ holds (where q_0 is the start state of the machine), and whenever $P(q)$ is true of a state q, and $q \to r$ for some state r, then $P(r)$ holds.*

Since, by definition, an invariant holds for the start state and for all transitions of the state machine, then we know (by induction) that the invariant must hold for all reachable states. Therefore, if we know some property to be an invariant, and we know that a certain state violates that property, then we can say that this state is unreachable.

Now we'll show that no monk can ever escape the Temple of Forever because the state $(5, 5)$ violates an invariant of the Temple of Forever machine.

Theorem 1. *No one ever leaves the Temple of Forever.*

Prove this theorem by induction. Begin by searching for an invariant that holds initially and is maintained by each operation, but would be violated by the condition required for departure.

Solution. *Proof.* We use induction on the number of gong rings. Let $P(n)$ be the proposition that after n rings, the number of red beads in the monk's bowl minus the number of green beads is equal to $5k + 2$ or $5k + 3$ for some integer k.

Base case: $P(0)$ is true because initially (after zero rings) the number of red beads minus the number of green beads is $15 - 12 = 5 \cdot 0 + 3$.

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. □

■

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:

1. If $r \geq 3$, then the monk may have exchanged 3 red beads for 2 green beads. Thus, the number of red beads plus the number of green becomes:

$$(r - 3) + (g + 2) = (r + g) - 1$$

   Since, by the inductive hypothesis, $r + g \leq 27$, it follows that $(r + g) - 1 \leq 27$, and 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 plus the number of greens becomes $g+r = r+g$. Thus, $P(n + 1)$ is again true by the inductive hypothesis.

□

We now prove the theorem by showing that there is a finite upper bound on the number of reachable states.

Proof. We give a direct argument. Lemma 3 tells us $r + g \leq 27$. This implies that there is an upper bound on the number of reachable states, since there can be at most 28 ways for r and g to sum to 27, 27 ways to sum to 26, and so on. Therefore, there can be at most $28 + 27 + \ldots + 1 = \frac{29 \cdot 28}{2} = 406$ states in the Temple of Forever machine. Hence, there is a finite number of reachable states in the Temple of Forever machine. □

■

Inside the Temple of Forever, the Gong of Time rings on. As you may well imagine, the monks begin to recognize that no matter how many ways they try to exchange or swap their beads, they always seem to end up in some state they've already been in before! For one or two monks, this realization is all they need to propel them instantly into a state of enlightenment. For the overwhelming majority, however, this knowledge does nothing but weaken their resolve. They just get depressed. Taking note of the mental state of this second group, the Keeper of the Temple makes an unannounced appearance and proclaims to the group, "From now on, any monk who is able to visit 108 (108 being the mystical number that encompasses all of existence[1]) unique states will be allowed to leave the Temple of Forever."

Do the monks have any chance of leaving the Temple of Forever?

Theorem 4. *It is not possible to visit 108 unique states in the Temple of Forever machine.*

Prove this theorem. (Hint: Consider a proof by contradiction.)

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

[1] See http://astrologyforthesoul.com/vp/mysticalnumber108.html. Also consider: $42 + 24 + 42 = 108$.

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.

■

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Problems 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.

(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?

2 Problem: The Fibonacci numbers.

The Fibonacci numbers are defined as follows:

$$F_0 = 0 \qquad F_1 = 1 \qquad F_n = F_{n-1} + F_{n-2} \quad (\text{for } n \geq 2).$$

Give an inductive proof that the Fibonacci numbers F_n and F_{n+1} are relatively prime for all $n \geq 0$.

3 Extra Problem: The power of 3.[1]

Let N be a number whose decimal expansion consists of 3^n identical digits. Show by induction that $3^n \mid N$. For example:

$$3^2 \mid \underbrace{777777777}_{3^2 = 9 \text{ digits}}$$

Recall that 3 divides a number iff it divides the sum of its digits.

[1] Try this if you have time!

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Notes for Recitation 4

1 The Pulverizer

We saw in lecture that the greatest common divisor (GCD) of two numbers can be written as a linear combination of them.[1] That is, no matter which pair of integers a and b we are given, there is always a pair of integer coefficients s and t such that

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

However, the proof was *nonconstructive*: it didn't suggest a way for finding such s and t.

That job is tackled by a mathematical tool that dates to sixth-century India, where it was called *kuttak*, which means "The Pulverizer". Today, the Pulverizer is more commonly known as "the extended Euclidean GCD algorithm", but that's lame. We're sticking with "Pulverizer".

Euclid's algorithm for finding the GCD of two numbers relies on repeated application of the equation:

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

which was proved in lecture (see the notes "Number Theory I"). For example, we can compute the GCD of 259 and 70 as follows:

$$\begin{aligned}
\gcd(259, 70) &= \gcd(70, 49) && \text{since } \operatorname{rem}(259, 70) = 49 \\
&= \gcd(49, 21) && \text{since } \operatorname{rem}(70, 49) = 21 \\
&= \gcd(21, 7) && \text{since } \operatorname{rem}(49, 21) = 7 \\
&= \gcd(7, 0) && \text{since } \operatorname{rem}(21, 7) = 0 \\
&= 7.
\end{aligned}$$

The Pulverizer goes through the same steps, but requires some extra bookkeeping along the way: as we compute $\gcd(a, b)$, we keep track of how to write each of the remainders (49, 21, and 7, in the example) as a linear combination of a and b (this is worthwhile, because our objective is to write the last nonzero remainder, which is the GCD, as such a linear

[1] In fact, we proved that among all positive linear combinations of the numbers their GCD is the smallest.

combination). For our example, here is this extra bookkeeping:

x	y	$(\operatorname{rem}(x,y))$	$=$	$x - q \cdot y$
259	70	49	$=$	$259 - 3 \cdot 70$
70	49	21	$=$	$70 - 1 \cdot 49$
			$=$	$70 - 1 \cdot (259 - 3 \cdot 70)$
			$=$	$-1 \cdot 259 + 4 \cdot 70$
49	21	7	$=$	$49 - 2 \cdot 21$
			$=$	$(259 - 3 \cdot 70) - 2 \cdot (-1 \cdot 259 + 4 \cdot 70)$
			$=$	$\boxed{3 \cdot 259 - 11 \cdot 70}$
21	7	0		

We began by initializing two variables, $x = a$ and $y = b$. In the first two columns above, we carried out Euclid's algorithm. At each step, we computed $\operatorname{rem}(x, y)$, which can be written in the form $x - q \cdot y$. (Remember that the Division Algorithm says $x = q \cdot y + r$, where r is the remainder. We get $r = x - q \cdot y$ by rearranging terms.) Then we replaced x and y in this equation with equivalent linear combinations of a and b, which we already had computed. After simplifying, we were left with a linear combination of a and b that was equal to the remainder as desired. The final solution is boxed.

2 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?

Solution. We go through the steps as described in Section 1 (see the table below) to get that $1 = 8 \cdot 50 - 19 \cdot 21$, or $1 = -(-8) \cdot 50 + (-19) \cdot 21$. Hence, $c = -8$ and $j = -19$, which makes little sense. What does it mean for the frog to make -19 jumps?

The point is that the Pulverizer is guaranteed to give us the integer coefficients of a linear combination that equals the GCD, but it promises nothing about the signs of those coefficients (in this case we wanted them to be $-$ and $+$). To get coefficients of the desired sign, we have to think more.

One way to fix it is as explained in lecture. That is, subtract 21 from 8 and add 50 to -19.

Here is another way: We know $1 = 8 \cdot 50 - 19 \cdot 21$. Or, to obtain meaningful signs for the numbers, $19 \cdot 21 = 8 \cdot 50 - 1$. That is, after 19 jumps the frog will land 1 pebble short of

8 full cycles. So, he will be right next to his original pebble, but in the counter-clockwise direction. Given this, how can he reach the pebble he is after?

Well, if he makes 19 more jumps, he will land 2 pebbles away from his original position in the counter-clockwise direction. Another 19 jumps will lead him 3 pebbles away, and so on. After a total of 49 sets of 19 jumps, he will be 49 pebbles away from its original position in the counter-clockwise direction, which is of course the worm's pebble. Then, the frog will have made $49 * 19 = 931$ jumps.

Here is the table produced by the Pulverizer:

x	y	$(\operatorname{rem}(x,y))$	$=$	$x - q \cdot y$
50	21	8	$=$	$50 - 2 \cdot 21$
21	8	5	$=$	$21 - 2 \cdot 8$
			$=$	$21 - 2 \cdot (50 - 2 \cdot 21)$
			$=$	$-2 \cdot 50 + 5 \cdot 21$
8	5	3	$=$	$8 - 1 \cdot 5$
			$=$	$(50 - 2 \cdot 21) - 1 \cdot (-2 \cdot 50 + 5 \cdot 21)$
			$=$	$3 \cdot 50 - 7 \cdot 21$
5	3	2	$=$	$5 - 1 \cdot 3$
			$=$	$(-2 \cdot 50 + 5 \cdot 21) - 1 \cdot (3 \cdot 50 - 7 \cdot 21)$
			$=$	$-5 \cdot 50 + 12 \cdot 21$
3	2	1	$=$	$3 - 1 \cdot 2$
			$=$	$(3 \cdot 50 - 7 \cdot 21) - 1 \cdot (-5 \cdot 50 + 12 \cdot 21)$
			$=$	$\boxed{8 \cdot 50 - 19 \cdot 21}$
2	1	0		

■

3 Problem: The Fibonacci numbers.

The Fibonacci numbers are defined as follows:

$$F_0 = 0 \qquad F_1 = 1 \qquad F_n = F_{n-1} + F_{n-2} \quad (\text{for } n \geq 2).$$

Give an inductive proof that the Fibonacci numbers F_n and F_{n+1} are relatively prime for all $n \geq 0$.

Solution. We use induction on n. Let $P(n)$ be the proposition that F_n and F_{n+1} are relatively prime.

Base case: $P(0)$ is true because $F_0 = 0$ and $F_1 = 1$ are relatively prime.

Inductive step: We show that, for all $n \geq 0$, $P(n)$ implies $P(n+1)$. So, fix some $n \geq 0$ and assume that $P(n)$ is true, that is, F_n and F_{n+1} are relatively prime. We will show that F_{n+1} and F_{n+2} are relatively prime as well. We will do this by contradiction.

Suppose F_{n+1} and F_{n+2} are not relatively prime. Then they have a common divisor $d > 1$. But then d also divides the linear combination $F_{n+2} - F_{n+1}$, which actually equals F_n. Hence, $d > 1$ divides both F_n and F_{n+1}. Which implies F_n, F_{n+1} are not relatively prime, a contradiction to the inductive hypothesis.

Therefore, F_{n+1} and F_{n+2} are relatively prime. That is, $P(n+1)$ is true.

The theorem follows by induction. ■

4 Extra Problem: The power of 3.[2]

Let N be a number whose decimal expansion consists of 3^n identical digits. Show by induction that $3^n \mid N$. For example:

$$3^2 \mid \underbrace{777777777}_{3^2 = 9 \text{ digits}}$$

Recall that 3 divides a number iff it divides the sum of its digits.

Solution. We proceed by induction on n. Let $P(n)$ be the proposition that $3^n \mid N$, where the decimal expansion of N consists of 3^n identical digits.

Base case. $P(0)$ is true because $3^0 = 1$ divides every number.

Inductive step. Now we show that, for all $n \geq 0$, $P(n)$ implies $P(n+1)$. Fix any $n \geq 0$ and assume $P(n)$ is true. Consider a number whose decimal expansion consists of 3^{n+1} copies of the digit a:

$$\begin{aligned}\underbrace{aaaaaa\ldots aaaaaa}_{3^{n+1}\text{ digits}} &= \underbrace{aaa\ldots aaa}_{3^n\text{ digits}}\underbrace{aaa\ldots aaa}_{3^n\text{ digits}}\underbrace{aaa\ldots aaa}_{3^n\text{ digits}} \\ &= \underbrace{aaa\ldots aaa}_{3^n\text{ digits}} \quad \cdot \quad 1\underbrace{000\ldots 001}_{3^n\text{ digits}}\underbrace{000\ldots 001}_{3^n\text{ digits}}\end{aligned}$$

Now 3^n divides the first term by the assumption $P(n)$, and 3 divides the second term since the digits sum to 3. Therefore, the whole expression is divisible by 3^{n+1}. This proves $P(n+1)$.

By the principle of induction $P(n)$ is true for all $n \geq 0$. ■

[2] Try this if you have time!

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Problems for Recitation 5

1 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.

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

   (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.

   (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!)
   - 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!

   When you're done, write your public key and group members' names on the board.

2. Now ask your recitation instructor for a message to encrypt and send to another team using *their* public key.

   The messages m correspond to statements from the codebook below:

   2 = Greetings and salutations!

   3 = Wassup, yo?

   4 = You guys are slow!

   5 = All your base are belong to us.

   6 = Someone on *our* team thinks someone on *your* team is kinda cute.

   7 = You are the weakest link. Goodbye.

3. **Encode** the message you were given using another team's public key.
4. Now **decrypt** the message sent to you and verify that you received what the other team sent!
5. Explain how you could read messages encrypted with RSA if you could quickly factor large numbers.

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:

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

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

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

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Notes for Recitation 5

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:

$$a_1 \cdot a_2 \cdot \ldots \cdot a_n \equiv \text{rem}\,(a_1, c) \cdot \text{rem}\,(a_2, c) \cdot \ldots \cdot \text{rem}\,(a_n, c) \pmod{c}$$

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 $m = \operatorname{rem}(m^e, n)$ 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 $n = 17$.

We are trying to find $\operatorname{rem}(m^e, n)$. We know that $m^e = 5^6 = 5 \cdot 5 \cdot 5 \cdot 5 \cdot 5 \cdot 5$.

$$\begin{aligned}
5^2 &\equiv 8 \pmod{17} \\
5^3 &\equiv 8 \cdot 5 \equiv 6 \pmod{17} \\
5^4 &\equiv 6 \cdot 5 \equiv 13 \pmod{17} \\
5^5 &\equiv 13 \cdot 5 \equiv 14 \pmod{17} \\
5^6 &\equiv 14 \cdot 5 \equiv 2 \pmod{17}
\end{aligned}$$

OK, that's nice, but for large e, $e-1$ is still a lot of multiplications! As we promised earlier, there's a yet more efficient way to do the exponentiation. It's called *repeated squaring*.

Example: Encrypt a message $m = 5$ with $e = 149$ and $n = 17$.

Note that the binary expansion of 149 is 10010101, so one can compute $\operatorname{rem}(5^{149}, 17)$ by computing $\operatorname{rem}(5^{128+16+4+1}, 17)$.

$$\begin{aligned}
5^2 &\equiv 8 \pmod{17} \\
5^4 &\equiv 8 \cdot 8 \equiv 13 \pmod{17} \\
5^8 &\equiv 13 \cdot 13 \equiv 16 \pmod{17} \\
5^{16} &\equiv 16 \cdot 16 \equiv 1 \pmod{17} \\
5^{32} &\equiv 1 \cdot 1 \equiv 1 \pmod{17} \\
5^{64} &\equiv 1 \cdot 1 \equiv 1 \pmod{17} \\
5^{128} &\equiv 1 \cdot 1 \equiv 1 \pmod{17}
\end{aligned}$$

We used only 7 multiplications to find the remainders of $5^{2k} \pmod{17}$ by repeatedly squaring each previous output and taking the remainder. Then, with only 3 additional multiplications to combine these products, we can compute $5^{128} \cdot 5^{16} \cdot 5^4 \cdot 5^1 \equiv 1 \cdot 1 \cdot 13 \cdot 5 \equiv 14 \pmod{13}$. This saved us $(149-1) - (7+3) = 138$ multiplications!

You may notice that in this particular case, $5^{16} \equiv 1 \pmod{17}$, so we could have even stopped our squaring at 5^{16} and reduced the problem to computing $\operatorname{rem}(5^{16 \cdot 9+4+1}, 17) \equiv (5^{16})^9 \cdot 5^4 \cdot 5 \equiv 1^9 \cdot 13 \cdot 5 \equiv 14 \pmod{17}$. For this we only needed $(4+2) = 6$ multiplications!

You may find this technique very useful in the next problem.

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 $13^2 = 169 \equiv 49$

(mod 60). We square our result to get $13^4 = (13^2)^2 \equiv 49^2 = 2401 \equiv 1 \pmod{60}$.

Once we know $13^4 \equiv 1 \pmod{60}$, our job is much easier. $13^{15} = (13^4)^3 \cdot 13^2 \cdot 13 \equiv 1^3 \cdot 49 \cdot 13 = 637 \equiv 37 \pmod{60}$. This matches our answer from the Pulverizer. Which method is easier depends on the particular numbers that we've chosen. ■

When you're done, write your public key and group members' names on the board.

2. Now ask your recitation instructor for a message to encrypt and send to another team using *their* public key.

   The messages m correspond to statements from the codebook below:

   2 = Greetings and salutations!

   3 = Wassup, yo?

   4 = You guys are slow!

   5 = All your base are belong to us.

   6 = Someone on *our* team thinks someone on *your* team is kinda cute.

   7 = You are the weakest link. Goodbye.

3. **Encode** the message you were given using another team's public key.

   Solution. Let's say our message was $m = 3$ and the other team's public key was $(e, n) = (11, 35)$. The encrypted message would then be $m' = \text{rem}\,(3^{11}, 35)$. Using repeated squaring, we see that $3^{11} = 3^{8+2+1}$. We compute $3^2 = 9 \pmod{35}$, $3^4 = 81 \equiv 11 \pmod{35}$, $3^8 = (3^4)^2 \equiv 11^2 = 121 \equiv 16 \pmod{35}$. Therefore $3^{11} \equiv 16 \cdot 9 \cdot 3 = 432 \equiv 12 \pmod{35}$, so our message is $m' = 12$. ■

4. Now **decrypt** the message sent to you and verify that you received what the other team sent!

   Solution. Let's say the other team sent you the encrypted message $m' = 26$. In our case, our private key was $(d, n) = (37, 77)$. The decrypted original message would then be $m = \text{rem}\,(26^{37}, 77)$. Using repeated squaring, we find $m = 5$. ■

5. Explain how you could read messages encrypted with RSA if you could quickly factor large numbers.

   Solution. Suppose you see a public key (e, n). If you can factor n to obtain p and q, then you can compute d using the Pulverizer or Euler's Theorem. This gives you the secret key (d, n), and so you can decode messages as well as the intended recipient. ■

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:

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

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

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

MIT OpenCourseWare
http://ocw.mit.edu

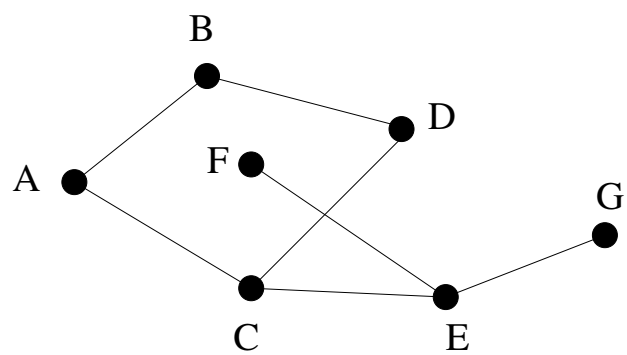
6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

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

$$v_1,\ v_2,\ v_3,\ \ldots,\ v_n$$

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.

MIT OpenCourseWare
http://ocw.mit.edu

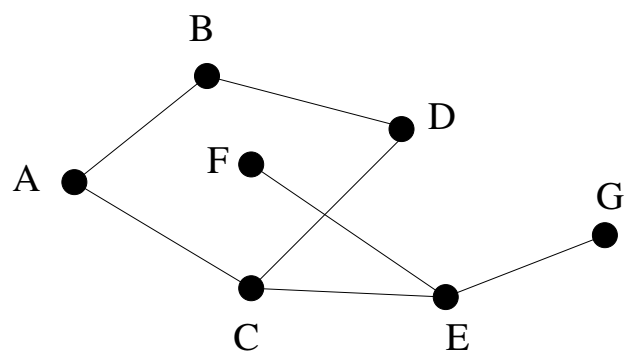
6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Notes 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

$$v_1,\ v_2,\ v_3,\ \ldots,\ v_n$$

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.)

Solution. We use induction on n, the number of vertices. Let $P(n)$ be the proposition that every graph with width w is $(w+1)$ colorable.

Base case: Every graph with $n = 1$ vertex has width 0 and is $0+1 = 1$ colorable. Therefore, $P(1)$ is true.

Inductive step: Now we assume $P(n)$ in order to prove $P(n+1)$. Let G be an $(n+1)$-vertex graph with width w. This means that the vertices can be arranged in a sequence

$$v_1, v_2, v_3, \ldots, v_n, v_{n+1}$$

such that each vertex v_i is connected to at most w preceding vertices. Removing vertex v_{n+1} and all incident edges gives a graph G' with n vertices and width at most w. (If original sequence is retained, then removing v_{n+1} does not increase the number of edges from a vertex v_i to a preceding vertex.) Thus, G' is $(w+1)$-colorable by the assumption $P(n)$. Now replace vertex v_{n+1} and its incident edges. Since v_{n+1} is joined by an edge to at most w preceding vertices, we can color v_{n+1} differently from all of these. Therefore, $P(n+1)$ is true.

The theorem follows by the principle of induction. ■

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.

   Solution. The edge set of any subgraph will be a subset of the set of edges in the original planar graph. This means that since edges in the original graphs do not cross, edges in a subset of the original set of edges also do not cross. ■

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.

   Solution. We prove by induction. First, let n be the number of nodes in the graph. Then define

   $$P(n) = \text{Any planar graph with } n \text{ nodes is 6-colorable.}$$

   Base case, $P(1)$: Every graph with $n = 1$ vertex is 6-colorable. Clearly true since it's actually 1-colorable.

Inductive step, $P(n) \to P(n+1)$: Take a graph G with $n+1$ nodes. Then take a node v with degree at most 5 (which we know exists because we know any planar graph has a node of degree ≤ 5), and remove it. We know that the induced subgraph G' formed in this way has n nodes, so by our inductive hypothesis, G' is 6-colorable. But v is adjacent to at most 5 other nodes, which can have at most 5 different colors between them. We then choose v to have an unused color (from the 6 colors), and as we have constructed a 6-coloring for G, we are done with the inductive step.

Because we have shown the base case and the inductive step, we have proved

$$\forall n \in \mathbb{Z}_+ : P(n)$$

(Note: $\mathbb{Z}_+$ refers to the set of positive integers.) ■

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Problems for Recitation 7

1 A Protocol for College Admission

Next, we are going to talk about a generalization of the stable marriage problem. Recall that we have some horses and we'd like to pair them with stables so that there is no incentive for two horses to swap stables. Oh wait, that's a different problem.

The problem we're going to talk about is a generalization of the one done in lecture. In the new problem, there are N students $s_1, s_2, \ldots, s_N$ and M universities $u_1, u_2, \ldots, u_M$. University u_i has n_i slots for students, and we're guaranteed that $\sum_{i=1}^{M} n_i = N$. Each student ranks all universities (no ties) and each university ranks all students (no ties).

Design an algorithm to assign students to universities with the following properties

1. Every student is assigned to one university.
2. $\forall i$, u_i gets assigned n_i students.
3. There does not exist s_i, s_j, u_k, u_ℓ where s_i is assigned to u_k, s_j is assigned to u_ℓ, s_j prefers u_k to u_ℓ, and u_k prefers s_j to s_i.
4. It is student-optimal. This means that of all possible assignments satisfying the first three properties, the students get their top choice of university amongst these assignments.

The algorithm will be a slight modification of the mating algorithm given in lecture. For your convenience, we have provided a copy of the mating algorithm on the next page.

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).

1. Before we can say anything about our algorithm, we need to show that it terminates. Show that the algorithm terminates after $NM + 1$ days.

2. Next, we will show that the four properties stated earlier are true of our algorithm. To start, let's show the following: if during some day a university u_j has at least n_j applicants, then when the algorithm terminates it accepts exactly n_j students.

3. Next, show that every student is assigned to one university.

4. Next, show that for all i, u_i gets assigned n_i students.

5. Before continuing, we need to establish the following property. Suppose that on some day a university u_j has at least n_j applicants. Define the *rank* of an applicant s_i with respect to a university u_j as s_i's location on u_j's preference list. So, for example, u_j's favorite student has rank 1. Show that the rank of u_j's least favorite applicant that it says "Maybe, ..." to cannot decrease (e.g., going from 1000 to 1005 is decreasing) on any future day. Note that u_j's least favorite applicant might change from one day to the next.

6. Next, show there does not exist s_i, s_j, u_k, and u_ℓ where s_i is assigned to u_k, s_j is assigned to u_ℓ, s_j prefers u_k to u_ℓ, and u_k prefers s_j to s_i. Note that this is analogous to a "rogue couple" considered in lecture.

7. Finally, we show in a very precise sense that this algorithm is *student-optimal*. As in lecture, define the *realm of possibility* of a student to be the set of all universities u, for which there exists some assignment satisfying the first three properties above, in which the student is assigned to u. Of all universities in the realm of possibility of a student we say that the student's favorite is *optimal* for that student.

   Show that each student is assigned to its optimal university.

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Notes for Recitation 7

1 A Protocol for College Admission

Next, we are going to talk about a generalization of the stable marriage problem. Recall that we have some horses and we'd like to pair them with stables so that there is no incentive for two horses to swap stables. Oh wait, that's a different problem.

The problem we're going to talk about is a generalization of the one done in lecture. In the new problem, there are N students $s_1, s_2, \ldots, s_N$ and M universities $u_1, u_2, \ldots, u_M$. University u_i has n_i slots for students, and we're guaranteed that $\sum_{i=1}^{M} n_i = N$. Each student ranks all universities (no ties) and each university ranks all students (no ties).

Design an algorithm to assign students to universities with the following properties

1. Every student is assigned to one university.
2. $\forall i$, u_i gets assigned n_i students.
3. There does not exist s_i, s_j, u_k, u_ℓ where s_i is assigned to u_k, s_j is assigned to u_ℓ, s_j prefers u_k to u_ℓ, and u_k prefers s_j to s_i.
4. It is student-optimal. This means that of all possible assignments satisfying the first three properties, the students get their top choice of university amongst these assignments.

The algorithm will be a slight modification of the mating algorithm given in lecture. For your convenience, we have provided a copy of the mating algorithm on the next page.

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).

Solution. Each Day:

- Morning:
  - Each university asks which students are interested in applying.
  - Each student applies to his/her favorite university that has not yet rejected him/her. If there are no universities left on the student's list, the student takes some time off to think about life, the future, and the unchangeable past.
- Afternoon:
  - Universities u_i do the following:
  - u_i tells its favorite n_i applicants "Maybe, we are still processing your application." If u_i has less than n_i applicants, it tells all of its applicants this message.
  - If u_i has more than n_i applicants, it tells the remaining ones "Sorry, there were a large number of very qualified students applying this year, yet we can only accept a very limited number. We regret to inform you that you were not accepted. Thank you for applying to our university."
- Evening:
  - Any student who hears "Sorry, ..." from some university, crosses off that university from his/her list.

Termination Condition: If there is a day when each university u_i has at most n_i applicants, we stop and each university accepts all of its applicants (if any). ■

1. Before we can say anything about our algorithm, we need to show that it terminates. Show that the algorithm terminates after $NM + 1$ days.

   Solution. On each day, if the algorithm has not terminated, then some university u_i has more than n_i applicants. It follows that in the afternoon, at least one student s_j hears "Sorry, ...", and thus in the evening s_j crosses off u_i from his/her list. As there are N students and M universities, it follows that the algorithm must terminate after $NM + 1$ days, as otherwise there would be no university left for any student to cross off. ■

2. Next, we will show that the four properties stated earlier are true of our algorithm. To start, let's show the following: if during some day a university u_j has at least n_j applicants, then when the algorithm terminates it accepts exactly n_j students.

   Solution. At this day, each of the students applying to u_j has u_j as their favorite university that has not yet rejected him/her. Therefore, if u_j tells a student "Maybe, ...", that student will come back the next day. Since there are at least n_j applicants, it follows that u_j will tell its favorite n_j "Maybe, ...". It follows by induction that every day after this day, u_j will have at least n_j applicants. Thus, this holds when the algorithm terminates. Since when the algorithm terminates there are at most n_j applicants, it follows that exactly n_j students are assigned to u_j. ■

3. Next, show that every student is assigned to one university.

   Solution. It is clear that no student can apply to more than one university at once since a student applies to at most one university on any given day, so this means the students can be assigned to at most one university. So we just need to show that each student is assigned to at least one university. We argue by contradiction.

   Suppose not, and let s_j be a student not assigned to any university. Then, since the algorithm terminates, and when the algorithm terminates each university u_i accepts at most n_i students, it follows that some university u_i accepts less than n_i students. By the previous problem, it follows that in every day, u_i had less than n_i applicants. But then consider the day that s_j applied to u_i. Since there were less than n_i applicants to u_i that day, it follows that u_i would have told s_j "Maybe, ..." in that day, and thus in every future day. Thus, s_j would be assigned u_i when the algorithm terminates. This is a contradiction. ■

4. Next, show that for all i, u_i gets assigned n_i students.

   Solution. Since the algorithm terminates, on some day each u_i gets assigned at most n_i students. Suppose some u_i got assigned strictly less than n_i students. Since $\sum_{i=1}^{M} u_i = N$, this means that some student is not assigned. This contradicts the previous problem. ■

5. Before continuing, we need to establish the following property. Suppose that on some day a university u_j has at least n_j applicants. Define the *rank* of an applicant s_i with respect to a university u_j as s_i's location on u_j's preference list. So, for example, u_j's favorite student has rank 1. Show that the rank of u_j's least favorite applicant that it says "Maybe, ..." to cannot decrease (e.g., going from 1000 to 1005 is decreasing) on any future day. Note that u_j's least favorite applicant might change from one day to the next.

   Solution. On the next day, there are two cases: u_j either says "Maybe, ..." to its least favorite applicant s_i from the previous day, or it says "Sorry, ..." to s_i. In the first case, this means that all of the $n_j - 1$ applicants u_j liked more than s_i on the previous day will also be told "Maybe, ...", and so s_i will again be u_j's least favorite applicant it did not reject. Thus, the rank of its least favorite applicant did not decrease. In the second case, this means that there were at least n_j applicants that u_j preferred to s_i, and thus the rank of its least favorite applicant it said "Maybe, ..." to did not decrease. As shown above, on any future day u_j has at least n_j applicants, and so by applying this analysis again, we conclude that the rank of u_j's least favorite applicant that it says "Maybe, ..." to cannot decrease on any future day. ■

6. Next, show there does not exist s_i, s_j, u_k, and u_ℓ where s_i is assigned to u_k, s_j is assigned to u_ℓ, s_j prefers u_k to u_ℓ, and u_k prefers s_j to s_i. Note that this is analogous to a "rogue couple" considered in lecture.

Solution. Assume, towards a contradiction, that such an s_i, s_j, u_k and u_ℓ existed. Since s_j prefers u_k to u_ℓ, but is assigned to u_ℓ, on some day u_k told s_j "Sorry, ...". On that day, there must have been more than n_k applicants to u_k. If s_i was also an applicant to u_k on that day, then s_i would have also been rejected since u_k prefers s_j to s_i, and thus s_i could not have been assigned to u_k. On the other hand, if at any later day s_i were to apply to u_k, it would have been rejected since s_i's rank is less than s_j's with respect to u_k, since by the previous problem we know that the rank of the least favorite applicant that u_k says "Maybe, ..." to, cannot decrease. Thus, it is impossible for s_i to be assigned to u_k, which is a contradiction. ■

7. Finally, we show in a very precise sense that this algorithm is *student-optimal*. As in lecture, define the *realm of possibility* of a student to be the set of all universities u, for which there exists some assignment satisfying the first three properties above, in which the student is assigned to u. Of all universities in the realm of possibility of a student we say that the student's favorite is *optimal* for that student.

   Show that each student is assigned to its optimal university.

   Solution. We argue by contradiction. Consider the first (in time) student s_i not assigned to its optimal university, and suppose for s_i this university is u_k. Then on the date s_i is rejected from u_k, there was another student s_j which u_k preferred to s_i. Since u_k is in the realm of possibility of s_i, there is an assignment M of students to universities assigning s_i to u_k with the properties above. Suppose M assigns s_j to u_ℓ. Suppose u_{opt} is s_j's optimal university. Then, since s_i was the first student not assigned its optimal university, s_j prefers u_k to u_{opt}, though u_k may equal u_{opt}. On the other hand, s_j prefers u_{opt} to u_ℓ, since u_{opt} is its favorite university in its realm of possibility, and u_ℓ occurs in its realm of possibility. It follows that s_j prefers u_k to u_ℓ. But then in the assignment M we have found an s_i, s_j, u_k, and u_ℓ with s_i assigned to u_k, s_j assigned to u_ℓ, s_j prefers u_k to u_ℓ, and u_k prefers s_j to s_i. This is a contradiction to the property of M established in the previous problem. ■

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Problems for Recitation 8

1 Build-up error

Recall a graph is ***connected*** iff there is a path between every pair of its vertices.

False Claim. *If every vertex in a graph has positive degree, then the graph is connected.*

1. Prove that this Claim is indeed false by providing a counterexample.

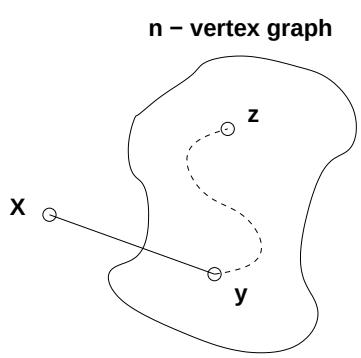
2. Since the Claim is false, there must be a logical mistake in the following bogus proof. Pinpoint the *first* logical mistake (unjustified step) in the proof.

   Proof. We prove the Claim above by induction. Let $P(n)$ be the proposition that if every vertex in an n-vertex graph has positive degree, then the graph is connected.

   Base cases: ($n \leq 2$). In a graph with 1 vertex, that vertex cannot have positive degree, so $P(1)$ holds vacuously.

   $P(2)$ holds because there is only one graph with two vertices of positive degree, namely, the graph with an edge between the vertices, and this graph is connected.

   Inductive step: We must show that $P(n)$ implies $P(n+1)$ for all $n \geq 2$. Consider an n-vertex graph in which every vertex has positive degree. By the assumption $P(n)$, this graph is connected; that is, there is a path between every pair of vertices. Now we add one more vertex x to obtain an $(n+1)$-vertex graph:



   All that remains is to check that there is a path from x to every other vertex z. Since x has positive degree, there is an edge from x to some other vertex, y. Thus, we can

obtain a path from x to z by going from x to y and then following the path from y to z. This proves $P(n+1)$.

By the principle of induction, $P(n)$ is true for all $n \geq 0$, which proves the Claim.

□

2 The Grow Algorithm

Yesterday in lecture, we saw the following algorithm for constructing a minimum-weight spanning tree (MST) from an edge-weighted N-vertex graph G.

ALG-GROW:

1. Label the edges of the graph $e_1, e_2, \ldots, e_t$ so that $wt(e_1) \leq wt(e_2) \ldots \leq wt(e_t)$.
2. Let S be the empty set.
3. For $i = 1 \ldots t$, if $S \cup \{e_i\}$ does not contain a cycle, then extend S with the edge e_i.
4. Output S.

2.1 Analysis of ALG-GROW

In this problem you may assume the following lemma from the problem set:

Lemma 1. *Suppose that $T = (V, E)$ is a simple, connected graph. Then T is a tree iff $|E| = |V| - 1$.*

In this exercise you will prove the following theorem.

Theorem. *For any connected, weighted graph G, ALG-GROW produces an MST of G.*

(a) Prove the following lemma.

Lemma 2. *Let $T = (V, E)$ be a tree and let e be an edge not in E. Then, $G = (V, E \cup \{e\})$ contains a cycle.*

(Hint: Suppose G does *not* contain a cycle. Is G a tree?)

(b) Prove the following lemma.

Lemma 3. *Let $T = (V, E)$ be a spanning tree of G and let e be an edge not in E. Then there exists an edge $e' \neq e$ in E such that $T^* = (V, E - \{e'\} \cup \{e\})$ is a spanning tree of G.*

(Hint: Adding e to E introduces a cycle in $(V, E \cup \{e\})$.)

(c) Prove the following lemma.

Lemma 4. *Let $T = (V, E)$ be a spanning tree of G, let e be an edge not in E and let $S \subseteq E$ such that $S \cup \{e\}$ does not contain a cycle. Then there exists an edge $e' \neq e$ in $E - S$ such that $T^* = (V, E - \{e'\} \cup \{e\})$ is a spanning tree of G.*

(Hint: Modify your proof to part (b). Of all possible edges $e' \neq e$ that can be removed to construct T^*, at least one is not in S.)

(d) Prove the following lemma.

Lemma 5. *Define S_m to be the set consisting of the first m edges selected by ALG-GROW from a connected graph G. Let $P(m)$ be the predicate that if $m \leq |V|$ then $S_m \subseteq E$ for some MST $T = (V, E)$ of G. Then $\forall m \,.\, P(m)$.*

(Hint: Use induction. There are two cases: $m + 1 > |V|$ and $m + 1 \leq |V|$. In the second case, there are two subcases.)

(e) Prove the theorem. (Hint: Lemma 5 says there exists an MST $T = (V, E)$ for G such that $S \subseteq E$. Use contradiction to rule out the case in which S is a proper subset of E.)

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Notes for Recitation 8

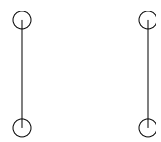
1 Build-up error

Recall a graph is ***connected*** iff there is a path between every pair of its vertices.

False Claim. *If every vertex in a graph has positive degree, then the graph is connected.*

1. Prove that this Claim is indeed false by providing a counterexample.

   Solution. There are many counterexamples; here is one:



■

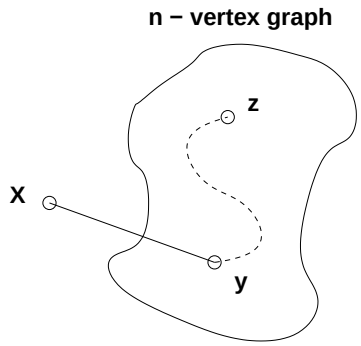
2. Since the Claim is false, there must be a logical mistake in the following bogus proof. Pinpoint the *first* logical mistake (unjustified step) in the proof.

   Proof. We prove the Claim above by induction. Let $P(n)$ be the proposition that if every vertex in an n-vertex graph has positive degree, then the graph is connected.

   Base cases: ($n \leq 2$). In a graph with 1 vertex, that vertex cannot have positive degree, so $P(1)$ holds vacuously.

   $P(2)$ holds because there is only one graph with two vertices of positive degree, namely, the graph with an edge between the vertices, and this graph is connected.

   Inductive step: We must show that $P(n)$ implies $P(n+1)$ for all $n \geq 2$. Consider an n-vertex graph in which every vertex has positive degree. By the assumption $P(n)$, this graph is connected; that is, there is a path between every pair of vertices. Now we add one more vertex x to obtain an $(n+1)$-vertex graph:



All that remains is to check that there is a path from x to every other vertex z. Since x has positive degree, there is an edge from x to some other vertex, y. Thus, we can obtain a path from x to z by going from x to y and then following the path from y to z. This proves $P(n+1)$.

By the principle of induction, $P(n)$ is true for all $n \geq 0$, which proves the Claim.

□

Solution. This one is tricky: the proof is actually a good proof of something else. The first error in the proof is only in the final statement of the inductive step: "This proves $P(n+1)$".

The issue is that to prove $P(n+1)$, *every* $(n+1)$-vertex positive-degree graph must be shown to be connected. But the proof doesn't show this. Instead, it shows that every $(n+1)$-vertex positive-degree graph *that can be built up by adding a vertex of positive degree to an n-vertex connected graph*, is connected.

The problem is that *not every* $(n+1)$-vertex positive-degree graph can be built up in this way. The counterexample above illustrates this: there is no way to build that 4-vertex positive-degree graph from a 3-vertex positive-degree graph.

More generally, this is an example of "buildup error". This error arises from a faulty assumption that every size $n+1$ graph with some property can be "built up" in some particular way from a size n graph with the same property. (This assumption is correct for some properties, but incorrect for others— such as the one in the argument above.)

One way to avoid an accidental build-up error is to use a "shrink down, grow back" process in the inductive step: start with a size $n+1$ graph, remove a vertex (or edge), apply the inductive hypothesis $P(n)$ to the smaller graph, and then add back the vertex (or edge) and argue that $P(n+1)$ holds. Let's see what would have happened if we'd tried to prove the claim above by this method:

Inductive step: We must show that $P(n)$ implies $P(n+1)$ for all $n \geq 1$. Consider an $(n+1)$-vertex graph G in which every vertex has degree at least 1. Remove an arbitrary vertex v, leaving an n-vertex graph G' in which every vertex has degree... uh-oh!

The reduced graph G' might contain a vertex of degree 0, making the inductive hypothesis $P(n)$ inapplicable! We are stuck— and properly so, since the claim is false! ■

2 The Grow Algorithm

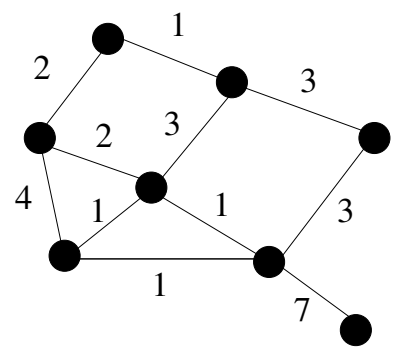
Yesterday in lecture, we saw the following algorithm for constructing a minimum-weight spanning tree (MST) from an edge-weighted N-vertex graph G.

ALG-GROW:

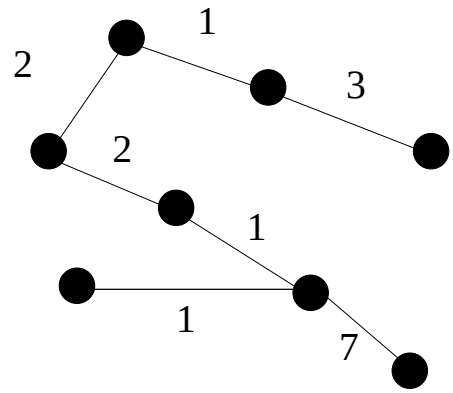
1. Label the edges of the graph $e_1, e_2, \ldots, e_t$ so that $wt(e_1) \leq wt(e_2) \ldots \leq wt(e_t)$.
2. Let S be the empty set.
3. For $i = 1 \ldots t$, if $S \cup \{e_i\}$ does not contain a cycle, then extend S with the edge e_i.
4. Output S.

In summary, ALG-GROW selects edges one at a time, always choosing the minimum weight edge that does not create a cycle with previously selected edges. Notice that as edges are added S may not be connected. When the algorithm terminates, S contains $N-1$ edges. If it is connected, then it is a spanning tree.

Consider, for example, the following edge-weighted graph.



Now suppose we run ALG-GROW on our graph. We may choose the weight 1 edge on the bottom of the triangle of weight 1 edges in our graph. In the next step, we may choose the weight 1 edge on the top of the graph. Note that this edge still has minimum weight, and does not cause us to form a cycle, so ALG-GROW can choose it. We will then choose one of the remaining weight 1 edges. Note that neither causes us to form a cycle. Continuing the algorithm, we may end up with the same spanning tree shown below.



In this recitation, we will analyze ALG-GROW.

3 Analysis of ALG-GROW

In this problem you may assume the following lemma from the problem set:

Lemma 1. *Suppose that $T = (V, E)$ is a simple, connected graph. Then T is a tree iff $|E| = |V| - 1$.*

In this exercise you will prove the following theorem.

Theorem. *For any connected, weighted graph G, ALG-GROW produces an MST of G.*

(a) Prove the following lemma.

Lemma 2. *Let $T = (V, E)$ be a tree and let e be an edge not in E. Then, $G = (V, E \cup \{e\})$ contains a cycle.*

(Hint: Suppose G does *not* contain a cycle. Is G a tree?)

Solution. *Proof.* (by contradiction) Suppose G does not contain a cycle. By the definition of a tree, T is connected. Notice that T is a subgraph of G. Because any two nodes in G are connected by a path in T, G is a connected graph. So G is connected and acyclic and therefore a tree by definition. Both G and T are trees and have the same number of nodes. Therefore, they have the same number of edges (by Lemma 1). This is a contradiction because G has one more edge than T. □

■

(b) Prove the following lemma.

Lemma 3. *Let $T = (V, E)$ be a spanning tree of G and let e be an edge not in E. Then there exists an edge $e' \neq e$ in E such that $T^* = (V, E - \{e'\} \cup \{e\})$ is a spanning tree of G.*

(Hint: Adding e to E introduces a cycle in $(V, E \cup \{e\})$.)

Solution. *Proof.* By Lemma 2, we know that the set of edges $E \cup \{e\}$ contains a cycle. If this cycle does not contain the edge e, then this cycle is a subset of E. Since E is the set of edges of a tree, this cannot occur. So, this cycle contains e. If e' is another edge distinct from e in this cycle, then the graph T^* that results after removing e' from $E \cup \{e\}$ is still connected. The number of edges in T^* is equal to the number of edges in T, which is equal to $|V| - 1$ by Lemma 1. Since T^* is connected, T^* is a tree by Lemma 1. Since T^* is a subgraph of G with vertices V, it spans G. □

■

(c) Prove the following lemma.

Lemma 4. *Let $T = (V, E)$ be a spanning tree of G, let e be an edge not in E and let $S \subseteq E$ such that $S \cup \{e\}$ does not contain a cycle. Then there exists an edge $e' \neq e$ in $E - S$ such that $T^* = (V, E - \{e'\} \cup \{e\})$ is a spanning tree of G.*

(Hint: Modify your proof to part (b). Of all possible edges $e' \neq e$ that can be removed to construct T^*, at least one is not in S.)

Solution. *Proof.* We need to change the proof in part (b) slightly. The proof of part (b) holds for any edge $e' \neq e$ in the cycle. We need to show that we can select an edge $e' \neq e$ that is in the cycle but not in S. We will prove this by contradiction. Suppose that all the edges not equal to e that are in the cycle are in S. Then, $S \cup e$ is a cycle. This contradicts the assumption of the lemma. □

■

(d) Prove the following lemma.

Lemma 5. *Define S_m to be the set consisting of the first m edges selected by ALG-GROW from a connected graph G. Let $P(m)$ be the predicate that if $m \leq |V|$ then $S_m \subseteq E$ for some MST $T = (V, E)$ of G. Then $\forall m \,.\, P(m)$.*

(Hint: Use induction. There are two cases: $m + 1 > |V|$ and $m + 1 \leq |V|$. In the second case, there are two subcases.)

Solution. *Proof.* (By induction.) Let $P(m)$ be the predicate as defined above.

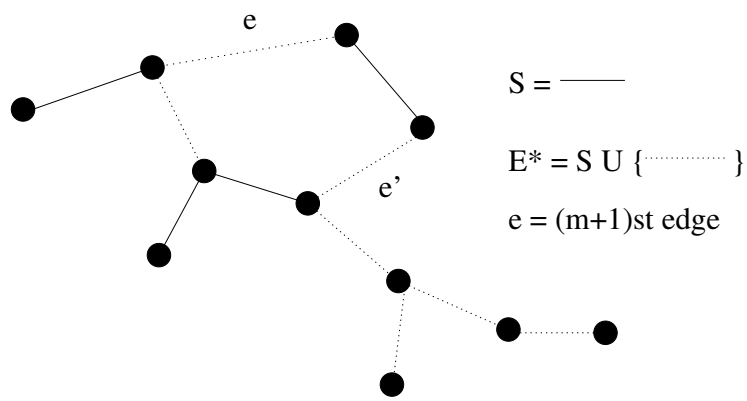
Base Case: S_0 contains 0 edges and is equal to the empty set, which is a subset of any set of edges E.

Inductive Step: Assume $P(m)$ in order to prove $P(m + 1)$.

If $m \geq |V|$ then $m+1 > |V|$ and $P(m+1)$ holds vacuously. Otherwise, if $m < |V|$ then let e denote the $(m+1)$th edge selected by ALG-GROW. By the inductive hypothesis, there exists an MST $T = (V, E)$ such that $S_m \subseteq E$. There are now two cases.

In the first case, $e \in E$ which case $S_m \cup \{e\} \subseteq E$, and thus $P(m + 1)$ holds.

In the second case, $e \notin E$, as illustrated by the following diagram. Now we need to find a different MST that contains $S_m \cup \{e\}$.



What happens when we add e to T? By the description of ALG-GROW, $S_m \cup \{e\}$ does not contain a cycle. Therefore, by Lemma 4, there exists an edge $e' \neq e$ in $E - S_m$ such that $T^* = (V, E - \{e'\} \cup \{e\})$ is a spanning tree for G.

In order to prove that T^* is a MST, we need to show that $wt(e) \leq wt(e')$. We will prove this by contradiction. Suppose that $wt(e') < wt(e)$. Since $e' \in E$, which is the set of edges of the MST T, and $S_m \subseteq E$, the set of edges $S_m \cup \{e'\}$, does not contain a cycle. Therefore e' would have already been added to S_m in a previous iteration of ALG-GROW as one of the first m edges. However, e' is in $E - S_m$. This is a contradiction. □

■

(e) Prove the theorem. (Hint: Lemma 5 says there exists an MST $T = (V, E)$ for G such that $S \subseteq E$. Use contradiction to rule out the case in which S is a proper subset of E.)

Solution. *Proof.* (by contradiction) Let S be the set of edges produced by ALG-GROW. By Lemma 5, there exists an MST $T = (V, E)$ for G such that $S \subseteq E$. If $S = E$, then ALG-GROW outputs the edges of the MST T.

We will show that the other case, $S \neq E$, leads to a contradiction. Suppose $S \neq E$. Then there exists an edge $e \in E - S$. This implies that $S \cup \{e\} \subseteq E$. Since E is the set of edges of a tree, $S \cup \{e\}$ does not contain a cycle. Therefore, e would be added to S by ALG-GROW. So $e \in S$, and this contradicts $e \in E - S$.

□

■

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Problems for Recitation 9

1 Traveling Salesperson Problem

Now we're going to talk about a famous optimization problem known as the Traveling Salesperson Problem[1] (TSP). Given a number of cities and the costs of traveling from any city to any other city, what is the cheapest round-trip route that visits each city exactly once, and returns to the starting city?

One special, though very natural, case of this problem is when the costs of traveling between cities obey the *triangle inequality*. That is, if a, b, and c are distinct cities, then $d(a, c) \leq d(a, b) + d(b, c)$. This corresponds to our intuitive notion that the distance of traveling to a from c should be no larger than that of first traveling to a from b, and then from b to c. The triangle inequality holds if the cities correspond to points in the plane, since in this case the line segments joining a to b, b to c, and a to c form a triangle, and in high school we learned that the sum of any two edges of a triangle is larger than the third edge.

Let us abstract away the distractions and formulate the problem as follows. The cities will be vertices of a graph. We will then consider the complete graph on these vertices, and give edge (a, b) a weight of $d(a, b)$. We want an ordering of the vertices $(v_1, \ldots, v_n)$ that minimizes

$$COST = \sum_{i=1}^{n-1} d(v_i, v_{i+1}) + d(v_n, v_1).$$

Consider the following greedy algorithm for TSP. This might be the first thing you'd think of. It is called greedy since at each step we are choosing the locally-optimal best way to continue, though our overall actions may not be collectively optimal.

1. Start at an arbitrary city.
2. While there is still an unvisited city, go to the city with smallest distance from the current city. If there are no more unvisited cities, return to the starting city.

Even when the cities are points in the plane, the greedy algorithm sometimes outputs a suboptimal solution.

[1] Note that this is sometimes referred to as the Traveling Salesman Problem, but we take a gender-neutral stance.

For example, take $4N$ points on the perimeter of an $N \times N$ square (each at distance one from the next along the perimeter) and add one point at distance 2 (outwards) from the perimeter at each of the corners (for a total of $4N + 4$ points). (To make it consistent with our abstraction, one would take these points on the plane and consider the complete graph with their pairwise distances as weights, but its easier to think of it directly in this case). The Greedy Algorithm will go around the perimeter and then have to go around again to pick up the four corners (which will be skipped the first time around). For large N, this will be nearly a factor of 2 from optimal (which will pick up the corners as it goes around the first time).

It turns out that if you can design an efficient algorithm which solves TSP on every possible input, you can win a million dollars. Many researchers believe that in fact this is impossible, though no proof of this fact is known (this proof would also get you a million dollars). Thus, researchers have contented themselves with finding a good *approximation algorithm* to TSP. More precisely, if OPT is the cost of the optimal solution (the total distance traveled on an optimal route), then an algorithm outputs an α-approximation to TSP provided its output, on any input, is a valid round-trip route visiting each city once, and has cost at most $\alpha \cdot OPT$.

In what follows, we will describe a 3/2-approximation algorithm ALG for TSP, in the case where the triangle inequality holds and the distances are non-negative.

3/2-Approximation Algorithm ALG

1. Construct a graph G whose vertices are the N cities with an edge between every pair of cities $A \neq B$ with corresponding weight $d(A, B)$, where $d(A, B)$ is the distance between A and B in the plane.

2. Compute an MST T of G (Recall that a *minimum spanning tree* (MST) of a graph G is a spanning tree whose sum of edge weights is as small as possible).

3. Compute, for each city $A \in G$, the degree d_A of A in T.

4. Let $S = \{A \in G : d_A \text{ is odd}\}$.

5. Compute a minimum weight perfect matching M on the vertices in S (using the distances $d(\cdot, \cdot)$ as weights).

6. Compute a new set of edges $E' = M \cup T$. Note that the resulting graph $G' = (V, E')$ is not necessarily a simple graph since it might contain multiple edges.

7. Take the subgraph $G' = (V, E')$, and compute an Euler circuit on it.

8. Use the Euler circuit to give an induced ordering of the vertices (i.e., the order in which the vertices appear for the first time), and do a TSP tour on this order.

Try to run this on the $4N + 4$ vertex example and see what it gives.

Note that this algorithm is reasonably efficient, even for large graphs, since computing an MST, computing a minimum weight perfect matching, and computing Euler circuits can be done efficiently. Since the Euler circuit visits all vertices and then returns to the starting vertex, the output of our algorithm visits every city once, and returns to the starting city, as needed. Now let's show that the algorithm is a 3/2-approximation.

Let OPT be the optimum TSP tour cost.

1. Show that the cost of any tour is at least the cost of an MST of G. Hence conclude that the cost of an MST is at most OPT.
2. Prove that the size of S is even.
3. Prove that the weight of the min weight perfect matching is at most $OPT/2$. *Hint: Consider a minimum cost TSP tour on just the vertices in S.*
4. Prove that G' has an Euler circuit, and its cost is at most $3OPT/2$.
5. Show that the length of the TSP tour is at most $3OPT/2$.
6. Conclude that our algorithm outputs a 3/2-approximation.

In fact, for any constant $\epsilon > 0$, it is known how to efficiently achieve a $(1+\epsilon)$-approximation in the plane, but it is a **lot** more complicated.

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Notes for Recitation 9

1 Traveling Salesperson Problem

Now we're going to talk about a famous optimization problem known as the Traveling Salesperson Problem[1] (TSP). Given a number of cities and the costs of traveling from any city to any other city, what is the cheapest round-trip route that visits each city exactly once, and returns to the starting city?

One special, though very natural, case of this problem is when the costs of traveling between cities obey the *triangle inequality*. That is, if a, b, and c are distinct cities, then $d(a, c) \leq d(a, b) + d(b, c)$. This corresponds to our intuitive notion that the distance of traveling to a from c should be no larger than that of first traveling to a from b, and then from b to c. The triangle inequality holds if the cities correspond to points in the plane, since in this case the line segments joining a to b, b to c, and a to c form a triangle, and in high school we learned that the sum of any two edges of a triangle is larger than the third edge.

Let us abstract away the distractions and formulate the problem as follows. The cities will be vertices of a graph. We will then consider the complete graph on these vertices, and give edge (a, b) a weight of $d(a, b)$. We want an ordering of the vertices $(v_1, \ldots, v_n)$ that minimizes

$$COST = \sum_{i=1}^{n-1} d(v_i, v_{i+1}) + d(v_n, v_1).$$

Consider the following greedy algorithm for TSP. This might be the first thing you'd think of. It is called greedy since at each step we are choosing the locally-optimal best way to continue, though our overall actions may not be collectively optimal.

1. Start at an arbitrary city.
2. While there is still an unvisited city, go to the city with smallest distance from the current city. If there are no more unvisited cities, return to the starting city.

Even when the cities are points in the plane, the greedy algorithm sometimes outputs a suboptimal solution.

[1] Note that this is sometimes referred to as the Traveling Salesman Problem, but we take a gender-neutral stance.

For example, take $4N$ points on the perimeter of an $N \times N$ square (each at distance one from the next along the perimeter) and add one point at distance 2 (outwards) from the perimeter at each of the corners (for a total of $4N + 4$ points). (To make it consistent with our abstraction, one would take these points on the plane and consider the complete graph with their pairwise distances as weights, but its easier to think of it directly in this case). The Greedy Algorithm will go around the perimeter and then have to go around again to pick up the four corners (which will be skipped the first time around). For large N, this will be nearly a factor of 2 from optimal (which will pick up the corners as it goes around the first time).

It turns out that if you can design an efficient algorithm which solves TSP on every possible input, you can win a million dollars. Many researchers believe that in fact this is impossible, though no proof of this fact is known (this proof would also get you a million dollars). Thus, researchers have contented themselves with finding a good *approximation algorithm* to TSP. More precisely, if OPT is the cost of the optimal solution (the total distance traveled on an optimal route), then an algorithm outputs an α-approximation to TSP provided its output, on any input, is a valid round-trip route visiting each city once, and has cost at most $\alpha \cdot OPT$.

In what follows, we will describe a 3/2-approximation algorithm ALG for TSP, in the case where the triangle inequality holds and the distances are non-negative.

3/2-Approximation Algorithm ALG

1. Construct a graph G whose vertices are the N cities with an edge between every pair of cities $A \neq B$ with corresponding weight $d(A, B)$, where $d(A, B)$ is the distance between A and B in the plane.

2. Compute an MST T of G (Recall that a *minimum spanning tree* (MST) of a graph G is a spanning tree whose sum of edge weights is as small as possible).

3. Compute, for each city $A \in G$, the degree d_A of A in T.

4. Let $S = \{A \in G : d_A \text{ is odd}\}$.

5. Compute a minimum weight perfect matching M on the vertices in S (using the distances $d(\cdot, \cdot)$ as weights).

6. Compute a new set of edges $E' = M \cup T$. Note that the resulting graph $G' = (V, E')$ is not necessarily a simple graph since it might contain multiple edges.

7. Take the subgraph $G' = (V, E')$, and compute an Euler circuit on it.

8. Use the Euler circuit to give an induced ordering of the vertices (i.e., the order in which the vertices appear for the first time), and do a TSP tour on this order.

Try to run this on the $4N + 4$ vertex example and see what it gives.

Note that this algorithm is reasonably efficient, even for large graphs, since computing an MST, computing a minimum weight perfect matching, and computing Euler circuits can be done efficiently. Since the Euler circuit visits all vertices and then returns to the starting vertex, the output of our algorithm visits every city once, and returns to the starting city, as needed. Now let's show that the algorithm is a 3/2-approximation.

Let OPT be the optimum TSP tour cost.

1. Show that the cost of any tour is at least the cost of an MST of G. Hence conclude that the cost of an MST is at most OPT.

   Solution. Consider any route of the traveler. It corresponds to a cycle C in G containing every vertex once, since every city must be visited and the traveler must end in the city in which they started. The cost of the route is then just the sum of the weights of edges along C. Now if we delete any edge of C, we are left with a spanning tree of G, and the cost cannot increase. Therefore, the cost of a minimum spanning tree (MST) is a lower bound on the cost of any route C. ■

2. Prove that the size of S is even.

   Solution. Since an edge connects two vertices, every edge added to a graph increases the total degree over all vertices by exactly two. Therefore, the total degree in any graph must be even. By definition, S contains vertices with odd degree. Since the total degree in G is equal to the sum of the total degree of even-degree vertices and odd-degree vertices, we have that the total degree of odd-degree vertices is even (since the total degree of even-degree vertices is even). The total degree of odd-degree vertices is the sum of $|S|$ odd numbers, so to have it be even requires that $|S|$ be even. ■

3. Prove that the weight of the min weight perfect matching is at most $OPT/2$. *Hint: Consider a minimum cost TSP tour on just the vertices in S.*

   Solution. We will use the following lemma.

   Lemma 1. *If the triangle inequality holds in a graph with positive weights, then $d(a, b)$ is at most the weight of any path from a to b.*

   Proof. We prove the lemma by induction. Let the inductive hypothesis $P(n)$ be the statement "If n is the length of a path p from a to b, for any vertices a and b, then $d(a, b)$ is at most the weight of p."

   This holds for the base case of $n = 1$, since the minimum cost of traveling from a to b cannot be strictly greater than the weight of the direct path from a to b (otherwise, we could simply take the direct path for a smaller cost).

   For the induction step, we assume that $P(n)$ holds for an arbitrary $n \geq 1$. Then, for $n + 1$, we have an $(n + 1)$-length path from a to b. Let b' be the vertex preceding b on

this path. By our inductive assumption, we know that $d(a, b')$ is at most the weight of this path. Then by the triangle inequality, we have $d(a, b) \leq d(a, b') + d(b', b)$, which is less than or equal to the weight of the path from a to b' added to the cost of the b'-to-b edge, which is the weight of the path from a to b. So the inequality holds for the $n+1$ case. □

Consider the optimal TSP tour on the entire graph G. This has cost OPT. This is also a tour on the vertices in S. "Shortcut" this tour by using the lemma: i.e., for every segment of the tour between vertices $s_1, s_2 \in S$ with no other vertices of S en route, substitute in the edge from s_1 to s_2. By the lemma, these substitutions can only decrease the length of the tour, and thus it now has cost at most OPT. Now, since there are an even number of vertices in the tour, one can consider 2 perfect matchings formed by taking alternate edges of this tour. Clearly, the sum of their costs is at most OPT, so one of them must have a cost of at most $OPT/2$. Thus the min weight perfect matching on S has weight at most $OPT/2$. ■

4. Prove that G' has an Euler circuit, and its cost is at most $3OPT/2$.

   Solution. Every vertex has even degree in G' (by choice of S and M). Thus it has an Euler circuit. Every edge of $T \cup M$ is traversed once, and so the total cost is $cost(T) + cost(M) \leq OPT + OPT/2 = 3OPT/2$. ■

5. Show that the length of the TSP tour is at most $3OPT/2$.

   Solution. The induced TSP tour from is a "shortcutting" of the Euler circuit (replacing paths by single edges), and hence, by the triangle inequality lemma from above, its length is at most the length of the Euler circuit. ■

6. Conclude that our algorithm outputs a 3/2-approximation.

   Solution. Indeed it does. ■

In fact, for any constant $\epsilon > 0$, it is known how to efficiently achieve a $(1+\epsilon)$-approximation in the plane, but it is a **lot** more complicated.

MIT OpenCourseWare
http://ocw.mit.edu

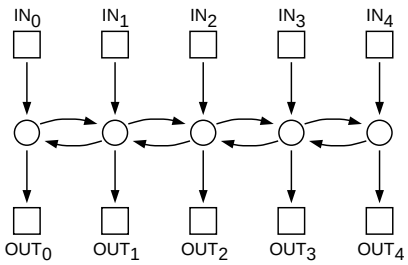
6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

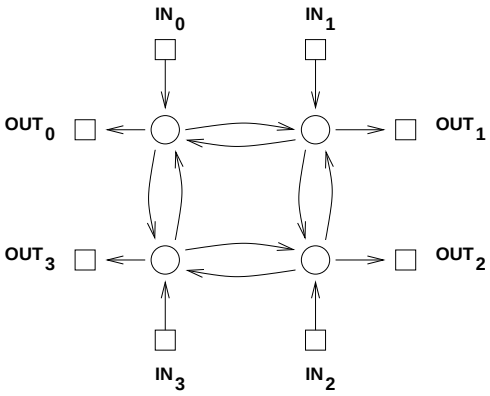
Problems for Recitation 10

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.

Routing in a Beneš Network

In lecture, we saw that the Beneš network has a max congestion of 1; that is, every permutation can be routed in such a way that a single packet passes through each switch. Let's work through an example. A Beneš network of size $N = 8$ is attached.

1. Within the Beneš network of size $N = 8$, there are two subnetworks of size $N = 4$. Put boxes around these. Hereafter, we'll refer to these as the *upper* and *lower* subnetworks.

2. Now consider the following permutation routing problem:

$$\begin{aligned} \pi(0) &= 3 & \qquad \pi(4) &= 2 \\ \pi(1) &= 1 & \qquad \pi(5) &= 0 \\ \pi(2) &= 6 & \qquad \pi(6) &= 7 \\ \pi(3) &= 5 & \qquad \pi(7) &= 4 \end{aligned}$$

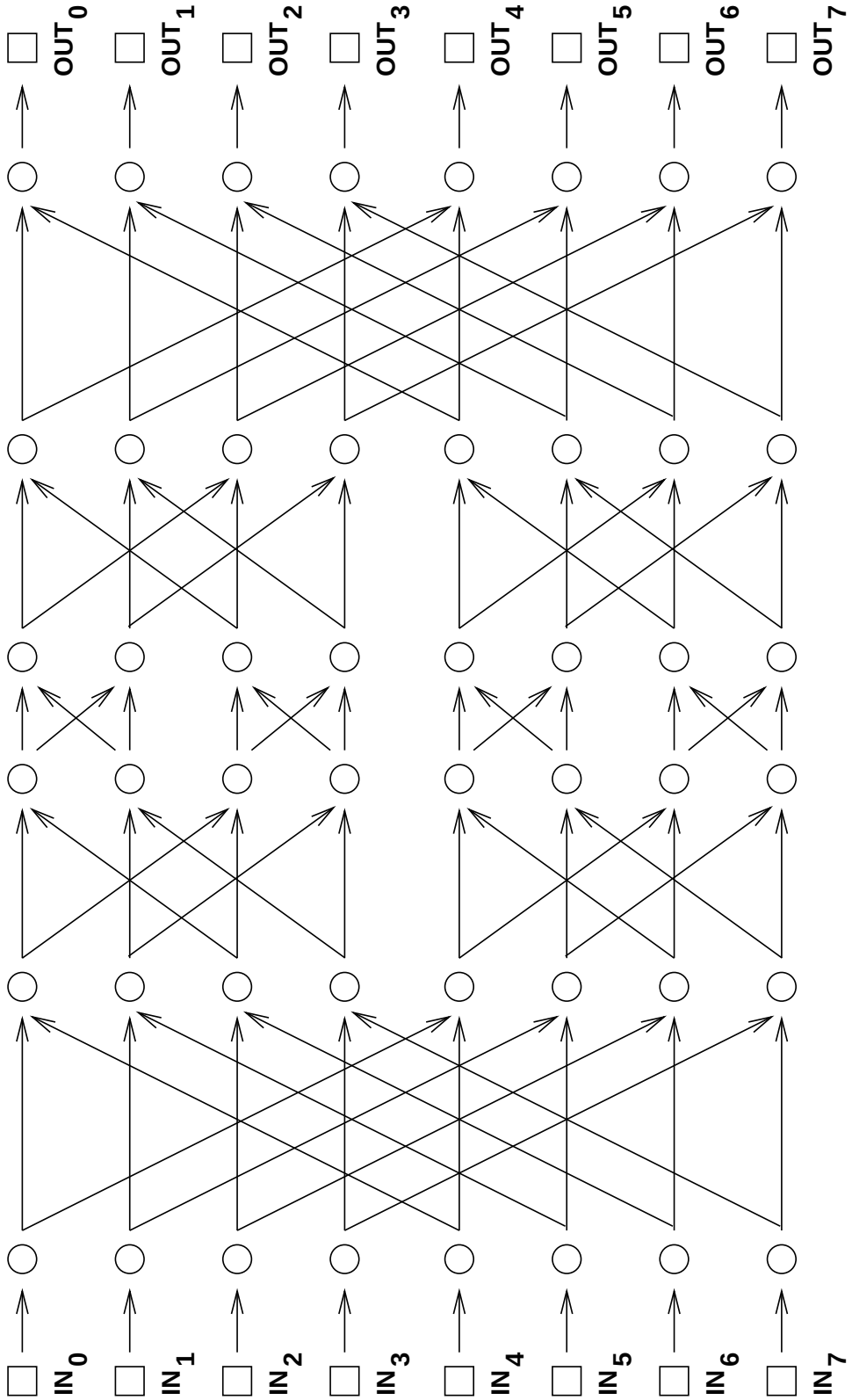
   Each packet must be routed through either the upper subnetwork or the lower subnetwork. Construct a graph with vertices 0, 1, ..., 7 and draw a *dashed* edge between each pair of packets that can not go through the same subnetwork because a collision would occur in the second column of switches.

3. Add a *solid* edge in your graph between each pair of packets that can not go through the same subnetwork because a collision would occur in the next-to-last column of switches.

4. Color (i.e., label) the vertices of your graph red and blue so that adjacent vertices get different colors. Why must this be possible, regardless of the permutation π?

5. Suppose that red vertices correspond to packets routed through the upper subnetwork and blue vertices correspond to packets routed through the lower subnetwork. On the attached copy of the Beneš network, highlight the first and last edge traversed by each packet.

6. All that remains is to route packets through the upper and lower subnetworks. One way to do this is by applying the procedure described above recursively on each subnetwork. However, since the remaining problems are small, see if you can complete all the paths on your own.


IN0
IN1
IN2
IN3
IN4
IN5
IN6
IN7
OUT0
OUT1
OUT2
OUT3
OUT4
OUT5
OUT6
OUT7

MIT OpenCourseWare
http://ocw.mit.edu

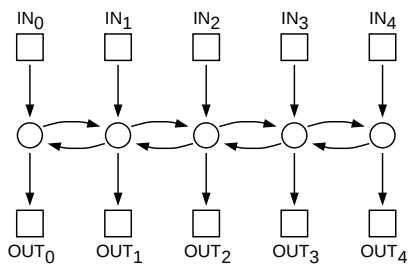
6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

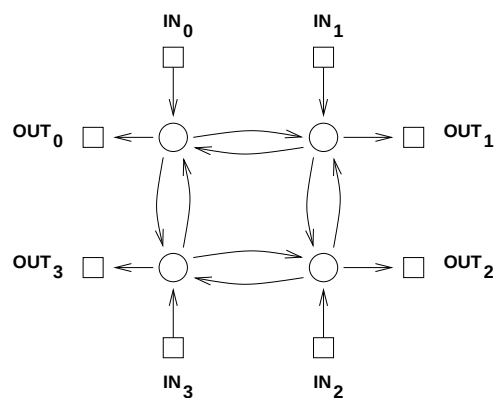
Notes for Recitation 10

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	5	3×3	6	5
4-cycle	4	3×3	4	3

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.

Solution. The congestion of the 5-path is at least 5, because every path must contain the central switch when $\pi(i) = 4 - i$. The congestion is at most 5, because there are only 5 paths in total.

The congestion of the 4-cycle is at least 3. Consider the permutation routing problem in which each input sends a packet to the diagonally opposite output: $\pi(0) = 2$, $\pi(1) = 3$,

$\pi(2) = 0$, $\pi(3) = 1$. Packets 0 and 2 must pass through the switches on the upper left and lower right in order to access the appropriate input and output terminals. Packet 1 must pass through one of these switches as well, so at least three packets pass through either the upper-left switch or the lower-left switch.

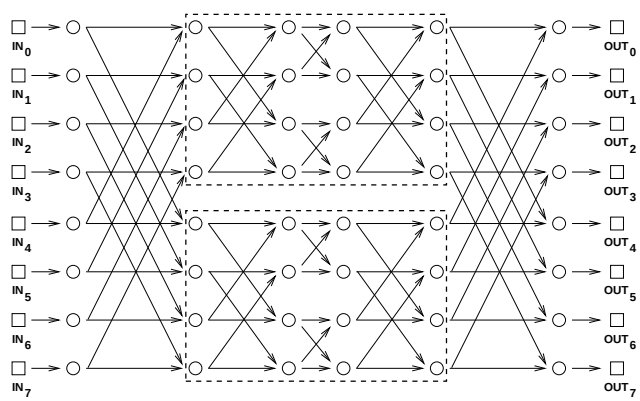
The congestion of the 4-cycle is at most 3. Suppose we route each packet by the shortest path and break ties by routing clockwise around the cycle. Now consider any particular switch, say the one in the upper right. At worst, this switch sees three packets: the one from input 1, the one destined for output 1, and one passing through from input 0 to output 2. By symmetry, every switch sees at most 3 packets. ■

Routing in a Beneš Network

In lecture, we saw that the Beneš network has a max congestion of 1; that is, every permutation can be routed in such a way that a single packet passes through each switch. Let's work through an example. A Beneš network of size $N = 8$ is attached.

1. Within the Beneš network of size $N = 8$, there are two subnetworks of size $N = 4$. Put boxes around these. Hereafter, we'll refer to these as the *upper* and *lower* subnetworks.

   Solution.



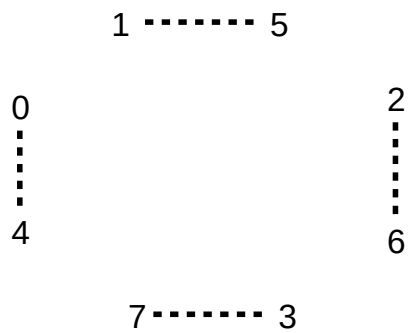
   ■

2. Now consider the following permutation routing problem:

$$\pi(0) = 3 \qquad \pi(4) = 2$$
$$\pi(1) = 1 \qquad \pi(5) = 0$$
$$\pi(2) = 6 \qquad \pi(6) = 7$$
$$\pi(3) = 5 \qquad \pi(7) = 4$$

   Each packet must be routed through either the upper subnetwork or the lower subnetwork. Construct a graph with vertices 0, 1, ..., 7 and draw a *dashed* edge between each pair of packets that can not go through the same subnetwork because a collision would occur in the second column of switches.

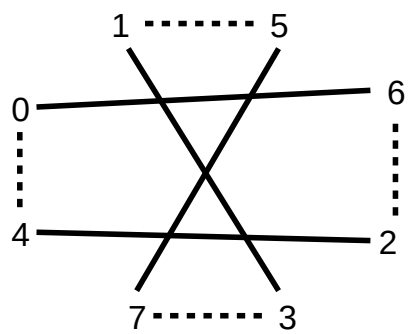
Solution.



■

3. Add a *solid* edge in your graph between each pair of packets that can not go through the same subnetwork because a collision would occur in the next-to-last column of switches.

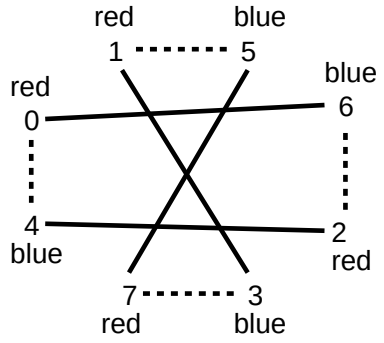
Solution.



■

4. Color (i.e., label) the vertices of your graph red and blue so that adjacent vertices get different colors. Why must this be possible, regardless of the permutation π?

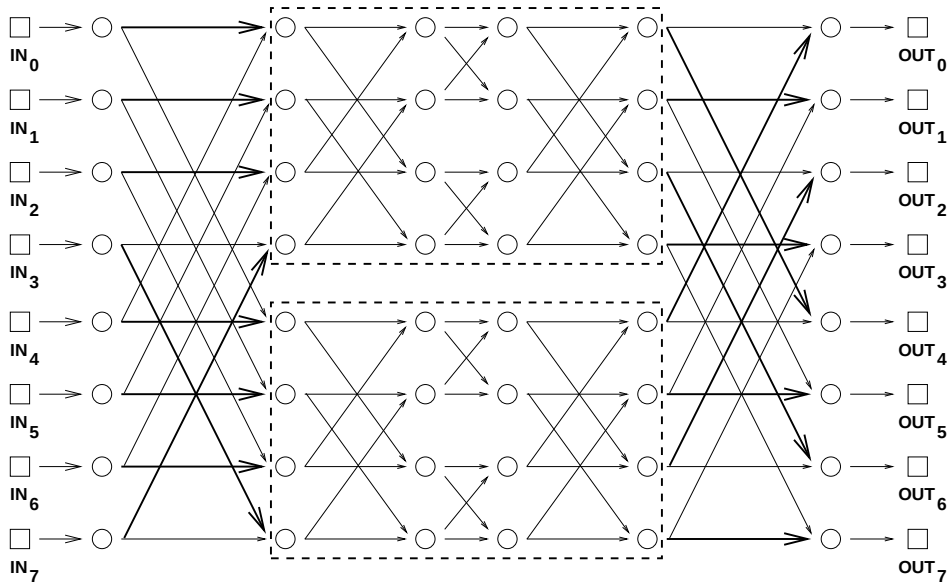
Solution. This must be possible, because the dashed edges form a matching and the solid edges form another matching. Because of the result you proved in homework, when you combine the edges, the result is a bipartite graph, which must be 2-colorable.



■

5. Suppose that red vertices correspond to packets routed through the upper subnetwork and blue vertices correspond to packets routed through the lower subnetwork. On the attached copy of the Beneš network, highlight the first and last edge traversed by each packet.

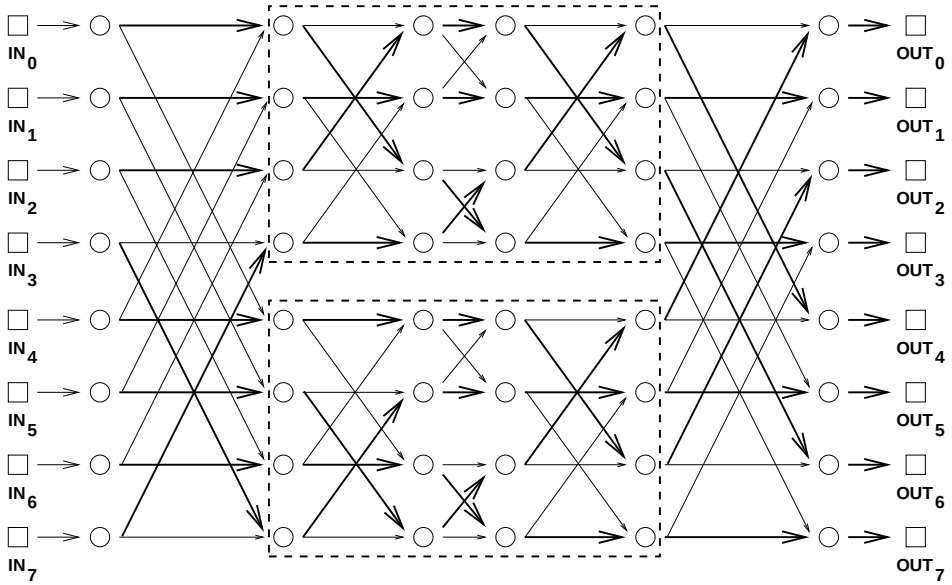
Solution.



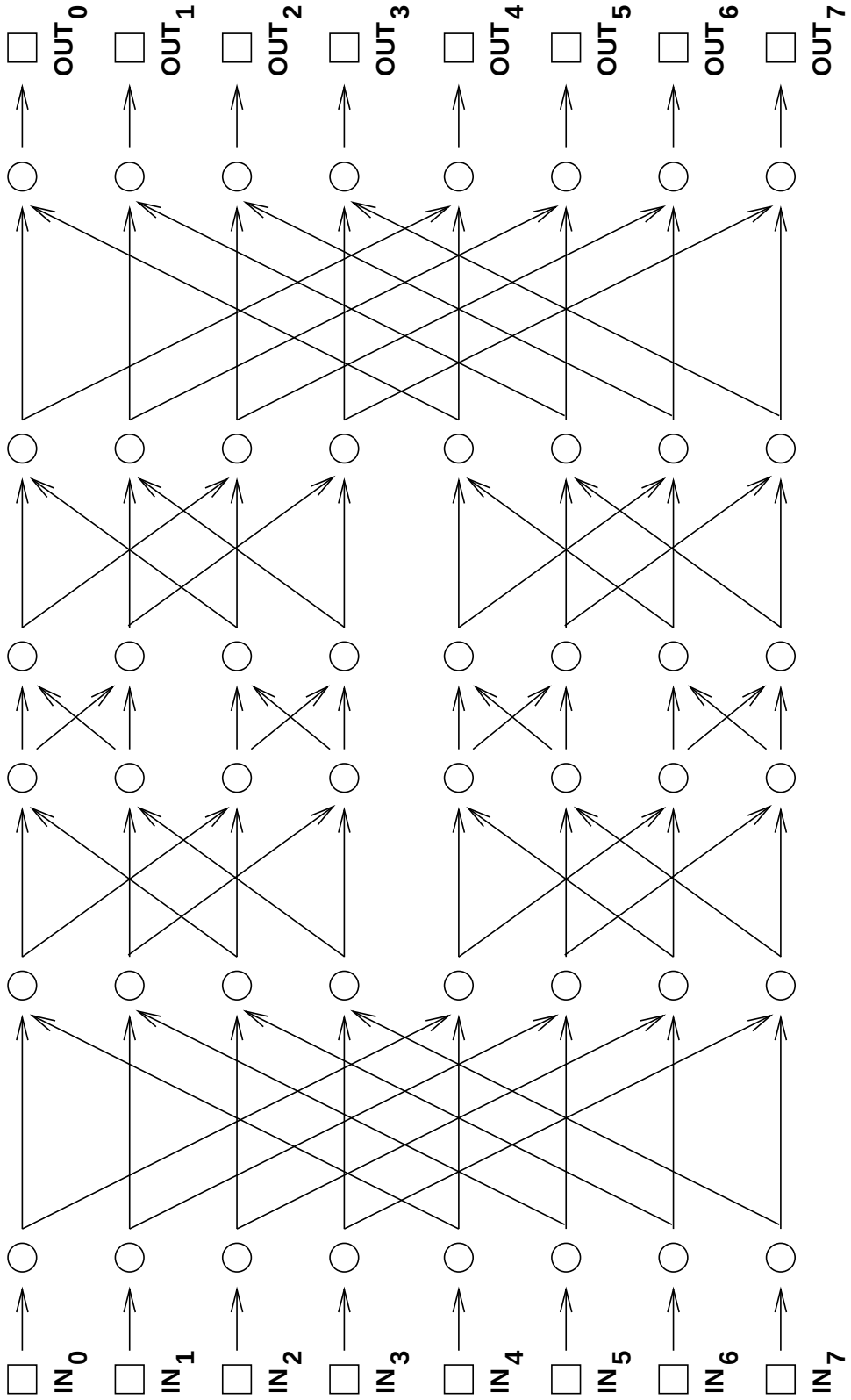
■

6. All that remains is to route packets through the upper and lower subnetworks. One way to do this is by applying the procedure described above recursively on each subnetwork. However, since the remaining problems are small, see if you can complete all the paths on your own.

Solution.


IN0
IN1
IN2
IN3
IN4
IN5
IN6
IN7
OUT0
OUT1
OUT2
OUT3
OUT4
OUT5
OUT6
OUT7

■


IN0
IN1
IN2
IN3
IN4
IN5
IN6
IN7
OUT0
OUT1
OUT2
OUT3
OUT4
OUT5
OUT6
OUT7

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Problems for Recitation 11

1. Give a description of the equivalence classes associated with each of the following equivalence relations.

   (a) Integers x and y are equivalent if $x \equiv y \pmod 3$.

   (b) Real numbers x and y are equivalent if $\lceil x \rceil = \lceil y \rceil$, where $\lceil z \rceil$ denotes the smallest integer greater than or equal to z.

2. Show that neither of the following relations is an equivalence relation by identifying a missing property (reflexivity, symmetry, or transitivity).

   (a) The "divides" relation on the positive integers.

   (b) The "implies" relation on propositional formulas.

3. Here is prerequistite information for some MIT courses:

$18.01 \to 6.042$	$18.01 \to 18.02$
$18.01 \to 18.03$	$6.046 \to 6.840$
$8.01 \to 8.02$	$6.01 \to 6.034$
$6.042 \to 6.046$	$18.03, 8.02 \to 6.02$
$6.01, 6.02 \to 6.003$	$6.01, 6.02 \to 6.004$
$6.004 \to 6.033$	$6.033 \to 6.857$

(a) Draw a Hasse diagram for the corresponding partially-ordered set. (A ***Hasse diagram*** is a way of representing a poset $(A, \preceq)$ as a directed acyclic graph. The vertices are the element of A, and there is generally an edge $u \to v$ if $u \preceq v$. However, self-loops and edges implied by transitivity are omitted.) You'll need this diagram for all the subsequent problem parts, so be neat!

(b) Identify a largest chain. (A ***chain*** in a poset $(S, \preceq)$ is a subset $C \subseteq S$ such that for all $x, y \in C$, either $x \preceq y$ or $y \preceq x$.)

(c) Suppose that you want to take all the courses. What is the minimum number of terms required to graduate, if you can take as many courses as you want per term?

(d) Identify a largest **antichain**. (An ***antichain*** in a poset $(S, \preceq)$ is a subset $A \subseteq S$ such that for all $x, y \in A$ with $x \neq y$, neither $x \preceq y$ nor $y \preceq x$.)

(e) What is the maximum number of classes that you could possibly take at once?

(f) Identify a topological sort of the classes. (A ***topological sort*** of a poset $(A, \preceq)$ is a total order of all the elements such that if $a_i \preceq a_j$ in the partial order, then a_i precedes a_j in the total order.)

(g) Suppose that you want to take all of the courses, but can handle only two per term. How many terms are required to graduate?

(h) What if you could take three courses per term?

(i) Stanford's computer science department offers n courses, limits students to at most k classes per term, and has its own complicated prerequisite structure. Describe two different lower bounds on the number of terms required to complete all the courses. One should be based on your answers to parts (b) and (c) and a second should be based on your answer to part (g).

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Notes for Recitation 11

1. Give a description of the equivalence classes associated with each of the following equivalence relations.

   (a) Integers x and y are equivalent if $x \equiv y \pmod 3$.

   Solution.

$$\{\ldots, -6, -3, 0, 3, 6, \ldots\}$$
$$\{\ldots, -5, -2, 1, 4, 7, \ldots\}$$
$$\{\ldots, -4, -1, 2, 5, 8, \ldots\}$$

■

   (b) Real numbers x and y are equivalent if $\lceil x \rceil = \lceil y \rceil$, where $\lceil z \rceil$ denotes the smallest integer greater than or equal to z.

   Solution. For each integer n, all the real numbers r such that $n - 1 < r \leq n$ form an equivalence class. ■

2. Show that neither of the following relations is an equivalence relation by identifying a missing property (reflexivity, symmetry, or transitivity).

   (a) The "divides" relation on the positive integers.

   Solution. This relation is reflexive (since $a \mid a$) and transitive (since $a \mid b$ and $b \mid c$ implies $a \mid c$), but not symmetric (since $3 \mid 6$, but not $6 \mid 3$). ■

   (b) The "implies" relation on propositional formulas.

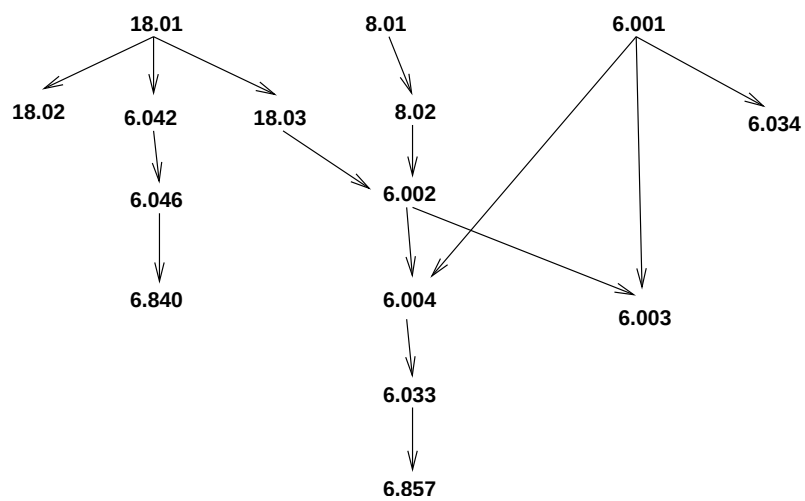
   Solution. This relation is reflexive since $p \Rightarrow p$. It is also transitive, since if $p \Rightarrow q$ and $q \Rightarrow r$, then $p \Rightarrow r$. However, it isn't symmetric since, for example, `false` $\Rightarrow$ `true`, but not `true` $\Rightarrow$ `false`. ■

3. Here is prerequistite information for some MIT courses:

$$
\begin{aligned}
18.01 &\to 6.042 & 18.01 &\to 18.02 \\
18.01 &\to 18.03 & 6.046 &\to 6.840 \\
8.01 &\to 8.02 & 6.01 &\to 6.034 \\
6.042 &\to 6.046 & 18.03, 8.02 &\to 6.02 \\
6.01, 6.02 &\to 6.003 & 6.01, 6.02 &\to 6.004 \\
6.004 &\to 6.033 & 6.033 &\to 6.857
\end{aligned}
$$

(a) Draw a Hasse diagram for the corresponding partially-ordered set. (A ***Hasse diagram*** is a way of representing a poset $(A, \preceq)$ as a directed acyclic graph. The vertices are the element of A, and there is generally an edge $u \to v$ if $u \preceq v$. However, self-loops and edges implied by transitivity are omitted.) You'll need this diagram for all the subsequent problem parts, so be neat!

Solution.



■

(b) Identify a largest chain. (A ***chain*** in a poset $(S, \preceq)$ is a subset $C \subseteq S$ such that for all $x, y \in C$, either $x \preceq y$ or $y \preceq x$.)

Solution. There are two largest chains:

$$8.01 \preceq 8.02 \preceq 6.02 \preceq 6.004 \preceq 6.033 \preceq 6.857$$

and

$$18.01 \preceq 18.03 \preceq 6.02 \preceq 6.004 \preceq 6.033 \preceq 6.857$$

■

(c) Suppose that you want to take all the courses. What is the minimum number of terms required to graduate, if you can take as many courses as you want per term?

Solution. Six terms are necessary, because at most one course in the longest chain can be taken each term. Six terms are sufficient by a theorem proved in lecture. ■

(d) Identify a largest **antichain**. (An ***antichain*** in a poset $(S, \preceq)$ is a subset $A \subseteq S$ such that for all $x, y \in A$ with $x \neq y$, neither $x \preceq y$ nor $y \preceq x$.)

Solution. There are several five-element antichains. One is:

$$\{18.02, 6.042, 18.03, 8.01, 6.01\}$$

■

(e) What is the maximum number of classes that you could possibly take at once?

Solution. Classes you are taking simultaneously must form an antichain, so you can take at most five at once. ■

(f) Identify a topological sort of the classes. (A ***topological sort*** of a poset $(A, \preceq)$ is a total order of all the elements such that if $a_i \preceq a_j$ in the partial order, then a_i precedes a_j in the total order.)

Solution. Many answers are possible. One is 18.01, 8.01, 6.01, 18.02, 6.042, 18.03, 8.02, 6.034, 6.046, 6.02, 6.840, 6.004, 6.003, 6.033, 6.857. ■

(g) Suppose that you want to take all of the courses, but can handle only two per term. How many terms are required to graduate?

Solution. There are 15 courses, so at least 8 terms are necessary. The schedule below shows that 8 terms are sufficient as well:

1:	18.01	8.01
2:	6.01	18.02
3:	6.042	18.03
4:	8.02	6.034
5:	6.046	6.02
6:	6.840	6.004
7:	6.003	6.033
8:	6.857	

■

(h) What if you could take three courses per term?

Solution. In part (c) we argued that six terms are required even if there is no limit on the number of courses per term. Six terms are also sufficient, as the following schedule shows:

1:	18.01	8.01	6.01
2:	6.042	18.03	8.02
3:	18.02	6.046	6.02
4:	6.004	6.003	6.034
5:	6.840	6.033	
6:	6.857		

■

(i) Stanford's computer science department offers n courses, limits students to at most k classes per term, and has its own complicated prerequisite structure. Describe two different lower bounds on the number of terms required to complete all the courses. One should be based on your answers to parts (b) and (c) and a second should be based on your answer to part (g).

Solution. One lower bound is the length of the longest chain and another is $\lceil n/k \rceil$. ■

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Problems for Recitation 12

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 $\frac{3x-3}{2} \leq k \leq \frac{3x-1}{2}$.

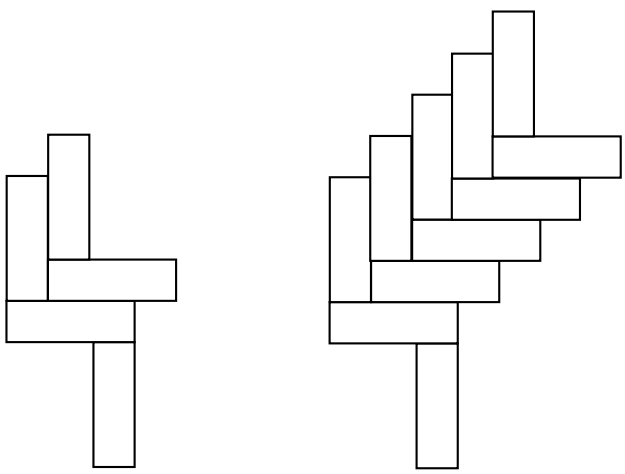


Figure 1: Too few or too many L shapes make the tower unstable

2 Double Sums

Sometimes we have to evaluate sums of sums, otherwise known as *double summations*. It's good to know how to tame these beasts! Here's an example of a double summation:

$$\sum_{i=1}^{n} \sum_{j=1}^{i} j$$

It looks ferocious...all those sharp teeth! But actually, this double summation is just a sheep in wolf's clothing: to evaluate it, we can just evaluate the inner sum, replace it with a closed form we already know, and then evaluate the outer sum which no longer has a summation inside it.

- Evaluate the summation. (*Hint:* $\sum(a+b) = \sum a + \sum b$.)

Unfortunately, not all summations are so docile. Fortunately, we have ways to deal with this. There's a special trick that is often extremely useful for sums, and that is to *exchange the order of summation.* We'll go through an example here.

For the remainder of the problem we'll wrestle with the sum of the harmonic numbers:

$$\sum_{k=1}^{n} H_k$$

At first glance, it looks like just a single summation, but do not be deceived.

- First, write it as a double summation.

- Now try to gain some intuition for exactly what you're up against by integrating the summation in its less threatening single-summation form. You may use $H_k \approx \ln k$.

- Finally, we'll look for an exact answer. 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. The first two rows have been filled in for you. Complete the remaining three rows to see the pattern.

j	1	2	3	4	...	n
k						
1	1					
2	1	1/2				
3						
4						
	...					
n					...	

- 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.

- Now simplify the summation to derive a closed form in terms of n and H_n.

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Notes for Recitation 12

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 $\frac{3x-3}{2} \leq k \leq \frac{3x-1}{2}$.

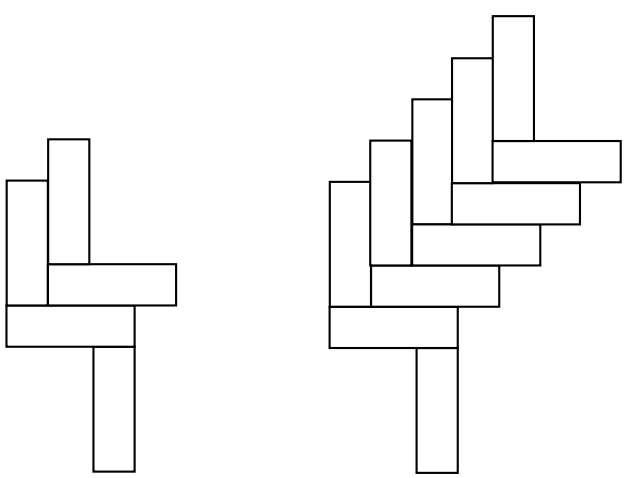


Figure 1: Too few or too many L shapes make the tower unstable

Solution. As indicated in the description, an arbitrary structure is considered stable if and only if its center of gravity lies on top of some form of support. For our L-towers this implies several conditions that must hold simultaneously. Consider the case $k = 3$,

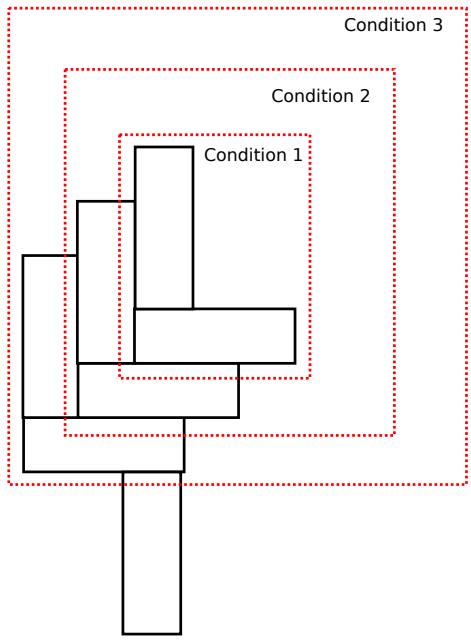


Figure 2: Three necessary conditions for stability, when $k = 3$. Put together, we also consider them sufficient.

illustrated in Figure 2. For the structure to be stable, we need the topmost L-shape to be stable on top of the second L-shape. This is Condition 1. We also need the structure formed by the topmost 2 L-shapes to be overall stable on top of the lowest L-shape. This is Condition 2. And finally, we need the three L-shapes to be stable on top of the single standing block. Call this Condition 3. In general, if there are k L-shapes, there are k conditions that must be satisfied for the structure to be stable. We will explicitly describe these constraints below.

Throughout the solution we are only concerned with the x-coordinates of the center of mass. The center of mass of a set of L-shapes can be computed by partitioning the figure into individual L shapes, and then averaging their individual centers.

Let c_1 be the position of the center of mass for an individual L-shape starts at 0. The vertical block has its center at $1/2$, the horizontal block at $x/2$. Hence, the center of a single L-shape is $(1/2 + x/2)/2 = (x + 1)/4$ to the right of where it starts.

Now let c_h be the position for the center of mass of h L-shapes stacked together, where we have shifted our reference frame so that the structure starts at 0. This new center of mass can be obtained from averaging the center of a single L shaped shifted by 1 every time. So we get $c_h = [c_1 + (c_1 + 1) + \cdots + (c_1 + (h - 1))]/h$. Regrouping, we get $c_h =$

$[hc_1 + (1 + 2 + \cdots + (h-1))]/h = [hc_1 + h(h-1)/2]/h$, from where we arrive to: $c_h = c_1 + (h-1)/2$. It is helpful to think visually about this equation. When $h = 1$, we naturally get $c_h = c_1$. As h increases, the center shifts by $1/2$ to the right every time.

Now we can write explicitly all the conditions for a structure with k L-shapes to be stable. We need every structure to stand stable on top of the next L. If we imagine we have a stack of i L-shapes starting at 0 on top of another L-shape starting at -1, we require:

$$c_i \leq x - 1$$

This simply means that if we have a stack of i L-shapes starting right on the origin, we need its center of gravity to fall within the next L. The next L ends at $x - 1$.

If we have k L-shapes, this constraint applies to the first $k - 1$ stacks. So we get the conditions:

$$\begin{gathered} c_1 \leq x - 1 \\ c_2 \leq x - 1 \\ \vdots \\ c_{k-1} \leq x - 1 \end{gathered}$$

Finally, the overall stack of k L-shapes has a slightly different requirement, since it stands on top of the vertical block:

$$x - 1 \leq c_k \leq x$$

Now we need to find whether some k satisfies all of these constraints. Recall $c_{k-1} > c_{k-2} > \cdots > c_1$, so we only need to check the last two constraints $c_{k-1} \leq x - 1$ and $x - 1 \leq c_k \leq x$. Substituting $c_k = (x+1)/4 + (k-1)/2$ and $c_{k-1} = c_k - 1/2 = (x+1)/4 + (k-1)/2 - 1/2$, we get the inequalities:

$$\begin{gathered} (x+1)/4 + (k-1)/2 \leq x - 1/2 \\ x - 1 \leq (x+1)/4 + (k-1)/2 \leq x \end{gathered}$$

The first inequality is stronger on the right than the second, so consolidating both we get:

$$x - 1 \leq (x+1)/4 + (k-1)/2 \leq x - 1/2$$

which solved for k yields:

$$\frac{3x-3}{2} \leq k \leq \frac{3x-1}{2}$$

There is a gap between the upper and lower bound of exactly 1. Hence either there is an integer right between them, or both are integers. For $(3x-3)/2$ to be an integer we need x to be an odd integer. Hence in the case where x is an odd integer we have two choices of k but on each the structure will be barely stable.

■

2 Double Sums

Sometimes we have to evaluate sums of sums, otherwise known as *double summations*. It's good to know how to tame these beasts! Here's an example of a double summation:

$$\sum_{i=1}^{n}\sum_{j=1}^{i} j$$

It looks ferocious...all those sharp teeth! But actually, this double summation is just a sheep in wolf's clothing: to evaluate it, we can just evaluate the inner sum, replace it with a closed form we already know, and then evaluate the outer sum which no longer has a summation inside it.

- Evaluate the summation. (*Hint:* $\sum(a+b) = \sum a + \sum b$.)

  Solution.

$$\begin{aligned}
\sum_{i=1}^{n}\sum_{j=1}^{i} j &= \sum_{i=1}^{n}\frac{i(i+1)}{2} \\
&= 1/2\sum_{i=1}^{n}(i^2+i) \\
&= 1/2(\sum_{i=1}^{n} i^2 + \sum_{i=1}^{n} i) \\
&= 1/2(\frac{n^3}{3}+\frac{n^2}{2}+\frac{n}{6}+\frac{n(n+1)}{2}) \\
&= 1/2(\frac{n^3+3n^2+2n}{3})
\end{aligned}$$

■

Unfortunately, not all summations are so docile. Fortunately, we have ways to deal with this. There's a special trick that is often extremely useful for sums, and that is to *exchange the order of summation.* We'll go through an example here.

For the remainder of the problem we'll wrestle with the sum of the harmonic numbers:

$$\sum_{k=1}^{n} H_k$$

At first glance, it looks like just a single summation, but do not be deceived.

- First, write it as a double summation.

Solution.

$$\sum_{k=1}^{n} H_k = \sum_{k=1}^{n} \sum_{j=1}^{k} 1/j$$

■

- Now try to gain some intuition for exactly what you're up against by integrating the summation in its less threatening single-summation form. You may use $H_k \approx \ln k$.

Solution.

$$\sum_{k=1}^{n} H_k \approx \int_{k=1}^{n} \ln n = n \ln n - n + 1.$$

■

- Finally, we'll look for an exact answer. 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. The first two rows have been filled in for you. Complete the remaining three rows to see the pattern.

j	1	2	3	4	...	n
k						
1	1					
2	1	1/2				
3						
4						
	...					
n					...	

Solution.

j	1	2	3	4	...	n
k						
1	1					
2	1	1/2				
3	1	1/2	1/3			
4	1	1/2	1/3	1/4		
	...					
n	1	1/2	1/3	1/4	...	$1/n$

■

- 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.

Solution.

$$\begin{aligned}\sum_{k=1}^{n} H_k &= \sum_{k=1}^{n}\sum_{j=1}^{k} 1/j \\ &= \sum_{j=1}^{n}\sum_{k=j}^{n} 1/j\end{aligned}$$

■

- Now simplify the summation to derive a closed form in terms of n and H_n.

Solution.

$$\begin{aligned}\sum_{k=1}^{n} H_k &= \sum_{j=1}^{n}\sum_{k=j}^{n} 1/j \\ &= \sum_{j=1}^{n} 1/j \sum_{k=j}^{n} 1 \\ &= \sum_{j=1}^{n} \frac{1}{j}(n-j+1) \\ &= \sum_{j=1}^{n} \frac{n-j+1}{j} \\ &= \sum_{j=1}^{n} \frac{n+1}{j} - \sum_{j=1}^{n} \frac{j}{j} \\ &= (n+1)\sum_{j=1}^{n} \frac{1}{j} - \sum_{j=1}^{n} 1 \\ &= (n+1)H_n - n\end{aligned}$$

Notice that the exact solution is very similar in form to the approximation we generated above. ■

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Problems for Recitation 13

1 Asymptotic Notation

Which of these symbols

$$\Theta \qquad O \qquad \Omega \qquad o \qquad \omega$$

can go in these boxes? (List all that apply.)

$$2n + \log n \quad = \quad \boxed{}\ (n)$$

$$\log n \quad = \quad \boxed{}\ (n)$$

$$\sqrt{n} \quad = \quad \boxed{}\ (\log^{300} n)$$

$$n2^n \quad = \quad \boxed{}\ (n)$$

$$n^7 \quad = \quad \boxed{}\ (1.01^n)$$

2 Asymptotic Equivalence

Suppose $f, g : \mathbb{Z}^+ \to \mathbb{Z}^+$ and $f \sim g$.

1. Prove that $2f \sim 2g$.
2. Prove that $f^2 \sim g^2$.
3. Give examples of f and g such that $2^f \not\sim 2^g$.
4. Show that $\sim$ is an equivalence relation
5. Show that Θ is an equivalence relation

3 More Asymptotic Notation

1. Show that
$$(an)^{b/n} \sim 1.$$
where a, b are positive constants and $\sim$ denotes asymptotic equality. Hint $an = a2^{\log_2 n}$.
2. You may assume that if $f(n) \geq 1$ and $g(n) \geq 1$ for all n, then $f \sim g \Rightarrow f^{\frac{1}{n}} \sim g^{\frac{1}{n}}$. Show that
$$\sqrt[n]{n!} = \Theta(n).$$

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Notes for Recitation 13

1 Asymptotic Notation

Which of these symbols

$$\Theta \qquad O \qquad \Omega \qquad o \qquad \omega$$

can go in these boxes? (List all that apply.)

$$2n + \log n \quad = \quad \boxed{}\ (n)$$

$$\Theta, O, \Omega$$

$$\log n \quad = \quad \boxed{}\ (n)$$

$$O, o$$

$$\sqrt{n} \quad = \quad \boxed{}\ (\log^{300} n)$$

$$\Omega, \omega$$

$$n2^n \quad = \quad \boxed{}\ (n)$$

$$\Omega, \omega$$

$$n^7 \quad = \quad \boxed{}\ (1.01^n)$$

$$O, o$$

2 Asymptotic Equivalence

Suppose $f, g : \mathbb{Z}^+ \to \mathbb{Z}^+$ and $f \sim g$.

1. Prove that $2f \sim 2g$.

   Solution.
   $$\frac{2f}{2g} = \frac{f}{g},$$
   so they have the same limit as $n \to infty$. ■

2. Prove that $f^2 \sim g^2$.

   Solution.
   $$\lim_{n\to\infty} \frac{f(n)^2}{g(n)^2} = \lim_{n\to\infty} \frac{f(n)}{g(n)} \cdot \frac{f(n)}{g(n)} = \lim_{n\to\infty} \frac{f(n)}{g(n)} \cdot \lim_{n\to\infty} \frac{f(n)}{g(n)} = 1 \cdot 1 = 1.$$
   ■

3. Give examples of f and g such that $2^f \not\sim 2^g$.

   Solution.
   $$f(n) = n + 1$$
   $$g(n) = n.$$
   Then $f \sim g$ since $\lim(n+1)/n = 1$, but $2^f = 2^{n+1} = 2 \cdot 2^n = 2 \cdot 2^g$ so
   $$\lim \frac{2^f}{2^g} = 2 \neq 1.$$
   ■

4. Show that $\sim$ is an equivalence relation

   Solution. (a) Reflexive: $f \sim f$ since $f(x)/f(x) = 1$ for all x (assuming $f(x) \neq 0$), so $\lim_{x\to\infty} f(x)/f(x) = 1$

   (b) Symmetric: $f \sim g$ implies $g \sim f$ since if $\lim_{x\to\infty} f(x)/g(x) = 1$, then by the laws of limits $lim_{x\to\infty} g(x)/f(x) = 1$

   (c) Transitive: $f \sim g$ and $g \sim h$ implies $f \sim h$: if $\lim_{x\to\infty} f(x)/g(x) = 1$, and $\lim_{x\to\infty} g(x)/h(x) = 1$, then multiplying the limits we get
   $$\lim_{x\to\infty} f(x)/h(x) = \lim_{x\to\infty} \frac{f(x)}{g(x)} \times \frac{g(x)}{h(x)} = 1$$
   ■

5. Show that Θ is an equivalence relation

Solution. (a) Reflexive: $\lim_{x\to\infty} f(x)/f(x) = 1 < \infty$, trivial.

(b) Symmetric: If $f = \Theta(g)$, we wish to show $g = \Theta(f)$. From the definiton: $\lim_{x\to\infty} f(x)/g(x) = c$ for some non-zero finite constant c. Hence $\lim_{x\to\infty} g(x)/f(x) = 1/c$. Also a non-zero finite constant, so $g = \Theta(f)$.

(c) Transitive: Want to show $f = \Theta(g), g = \Theta(h)$ then $f = \Theta(h)$. Let $\lim_{x\to\infty} f(x)/g(x) = c_1$ and $\lim_{x\to\infty} g(x)/h(x) = c_2$. Then $\lim_{x\to\infty} f(x)/h(x) = \lim_{x\to\infty} f(x)/g(x) \times g(x)/h(x) = c_1 \times c_2$. Since both c_1 and c_2 are non-zero and finite, so is $c_1 \times c_2$.

■

3 More Asymptotic Notation

1. Show that

$$(an)^{b/n} \sim 1.$$

where a, b are positive constants and $\sim$ denotes asymptotic equality. Hint $an = a2^{\log_2 n}$.

Solution.

$$(an)^{b/n} = \left(a^b\right)^{1/n} \cdot 2^{(b\log_2 n)/n} \to 1 \cdot 2^0 = 1,$$

as $n \to \infty$. ■

2. You may assume that if $f(n) \geq 1$ and $g(n) \geq 1$ for all n, then $f \sim g \Rightarrow f^{\frac{1}{n}} \sim g^{\frac{1}{n}}$. Show that

$$\sqrt[n]{n!} = \Theta(n).$$

Solution.

$$\begin{aligned}
\sqrt[n]{n!} &\sim \left((2\pi n)^{\frac{1}{2}} \left(\frac{n}{e}\right)^n\right)^{1/n} && \text{(Stirling)} \\
&\sim (2\pi n)^{\frac{1}{2n}} \frac{n}{e} \\
&\sim 1 \cdot \frac{n}{e} && \text{part (a)} \\
&= \Theta(n)
\end{aligned}$$

■

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Problems for Recitation 14

1 TriMergeSort

We noted in lecture that reducing the size of subproblems is much more important to the speed of an algorithm than reducing the number of additional steps per call. Let's see if we can improve the $\Theta(n \log n)$ bound on `MergeSort` from lecture.

Let's consider a new version of `MergeSort` called `TriMergeSort`, where the size n list is now broken into *three* sublists of size $n/3$, which are sorted recursively and then merged. Since we know that floors and ceilings do not affect the asymptotic solution to a recurrence, let's assume that n is a power of 3.

1. How many comparisons are needed to merge three lists of 1 item each?

2. In the worst case, how many comparisons are needed to merge three lists of $n/3$ items, where n is a power of 3?

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?

Appendix

Theorem 1 (Akra-Bazzi, strong form)**.** *Suppose that:*

$$T(x) = \begin{cases} \textit{is defined} & \textit{for } 0 \leq x \leq x_0 \\ \displaystyle\sum_{i=1}^{k} a_i T(b_i x + h_i(x)) + g(x) & \textit{for } x > x_0 \end{cases}$$

where:

- $a_1, \ldots, a_k$ *are positive constants*
- $b_1, \ldots, b_k$ *are constants between 0 and 1*
- x_0 *is "large enough" in a technical sense we leave unspecified*
- $|g'(x)| = O(x^c)$ *for some* $c \in \mathbb{N}$
- $|h_i(x)| = O(x/\log^2 x)$

Then:

$$T(x) = \Theta\left(x^p \left(1 + \int_1^x \frac{g(u)}{u^{p+1}}\, du\right)\right)$$

where p *satisfies the equation* $\sum_{i=1}^{k} a_i b_i^p = 1$.

Linear Recurrences

Find closed-form solutions to the following linear recurrences.

1. $$\begin{aligned} T_0 &= 0 \\ T_1 &= 1 \\ T_n &= T_{n-1} + T_{n-2} + 1 \end{aligned}$$

2. $$\begin{aligned} S_0 &= 0 \\ S_1 &= 1 \\ S_n &= 6S_{n-1} - 9S_{n-2} \end{aligned}$$

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Notes for Recitation 14

1 The Akra-Bazzi Theorem

Theorem 1 (Akra-Bazzi, strong form)**.** *Suppose that:*

$$T(x) = \begin{cases} \textit{is defined} & \textit{for } 0 \leq x \leq x_0 \\ \displaystyle\sum_{i=1}^{k} a_i T(b_i x + h_i(x)) + g(x) & \textit{for } x > x_0 \end{cases}$$

where:

- $a_1, \ldots, a_k$ *are positive constants*
- $b_1, \ldots, b_k$ *are constants between 0 and 1*
- x_0 *is "large enough" in a technical sense we leave unspecified*
- $|g'(x)| = O(x^c)$ *for some* $c \in \mathbb{N}$
- $|h_i(x)| = O(x/\log^2 x)$

Then:

$$T(x) = \Theta\left(x^p \left(1 + \int_1^x \frac{g(u)}{u^{p+1}}\, du\right)\right)$$

where p *satisfies the equation* $\sum_{i=1}^{k} a_i b_i^p = 1$.

The only difference between the strong and weak forms of Akra-Bazzi is the appearance of this $h_i(x)$ term in the recurrence, where $h_i(x)$ represents a small change in the size of the subproblems ($O(x/\log^2 x)$). Notice that, despite the change in the recurrence, the solution $T(x)$ remains the same in both the strong and weak forms, with no dependence on $h_i(x)$! In algorithmic terms, this means that *small* changes in the size of subproblems have no impact on the asymptotic running time.

Example: Let's compare the Θ bounds for the following divide-and-conquer recurrences.

$$T_a(n) = 3T\left(\frac{n}{3}\right) + n \qquad\qquad T_b(n) = 3T\left(\left\lfloor \frac{n}{3} \right\rfloor\right) + n$$

For $T_a(n)$ we have $a_1 = 3$, $b_1 = 1/3$, $g(n) = n$, $p = 1$.

For $T_b(n)$ we have the same parameters as for $T_a(n)$, plus $h_1(n) = \lfloor n/3 \rfloor - n/3$.

Using the strong Akra-Bazzi form, the $h_1(n)$ falls out of the equation:

$$T_a(n) = T_b(n) = \Theta(n(1 + \int_1^n \frac{u}{u^2} du)) = \Theta(n \log n).$$

The addition of the ceiling operator changes the value of $n/3$ by at most 1, which is easily $O(n/\log^2 n)$. So floor and ceiling operators have no impact on the asymptotic solution to a recurrence.

2 TriMergeSort

We noted in lecture that reducing the size of subproblems is much more important to the speed of an algorithm than reducing the number of additional steps per call. Let's see if we can improve the $\Theta(n \log n)$ bound on `MergeSort` from lecture.

Let's consider a new version of `MergeSort` called `TriMergeSort`, where the size n list is now broken into *three* sublists of size $n/3$, which are sorted recursively and then merged. Since we know that floors and ceilings do not affect the asymptotic solution to a recurrence, let's assume that n is a power of 3.

1. How many comparisons are needed to merge three lists of 1 item each?

   Solution. 3. For example, a merge of lists {4},{5}, and {2} compares 4 with 5 and 4 with 2, adding 2 to the final list. Then it compares 4 with 5 and adds 4 to the final list. Finally, it appends the remaining 5. (This could be made more efficient, but let's not worry about that here.) ■

2. In the worst case, how many comparisons are needed to merge three lists of $n/3$ items, where n is a power of 3?

   Solution. $2(n-2)+1$. The worst case occurs if the first list empties when there is exactly 1 item in each of the other two. Prior to this, each of the other $n-2$ numbers requires 2 comparisons before going into the big list. After this, we only need 1 more comparison between the 2 leftover items. ■

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.

   Solution. $T(n) = 3T(n/3) + 2n - 3$. ■

4. We could analyze the running time of this using plug-and-chug, but let's try Akra-Bazzi. First, what is p?

Solution. $p = 1$. Using $\sum_{i=1}^{k} a_i b_i^p = 1$ with $a_1 = 3$ and $b_1 = 1/3$, we get the constraint that $3(1/3)^p = 1$. ■

5. Does the condition $|g'(x)| = O(x^c)$ hold for some $c \in N$?

   Solution. Yes. $g(n) = 2n - 3$, so $|g'(n)| = 2 = O(n^c)$ for $c = 0$. ■

6. Determine the theta bound on T(n) by integration.

   Solution.

$$\begin{aligned}
T(n) =& \Theta\left(n^p\left(1 + \int_1^n \frac{g(u)}{u^{p+1}}\, du\right)\right) \\
=& \Theta\left(n\left(1 + \int_1^n \frac{2u-3}{u^2}\, du\right)\right) \\
=& \Theta\left(n\left(1 + \int_1^n \frac{2}{u} - \int_1^n \frac{3}{u^2}\, du\right)\right) \\
=& \Theta\left(n\left(1 + 2\log u\Big|_1^n + \frac{3}{u}\Big|_1^n\right)\right) \\
=& \Theta\left(n\left(1 + 2\log n + \frac{3}{n} - 3\right)\right) \\
=& \Theta\left(2n\log n - 2n + 3\right) \\
=& \Theta\left(n\log n\right)
\end{aligned}$$

■

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?

   Solution. Given a constant size partition, the recurrence will always be in the form:

$$T(n) = cT(n/c) + \left[(c-1)(n-(c-1)) + \sum_{i=1}^{c-2} i\right]$$

   The first term creates the constraint that $c(1/c)^p = 1$, which always gives $p = 1$. The second term will always be $\Theta(n)$, which dominates the final bound after integration. Therefore, no matter what $c > 1$ we choose, $T(n) = \Theta(n \log n)$. ■

Guessing a Particular Solution

A general linear recurrence has the form:

$$f(n) = b_1 f(n-1) + b_2 f(n-2) + \ldots + b_d f(n-d) + g(n)$$

One step in solving this recurrence is finding a *particular solution*. This is a function $f(n)$ that satisfies the recurrence equation, but may not be consistent with the boundary conditions. Here's a recipe to help you guess a particular solution:

- If $g(n)$ is a constant, guess that $f(n)$ is some constant c. Plug this into the recurrence equation and see if any constant actually works. If not, try $f(n) = bn + c$, then $f(n) = an^2 + bn + c$, etc.
- More generally, if $g(n)$ is a polynomial, try a polynomial of the same degree. If that fails, try 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 $an^2 3^n + bn3^n + c3^n$, etc.

In practice, your first or second guess will almost always work.

Linear Recurrences

Find closed-form solutions to the following linear recurrences.

1. $$\begin{aligned} T_0 &= 0 \\ T_1 &= 1 \\ T_n &= T_{n-1} + T_{n-2} + 1 \end{aligned}$$

Solution. Following the guide to solving linear recurrences:

(a) First, we find the general solution to the homogeneous recurrence. The characteristic equation is $r^2 - r - 1 = 0$. The roots of this equation are:

$$\begin{aligned} r_1 &= \frac{1+\sqrt{5}}{2} \\ r_2 &= \frac{1-\sqrt{5}}{2} \end{aligned}$$

(b) Using the roots, we write down the homogeneous solution in the form

$$T_n = A\left(\frac{1+\sqrt{5}}{2}\right)^n + B\left(\frac{1-\sqrt{5}}{2}\right)^n.$$

(c) Next, we need a particular solution to the inhomogenous recurrence. Since the inhomogenous term is constant, we guess a constant solution, c. So replacing the $T(n)$ terms in $T_n = T_{n-1} + T_{n-2} + 1$ by c, we require

$$c = c + c + 1,$$

namely, $c = -1$. That is, $T_n = -1$ is a particular solution to the equation.

(d) Putting it together, the complete solution to the recurrence is the homogenous solution plus the particular solution:

$$T_n = A\left(\frac{1+\sqrt{5}}{2}\right)^n + B\left(\frac{1-\sqrt{5}}{2}\right)^n - 1$$

(e) All that remains is to find the constants A and B. Substituting the initial conditions gives a system of linear equations.

$$\begin{aligned} 0 &= A + B - 1 \\ 1 &= A\left(\frac{1+\sqrt{5}}{2}\right) + B\left(\frac{1-\sqrt{5}}{2}\right) - 1 \end{aligned}$$

The solution to this linear system is:

$$\begin{aligned} A &= \frac{5+3\sqrt{5}}{10} \\ B &= \frac{5-3\sqrt{5}}{10} \end{aligned}$$

Therefore, the complete solution to the recurrence is

$$T_n = \left(\frac{5+3\sqrt{5}}{10}\right) \cdot \left(\frac{1+\sqrt{5}}{2}\right)^n + \left(\frac{5-3\sqrt{5}}{10}\right) \cdot \left(\frac{1-\sqrt{5}}{2}\right)^n - 1.$$

■

2. $S_0 = 0$

$S_1 = 1$

$S_n = 6S_{n-1} - 9S_{n-2}$

Solution. The characteristic polynomial is $r^2 - 6r + 9 = (r-3)^2$, so the solution is of the form $A3^n + Bn3^n$ for some constants A and B. Setting $n = 0$, we have $0 = S_0 = A3^0 + B \cdot 0 \cdot 3^0 = A$. Setting $n = 1$, we have $1 = S_1 = A3^1 + B \cdot 1 \cdot 3^1 = 3B$, so $B = 1/3$. That is,

$$S_n = 0 \cdot 3^n + \frac{1}{3} \cdot n3^n = n3^{n-1}.$$

■

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

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 $POKE$?

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.

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.

O_2BO_1K	
KO_2BO_1	$BOOK$
O_1BO_2K	$OBOK$
KO_1BO_2	$KOBO$
BO_1O_2K	$\ldots$
BO_2O_1K	
$\ldots$	

4. What kind of mapping is this, young grasshopper?

5. In light of the Division Rule, how many arrangements are there of $BOOK$?

6. Very good, young master! How many arrangements are there of the letters in $KE_1E_2PE_3R$?

7. Suppose we map each arrangement of $KE_1E_2PE_3R$ to an arrangement of $KEEPER$ by erasing subscripts. List all the different arrangements of $KE_1E_2PE_3R$ that are mapped to $REPEEK$ in this way.

8. What kind of mapping is this?

9. So how many arrangements are there of the letters in $KEEPER$?

10. *Now you are ready to face the BOOKKEEPER!*

    How many arrangements of $BO_1O_2K_1K_2E_1E_2PE_3R$ are there?

11. How many arrangements of $BOOK_1K_2E_1E_2PE_3R$ are there?

12. How many arrangements of $BOOKKE_1E_2PE_3R$ are there?

13. How many arrangements of $BOOKKEEPER$ are there?

14. How many arrangements of $VOODOODOLL$ are there?

15. **(IMPORTANT)** How many n-bit sequences contain k zeros and $(n-k)$ ones?

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.

2 Pigeonhole Principle

Solve the following problems using the pigeonhole principle. For each problem, try to identify the *pigeons*, the *pigeonholes*, and a *rule* assigning each pigeon to a pigeonhole.

1. In a room of 500 people, there exist two who share a birthday.

2. Suppose that each of the 115 students in 6.042 sums the nine digits of his or her ID number. Must two people arrive at the same sum?

3. In every set of 100 integers, there exist two whose difference is a multiple of 37.

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 $\{x_1, x_2, \ldots, x_n\}$?

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.)

3. An independent living group is hosting eight pre-frosh, affectionately known as $P_1, \ldots, P_8$ 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 $P_1, \ldots, P_8$ be put to productive use?

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?

4 Fun with Phonology: Hawaiian

The Hawaiian language is rich in vowels: it contains 8 consonants and 25 vowels[1]! In addition, every word in Hawaiian must end in a vowel and must not contain two consonants in a row. Let's assume that all combinations of vowels and consonants that satisfy these constraints are valid.

We'd like to know how many n-phoneme words there are in Hawaiian. (A *phoneme* is either a single vowel or a single consonant. Assume no phoneme can be both a vowel and a consonant.) For simplicity, let's assume n is even.

1. Before tackling the general problem, work out how many different words there are with exactly 4 phonemes. (Which distributions of vowels and consonants are possible?)

2. Now for the general case. Let A be the set of all n-phoneme words, and let A_k be the set of all n-phoneme words with exactly k consonants. Express $|A|$ in terms of $|A_k|$ for all possible k.

3. Now let's find $|A_k|$ for an arbitrary k. For simplicity's sake, assume Hawaiian has only one consonant and only one vowel. Find a bijection between A_k and a set of arbitrary sequences of `0` and `1` of length p. What is p?

4. Using this bijection, compute $|A_k|$.

5. How would you change your expression for $|A_k|$ to allow for 8 consonants and 25 vowels, not just one of each?

[1] Counting long vowels and diphthongs. For this problem, treat each of the 25 vowels as a unique single vowel.

6. How many n-phoneme words are there in Hawaiian? (You don't have to find a closed form for your expression.)

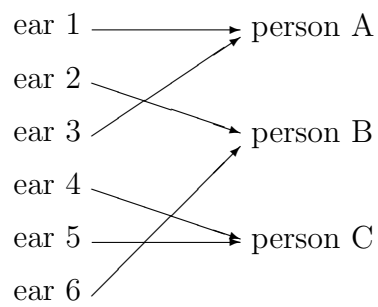
Appendix: Basic Counting Notions

Rule 1 (Bijection Rule)**.** *If there exists a bijection $f : A \to B$, then $|A| = |B|$.*

Rule 2 (Generalized Pigeonhole Principle)**.** *If $|X| > k \cdot |Y|$, then for every function $f : X \to Y$ there exist $k + 1$ different elements of X that are mapped to the same element of Y.*

> "If more than n pigeons are assigned to n holes, then there must exist two pigeons assigned to the same hole."

A ***k-to-1 function*** maps exactly k elements of the domain to every element of the range. For example, the function mapping each ear to its owner is 2-to-1:



Rule 3 (Division Rule)**.** *If $f : A \to B$ is k-to-1, then $|A| = k \cdot |B|$.*

Rule 4 (Product Rule)**.** *If $P_1, P_2, \ldots P_n$ are sets, then:*

$$|P_1 \times P_2 \times \cdots \times P_n| = |P_1| \cdot |P_2| \cdots |P_n|$$

Rule 5 (Generalized Product Rule)**.** *Let S be a set of length-k sequences. If there are:*

- *n_1 possible first entries,*
- *n_2 possible second entries for each first entry,*
- *n_3 possible third entries for each combination of first and second entries, etc.*

then:

$$|S| = n_1 \cdot n_2 \cdot n_3 \cdots n_k$$

Rule 6 (Sum Rule)**.** *If $A_1, \ldots, A_n$ are disjoint sets, then:*

$$|A_1 \cup \cdots \cup A_n| = \sum_{k=1}^{n} |A_k|$$

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Notes 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 $POKE$?

   Solution. There are 4! arrangements corresponding to the 4! permutations of the set $\{P, O, K, E\}$. ■

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.

   Solution. There are 4! arrangements corresponding to the 4! permutations of the set $\{B, O_1, O_2, K\}$. ■

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.

O_2BO_1K	
KO_2BO_1	$BOOK$
O_1BO_2K	$OBOK$
KO_1BO_2	$KOBO$
BO_1O_2K	...
BO_2O_1K	
...	

4. What kind of mapping is this, young grasshopper?

   Solution. 2-to-1 ■

5. In light of the Division Rule, how many arrangements are there of $BOOK$?

   Solution. 4!/2 ■

6. Very good, young master! How many arrangements are there of the letters in $KE_1E_2PE_3R$?

Solution. 6! ■

7. Suppose we map each arrangement of $KE_1E_2PE_3R$ to an arrangement of $KEEPER$ by erasing subscripts. List all the different arrangements of $KE_1E_2PE_3R$ that are mapped to $REPEEK$ in this way.

   Solution. $RE_1PE_2E_3K$, $RE_1PE_3E_2K$, $RE_2PE_1E_3K$, $RE_2PE_3E_1K$, $RE_3PE_1E_2K$, $RE_3PE_2E_1K$ ■

8. What kind of mapping is this?

   Solution. 3!-to-1 ■

9. So how many arrangements are there of the letters in $KEEPER$?

   Solution. 6!/3! ■

10. *Now you are ready to face the BOOKKEEPER!*

    How many arrangements of $BO_1O_2K_1K_2E_1E_2PE_3R$ are there?

    Solution. 10! ■

11. How many arrangements of $BOOK_1K_2E_1E_2PE_3R$ are there?

    Solution. 10!/2! ■

12. How many arrangements of $BOOKKE_1E_2PE_3R$ are there?

    Solution. $10!/(2! \cdot 2!)$ ■

13. How many arrangements of $BOOKKEEPER$ are there?

    Solution. $10!/(2! \cdot 2! \cdot 3!)$ ■

14. How many arrangements of $VOODOODOLL$ are there?

    Solution. $10!/(2! \cdot 2! \cdot 5!)$ ■

15. **(IMPORTANT)** How many n-bit sequences contain k zeros and $(n-k)$ ones?

    Solution. $n!/(k! \cdot (n-k)!)$ ■

    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.

2 Pigeonhole Principle

Solve the following problems using the pigeonhole principle. For each problem, try to identify the *pigeons*, the *pigeonholes*, and a *rule* assigning each pigeon to a pigeonhole.

1. In a room of 500 people, there exist two who share a birthday.

   Solution. The pigeons are the 500 people. The pigeonholes are 366 possible birthdays. Map each person to his or her own birthday. Since there 500 people and 366 birthdays, some two people must have the same birthday by the Pigeonhole Principle. ■

2. Suppose that each of the 115 students in 6.042 sums the nine digits of his or her ID number. Must two people arrive at the same sum?

   Solution. Yes. The students are the pigeons, the possible sums are the pigeonholes, and we map each student to the sum of the digits in his or her MIT ID number. Every sum is in the range from 0 to $9 \cdot 9 = 81$, which means that there are 82 pigeonholes. Since there are more pigeons than pigeonholes, there must be two pigeons in the same pigeonhole; in other words, there must be two students with the same sum. ■

3. In every set of 100 integers, there exist two whose difference is a multiple of 37.

   Solution. The pigeons are the 100 integers. The pigeonholes are the numbers 0 to 36. Map integer k to k rem 37. Since there are 100 pigeons and only 37 pigeonholes, two pigeons must go in the same pigeonhole. This means k_1 rem 37 $=$ k_2 rem 37, which implies that $k_1 - k_2$ is a multiple of 37. ■

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 $\{x_1, x_2, \ldots, x_n\}$?

   Solution. There is a bijection from n-bit sequences with k ones and $n-k$ zeros. The sequence $(b_1, \ldots, b_n)$ maps to the subset that contains x_i if and only if $b_i = 1$. Therefore, the number of such subsets is $\binom{n}{k}$. ■

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.)

   Solution. There is a bijection from selections of a dozen donuts to 16-bit sequences with exactly 4 ones. In particular, suppose that the varieties are glazed, chocolate, lemon, sugar, and Boston creme. Then a selection of g glazed, c chocolate, l lemon, s sugar, and b Boston creme maps to the sequence:

   $$(g\ 0's)\ 1\ (c\ 0's)\ 1\ (l\ 0's)\ 1\ (s\ 0's)\ 1\ (b\ 0's)$$

   Therefore, the number of selections is equal to the number of 16-bit sequences with exactly 4 ones, which is:

   $$\frac{16!}{4!\ 12!} = \binom{16}{4}$$

   ■

3. An independent living group is hosting eight pre-frosh, affectionately known as $P_1, \ldots, P_8$ 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 $P_1, \ldots, P_8$ be put to productive use?

   Solution. There is a bijection from sequences containing two P's, two K's, a B, a C, and two D's. In particular, the sequence $(t_1, \ldots, t_8)$ corresponds to assigning P_i to washing pots if $t_i = P$, to cleaning the kitchen if $t_i = K$, to cleaning the bathrooms if $t_i = B$, etc. Therefore, the number of possible assignments is:

   $$\frac{8!}{2!\ 2!\ 1!\ 1!\ 2!}$$

   ■

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?

   Solution. The number of sequences of the 104 cards containing 2 of each card is $104!/(2!)^{52}$. ■

4 Fun with Phonology: Hawaiian

The Hawaiian language is rich in vowels: it contains 8 consonants and 25 vowels[1]! In addition, every word in Hawaiian must end in a vowel and must not contain two consonants in a row. Let's assume that all combinations of vowels and consonants that satisfy these constraints are valid.

We'd like to know how many n-phoneme words there are in Hawaiian. (A *phoneme* is either a single vowel or a single consonant. Assume no phoneme can be both a vowel and a consonant.) For simplicity, let's assume n is even.

1. Before tackling the general problem, work out how many different words there are with exactly 4 phonemes. (Which distributions of vowels and consonants are possible?)

   Solution. Since a consonant cannot go at the end of a word and no consonant can directly follow another (or equivalently, each consonant must be followed by a vowel), we have these possibilities for vowel/consonant distributions:

   $$VVVV \quad VVCV \quad VCVV \quad CVVV \quad CVCV$$

   Since these are mutually exclusive, we can find the number of words for each of the five types and sum them together. Using the product rule for each type, we find that the total number of n-phoneme words is

   $$\begin{aligned} 25^4 + 25^2 \cdot 8 \cdot 25 + 25 \cdot 8 \cdot 25^2 + 8 \cdot 25^3 + 8 \cdot 25 \cdot 8 \cdot 25 &= 25^4 + 3 \cdot 25^3 \cdot 8 + 8^2 \cdot 25^2 \\ &= 805625 \end{aligned}$$

   ■

2. Now for the general case. Let A be the set of all n-phoneme words, and let A_k be the set of all n-phoneme words with exactly k consonants. Express $|A|$ in terms of $|A_k|$ for all possible k.

   Solution. k can range from 0 to $n/2$ since every consonant is followed by a vowel. Since the set of words with k consonants and the set of words with j consonants where $j \neq k$ are disjoint, we can use the sum rule to compute $|A|$:

   $$|A| = \sum_{k=0}^{n/2} |A_k|$$

   ■

[1] Counting long vowels and diphthongs. For this problem, treat each of the 25 vowels as a unique single vowel.

3. Now let's find $|A_k|$ for an arbitrary k. For simplicity's sake, assume Hawaiian has only one consonant and only one vowel. Find a bijection between A_k and a set of arbitrary sequences of `0` and `1` of length p. What is p?

   Solution. Since every consonant must be followed by a vowel, we can group each consonant and the vowel after it into a cluster. If there are k consonants, then there are k clusters. Since there are no further constraints on the distribution of these clusters, we can map each cluster to `0` and each remaining vowel to `1`. Since the clustering removes k vowels from consideration, and the number of consonants is equal to the number of clusters, the resulting sequences of `0` and `1` have length $n-k$. ■

4. Using this bijection, compute $|A_k|$.

   Solution. The number of sequences of k `0`'s and $n-2k$ `1`'s is

$$\binom{n-k}{k}$$

■

5. How would you change your expression for $|A_k|$ to allow for 8 consonants and 25 vowels, not just one of each?

   Solution. Each word in A_k is a sequence of V's and C's, where each V can represent any vowel and each C can represent any consonant. The total number of these sequences is $\binom{n-k}{k}$, as derived in the previous part.

   Since each sequence has k C's and $n-k$ V's, there are $8^k \cdot 25^{n-k}$ distinct words that map to the same sequence of V's and C's. In other words, this mapping is $(8^k \cdot 25^{n-k})$-to-1, so

$$|A_k| = \binom{n-k}{k} \cdot 8^k \cdot 25^{n-k}$$

■

6. How many n-phoneme words are there in Hawaiian? (You don't have to find a closed form for your expression.)

   Solution. Plugging this into the summation, we get

$$\begin{aligned} |A| &= \sum_{k=0}^{n/2} |A_k| \\ &= \sum_{k=0}^{n/2} \binom{n-k}{k} \cdot 8^k \cdot 25^{n-k} \end{aligned}$$

■

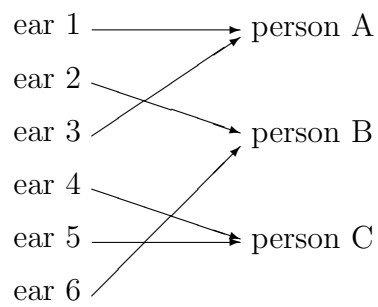
Appendix: Basic Counting Notions

Rule 1 (Bijection Rule)**.** *If there exists a bijection $f : A \to B$, then $|A| = |B|$.*

Rule 2 (Generalized Pigeonhole Principle)**.** *If $|X| > k \cdot |Y|$, then for every function $f : X \to Y$ there exist $k + 1$ different elements of X that are mapped to the same element of Y.*

> "If more than n pigeons are assigned to n holes, then there must exist two pigeons assigned to the same hole."

A ***k-to-1 function*** maps exactly k elements of the domain to every element of the range. For example, the function mapping each ear to its owner is 2-to-1:



Rule 3 (Division Rule)**.** *If $f : A \to B$ is k-to-1, then $|A| = k \cdot |B|$.*

Rule 4 (Product Rule)**.** *If $P_1, P_2, \ldots P_n$ are sets, then:*

$$|P_1 \times P_2 \times \cdots \times P_n| = |P_1| \cdot |P_2| \cdots |P_n|$$

Rule 5 (Generalized Product Rule)**.** *Let S be a set of length-k sequences. If there are:*

- *n_1 possible first entries,*
- *n_2 possible second entries for each first entry,*
- *n_3 possible third entries for each combination of first and second entries, etc.*

then:

$$|S| = n_1 \cdot n_2 \cdot n_3 \cdots n_k$$

Rule 6 (Sum Rule)**.** *If $A_1, \ldots, A_n$ are disjoint sets, then:*

$$|A_1 \cup \cdots \cup A_n| = \sum_{k=1}^{n} |A_k|$$

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Problems for Recitation 16

1 Combinatorial Proof

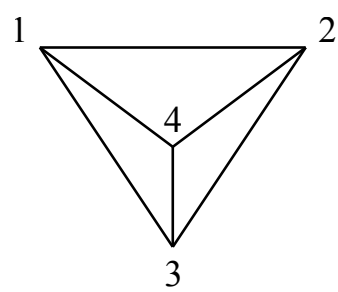
A ***combinatorial proof*** is an argument that establishes an algebraic fact by relying on counting principles. Many such proofs follow the same basic outline:

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 $T = \{X_1, \ldots, X_t\}$ be a set whose elements X_i are themselves sets such that each X_i has size 3 and is $\subseteq \{1, 2, \ldots, n\}$. We call the elements of T "triangles". Suppose that for all "edges" $E \subseteq \{1, 2, \ldots, n\}$ with $|E| = 2$ there are exactly λ 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:

$$\{1, 2, 3\}, \{1, 2, 4\}, \{1, 3, 4\}, \{2, 3, 4\}$$

Prove

$$\lambda \cdot \frac{n(n-1)}{2} = 3t$$

by counting the set

$$C = \{(E, X) : X \in T, E \subseteq X, |E| = 2\}$$

in two different ways.

3 Counting, counting, counting

Learning to count takes practice! Briefly justify your answers to the following questions. Not every problem can be solved with a cute formula; you may have to fall back on case analysis, explicit enumeration, or ad hoc methods. Do as many problems as you can and save the rest to study for Quiz II. You may leave factorials and binomial coefficients in your answers.

1. How many different arrangements are there of the letters in $BANANA$?

2. How many different paths are there from point $(0, 0, 0)$ to point $(10, 20, 30)$ if every step increments one coordinate and leaves the other two unchanged?

3. Find the number of 5-card hands with exactly three aces.

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?

10. In how many different ways can 10 indistinguishable balls be placed in four distinguishable boxes, such that every box gets 1, 2, 3, or 4 balls?

11. In how many different ways can Blockbuster arrange 64 copies of *Cat in the Hat*, 96 copies of *Matrix Revolutions*, and 1 copy of *Amelie* on 5 shelves?

4 There's more than one way...

In the beginning of today's recitation, we gave a combinatorial proof of the following theorem:

Theorem.

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

We can also prove this theorem using induction. Give such a proof.

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Notes 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 $n = m$.

Consider the following theorem:

Theorem.

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

We can prove it with a combinatorial approach:

Proof. We give a combinatorial proof. Let S be the set of all binary sequences with exactly n zeroes and $k+1$ ones.

On the one hand, we know from a previous recitation that the number of such sequences is equal to $\binom{k+n+1}{k}$.

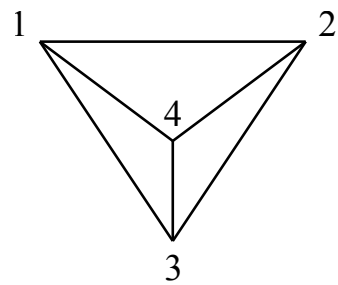
On the other hand, the number of zeroes i to the left of the rightmost one ranges from 0 to n. For a fixed value of i, there are $\binom{k+i}{k}$ possible choices for the sequence of bits before the rightmost one. If we sum over all possible i, we find that $|S| = \sum_{i=0}^{n} \binom{k+i}{k}$.

Equating these two expressions for $|S|$ proves the theorem. $\square$

2 Triangles

Let $T = \{X_1, \ldots, X_t\}$ be a set whose elements X_i are themselves sets such that each X_i has size 3 and is $\subseteq \{1, 2, \ldots, n\}$. We call the elements of T "triangles". Suppose that for all "edges" $E \subseteq \{1, 2, \ldots, n\}$ with $|E| = 2$ there are exactly λ 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:

$$\{1, 2, 3\}, \{1, 2, 4\}, \{1, 3, 4\}, \{2, 3, 4\}$$

Prove

$$\lambda \cdot \frac{n(n-1)}{2} = 3t$$

by counting the set

$$C = \{(E, X) : X \in T, E \subseteq X, |E| = 2\}$$

in two different ways.

Solution. We give a combinatorial proof. Let C be $\{(E, X) : X \in T, E \subseteq X, |E| = 2\}$.

On the one hand, there are $\binom{n}{2}$ sets $E \subseteq \{1, \ldots, n\}$ of size $|E| = 2$. For each such E there are exactly λ triangles $X \in T$ with $E \subseteq X$. So, $|C| = \lambda\binom{n}{2} = \lambda \cdot \frac{n(n-1)}{2}$.

On the other hand, there are t triangles. Each triangle has exactly $\binom{3}{2} = 3$ subsets E of size 2. So, $|C| = 3t$.

Equating these two expressions for $|C|$ proves the theorem. ■

3 Counting, counting, counting

Learning to count takes practice! Briefly justify your answers to the following questions. Not every problem can be solved with a cute formula; you may have to fall back on case analysis, explicit enumeration, or ad hoc methods. Do as many problems as you can and save the rest to study for Quiz II. You may leave factorials and binomial coefficients in your answers.

1. How many different arrangements are there of the letters in $BANANA$?

   Solution. By the Bookkeeper Rule, there are:

$$\frac{6!}{1!\ 3!\ 2!} = 60$$

■

2. How many different paths are there from point $(0, 0, 0)$ to point $(10, 20, 30)$ if every step increments one coordinate and leaves the other two unchanged?

   Solution. There is a bijection between the set of all such paths and the set of strings containing 10 X's, 20 Y's, and 30 Z's. In particular, we obtain a path by working through a string from left to right. An X corresponds to a step that increments the first coordinate, a Y increments the second coordinate, and a Z increments the third. The number of such strings is:

$$\frac{60!}{10!\ 20!\ 30!}$$

   Therefore, this is also the number of paths. ■

3. Find the number of 5-card hands with exactly three aces.

   Solution. We can choose the three aces in $\binom{4}{3}$ ways, and we can choose the remaining two cards in $\binom{48}{2}$ ways. Thus, there are $\binom{4}{3}\binom{48}{2}$ such hands. ■

4. Find the number of 5-card hands in which every suit appears at most twice.

   Solution. There are two cases. Either one suit appears twice or else two suits appear twice. The number of hands in which one suit appears twice is $\binom{13}{2} \cdot 13^3 \cdot 4$, since there are 4 ways to choose the doubly represented suit, $\binom{13}{2}$ ways to choose two values from this suit, and 13^3 ways to choose values for cards in the three remaining suits. Similarly, the number of hands in which two suits appear twice is $\binom{13}{2}^2 \cdot 13 \cdot \binom{4}{2} \cdot 2$. Therefore, there are a total of

$$\binom{13}{2} \cdot 13^3 \cdot 4 + \binom{13}{2}^2 \cdot 13 \cdot \binom{4}{2} \cdot 2$$

   such hands. ■

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.

Solution.

$$\binom{15}{0} + \binom{15}{1} + \binom{15}{2} + \binom{15}{3}$$

■

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.

   Solution. We fall back on explicit enumeration. Let R, G, and B be the values shown on the three dice.

$$\begin{array}{llll} R & = 1, & B+G & = 14 & \to 0 \text{ ways} \\ R & = 2, & B+G & = 13 & \to 0 \text{ ways} \\ R & = 3, & B+G & = 12 & \to 1 \text{ ways} \\ R & = 4, & B+G & = 11 & \to 2 \text{ ways} \\ R & = 5, & B+G & = 10 & \to 3 \text{ ways} \\ R & = 6, & B+G & = 9 & \to 4 \text{ ways} \end{array}$$

   A total of 10 ways.

   Another approach (suggested by a student in recitation) begins by observing that the number of ways the dice can sum to 15 is the same as the number of positive integer solutions to the equation

$$x_1 + x_2 + x_3 = 15$$

   such that $x_i \leq 6$. In general, counting solutions with inequality constraints on the variables involves a tedious case analysis, but in this case there's a trick to remove the constraints: let $y_i = 6 - x_i$. Now the number of desired x_i solutions is the same as the number of nonnegative integer solutions to

$$y_1 + y_2 + y_3 = 3 \tag{1}$$

   such that $y_i \leq 5$. But since the sum of the y_i's must be three, the constraint that $y_i \leq 5$ will be met by every nonnegative integer solution to (1). So we need only count the number of nonnegative integer solutions to (1), which we already know is the same as the number of binary sequences of two zeros and three ones, namely

$$\binom{2+3}{3} = 10.$$

   ■

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?

   Solution. There is a bijection from 13-bit strings with exactly 3 ones. In particular, the string $0^a10^b10^c10^d$ corresponds to to storing $2a$ prefrosh in the first crate, $2b$ in the second, $2c$ in the third, and $2d$ in the fourth. Therefore, the number of ways to store the pre-frosh is equal to the number of 13-bit strings with exactly 3 ones, which is $\binom{13}{3}$. ■

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 $(20,20)$?

   Solution. We use inclusion-exclusion. The total number of paths is $\binom{100}{50}$, but we must subtract off the obstructed paths. There are $\binom{20}{10} \cdot \binom{80}{40}$ paths through the first boulder, since there are $\binom{20}{10}$ paths from the start to the boulder and $\binom{80}{40}$ paths from the boulder to the finish. Similarly, there are $\binom{40}{20} \cdot \binom{60}{30}$ paths through the second boulder. However, we must add back in paths going through both boulders, and there are $\binom{20}{10} \cdot \binom{20}{10} \cdot \binom{60}{30}$ of those. Therefore, the total number of paths is:

$$\binom{100}{50} - \binom{20}{10} \cdot \binom{80}{40} - \binom{40}{20} \cdot \binom{60}{30} + \binom{20}{10} \cdot \binom{20}{10} \cdot \binom{60}{30}$$

■

9. In how many ways can the 180 students in 6.042 be divided into 36 groups of 5?

   Solution. We can group the students using the following procedure: line up the students in some order. Group the first five students, the sixth through tenth students, the eleventh through fifteenth students, and so on. The students can be lined up in 180! ways. However, this overcounts by a factor of $(5!)^{36}$, since the students within each of the 36 groups can be ordered in 5! ways. We are also overcounting by an additional factor of 36!, since the 36 groups can be ordered in 36! ways. Thus, the number of groupings is

$$\frac{180!}{(5!)^{36} \cdot 36!}$$

■

10. In how many different ways can 10 indistinguishable balls be placed in four distinguishable boxes, such that every box gets 1, 2, 3, or 4 balls?

    Solution. First, we might as well put 1 ball in every box. Now the problem is to put the remaining 6 balls into 4 boxes so that no box gets more than 3 balls. Now we turn to case analysis. For example, we could put 3 balls in two boxes and 0 balls in the other two boxes. There are $4!/(2!\ 2!) = 6$ ways to do this. All cases are listed below:

distribution of balls	# of ways	
$3,3,0,0$	$\frac{4!}{2!\ 2!}$	$= 6$
$3,2,1,0$	$\frac{4!}{1!\ 1!\ 1!\ 1!}$	$= 24$
$3,1,1,1$	$\frac{4!}{3!\ 1!}$	$= 4$
$2,2,2,0$	$\frac{4!}{3!\ 1!}$	$= 4$
$2,2,1,1$	$\frac{4!}{2!\ 2!}$	$= 6$

■

11. In how many different ways can Blockbuster arrange 64 copies of *Cat in the Hat*, 96 copies of *Matrix Revolutions*, and 1 copy of *Amelie* on 5 shelves?

    Solution. This is the number of ways to arrange 64 C's (Cat in the Hat), 96 M's (Matrix), 1 A's (Amelie), and 4 X's (dividers between shelves). This is equal to:

$$\frac{(64+96+1+4)!}{64!\ 96!\ 1!\ 4!}$$

■

4 There's more than one way...

In the beginning of today's recitation, we gave a combinatorial proof of the following theorem:

Theorem.

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

We can also prove this theorem using induction. Give such a proof.

Solution. *Proof.* The proof is by induction on n. Let $P(n)$ be the proposition that $\forall k > 0$, $\sum_{i=0}^{n} \binom{k+i}{k} = \binom{k+n+1}{k+1}, n \geq 0$.

Base case: $P(0)$ is true because

$$\forall k > 0, \sum_{i=0}^{0} \binom{k+i}{k} = \binom{k}{k} = \binom{k+1}{k+1} = 1$$

Inductive step: Assume $P(n)$ is true. Show $P(n+1)$ must be true $\forall n > 0$.

$$\begin{aligned}
\forall k > 0, \sum_{i=0}^{n+1} \binom{k+i}{k} &= \binom{k+n+1}{k} + \sum_{i=0}^{n} \binom{k+i}{k} && (2)\\
&= \binom{k+n+1}{k} + \binom{k+n+1}{k+1} && (3)\\
&= \frac{(k+n+1)!}{k!(n+1)!} + \frac{(k+n+1)!}{(k+1)!n!} && (4)\\
&= \frac{(k+1)(k+n+1)! + (k+n+1)!(n+1)}{(k+1)!(n+1)!} && (5)\\
&= \frac{(k+n+1)!(k+n+2)}{(k+1)!(n+1)!} && (6)\\
&= \frac{(k+n+2)!}{(k+1)!(n+1)!} && (7)\\
&= \binom{k+n+2}{k+1} && (8)
\end{aligned}$$

The inductive hypothesis is applied in Step 3. Step 4 follows by definition of choose, and the remaining steps are algebraic simplifications. □

It is always good to have more than one way to solve a problem! ■

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Problems for Recitation 17

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.)

1 The Four-Door Deal

Suppose that *Let's Make a Deal* is played according to different rules. Now there are <u>four</u> doors, with a prize hidden behind one of them. The contestant is allowed to pick a door. The host must then reveal a different door that has no prize behind it. The contestant is allowed to stay with his or her original door or to pick one of the other two that are still closed. If the contestant chooses the door concealing the prize in this second stage, then he or she wins.

1. Contestant Stu, a sanitation engineer from Trenton, New Jersey, stays with his original door. What is the probability that he wins the prize?

   The tree diagram is awkwardly large. This often happens; in fact, sometimes you'll encounter *infinite* tree diagrams! Try to draw enough of the diagram so that you understand the structure of the remainder.

2. Contestant Zelda, an alien abduction researcher from Helena, Montana, switches to one of the remaining two doors with equal probability. What is the probability that she wins the prize?

2 Earliest Door

Let's consider another variation of the four-doors problem. Say the doors are labeled A, B, C, and D. Suppose that Carol always opens the *earliest* door possible (the door whose label is earliest in the alphabet) with the restriction that she can neither reveal the prize nor open the door that the player picked.

This gives contestant Mergatroid— an engineering student from Cambridge, MA— just a little more information about the location of the prize. Suppose that Mergatroid always switches to the earliest door, excluding his initial pick and the one Carol opened. What is the probability that he wins the prize?

3 The 3 doors version revisited

3.1 Carol picks the smallest door

Suppose we are in the original game show with 3 doors. In our original analysis we assumed Carol picked the door randomly. In this case suppose Carol picks the smallest door, while still making sure of both i) it contains a goat and ii) it is not the contestants first choice. The contestant follows the switching strategy. What is the probability the contestant wins?

3.2 Carol picks the smallest door with probability p

This time, when Carol has a choice she chooses the smallest possible door with probability p and the other remaining door with probability $1 - p$. The contestant still follows the switching strategy. What is the probability the contestant wins, in terms of p?

3.3 Optimal strategery

So far we assumed the contestant always switches. We also know from lecture another strategy: the contestant always sticks to her original choice. We determined that the probability of winning with the "always stay" strategy is simple to calculate from the probability of winning with the "always switch" strategy, and that switching was better.

What if the contestant decides whether to switch or not on a case by case basis? That is, suppose the contestant makes a decision of whether to switch or to stay based on 1) Her original choice, and 2) Carol's choice of door. Suppose the doors are labelled A, B and C. Show "always switching" is optimal. (Hint: a strategy can be seen as a mapping that assigns a pair (D_1, D_2) of observations to a decision: switch to D_3 or stay in D_1. The strategy needs to be defined for all pairs (A, B), $(A, C) \cdots$. You can optimize the reaction for each observation individually.)

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Notes for Recitation 17

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.)

1 The Four-Door Deal

Suppose that *Let's Make a Deal* is played according to different rules. Now there are <u>four</u> doors, with a prize hidden behind one of them. The contestant is allowed to pick a door. The host must then reveal a different door that has no prize behind it. The contestant is allowed to stay with his or her original door or to pick one of the other two that are still closed. If the contestant chooses the door concealing the prize in this second stage, then he or she wins.

1. Contestant Stu, a sanitation engineer from Trenton, New Jersey, stays with his original door. What is the probability that he wins the prize?

   The tree diagram is awkwardly large. This often happens; in fact, sometimes you'll encounter *infinite* tree diagrams! Try to draw enough of the diagram so that you understand the structure of the remainder.

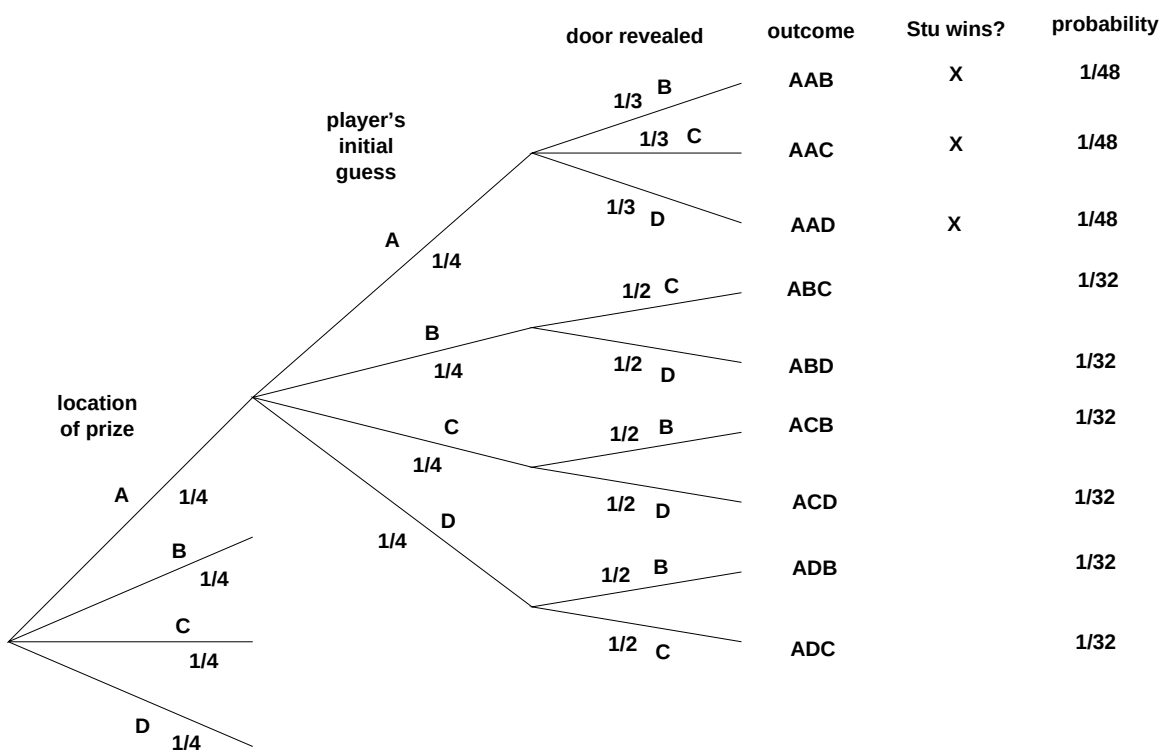
   Solution. Let's make the following assumptions:

   (a) The prize is equally likely to be behind each door.

   (b) The contestant is equally likely to pick each door initially, regardless of the prize's location.

   (c) The host is equally likely to reveal each door that does not conceal the prize and was not selected by the player.

   A partial tree diagram is shown below. The remaining subtrees are symmetric to the full-expanded subtree.



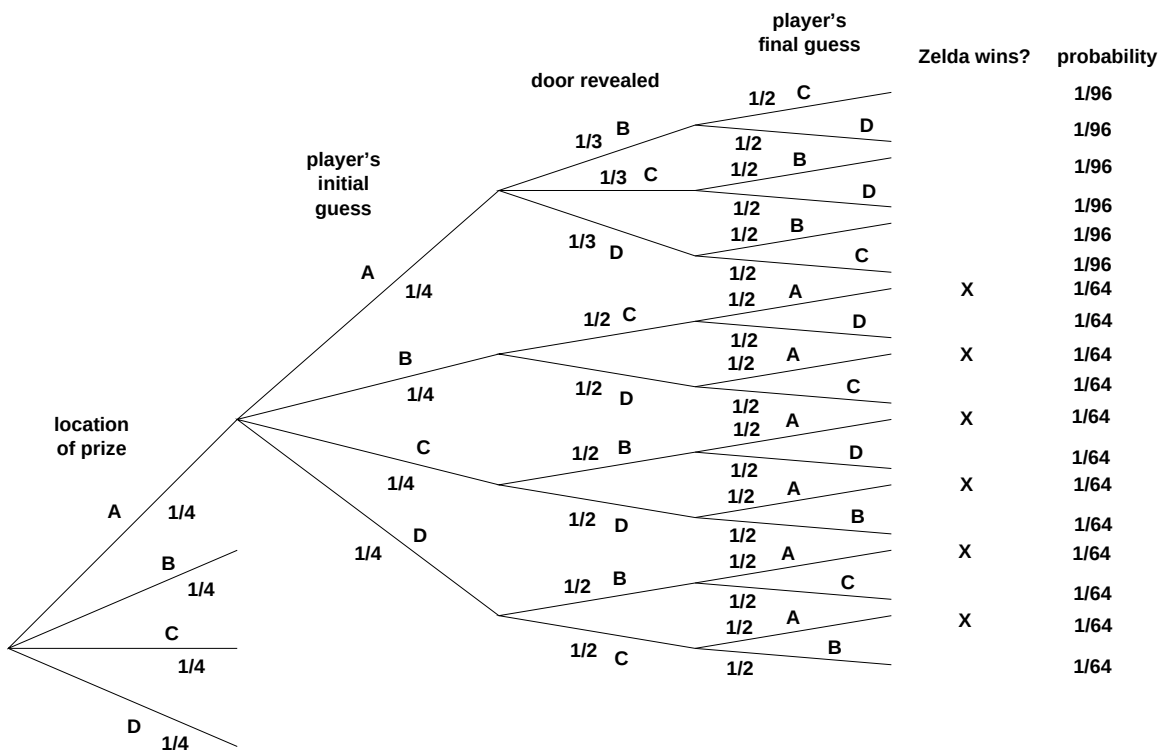
The probability that Stu wins the prize is:

$$\Pr\{\text{Stu wins}\} = 4 \cdot \left(\frac{1}{48} + \frac{1}{48} + \frac{1}{48}\right) = \frac{1}{4}$$

We multiply by 4 to account for the four subtrees, of which we've only drawn one.■

2. Contestant Zelda, an alien abduction researcher from Helena, Montana, switches to one of the remaining two doors with equal probability. What is the probability that she wins the prize?

Solution. A partial tree diagram is worked out below.



The probability that Zelda wins the prize is:

$$\Pr\{\text{Zelda wins}\} = 4 \cdot \left(\frac{1}{64} + \frac{1}{64} + \frac{1}{64} + \frac{1}{64} + \frac{1}{64} + \frac{1}{64}\right) = \frac{3}{8}$$

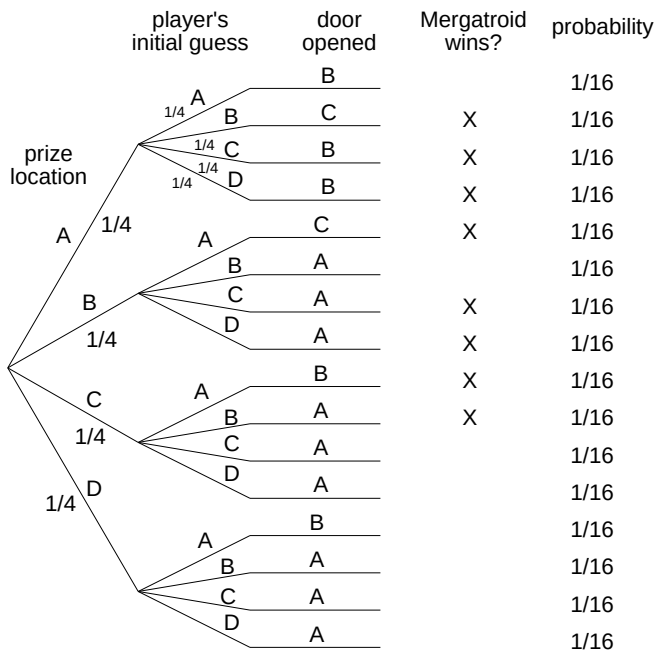
■

2 Earliest Door

Let's consider another variation of the four-doors problem. Say the doors are labeled A, B, C, and D. Suppose that Carol always opens the *earliest* door possible (the door whose label is earliest in the alphabet) with the restriction that she can neither reveal the prize nor open the door that the player picked.

This gives contestant Mergatroid— an engineering student from Cambridge, MA— just a little more information about the location of the prize. Suppose that Mergatroid always switches to the earliest door, excluding his initial pick and the one Carol opened. What is the probability that he wins the prize?

Solution. A tree diagram is worked out below.



The probability that Mergatroid wins is:

$$\Pr\{\text{win}\} = 8 \cdot \frac{1}{16} = \frac{1}{2}$$

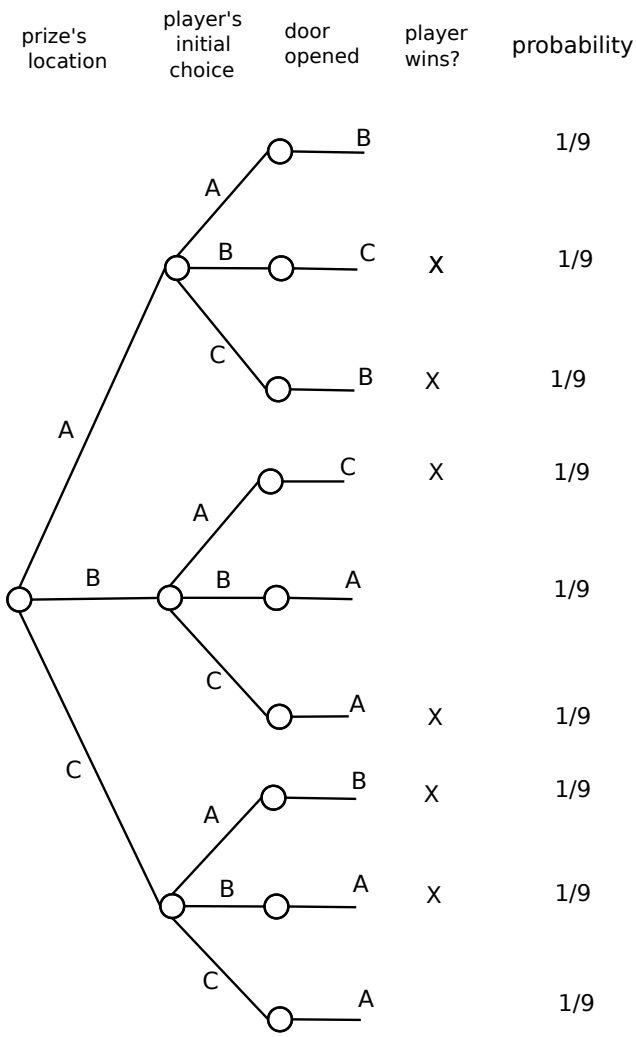
■

3 The 3 doors version revisited

3.1 Carol picks the smallest door

Suppose we are in the original game show with 3 doors. In our original analysis we assumed Carol picked the door randomly. In this case suppose Carol picks the smallest door, while still making sure of both i) it contains a goat and ii) it is not the contestants first choice. The contestant follows the switching strategy. What is the probability the contestant wins?

Solution. A tree diagram is worked out below.



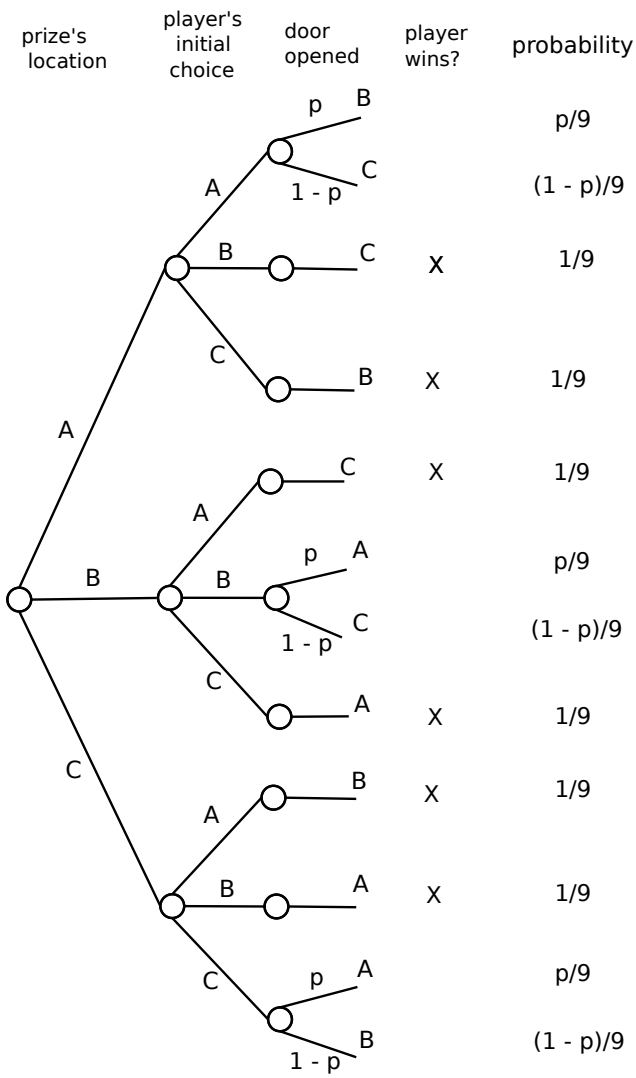
Hence the probability of winning using the switch strategy is 2/3, even when Carol picks the earliest door possible. This is also the case using Mergatroid's strategy of switching to the leftmost option. There is always only one option left over, so only one option to switch to anyway. Contrast this to the 4 door case worked out before.

■

3.2 Carol picks the smallest door with probability p

This time, when Carol has a choice she chooses the smallest possible door with probability p and the other remaining door with probability $1 - p$. The contestant still follows the switching strategy. What is the probability the contestant wins, in terms of p?

Solution. Observe the tree diagram below:



Therefore the probability remains 2/3. Notice the parameter p has no effect in the final formula, because if the contestant plays the "always switch" strategy then she will invariably lose in those situations where Carol has the option to choose among two doors. In the solution we could even allow the parameters p, q, and r rather than a single p, allowing Carol pick them separately, it would still make no difference.

To analyze the original Monty Hall problem, we made the assumption "Carol picks one of the remaining doors equally at random", but from the two previous problems we can see this assumption was not needed in order to conclude switching wins more than half of the time.

■

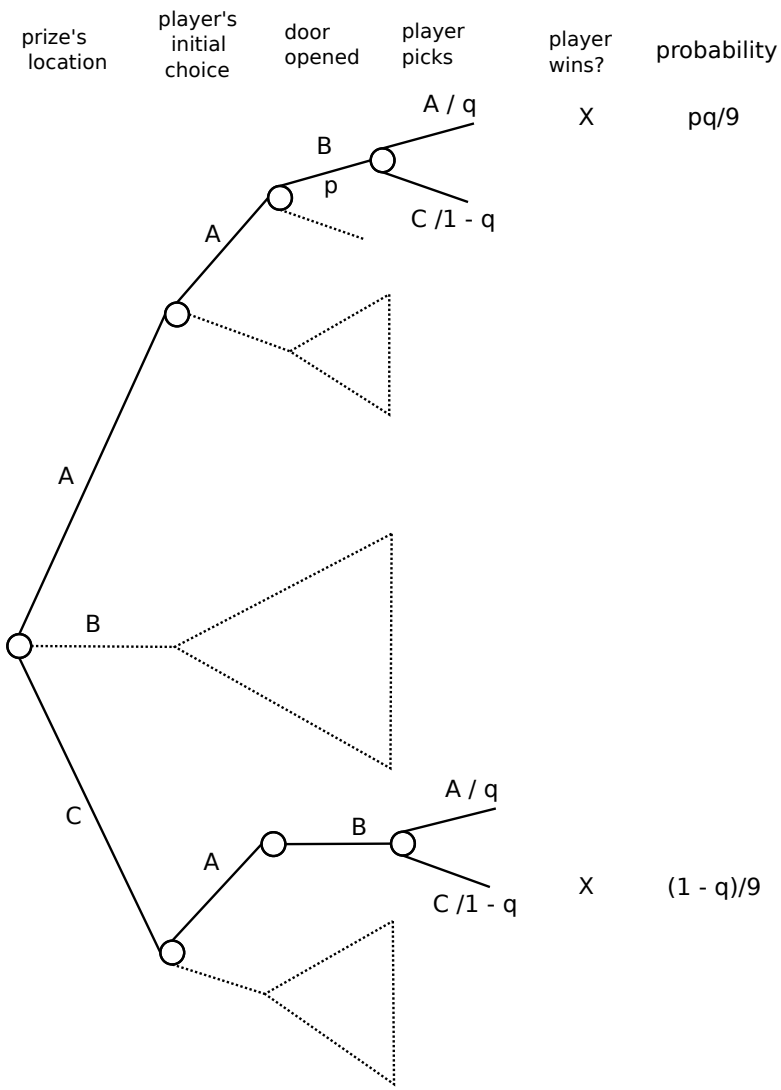
3.3 Optimal strategery

So far we assumed the contestant always switches, and analyzed the outcome of the game that way. In lecture we also discussed the strategy of always sticking to the original choice. We determined that the probability of winning with the "always stay" is larger than the one of winning with an "always switch" strategy, in fact they are complementary.

But, what if the contestant decides whether to switch or not on a case by case basis? That is, suppose the contestant makes a decision on whether to switch or to stay based on 1) Her original choice, and 2) Carol's choice of door.

Suppose the doors are labelled A, B and C. Show "always switching" is optimal. (Hint: a strategy can be seen as a mapping that assigns a pair (D_1, D_2) of observations to a decision: switch to D_3 or stay in D_1. The strategy needs to be defined for all pairs (A, B), $(A, C) \cdots$. You can optimize the reaction for each potential observation individually.)

Solution. Suppose without losing generality the contestant picks door A and is shown door B. He must decide on the basis of this information whether to switch to door C or stay in A. Picking A and observing B is consistent with the prize being behind A or C. As the tree diagram shows, we are only dealing with a subgraph of the full tree:



The diagram indicates Carol picks one door with probability p, and the other with probability $1-p$ when she has a choice. The contestant has a strategy summarized of switching with probablity q. When $q = 1$ we mean she will stay, when $q = 0$ we mean she will switch.

The total probability of winning will include two contributions from the two illustrated

branches, for a total of $(pq + 1 - q)/9$. For a given p, if the contestant decides to move q to 0, the contribution becomes $1/9$. If the contestant moves q towards 1, she gets a contribution of $p/9$. Since $p \leq 1$, then switching will at worst be as good. Notice also that any other value of q in between 0 and 1 will not be optimal (since it is an average of the two extremes). Hence, having $q = 0$ is the best choice, no matter what p is.

In this way, the contestant can optimize her choice for every possible pair of observations. By symmetry you can see the answers will be the same: switch. Hence the overall optimal strategy is "always switch". This result is actually a little stronger than what was suggested in the problem, in which you were only asked to consider the extreme values of $q = 1$ and $q = 0$.

Contrast this with the 4 door version of the problem, where Mergatroid has a strategy that performs better than switching. ■

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Problems for Recitation 18

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?

2 Barglesnort

A Barglesnort makes its lair in one of three caves:

The Barglesnort inhabits cave 1 with probability $\frac{1}{2}$, cave 2 with probability $\frac{1}{4}$, and cave 3 with probability $\frac{1}{4}$. A rabbit subsequently moves into one of the two unoccupied caves, selected with equal probability. With probability $\frac{1}{3}$, the rabbit leaves tracks at the entrance to its cave. (Barglesnorts are much too clever to leave tracks.) What is the probability that the Barglesnort lives in cave 3, given that there are no tracks in front of cave 2?

Use a tree diagram and the four-step method.

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.

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Notes for Recitation 18

The Total Probability Law is a handy tool for breaking down the computation of a probability into distinct cases.

Theorem 1 (Total Probability Law)**.** *Let E and X be events. Then*

$$\Pr\{E\} = \Pr\{E \mid X\} \cdot \Pr\{X\} + \Pr\left\{E \mid \overline{X}\right\} \cdot \Pr\left\{\overline{X}\right\}$$

provided $0 < \Pr\{X\} < 1$.

Proof. Let's simplify the right side.

$$\begin{aligned}
&\Pr\{E \mid X\} \cdot \Pr\{X\} + \Pr\left\{E \mid \overline{X}\right\} \cdot \Pr\left\{\overline{X}\right\} \\
&\qquad = \frac{\Pr\{E \cap X\}}{\Pr\{X\}} \cdot \Pr\{X\} + \frac{\Pr\left\{E \cap \overline{X}\right\}}{\Pr\left\{\overline{X}\right\}} \cdot \Pr\left\{\overline{X}\right\} \\
&\qquad = \Pr\{E \cap X\} + \Pr\left\{E \cap \overline{X}\right\} \\
&\qquad = \Pr\{E\}
\end{aligned}$$

The first step uses the definition of conditional probability. On the next-to-last line, we're adding the probabilities of all outcomes in E *and* X to the probabilities of all outcomes in E and *not* in X. Since every outcome in E is either in X or not in X, this is the sum of the probabilities of all outcomes in E, which equals $\Pr\{E\}$ by the definition of the probability of an event. □

The theorem generalizes as follows:

Theorem 2. *Let E be an event and let $X_1, \ldots, X_n$ be disjoint events whose union is the entire sample space. Then*

$$\Pr\{E\} = \sum_{i=1}^{n} \Pr\{E \mid X_i\} \cdot \Pr\{X_i\}$$

provided $0 < \Pr\{X_i\} < 1$.

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.

   Solution. By the Total Probability Law:

$$\begin{aligned}\Pr\{E\} &= \Pr\{E \mid D\} \cdot \Pr\{D\} + \Pr\left\{E \mid \overline{D}\right\} \cdot \Pr\left\{\overline{D}\right\} \\ &= 0.01 \cdot 0.001 + 0.03 \cdot 0.999 \\ &= 0.02998\end{aligned}$$

■

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.

   Solution. By the Total Probability Law:

$$\begin{aligned}\Pr\{F\} &= \Pr\{F \mid D\} \cdot \Pr\{D\} + \Pr\left\{F \mid \overline{D}\right\} \cdot \Pr\left\{\overline{D}\right\} \\ &= 0.999 \cdot 0.001 + 0.001 \cdot 0.999 \\ &= 0.001998\end{aligned}$$

■

3. Which doctor is more reliable?

   Solution. Doctor X makes more than 15 times as many errors as Doctor Y. ■

2 Barglesnort

A Barglesnort makes its lair in one of three caves:

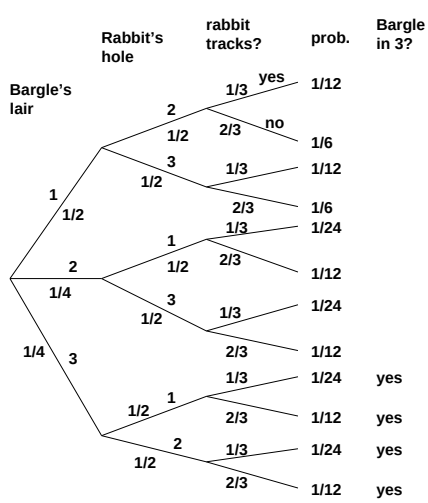
The Barglesnort inhabits cave 1 with probability $\frac{1}{2}$, cave 2 with probability $\frac{1}{4}$, and cave 3 with probability $\frac{1}{4}$. A rabbit subsequently moves into one of the two unoccupied caves, selected with equal probability. With probability $\frac{1}{3}$, the rabbit leaves tracks at the entrance to its cave. (Barglesnorts are much too clever to leave tracks.) What is the probability that the Barglesnort lives in cave 3, given that there are no tracks in front of cave 2?

Use a tree diagram and the four-step method.

Solution. A tree diagram is given below. Let B_3 be the event that the Barglesnort inhabits cave 3, and let T_2 be the event that there are tracks in front of cave 2. Taking data from the tree diagram, we can compute the desired probability as follows:

$$\begin{aligned}
\Pr\left\{B_3 \mid \overline{T_2}\right\} &= \frac{\Pr\left\{B_3 \cap \overline{T_2}\right\}}{\Pr\left\{\overline{T_2}\right\}} \\
&= \frac{\frac{1}{24} + \frac{1}{12} + \frac{1}{12}}{1 - \frac{1}{12} - \frac{1}{24}} \\
&= \frac{5}{21}
\end{aligned}$$

In the denominator, we apply the formula $\Pr\left\{\overline{T_2}\right\} = 1 - \Pr\{T_2\}$ for convenience.



■

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.

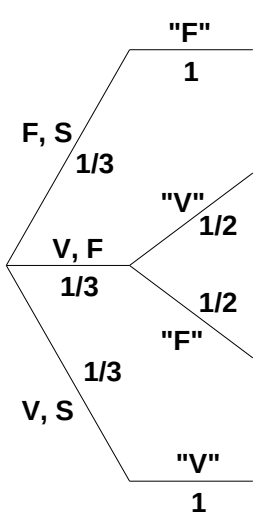
Using a tree diagram and the four-step method, either prove that the Dark Lord Sauron has reasoned correctly or prove that he is wrong. Assume that if the guard has a choice of naming either Voldemort or Foo-Foo (because both are to be released), then he names one of the two uniformly at random.

Solution. Sauron has reasoned incorrectly.

We can explain the issue in terms of conditional probability. Let S, F, and $"F"$ be the events:

$$\begin{aligned} S &= \text{Sauron is released,} \\ F &= \text{Foo-Foo is released,} \\ "F" &= \text{Guard says Foo-Foo is released.} \end{aligned}$$

We can use the tree diagram below to analyze the outcomes and probabilities we should consider. The first split in the tree is based on which two of the three prisoners are to be released, and the second is based on what the guard tells Sauron.



From the diagram we can see that Sauron has correctly observed that

$$\Pr\{S\} = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}.$$

And Sauron is correct in reasoning that if event "F" happens, then event F has also happened. And he's correct that the probability of his release given F shrinks to 1/2. Namely, from the diagram we have:

$$\begin{aligned}
\Pr\{F\} &= \frac{1}{3} + \frac{1}{6} + \frac{1}{6} = \frac{2}{3}, \\
\Pr\{S \cap F\} &= \frac{1}{3} \\
\Pr\{S \mid F\} &= \frac{\Pr\{S \cap F\}}{\Pr\{F\}} = \frac{1}{2}.
\end{aligned}$$

So he worries that if he lets "F" happen, then his probability of release will shrink from 2/3 to $\Pr\{S \mid F\} = 1/2$.

Sauron's confusion is not realizing that the events F and "F" are different, as the tree makes clear. These events even have different probabilities:

$$\Pr\{\text{"}F\text{"}\} = \frac{1}{3} + \frac{1}{6} = \frac{1}{2} \neq \frac{2}{3} = \Pr\{F\}.$$

Now in determining his probability of release, Sauron should use *all* the information available. He should be concerned about his chances given that *the Guard says* Foo-Foo is released, not merely given the weaker fact that Foo-Foo is released. That is, Sauron should care about $\Pr\{S \mid \text{"}F\text{"}\}$, not $\Pr\{S \mid F\}$.

We can see from the diagram that

$$\Pr\{S \mid \text{"}F\text{"}\} = \frac{\Pr\{S \cap \text{"}F\text{"}\}}{\Pr\{\text{"}F\text{"}\}} = \frac{1/3}{1/2} = \frac{2}{3} = \Pr\{S\}.$$

So Sauron's probability of release is not changed by hearing from the guard. ■

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Problems for Recitation 19

1 Bayes' Rule

Bayes' Rule says that if A and B are events with nonzero probabilities, then:

$$\Pr\{A \mid B\} \cdot \Pr\{B\} = \Pr\{B \mid A\} \cdot \Pr\{A\}$$

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:

$$\Pr\{\text{carries umbrella} \mid \text{rains}\} = 0.80$$

Suppose you see the weatherman walking to work, carrying an umbrella. What is the probability of rain? Use Bayes' Rule.

2 DNA Profiles

Suppose that we create a a national database of DNA profiles. Let's make some (overly) simplistic assumptions:

- Each person can be classified into one of <u>20 billion</u> different "DNA types". (For example, you might be type #13,646,572,661 and the person next to you might be type #2,785,466,098.) Let $T(x)$ denote the type of person x.
- Each DNA type is equally probable.
- The DNA types of Americans are mutually independent.

a. A congressman argues that there are only about 300 million Americans, so even if a profile for every American were stored in the database, the probability of even one coincidental match would be very small.

   Recall from lecture that if there are N days in a year and m people in a room, then the probability that no two people in the room have the same birthday is about $e^{-m^2/(2N)}$. Using this fact, what is the probability that two people's DNA profiles would match if every person's profile were stored in the database?

b. After this database is implemented, some DNA is found at a crime scene. The DNA is sequenced and a person with matching DNA is found through the database and accused of the crime. At the trial the defense attorney argues that, by the birthday principle, the probability that there are multiple people whose DNA is identical is a virtual certainty, and so the jury cannot conclude beyond a reasonable doubt that the defendant is the criminal.

   What is the flaw in this argument? Under what circumstances could you conclude based on DNA evidence alone that there is no doubt that the defendant committed the crime? (assuming no errors in the DNA tests, a comprehensive database, etc. etc.)

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!)

d. What is the probability of success if p is chosen in this way? What quantity does this approach when n, the number of Immortal Warriors, grows large?

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Notes for Recitation 19

1 Bayes' Rule

Bayes' Rule says that if A and B are events with nonzero probabilities, then:

$$\Pr\{A \mid B\} \cdot \Pr\{B\} = \Pr\{B \mid A\} \cdot \Pr\{A\}$$

a. Prove Bayes' Rule.

Solution. We reason as follows:

$$\begin{aligned}
\Pr\{A \cap B\} &= \Pr\{A \cap B\} \\
\frac{\Pr\{A \cap B\}}{\Pr\{B\}} \cdot \Pr\{B\} &= \frac{\Pr\{A \cap B\}}{\Pr\{A\}} \cdot \Pr\{A\} \\
\Pr\{A \mid B\} \cdot \Pr\{B\} &= \Pr\{B \mid A\} \cdot \Pr\{A\}
\end{aligned}$$

In the first step, we rewrite both sides using the facts that $\Pr\{A\}$ and $\Pr\{B\}$ are nonzero. The second step uses the definition of conditional probability. ■

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:

$$\Pr\{\text{carries umbrella} \mid \text{rains}\} = 0.80$$

Suppose you see the weatherman walking to work, carrying an umbrella. What is the probability of rain? Use Bayes' Rule.

Solution.

$$\begin{aligned}
\Pr\{\text{rains} \mid \text{carries umbrella}\} &= \Pr\{\text{carries umbrella} \mid \text{rains}\} \cdot \frac{\Pr\{\text{rains}\}}{\Pr\{\text{carries umbrella}\}} \\
&= 0.80 \cdot \frac{0.30}{0.40} \\
&= 0.60
\end{aligned}$$

We've turned around cause and effect! Risk of rain has the effect of making the weatherman carry his umbrella. Yet we've shown that if he carries his umbrella, it is pretty likely to rain! ■

2 DNA Profiles

Suppose that we create a a national database of DNA profiles. Let's make some (overly) simplistic assumptions:

- Each person can be classified into one of <u>20 billion</u> different "DNA types". (For example, you might be type #13,646,572,661 and the person next to you might be type #2,785,466,098.) Let $T(x)$ denote the type of person x.
- Each DNA type is equally probable.
- The DNA types of Americans are mutually independent.

a. A congressman argues that there are only about 300 million Americans, so even if a profile for every American were stored in the database, the probability of even one coincidental match would be very small.

   Recall from lecture that if there are N days in a year and m people in a room, then the probability that no two people in the room have the same birthday is about $e^{-m^2/(2N)}$. Using this fact, what is the probability that two people's DNA profiles would match if every person's profile were stored in the database?

   Solution. The probability of a match $P(\text{match}) = 1 - P(\text{nonmatch})$. By the birthday principle, this is:

$$P(\text{match}) = 1 - e^{-\frac{(3\cdot 10^8)^2}{2\cdot 2\cdot 10^{10}}} = 1 - e^{-\frac{9\cdot 10^6}{4}} \approx 1$$

■

b. After this database is implemented, some DNA is found at a crime scene. The DNA is sequenced and a person with matching DNA is found through the database and accused of the crime. At the trial the defense attorney argues that, by the birthday principle, the probability that there are multiple people whose DNA is identical is a virtual certainty, and so the jury cannot conclude beyond a reasonable doubt that the defendant is the criminal.

   What is the flaw in this argument? Under what circumstances could you conclude based on DNA evidence alone that there is no doubt that the defendant committed the crime? (assuming no errors in the DNA tests, a comprehensive database, etc. etc.)

   Solution. The birthday principle can be used to compute the probability that there are two people in the database who have the same DNA profile; this is analogous to the probability that there are two people in the room with the same birthday. It cannot be used directly to compute the probability that there is another person in the database with the same DNA profile as the criminal; this is analogous to the probability that there is a person in the room with the same birthday as a *particular person*.

So in this case, we are interested in the probability that the defendant is the criminal, given that the defendant DNA matches the crime scene profile. This could be computed if we knew how many people in the database share that profile: If there is only one person in the database with that profile, then you have your perpetrator. If there are n such people, then the probability that a person drawn at random is the perpetrator is $1/n$, which is less than $1/2$ for any $n > 1$ and thus not beyond a reasonable doubt.

We can also compute the probability that the defendant is guilty even if we didn't know the number of the people in the database who shared that profile by summing over all the possibilities. We'll look at how to describe this in a later lecture.

Note: Bear in mind that this is just a 6.042 problem and not a reasonable model for the way that DNA evidence is actually used in forensics; genomes are not independent, just for starters. ■

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?

Solution. Your intuition tells you that it is not to be trusted and that there *can be only one* way to solve this problem: do the math. ■

b. What is the probability that the experiment succeeds as a function of p and n?

Solution. The sample space consists of all possible results of n coin flips, which we can represent by the set $\{H, T\}^n$. Let E be the event that the experiment successfully

selects The One. Then E consists of the n outcomes which contain a single head. In general, the probability of an outcome with h heads and $n-h$ tails is:

$$p^h(1-p)^{n-h}$$

Summing the probabilities of the n outcomes in E gives the probability that the procedure succeeds:

$$\Pr\{E\} = np(1-p)^{n-1}$$

■

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!)

Solution. We compute the derivative of the success probability:

$$\frac{d}{dp}\, np(1-p)^{n-1} = n(1-p)^{n-1} - np(n-1)(1-p)^{n-2}$$

Now we set the right side equal to zero to find the best probability p:

$$\begin{aligned} n(1-p)^{n-1} &= np(n-1)(1-p)^{n-2} \\ (1-p) &= p(n-1) \\ p &= 1/n \end{aligned}$$

This answer makes sense, since we want the coin to come up heads exactly 1 time in n. ■

d. What is the probability of success if p is chosen in this way? What quantity does this approach when n, the number of Immortal Warriors, grows large?

Solution. Setting $p = 1/n$ in the formula for the probability that the experiment succeeds gives:

$$\Pr\{E\} = \left(1 - \frac{1}{n}\right)^{n-1}$$

In the limit, this tends to $1/e$. McLeod is right. ■

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Problems for Recitation 20

Suppose that a coin that comes up heads with probability p is flipped n times. Then for all $\alpha < p$

$$\Pr\{\#\text{ heads} \leq \alpha n\} \leq \frac{1-\alpha}{1-\alpha/p} \cdot \frac{2^{nH(\alpha)}}{\sqrt{2\pi\alpha(1-\alpha)n}} \cdot p^{\alpha n}(1-p)^{(1-\alpha)n}$$

where:

$$H(\alpha) = \alpha \log_2 \frac{1}{\alpha} + (1-\alpha)\log_2 \frac{1}{1-\alpha}$$

1 Approximating the Cumulative Binomial Distribution Function

A coin that comes up heads with probability p is flipped n times. Find an upper bound on

$$\Pr\{\#\text{ heads} \geq \beta n\}$$

where $\beta > p$. Think about the number of tails and plug into the monster formula above.

2 Gallup's Folly

A Gallup poll found that 45% of the adult population of the United States plan to vote Republican in the next election. Gallup polled 640 people and claims a margin of error of 3 percentage points.

Let's check Gallup's claim. Suppose that there are m adult Americans, of whom pm plan to vote Republican and $(1-p)m$ do not. Gallup polls n Americans selected uniformly and independently at random. Of these, qn plan to vote Republican and $(1-q)n$ do not. Gallup then estimates that the fraction of Americans who plan to vote Republican is q.

Note that the only randomization in this experiment is in who Gallup chooses to poll. So the sample space is all sequences of n adult Americans. The response of the i-th person polled is "yes" with probability p and "no" with probability $1-p$ since the person is selected uniformly at random. Furthermore, the n responses are mutually independent.

a. Give an upper bound on the probability that the poll's estimate will be 0.04 or more too low. Just write the expression; don't evaluate yet!

b. Give an upper bound on the probability that the poll's estimate will be 0.04 or more too high. Again, just write the expression.

c. The sum of these two answers is the probability that Gallup's poll will be off by 4 percentage points or more, one way or the other. Unfortunately, these expressions both depend on p— the unknown fraction of voters planning to vote Republican that Gallup is trying to estimate!

However, the sum of these two expressions is maximized when $p = 0.5$. So evaluate the sum with $p = 0.5$ and $n = 640$ to upper bound the probability that Gallup's error is 0.04 or more. Pollsters usually try to ensure that there is a 95% chance that the actual percentage p lies within the poll's error range, which is $q \pm 0.04$ in this case. Is Gallup's poll properly designed?

3 Noisy Channel

Suppose we are transmitting packets of data across a noisy channel. Each packet has probability .01 of being lost. Now suppose we are transmitting $10,000$ packets. What is the probability that at most 2% of the packets are lost?

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Notes for Recitation 20

Philosophy of Probability

Applying probability to real-world processes often involves a little bit of philosophy. Let's first consider this simple problem: What is the probability that

$$N = 2^{6972607} - 1$$

is a prime number? One might guess 1/10 or 1/100. Or one might get sophisticated and point out that the Prime Number Theorem implies that only about 1 in 5 million numbers in this range are prime. Or one can say that assigning a probability to this statement is nonsense because there is no randomness involved; the number is either prime or it isn't.

This question highlights the distinction between two philosophical approaches to probability. One school of thought says that probabilities can only be meaningfully applied to *repeatable processes* like rolling dice or flipping coins. In this view, the probability of an event represent the fraction of trials in which that event will occur. This view is sometimes called classical statistics, sampling theory, or the frequentist approach.

An alternate view is the Bayesian approach, in which a probability can be interpreted as a *degree of belief* in a proposition. A Bayesian would agree that the number above is either prime or composite, however would be perfectly willing to assign a probability to each possibility. The Bayesian approach is thus broader and willing to assign probabilities to any event, repeatable or not. One challenge with the Bayesian approach is coming up with reasonable *prior* probabilities for events that only occur once.

As an aside, it is not clear whether Bayes himself was Bayesian in this sense. However, a Bayesian would be willing to talk about the probability that Bayes was Bayesian while a sampling theorist would say that is nonsense because there is no repeatable process that generates Bayes' beliefs!

Getting back to prime numbers, there is a probabilistic primality test due to Rabin and Miller. If N is composite, there is at least a 3/4 chance that the test will discover this. (In the remaining 1/4 of the time, the test is inconclusive; it never produces a wrong answer.) Moreover, the test can be run again and again and the results are independent. So if N actually is composite, then the probability that $k = 100$ repetitions of the Rabin-Miller do not discover this is at most:

$$\left(\frac{1}{4}\right)^{100}$$

So 100 consecutive inconclusive answers would be extremely convincing evidence that N is prime! If you're comfortable using probability to describe your personal belief about primality after such an experiment, you might be a Bayesian. Otherwise, you might prefer more traditional views of probability.

The Bayesian/Frequentist divide is an interesting one philosophically, but relatively minor in practice; The mathematics of probability remains the same and either approach can lead one astray when modeling real-world processes if the model is based on unsound assumptions. This differences aren't relevant to the mathematics of probability that we teach in 6.042, but do come up in practical courses on statistics, estimation and decision theory.

A Sampling-Theory Approach to Polling

Consider a simple yes/no public opinion poll. In a classical view, every person in the population has a definite opinion and so we assume that there is some fraction p of the population would answer "yes" to the question and the remaining $1 - p$ fraction would answer "no". (Let's forget about the people who hang up on pollsters or launch into long stories about their dog — real pollsters have no such luxury!) Now, p is a fixed number, not a randomly-determined quantity. So trying to determine p by a random experiment is analogous to trying to determine whether N is prime or composite using a probabilistic primality test.

Probability slips into a poll since the pollster samples the opinions of a people selected uniformly and independently at random. The results are qualified by saying something like this:

> "One can say with 95% confidence that the maximum margin of sampling error is ± 4 percentage points."

This means that either the number, q, reported in the poll is within 0.04 of the actual fraction, p, or else an unlucky 1-in-20 event happened during the polling process; specifically, the pollster's random sample was not representative of the population at large. This is *not* the same thing as saying that there is a 95% chance that q is within 0.04 of p; it either is or it isn't, just as N is either prime or composite regardless of the Rabin-Miller test results.

Suppose that a coin that comes up heads with probability p is flipped n times. Then for all $\alpha < p$

$$\Pr\{\#\text{ heads} \leq \alpha n\} \leq \frac{1-\alpha}{1-\alpha/p} \cdot \frac{2^{nH(\alpha)}}{\sqrt{2\pi\alpha(1-\alpha)n}} \cdot p^{\alpha n}(1-p)^{(1-\alpha)n}$$

where:

$$H(\alpha) = \alpha \log_2 \frac{1}{\alpha} + (1-\alpha)\log_2 \frac{1}{1-\alpha}$$

1 Approximating the Cumulative Binomial Distribution Function

A coin that comes up heads with probability p is flipped n times. Find an upper bound on

$$\Pr\{\#\text{ heads} \geq \beta n\}$$

where $\beta > p$. Think about the number of tails and plug into the monster formula above.

Solution.

$$\Pr\{\#\text{ heads} \geq \beta n\} = \Pr\{\#\text{ tails} \leq (1-\beta)n\}$$

Now tails comes up with probability $1-p$. So the answer is the same as above with α replaced by $1-\beta$ and p replaced by $1-p$:

$$\Pr\{\#\text{ heads} > \beta n\} \leq \frac{\beta}{1-\frac{1-\beta}{1-p}} \cdot \frac{2^{nH(\beta)}}{\sqrt{2\pi\beta(1-\beta)n}} \cdot p^{\beta n}(1-p)^{(1-\beta)n}$$

Here we're using the fact that $H(1-\beta) = H(\beta)$. ■

2 Gallup's Folly

A Gallup poll found that 45% of the adult population of the United States plan to vote Republican in the next election. Gallup polled 640 people and claims a margin of error of 3 percentage points.

Let's check Gallup's claim. Suppose that there are m adult Americans, of whom pm plan to vote Republican and $(1-p)m$ do not. Gallup polls n Americans selected uniformly and independently at random. Of these, qn plan to vote Republican and $(1-q)n$ do not. Gallup then estimates that the fraction of Americans who plan to vote Republican is q.

Note that the only randomization in this experiment is in who Gallup chooses to poll. So the sample space is all sequences of n adult Americans. The response of the i-th person polled is "yes" with probability p and "no" with probability $1-p$ since the person is selected uniformly at random. Furthermore, the n responses are mutually independent.

a. Give an upper bound on the probability that the poll's estimate will be 0.04 or more too low. Just write the expression; don't evaluate yet!

   Solution. We can regard each response as a coin flip that is heads with probability p. In these terms, qn is the total number of heads flipped. So we have:

$$\Pr\{qn \le (p-0.04)n\}$$
$$\le \frac{1-(p-0.04)}{1-(p-0.04)/p} \cdot \frac{2^{nH(p-0.04)}}{\sqrt{2\pi(p-0.04)(1-(p-0.04))n}} \cdot p^{(p-0.04)n}(1-p)^{(1-(p-0.04))n}$$

■

b. Give an upper bound on the probability that the poll's estimate will be 0.04 or more too high. Again, just write the expression.

   Solution. Reasoning as before and using the answer to the preceding problem gives:

$$\Pr\{qn > (p+0.04)n\}$$
$$\le \frac{p+0.04}{1-\frac{1-(p+0.04)}{1-p}} \cdot \frac{2^{nH(p+0.04)}}{\sqrt{2\pi(p+0.04)(1-(p+0.04))n}} \cdot p^{(p+0.04)n}(1-p)^{(1-(p+0.04))n}$$

■

c. The sum of these two answers is the probability that Gallup's poll will be off by 4 percentage points or more, one way or the other. Unfortunately, these expressions both depend on p— the unknown fraction of voters planning to vote Republican that Gallup is trying to estimate!

   However, the sum of these two expressions is maximized when $p = 0.5$. So evaluate the sum with $p = 0.5$ and $n = 640$ to upper bound the probability that Gallup's error is 0.04 or more. Pollsters usually try to ensure that there is a 95% chance that the actual percentage p lies within the poll's error range, which is $q \pm 0.04$ in this case. Is Gallup's poll properly designed?

Solution. The probability that the error is 0.04 or more is at most

$$
\begin{aligned}
& \frac{.54}{1-.46/.5} \cdot \frac{2^{640\cdot H(.46)}}{\sqrt{2\pi\cdot .46\cdot .54\cdot 640}} \cdot .5^{.46\cdot 640}.5^{.54\cdot 640} \\
+ & \\
& \frac{.54}{1-.46/.5} \cdot \frac{2^{640\cdot H(.54)}}{\sqrt{2\pi\cdot .54\cdot .46\cdot 640}} \cdot .5^{.54\cdot 640}.5^{.46\cdot 640}. \\
= \; & 2\cdot \frac{.54}{1-.46/.5} \cdot \frac{2^{640\cdot H(.54)}}{\sqrt{2\pi\cdot .54\cdot .46\cdot 640}} \cdot .5^{.54\cdot 640}.5^{.46\cdot 640} \\
\leq \; & .427\cdot 2^{640\cdot H(.54)}(.5)^{640} \\
= \; & .427\cdot 2^{640\cdot H(.54)-640} \\
\leq \; & .427\cdot 2^{-3.008} \\
\leq \; & .054
\end{aligned}
$$

This means that p will lie within the error range of a polled fraction with probability 0.946. So our estimates suggest Gallup's poll is not quite large enough to meet the claimed 0.95 probability. Since Gallup is a professional, we expect he's got the poll size right, by using a more accurate numerical estimation formula – or he may have considered it legitimate to round a very slightly larger margin of error down to 0.04. ■

3 Noisy Channel

Suppose we are transmitting packets of data across a noisy channel. Each packet has probability .01 of being lost. Now suppose we are transmitting $10,000$ packets. What is the probability that at most 2% of the packets are lost?

Solution. Sending data over the noisy channel is analogous to flipping $10,000$ coins where the probability of heads is $p = 0.01$ (in this case, a coin coming up heads is equivalent to the packet being dropped). We want to know what the likelihood is of greater than $\alpha = .2$ of all the coins coming up heads. However, in this case, we have $a > p$, and cannot use the equation we developed earlier. However, we can ask ourselves the question in terms of number of tails, where the probability of tails is $p = 0.99$, and we want to know the probability of a at most $\alpha 0.98$ of them coming up tails.

Plugging this in to our equation, we find that the probability is approximately

$$\left(\frac{1-.98}{1-.98/.99}\right)\frac{2^{10000(.98log(\frac{.98}{.00})+.02log(\frac{.01}{.02}))}}{\sqrt{2\pi(.98)(1-.98)10000}} < 2^{-60}$$

■

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Problems for Recitation 21

Problem 1. [**points**] 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?

Problem 2. [**points**] The number of squares that a piece advances in one turn of the game Monopoly is determined as follows:

- Roll two dice, take the sum of the numbers that come up, and advance that number of squares.
- If you roll *doubles* (that is, the same number comes up on both dice), then you roll a second time, take the sum, and advance that number of additional squares.
- If you roll doubles a second time, then you roll a third time, take the sum, and advance that number of additional squares.
- However, as a special case, if you roll doubles a third time, then you go to jail. Regard this as advancing zero squares overall for the turn.

(a) [pts] What is the expected sum of two dice, given that the same number comes up on both?

(b) [pts] What is the expected sum of two dice, given that different numbers come up? (Use your previous answer and the Total Expectation Theorem.)

(c) [pts] 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.

(d) [pts] What is the expected number of squares that a piece advances in Monopoly?

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Notes for Recitation 21

1 Conditional Expectation and Total Expectation

There are conditional expectations, just as there are conditional probabilities. If R is a random variable and E is an event, then the conditional expectation $\text{Ex}(R \mid E)$ is defined by:

$$\text{Ex}(R \mid E) = \sum_{w \in S} R(w) \cdot \Pr\{w \mid E\}$$

For example, let R be the number that comes up on a roll of a fair die, and let E be the event that the number is even. Let's compute $\text{Ex}(R \mid E)$, the expected value of a die roll, given that the result is even.

$$\begin{aligned}
\text{Ex}(R \mid E) &= \sum_{w \in \{1,\ldots,6\}} R(w) \cdot \Pr\{w \mid E\} \\
&= 1 \cdot 0 + 2 \cdot \frac{1}{3} + 3 \cdot 0 + 4 \cdot \frac{1}{3} + 5 \cdot 0 + 6 \cdot \frac{1}{3} \\
&= 4
\end{aligned}$$

It helps to note that the conditional expectation, $\text{Ex}(R \mid E)$ is simply the expectation of R with respect to the probability measure $\Pr_E()$ defined in PSet 10. So it's linear:

$$\text{Ex}(R_1 + R_2 \mid E) = \text{Ex}(R_1 \mid E) + \text{Ex}(R_2 \mid E).$$

Conditional expectation is really useful for breaking down the calculation of an expectation into cases. The breakdown is justified by an analogue to the Total Probability Theorem:

Theorem 1 (Total Expectation)**.** *Let $E_1, \ldots, E_n$ be events that partition the sample space and all have nonzero probabilities. If R is a random variable, then:*

$$\text{Ex}(R) = \text{Ex}(R \mid E_1) \cdot \Pr\{E_1\} + \cdots + \text{Ex}(R \mid E_n) \cdot \Pr\{E_n\}$$

For example, let R be the number that comes up on a fair die and E be the event that result is even, as before. Then $\overline{E}$ is the event that the result is odd. So the Total Expectation theorem says:

$$\underbrace{\text{Ex}(R)}_{=\,7/2} = \underbrace{\text{Ex}(R \mid E)}_{=\,4} \cdot \underbrace{\Pr\{E\}}_{=\,1/2} + \underbrace{\text{Ex}\left(R \mid \overline{E}\right)}_{=\,?} \cdot \underbrace{\Pr\{E\}}_{=\,1/2}$$

The only quantity here that we don't already know is $\text{Ex}\left(R \mid \overline{E}\right)$, which is the expected die roll, given that the result is odd. Solving this equation for this unknown, we conclude that $\text{Ex}\left(R \mid \overline{E}\right) = 3$.

To prove the Total Expectation Theorem, we begin with a Lemma.

Lemma. *Let R be a random variable, E be an event with positive probability, and I_E be the indicator variable for E. Then*

$$\text{Ex}\left(R \mid E\right) = \frac{\text{Ex}\left(R \cdot I_E\right)}{\Pr\left\{E\right\}} \tag{1}$$

Proof. Note that for any outcome, s, in the sample space,

$$\Pr\left\{\{s\} \cap E\right\} = \begin{cases} 0 & \text{if } I_E(s) = 0, \\ \Pr\left\{s\right\} & \text{if } I_E(s) = 1, \end{cases}$$

and so

$$\Pr\left\{\{s\} \cap E\right\} = I_E(s) \cdot \Pr\left\{s\right\}. \tag{2}$$

Now,

$$\begin{aligned}
\text{Ex}\left(R \mid E\right) &= \sum_{s \in S} R(s) \cdot \Pr\left\{\{s\} \mid E\right\} && \text{(Def of } \text{Ex}\left(\cdot \mid E\right)\text{)} \\
&= \sum_{s \in S} R(s) \cdot \frac{\Pr\left\{\{s\} \cap E\right\}}{\Pr\left\{E\right\}} && \text{(Def of } \Pr\left\{\cdot \mid E\right\}\text{)} \\
&= \sum_{s \in S} R(s) \cdot \frac{I_E(s) \cdot \Pr\left\{s\right\}}{\Pr\left\{E\right\}} && \text{(by (2))} \\
&= \frac{\sum_{s \in S} (R(s) \cdot I_E(s)) \cdot \Pr\left\{s\right\}}{\Pr\left\{E\right\}} \\
&= \frac{\text{Ex}\left(R \cdot I_E\right)}{\Pr\left\{E\right\}} && \text{(Def of } \text{Ex}\left(R \cdot I_E\right)\text{)}
\end{aligned}$$

□

Now we prove the Total Expectation Theorem:

Proof. Since the E_i's partition the sample space,

$$R = \sum_i R \cdot I_{E_i} \tag{3}$$

for any random variable, R. So

$$\begin{aligned}
\operatorname{Ex}(R) &= \operatorname{Ex}\left(\sum_i R \cdot I_{E_i}\right) && \text{(by (3))} \\
&= \sum_i \operatorname{Ex}(R \cdot I_{E_i}) && \text{(linearity of Ex ())} \\
&= \sum_i \operatorname{Ex}(R \mid E_i) \cdot \Pr\{E_i\} && \text{(by (1))}
\end{aligned}$$

$\square$

Problem 1. [**points**] 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?

Solution. Let the random variable R be the amount of money won or lost by the player in a round. We can compute the expected value of R as follows:

$$\begin{aligned}
\mathrm{Ex}(R) &= -1 \cdot \Pr\{0 \text{ matches}\} + 1 \cdot \Pr\{1 \text{ match}\} + 2 \cdot \Pr\{2 \text{ matches}\} + 4 \cdot \Pr\{3 \text{ matches}\} \\
&= -1 \cdot \left(\frac{5}{6}\right)^3 + 1 \cdot 3\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^2 + 2 \cdot 3\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right) + 4 \cdot \left(\frac{1}{6}\right)^3 \\
&= \frac{-125 + 75 + 30 + 4}{216} \\
&= \frac{-16}{216}
\end{aligned}$$

You can expect to lose 16/216 of a dollar (about 7.4 cents) in every round. This is a horrible game! ■

Problem 2. [**points**] The number of squares that a piece advances in one turn of the game Monopoly is determined as follows:

- Roll two dice, take the sum of the numbers that come up, and advance that number of squares.
- If you roll *doubles* (that is, the same number comes up on both dice), then you roll a second time, take the sum, and advance that number of additional squares.
- If you roll doubles a second time, then you roll a third time, take the sum, and advance that number of additional squares.
- However, as a special case, if you roll doubles a third time, then you go to jail. Regard this as advancing zero squares overall for the turn.

(a) [pts] What is the expected sum of two dice, given that the same number comes up on both?

Solution. There are six equally-probable sums: 2, 4, 6, 8, 10, and 12. Therefore, the expected sum is:

$$\frac{1}{6} \cdot 2 + \frac{1}{6} \cdot 4 + \ldots + \frac{1}{6} \cdot 12 = 7$$

■

(b) [pts] What is the expected sum of two dice, given that different numbers come up? (Use your previous answer and the Total Expectation Theorem.)

Solution. Let the random variables D_1 and D_2 be the numbers that come up on the two dice. Let E be the event that they are equal. The Total Expectation Theorem says:

$$\text{Ex}\left(D_1 + D_2\right) = \text{Ex}\left(D_1 + D_2 \mid E\right) \cdot \Pr\left\{E\right\} + \text{Ex}\left(D_2 + D_2 \mid \overline{E}\right) \cdot \Pr\left\{\overline{E}\right\}$$

Two dice are equal with probability $\Pr\left\{E\right\} = 1/6$, the expected sum of two independent dice is 7, and we just showed that $\text{Ex}\left(D_1 + D_2 \mid E\right) = 7$. Substituting in these quantities and solving the equation, we find:

$$\begin{aligned} 7 &= 7 \cdot \frac{1}{6} + \text{Ex}\left(D_2 + D_2 \mid \overline{E}\right) \cdot \frac{5}{6} \\ \text{Ex}\left(D_2 + D_2 \mid \overline{E}\right) &= 7 \end{aligned}$$

■

(c) [pts] 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.

Solution. From the total expectation formula, we get:

$$\begin{aligned}\operatorname{Ex}(\text{advance}) = {} & \operatorname{Ex}\left(X_1 \mid \overline{E_1}\right) \cdot \Pr\left\{\overline{E_1}\right\} \\ & + \operatorname{Ex}\left(X_1 + X_2 \mid E_1 \cap \overline{E_2}\right) \cdot \Pr\left\{E_1 \cap \overline{E_2}\right\} \\ & + \operatorname{Ex}\left(X_1 + X_2 + X_3 \mid E_1 \cap E_2 \cap \overline{E_3}\right) \cdot \Pr\left\{E_1 \cap E_2 \cap \overline{E_3}\right\} \\ & + \operatorname{Ex}\left(0 \mid E_1 \cap E_2 \cap E_3\right) \cdot \Pr\left\{E_1 \cap E_2 \cap E_3\right\}\end{aligned}$$

Then using linearity of (conditional) expectation, we refine this to

$$\begin{aligned}&\operatorname{Ex}(\text{advance}) \\ &= \operatorname{Ex}\left(X_1 \mid \overline{E_1}\right) \cdot \Pr\left\{\overline{E_1}\right\} \\ &\quad + \left(\operatorname{Ex}\left(X_1 \mid E_1 \cap \overline{E_2}\right) + \operatorname{Ex}\left(X_2 \mid E_1 \cap \overline{E_2}\right)\right) \cdot \Pr\left\{E_1 \cap \overline{E_2}\right\} \\ &\quad + \left(\operatorname{Ex}\left(X_1 \mid E_1 \cap E_2 \cap \overline{E_3}\right) + \operatorname{Ex}\left(X_2 \mid E_1 \cap E_2 \cap \overline{E_3}\right) + \operatorname{Ex}\left(X_3 \mid E_1 \cap E_2 \cap \overline{E_3}\right)\right) \\ &\qquad \cdot \Pr\left\{E_1 \cap E_2 \cap \overline{E_3}\right\} \\ &\quad + 0.\end{aligned}$$

Using mutual independence of the rolls, we simplify this to

$$\begin{aligned}&\operatorname{Ex}(\text{advance}) \\ &= \operatorname{Ex}\left(X_1 \mid \overline{E_1}\right) \cdot \Pr\left\{\overline{E_1}\right\} \\ &\quad + \left(\operatorname{Ex}\left(X_1 \mid E_1\right) + \operatorname{Ex}\left(X_2 \mid \overline{E_2}\right)\right) \cdot \Pr\left\{E_1\right\} \cdot \Pr\left\{\overline{E_2}\right\} \\ &\quad + \left(\operatorname{Ex}\left(X_1 \mid E_1\right) + \operatorname{Ex}\left(X_2 \mid E_2\right) + \operatorname{Ex}\left(X_3 \mid \overline{E_3}\right)\right) \cdot \Pr\left\{E_1\right\} \cdot \Pr\left\{E_2\right\} \cdot \Pr\left\{\overline{E_3}\right\}\end{aligned} \tag{4}$$

■

(d) [pts] What is the expected number of squares that a piece advances in Monopoly?

Solution. We plug the values from parts (a) and (b) into equation (4):

$$\begin{aligned}\operatorname{Ex}(\text{advance}) &= 7 \cdot \frac{5}{6} + (7+7) \cdot \frac{1}{6} \cdot \frac{5}{6} + (7+7+7) \cdot \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{5}{6} \\ &= 8\frac{19}{72}\end{aligned}$$

■

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Problems for Recitation 22

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, $\text{Var}[R] = \text{E}[R^2] - \text{E}^2[R]$.

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.

c. Show that if R_1 and R_2 are independent random variables, then

$$\text{Var}[R_1 + R_2] = \text{Var}[R_1] + \text{Var}[R_2].$$

d. Give an example of random variables R_1 and R_2 for which

$$\text{Var}[R_1 + R_2] \neq \text{Var}[R_1] + \text{Var}[R_2].$$

e. Compute the variance and standard deviation of the Binomial distribution $H_{n,p}$ with parameters n and p.

f. Let's say we have a random variable T such that $T = \sum_{j=1}^{n} T_j$, where all of the T_j's are mutually independent and take values in the range $[0, 1]$. Prove that Var(T)≤Ex(T). We'll use this result in lecture tomorrow. *Hint: Upper bound* $\text{Var}[T_j]$ *with* $\text{E}[T_j]$ *using the definition of variance in part (a) and the rule for computing the expectation of a function of a random variable.*

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Notes for Recitation 22

1 Expected Value Rule for Functions of Random Variables

In lecture, we have computed the expectation of a function of a random variable without explicitly discussing the general rule. For example, yesterday we saw that the expectation of the square of the roll of a die is not equal to the square of the expectation of the roll. That is, if R is the outcome of a single roll of a die. Then $\mathrm{Ex}\left(R^2\right) \neq \mathrm{Ex}\left(R\right)^2$. We will now explicitly present the rule for computing the expection of a function of a random variable R.

Rule (Expected Value for the Function of a Random Variable)**.** *Let R be a random variable, and let $f(R)$ be a function of R. Then, the expected value of the random variable $f(R)$ is given by*

$$\mathrm{E}\left[f(R)\right] = \sum_{x \in Range(R)} f(x) \cdot \Pr\left\{R = x\right\}$$

2 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, $\mathrm{Var}\left[R\right] = \mathrm{E}\left[R^2\right] - \mathrm{E}^2\left[R\right]$.

Solution. Let $\mu = \mathrm{E}\left[R\right]$. Then

$$\begin{aligned}
\mathrm{Var}\left[R\right] &= \mathrm{E}\left[\left(R - \mathrm{E}\left[R\right]\right)^2\right] && \text{(Definition of variance)}\\
&= \mathrm{E}\left[\left(R - \mu\right)^2\right] && \text{(def. of } \mu\text{)}\\
&= \mathrm{E}\left[R^2 - 2\mu R + \mu^2\right]\\
&= \mathrm{E}\left[R^2\right] - 2\mu\,\mathrm{E}\left[R\right] + \mu^2 && \text{(linearity of expectation)}\\
&= \mathrm{E}\left[R^2\right] - 2\mu^2 + \mu^2 && \text{(def. of } \mu\text{)}\\
&= \mathrm{E}\left[R^2\right] - \mu^2\\
&= \mathrm{E}\left[R^2\right] - \mathrm{E}^2\left[R\right]. && \text{(def. of } \mu\text{)}
\end{aligned}$$

■

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.

Solution. We will transform the left side into the right side. The first step is to expand $\text{Var}[aR+b]$ using the alternate definition of variance.

$$\text{Var}[aR+b] = \text{E}\left[(aR+b)^2\right] - \text{E}^2[aR+b].$$

We will work on the first term and then the second term. For the first term, note that by linearity of expectation,

$$\text{E}\left[(aR+b)^2\right] = \text{E}\left[a^2R^2 + 2abR + b^2\right] = a^2\,\text{E}\left[R^2\right] + 2ab\,\text{E}[R] + b^2. \tag{1}$$

Similarly for the second term:

$$\text{E}^2[aR+b] = (a\,\text{E}[R]+b)^2 = a^2\text{E}^2[R] + 2ab\,\text{E}[R] + b^2. \tag{2}$$

Finally, we subtract the expanded second term from the first.

$$\begin{aligned}
\text{Var}[aR+b] &= \text{E}\left[(aR+b)^2\right] - \text{E}^2[aR+b] && \text{(previous part)}\\
&= a^2\,\text{E}\left[R^2\right] + 2ab\,\text{E}[R] + b^2 - \\
&\quad (a^2\text{E}^2[R] + 2ab\,\text{E}[R] + b^2) && \text{(by (??) and (??))}\\
&= a^2\,\text{E}\left[R^2\right] - a^2\text{E}^2[R]\\
&= a^2(\text{E}\left[R^2\right] - \text{E}^2[R])\\
&= a^2\,\text{Var}[R] && \text{(previous part)}
\end{aligned}$$

Since the standard deviation of a random variable is the square root of the variance, the standard deviation of $aR+b$ is $\sqrt{a^2\,\text{Var}[R]}$ which is just a times the standard deviation of R. ■

c. Show that if R_1 and R_2 are independent random variables, then

$$\text{Var}[R_1+R_2] = \text{Var}[R_1] + \text{Var}[R_2].$$

Solution. We will transform the left side into the right side. We begin by applying the alternate definition of variance.

$$\text{Var}[R_1+R_2] = \text{E}\left[(R_1+R_2)^2\right] - \text{E}^2[R_1+R_2].$$

We will work on the first term and then the second term separately. For the first term, note

$$\begin{aligned}
\text{E}\left[(R_1+R_2)^2\right] &= \text{E}\left[R_1^2 + 2R_1R_2 + R_2^2\right]\\
&= \text{E}\left[R_1^2\right] + \text{E}[2R_1R_2] + \text{E}\left[R_2^2\right]\\
&= \text{E}\left[R_1^2\right] + 2\,\text{E}[R_1]\,\text{E}[R_2] + \text{E}\left[R_2^2\right].
\end{aligned}$$

First, we multiply out the squared expression. The second step uses linearity of expectation. In the last step, we break the expectation of the product $R_1 R_2$ into a product of expectations; this is where we use the fact that R_1 and R_2 are independent. Now we work on the second term.

$$\begin{aligned} \mathrm{E}^2[R_1 + R_2] &= (\mathrm{E}[R_1] + \mathrm{E}[R_2])^2 \\ &= \mathrm{E}^2[R_1] + 2\,\mathrm{E}[R_1]\,\mathrm{E}[R_2] + \mathrm{E}^2[R_2]\,. \end{aligned}$$

The first step uses linearity of expectation, and in the second step we multiply out the squared expression. Now we subtract the (expanded) second term from the first. Cancelling and rearranging terms, we find that

$$\begin{aligned} \mathrm{Var}[R_1 + R_2] &= (\mathrm{E}\left[R_1^2\right] - \mathrm{E}^2[R_1]) + (\mathrm{E}\left[R_2^2\right]) - \mathrm{E}^2[R_2]) \\ &= \mathrm{Var}[R_1] + \mathrm{Var}[R_2]\,. \end{aligned}$$

■

d. Give an example of random variables R_1 and R_2 for which

$$\mathrm{Var}[R_1 + R_2] \neq \mathrm{Var}[R_1] + \mathrm{Var}[R_2]\,.$$

Solution. Suppose $R = R_1 = R_2$. If linearity of variance held, then $\mathrm{Var}[R + R] = \mathrm{Var}[R] + \mathrm{Var}[R]$. However, by part b, $\mathrm{Var}[R + R] = \mathrm{Var}[2R] = 4\,\mathrm{Var}[R]$. This is only possible if $\mathrm{Var}[R] = 0$. If, say, we choose R to be the outcome of a fair coin flip, $\mathrm{Var}[R] \neq 0$. In fact, any R which holds at least 2 distinct values each with positive probability will do. ■

e. Compute the variance and standard deviation of the Binomial distribution $H_{n,p}$ with parameters n and p.

Solution. We know that $H_{n,p} = \sum_{k=1}^{n} I_k$ where the I_k are mutually independent 0-1-valued variables with $\Pr\{I_k = 1\} = p$. The variance of I_k is $\mathrm{E}[I_k^2] - \mathrm{E}[I_k]^2 = \mathrm{E}[I_k] - \mathrm{E}[I_k]^2 = \mathrm{E}[I_k](1 - \mathrm{E}[I_k]) = p(1-p)$. Thus, by linearity of variance, we have $\mathrm{Var}[H_{n,p}] = n\,\mathrm{Var}[I_k] = np(1-p)$. Thus, the standard deviation of $H_{n,p}$ is $\sqrt{np(1-p)}$. ■

f. Let's say we have a random variable T such that $T = \sum_{j=1}^{n} T_j$, where all of the T_j's are mutually independent and take values in the range $[0, 1]$. Prove that Var(T)≤Ex(T). We'll use this result in lecture tomorrow. *Hint: Upper bound* $\mathrm{Var}[T_j]$ *with* $\mathrm{E}[T_j]$ *using the definition of variance in part (a) and the rule for computing the expectation of a function of a random variable.*

Solution. We know by linearity of variance for mutually-independent random variables that

$$\begin{aligned}\mathrm{Var}[T] &= \mathrm{Var}[T_1 + \ldots + T_n] \\ &= \mathrm{Var}[T_1] + \ldots + \mathrm{Var}[T_n]\end{aligned}$$

Now we evaluate the variance of T_j for any j. Using the definition of variance from part (a) above, we have

$$\mathrm{Var}[T_j] = \mathrm{E}\left[T_j^2\right] - \mathrm{E}[T_j]^2$$

By the rule for computing the expectation of a function of a random variable, we also know that

$$\mathrm{E}\left[T_j^2\right] = \sum_{x \in Range(T_j)} x^2 \Pr\{T_j = x\}$$

Now we can use the fact that T_j is in the range $[0, 1]$ to say that $x^2 \leq x$, and that therefore

$$\sum_{x \in Range(T_j)} x^2 \cdot \Pr\{T_j = x\} \leq \sum_{x \in Range(T_j)} x \cdot \Pr\{T_j = x\}$$

Thus, $\mathrm{E}\left[T_j^2\right] \leq \mathrm{E}[T_j]$ and

$$\begin{aligned}\mathrm{Var}[T_j] &\leq \mathrm{E}[T_j] - \mathrm{E}[T_j]^2 \\ &\leq \mathrm{E}[T_j]\end{aligned}$$

Since this holds for any j, we can now conclude that

$$\begin{aligned}\mathrm{Var}[T] &= \mathrm{Var}[T_1] + \ldots + \mathrm{Var}[T_n] \\ &\leq \mathrm{E}[T_1] + \ldots + \mathrm{E}[T_n] \\ &= \mathrm{E}[T]\end{aligned}$$

■

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Problems for Recitation 23

Theorem 1. *Let $E_1, \ldots, E_n$ be events, and let X be the number of these events that occur. Then:*

$$\mathrm{Ex}(X) = \Pr\{E_1\} + \Pr\{E_2\} + \ldots + \Pr\{E_n\}$$

Theorem 2 (Markov's Inequality)**.** *Let X be a nonnegative random variable. If $c > 0$, then:*

$$\Pr\{X \geq c\} \leq \frac{\mathrm{Ex}(X)}{c}$$

Theorem 3 (Chebyshev's Inequality)**.** *For all $x > 0$ and any random variable R,*

$$\Pr\{|R - \mathrm{Ex}(R)| \geq x\} \leq \frac{\mathrm{Var}[R]}{x^2}$$

Theorem 4 (Union Bound)**.** *For events $E_1, \ldots, E_n$:*

$$\Pr\{E_1 \cup \ldots \cup E_n\} \leq \Pr\{E_1\} + \ldots + \Pr\{E_n\}$$

Theorem 5 ("Murphy's Law")**.** *If events $E_1, \ldots E_n$ are mutually independent and X is the number of these events that occur, then:*

$$\Pr\{E_1 \cup \ldots \cup E_n\} \geq 1 - e^{-\mathrm{Ex}(X)}$$

Theorem 6 (Chernoff Bounds)**.** *Let $E_1, \ldots, E_n$ be a collection of mutually independent events, and let X be the number of these events that occur. Then:*

$$\Pr\{X \geq c\,\mathrm{Ex}(X)\} \leq e^{-(c \ln c - c + 1)\,\mathrm{Ex}(X)} \qquad \textit{when } c > 1$$

1 Getting Dressed

Sometimes I forget a few items when I leave the house in the morning.

For example, here are probabilities that I forget various pieces of footwear:

left sock	0.2
right sock	0.1
left shoe	0.1
right shoe	0.3

a. Let X be the number of these that I forget. What is $\text{Ex}(X)$?

b. Upper bound the probability that I forget one or more items. Make no independence assumptions.

c. Use the Markov Inequality to upper bound the probability that I forget 3 or more items.

d. Now suppose that I forget each item of footwear independently. Use Chebyshev's Inequality to upper bound the probability that I forget two or more items.

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 $\mathrm{Ex}(X) = 36$. 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.

2 A Financial Crisis

There are a lot of foreign words here, but don't be scared! We will be trying to understand why the subprime mortgage collapse happened!

For a more complete story of how the crisis happened, refer to section 19.5.3 of the text. The following is a set of vocabulary that we will be using:

- A **loan** is money lent to a borrower. If the borrower does not pay on the loan, the loan is said to be in **default**, and collateral is seized. In the case of mortgage loans, the borrower's home is used as collateral.
- A **bond** is a collection of loans, packaged into one entity. A bond can be divided into **tranches**, in some ordering, which tell us how to assign losses from defaults. Suppose a bond contains 1000 loans, and is divided into 10 tranches of 100 bonds each. Then, all the defaults must fill up the lowest tranche before the affect others. For example, suppose 150 defaults happened. Then, the first 100 defaults would occur in tranche 1, and the next 50 defaults would happen in tranche 2.
- The lowest tranche of a bond is called the **mezzanine tranche**.
- We can make a "super bond" of tranches called a **collateralized debt obligation (CDO)** by collecting mezzanine tranches from different bonds. This super bond can then be itself separated into tranches, which are again ordered to indicate how to assign losses.

Armed with this knowledge, we can now solve problems about the crisis!

1. Suppose that 1000 loans make up a bond, and the fail rate is 5% in a year. Assuming mutual independence, give an upper bound for the probability that there are one or more failures in the second-worst tranche. What is the probability that there are failures in the best Tranche?
2. Now, do not assume that the loans are independent. Give an upper bound for the probability that there are one or more failures in the second tranche. What is an upper bound for the probability that the entire bond defaults? Show that it is a tight bound. (Hint: Use Markov's theorem).
3. Given this setup (and assuming mutual independence between the loans), what is the expected failure rate in the mezzanine tranche?
4. We take the mezzanine tranches from 100 bonds and create a CDO. What is the expected number of underlying failures to hit the CDO?
5. We divide this CDO into 10 tranches of 1000 bonds each. Assuming mutual independence, give an upper bound on the probability of one or more failures in the best tranche. The third tranche?
6. Repeat the previous question without the assumption of mutual independence.

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Notes for Recitation 23

Theorem 1. *Let $E_1, \ldots, E_n$ be events, and let X be the number of these events that occur. Then:*

$$\mathrm{Ex}(X) = \Pr\{E_1\} + \Pr\{E_2\} + \ldots + \Pr\{E_n\}$$

Theorem 2 (Markov's Inequality)**.** *Let X be a nonnegative random variable. If $c > 0$, then:*

$$\Pr\{X \geq c\} \leq \frac{\mathrm{Ex}(X)}{c}$$

Theorem 3 (Chebyshev's Inequality)**.** *For all $x > 0$ and any random variable R,*

$$\Pr\{|R - \mathrm{Ex}(R)| \geq x\} \leq \frac{\mathrm{Var}[R]}{x^2}$$

Theorem 4 (Union Bound)**.** *For events $E_1, \ldots, E_n$:*

$$\Pr\{E_1 \cup \ldots \cup E_n\} \leq \Pr\{E_1\} + \ldots + \Pr\{E_n\}$$

Theorem 5 ("Murphy's Law")**.** *If events $E_1, \ldots E_n$ are mutually independent and X is the number of these events that occur, then:*

$$\Pr\{E_1 \cup \ldots \cup E_n\} \geq 1 - e^{-\mathrm{Ex}(X)}$$

Theorem 6 (Chernoff Bounds)**.** *Let $E_1, \ldots, E_n$ be a collection of mutually independent events, and let X be the number of these events that occur. Then:*

$$\Pr\{X \geq c\,\mathrm{Ex}(X)\} \leq e^{-(c \ln c - c + 1)\,\mathrm{Ex}(X)} \qquad \textit{when } c > 1$$

1 Getting Dressed

Sometimes I forget a few items when I leave the house in the morning.

For example, here are probabilities that I forget various pieces of footwear:

left sock	0.2
right sock	0.1
left shoe	0.1
right shoe	0.3

a. Let X be the number of these that I forget. What is $\text{Ex}(X)$?

Solution. By Theorem 1, the expected number of events that happen is the sum of the event probabilities. So,

$$\text{Ex}(X) = 0.2 + 0.1 + 0.1 + 0.3 = 0.7$$

■

b. Upper bound the probability that I forget one or more items. Make no independence assumptions.

Solution. By the Union Bound, the probability that I forget something is at most:

$$0.1 + 0.1 + 0.2 + 0.3 = 0.7$$

■

c. Use the Markov Inequality to upper bound the probability that I forget 3 or more items.

Solution.

$$\Pr\{X \geq 3\} \leq \frac{\text{Ex}(X)}{3} = \frac{7}{30}$$

■

d. Now suppose that I forget each item of footwear independently. Use Chebyshev's Inequality to upper bound the probability that I forget two or more items.

Solution. Let X_1 be the event I bring my left sock, X_2 my right sock, X_3 my left shoe, and X_4 my right shoe. Then $\text{Ex}(X_1) = .2$, $\text{Ex}(X_2) = .1$, $\text{Ex}(X_3) = .1$, and $\text{Ex}(X_4) = .3$. Moreover, since the X_i are Bernoulli random variables (binomial with $n = 1$), we have $\text{Var}[X_1] = .2(1 - .2) = .16, \text{Var}[X_2] = .1(1 - .1) = .09, \text{Var}[X_3] = .1(1 - .1) = .09$, and $\text{Var}[X_4] = .3(1 - .3) = .21$.

Let $X = \sum_{i=1}^{4} X_i$. Then $\mathrm{Ex}(X) = \sum_{i=1}^{4} \mathrm{Ex}(X_i) = .2 + .1 + .1 + .3 = .7$. Since the X_i are independent, $\mathrm{Var}[X] = \sum_{i=1}^{4} \mathrm{Var}[X_i] = .16 + .09 + .09 + .21 = .55$. Now by Chebyshev's Inequality,

$$\begin{aligned}
\Pr\{X \geq 2\} &\leq \Pr\{|X - .7| \geq 1.3\} \\
&= \Pr\{|X - \mathrm{Ex}(X)| \geq 1.3\} \\
&\leq \frac{\mathrm{Var}[X]}{1.3^2} \\
&= \frac{.55}{1.3^2} \\
&\leq .326
\end{aligned}$$

■

e. Use Theorem 5 to lower bound the probability that I forget one or more items.

Solution. Plugging into Theorem 5, the probability that I forget one or more items:

$$1 - e^{-\mathrm{Ex}(X)} = 1 - e^{-0.7} = 0.503\ldots$$

■

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 $\mathrm{Ex}(X) = 36$. Use Chernoff's Bound to give an upper bound on the probability that I remember 48 or more items.

Solution. By the Chernoff Bound,

$$\begin{aligned}
\Pr\{X \geq 48\} &= \Pr\{X \geq (1 + 1/3)\,\mathrm{Ex}(X)\} \\
&\leq e^{-(4/3 \ln 4/3 - 4/3 + 1)\cdot 36} \\
&\approx .1638
\end{aligned}$$

■

h. Give an upper bound on the probability that I remember 108 or more items.

Solution. By the Chernoff Bound,

$$\begin{aligned}
\Pr\{X \geq 108\} &= \Pr\{X \geq 3 \cdot \mathrm{Ex}(X)\} \\
&\leq e^{-(3 \ln 3 - 3 + 1)\cdot 36} \\
&\leq e^{-46} \approx 1.05 \times 10^{-20}.
\end{aligned}$$

■

2 A Financial Crisis

There are a lot of foreign words here, but don't be scared! We will be trying to understand why the subprime mortgage collapse happened!

For a more complete story of how the crisis happened, refer to section 19.5.3 of the text. The following is a set of vocabulary that we will be using:

- A **loan** is money lent to a borrower. If the borrower does not pay on the loan, the loan is said to be in **default**, and collateral is seized. In the case of mortgage loans, the borrower's home is used as collateral.
- A **bond** is a collection of loans, packaged into one entity. A bond can be divided into **tranches**, in some ordering, which tell us how to assign losses from defaults. Suppose a bond contains 1000 loans, and is divided into 10 tranches of 100 bonds each. Then, all the defaults must fill up the lowest tranche before the affect others. For example, suppose 150 defaults happened. Then, the first 100 defaults would occur in tranche 1, and the next 50 defaults would happen in tranche 2.
- The lowest tranche of a bond is called the **mezzanine tranche**.
- We can make a "super bond" of tranches called a **collateralized debt obligation (CDO)** by collecting mezzanine tranches from different bonds. This super bond can then be itself separated into tranches, which are again ordered to indicate how to assign losses.

Armed with this knowledge, we can now solve problems about the crisis!

1. Suppose that 1000 loans make up a bond, and the fail rate is 5% in a year. Assuming mutual independence, give an upper bound for the probability that there are one or more failures in the second-worst tranche. What is the probability that there are failures in the best Tranche?

   Solution. If we assume mutual independence, we can use Chebyshev's Theorem to give an upper bound to this. Because the fail rate is 5%, $EX[X] = 50$. From Chebyshev, we have that

   $$Pr(X \geq 2 \cdot 50) \leq e^{-(2\ln 2 - 2 + 1)\cdot 50}$$

   which evaluates to $1.759 * 10^{-6}$.

   In order for failure to reach the best Tranche, there must have been at least 900 failures. This corresponds to $c = 18$ in the Chernoff bound, so the bound is

   $$Pr(X \geq 18 \cdot 50) \leq e^{-(18\ln 18 - 18 + 1)\cdot 50}$$

   This evaluates to $e^{-2551.33}$, which evaluates to less than 10^{-1000}. ■

2. Now, do not assume that the loans are independent. Give an upper bound for the probability that there are one or more failures in the second tranche. What is an upper bound for the probability that the entire bond defaults? Show that it is a tight bound. (Hint: Use Markov's theorem).

   Solution. From the Markov bound, we have that

   $$Pr(X \geq 100) \leq \frac{50}{100} = 1/2$$

   Applying Markov's theorem again in the case of failures occuring in every bond, we have

   $$Pr(X \geq 1000) \leq \frac{50}{1000} = 1/20$$

   We can see that these bounds are tight by constructing situations where they occur. In the first case, consider if the bond consisted of 100 bonds that always defaulted half the time, and the other bonds never defaulted. In the second case, assume that all the bonds are completely correlated (they all default or all do not default), and default 5% of the time. ■

3. Given this setup (and assuming mutual independence between the loans), what is the expected failure rate in the mezzanine tranche?

   Solution. The expected number of failures per bond is 50. Because all losses are sent to the mezzanine, the expected failure rate is 50% (mighty high!). ■

4. We take the mezzanine tranches from 100 bonds and create a CDO. What is the expected number of underlying failures to hit the CDO?

   Solution. From the previous part, each mezzanine tranche has a 50% rate of failure. The CDO contains 10000 loans, at a 50% rate of failure, and so we expect 5000 failures in the CDO. ■

5. We divide this CDO into 10 tranches of 1000 bonds each. Assuming mutual independence, give an upper bound on the probability of one or more failures in the best tranche. The third tranche?

   Solution. We use a 50% failure rate to model the loans in this CDO. If we assume mutual independence, we use the Chernoff bound to bound this. One or more failures in the best tranche corresponds to over 9000 failures in the CDO. Using the Chernoff bound, we find that

   $$Pr(X \geq 9/5 * 5000) \leq e^{-(9/5 \ln(9/5) - 9/5 + 1) \cdot 5000}$$

   which is approximately zero.

For the third tranche, we make a similar calculation, finding the probability that there are over 7000 failures in the CDO.

$$Pr(X \geq 7/5 * 5000) \leq e^{-(7/5 \ln(7/5) - 7/5 + 1) \cdot 5000}$$

This is about 10^{-155}– that is, still vanishingly small.

■

6. Repeat the previous question without the assumption of mutual independence.

Solution. Without a mutual independence assumption, the best we can do is bounds from Markov's theorem, like we did in part b. Thus, the upper bound on failures in the top tranche, or over 9000 failures, is:

$$Pr(X \geq 9000) \leq \frac{5000}{9000} = 5/9$$

For failures in the third trance, Markov's theorem gives us:

$$Pr(X \geq 7000) \leq \frac{5000}{7000} = 5/7$$

To put it generously, these probabilities are not very low. ■

MIT OpenCourseWare
http://ocw.mit.edu

6.042J / 18.062J Mathematics for Computer Science
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.