Massachusetts Institute of Technology
6.042J/18.062J, Fall '02: Mathematics for Computer Science
Prof. Albert Meyer and Dr. Radhika Nagpal

Final Examination

Your name:____

Circle the name of your Tutorial Instructor:

Adrian Georgi Josh Karen Lee Min Nikos Tina

- The final is **closed book**. There is an **Appendix** with standard definitions.
- There are 12 problems totaling 150 points. Total time is 170 minutes.
- Put your name on the top of **every** page – *these pages may be separated for grading*.
- 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.
- GOOD LUCK!

DO NOT WRITE BELOW THIS LINE

Problem	Points	Grade	Grader
1	10		
2	10		
3	20		
4	10		
5	10		
6	15		
7	15		
8	10		
9	10		
10	20		
11	10		
12	10		
Total	150		

Problem 1 (10 points). For each of the following logical formulas with domain of discourse the natural numbers, $\mathbb{N}$, indicate whether it is a possible formulation of

- I: the **I**nduction Axiom,
- S: the **S**trong Induction Axiom,
- L: the **L**east Number Principle (called the Well-ordering Principle by Rosen), or
- N: **N**one of these.

For example, the Rule for the Least Number Principle in the appendix could be expressed with the following formula, so it gets labelled "L".

$$[\exists n P(n)] \longrightarrow \exists m\, [P(m) \land \forall n\, [P(n) \longrightarrow m \leq n]] \qquad \underline{\quad \textbf{L} \quad}$$

This is a multiple choice problem: do not explain your answer.

(a) (2 points) $[P(b) \land [\forall k \geq b\; P(k) \longrightarrow P(k+1)]] \longrightarrow \forall k \geq b\; P(k)$ ____

(b) (2 points) $[P(0) \land \forall k\, [\forall m \leq k\; P(m)] \longrightarrow P(k+1)] \longrightarrow \forall k\; P(k)$ ____

(c) (2 points) $[\forall n\, [\forall m < n\; P(m)] \longrightarrow P(n)] \longrightarrow \forall n\; P(n)$ ____

(d) (2 points) $[\exists n\; P(n)] \longrightarrow \exists n\, [P(n) \land \forall k < n\; \overline{P(k)}]$ ____

(e) (2 points) $\forall n\, [P(n) \longrightarrow \exists n\, [P(n) \land \forall k\; P(k) \longrightarrow n \leq k]]$ ____

Problem 2 (10 points). An integer, m, *divides* an integer, n, in symbols, $m \mid n$, iff there is an integer k such that $km = n$.

Claim. *For any prime, p, and positive integers $x_1, x_2, \ldots, x_n$, if $p \mid x_1 x_2 \cdots x_n$, then $p \mid x_i$ for some i between 1 and n.*

Underline the sentence where the following proof goes wrong and explain.

False proof. [By strong induction on n.]

The induction hypothesis is the Claim itself.

Base case ($n = 1$): When $n = 1$, we have $p \mid x_1$, therefore we can let $i = 1$ and conclude $p \mid x_i$.

Induction step: Now assuming the claim holds for all $k \leq n$, we must prove it for $n + 1$.

So suppose $p \mid x_1 x_2 \cdots x_{n+1}$. Let $y_n = x_n x_{n+1}$, so $x_1 x_2 \cdots x_{n+1} = x_1 x_2 \cdots x_{n-1} y_n$. Since the righthand side of this equality is a product of n terms, we have by induction that p divides one of them. If $p \mid x_i$ for some $i < n$, then we have the desired i. Otherwise $p \mid y_n$. But y_n is a product of the two terms x_n, x_{n+1}. Therefore, we have by strong induction that p divides one of them. So in this case $p \mid x_i$ for $i = n$ or $i = n + 1$. □

Problem 3 (20 points). Given a simple graph G, we apply the following operation to the graph: pick two vertices $u \neq v$ such that either

1. there is an edge of G between u and v and there is also a path from u to v which does *not* include this edge; in this case, delete the edge $\{u, v\}$.
2. or, there is no path from u to v; in which case, add the edge $\{u, v\}$.

We keep repeating these operations until it is no longer possible to find two vertices $u \neq v$ to which an operation applies.

Assume the vertices of G are the integers $1, 2, \ldots, n$ for some $n \geq 2$. This procedure can be modelled as a state machine whose states are all possible simple graphs with vertices $1, 2, \ldots, n$. The start state is G, and the final states are the graphs on which no operation is possible.

(a) (10 points) For any state, G', let e be the number of edges in G', c be the number of connected components it has, and s be the number of simple cycles. For each of the derived variables below, indicate the *strongest* of the properties that it is guaranteed to satisfy, no matter what the starting graph G is. The choices for properties are: *constant*, *strictly increasing*, *strictly decreasing*, *weakly increasing*, *weakly decreasing*, *none of these*. The derived variables are

(i) $e - s$ ______________________

(ii) $3c/2 + e$ ______________________

(iii) $c + s$ ______________________

(iv) (c, e), partially ordered coordinatewise ______________________

(v) (c, e), ordered lexicographically ______________________

(b) (6 points) Choose a derived variable from above and prove that it is strictly decreasing in some well-founded partial order. Conclude that the procedure terminates.

(c) (4 points) Prove that any final state must be a tree on the vertices.

Problem 4 (10 points). This problem is about game trees for two-player games of perfect information. Player 1 is the player who plays first. Player 1 wins iff Player 2 loses, but the game may end in a draw, or may continue forever.

For each of the parts below, indicate whether the statement is *true* or *false*. If the claim is false, briefly describe a counter-example.

(a) (2 points) If Player 1 has a winning strategy, then the game cannot continue forever.

(b) (2 points) If the game tree is finite-path, then the game cannot continue forever.

(c) (2 points) One of the players must have a non-losing strategy.

For the remaining parts, we only consider games with finite-path game trees.

(d) (2 points) If player 1 has a non-losing strategy, then so must player 2.

(e) (2 points) If the game can only end after an odd number of moves, then Player 1 has an advantage.

Problem 5 (10 points). Consider the seven letter word *armeyer*.

(a) (3 points) How many different sequences of these seven letters are there? __________

(b) (7 points) How many such sequences are there that do not contain either of the words *eye* or *ram*? _______________. Explain your answer.

Problem 6 (15 points). A pizza house is having a promotional sale. Their commercial reads:

> Buy 2 large pizzas at the regular price, and for each pizza get up to 11 different toppings from 11 possible absolutely free (no double toppings). That's $4,194,304$ different ways to design your order!

The ad writer was a former Harvard student who had figured out that $(2^{11})^2 = 4,194,304$. He came up with this number by reasoning that the number of ways to choose different toppings for one pizza is all the possible subsets of the set of 11 toppings, which is 2^{11}. Since there are two pizzas, the total possible combinations of pizzas are $(2^{11})^2$.

Unfortunately, the number $(2^{11})^2$ is actually wrong.

(a) (5 points) Explain what is wrong with the Harvard student's counting.

(b) (3 points) In how many ways can you choose toppings for the two pizzas? ____________

(c) (7 points) In how many ways can you choose toppings for n pizzas? ____________

Problem 7 (15 points). We consider a variation of Monty Hall's game. The contestant must pick one of *four* doors, with a prize randomly placed behind one door and goats behind the other three. Then, instead of always opening a door to reveal a goat, Carol *randomly* opens two of the three doors that the contestant hasn't picked. This means she may reveal two goats, or she may reveal the prize and a goat. If she reveals the prize, then the entire game is *restarted*, that is, the prize is again randomly placed behind some door, the contestant again picks a door, and so on until Carol finally reveals two goats. Then the contestant can choose to *stick* with his original choice of door or *switch* to the remaining unopened door. He wins if the prize is behind the door he last chooses.

(a) (6 points) Let R be the number of times the game is restarted before Carol picks two goats. What is $\mathrm{E}[R]$? __________

(b) (5 points) When Carol finally reveals two goats, the contestant has the choice of sticking or switching. Let's say that the contestant adopts the strategy of sticking. What is the probability that the contestant wins with this strategy? ______

(c) (4 points) For any final outcome where the contestant wins with a "stick" strategy, he would lose if he had used a "switch" strategy, and vice versa. In the original Monty Hall game, we concluded immediately that the probability that he would win with a "switch" strategy was $1 - \Pr\{\text{win with stick}\}$. Why isn't this conclusion quite as obvious for this new, restartable game? Briefly explain why this conclusion still sound.

Problem 8 (10 points). A *simple k-cycle* in a simple graph is an undirected path going through each of k distinct vertices exactly once and ending where it started (where $k > 2$). A simple k-cycle can be represented by the sequence of k distinct vertices $v_1, v_2, \ldots, v_k$ along the path.

Note that every simple k-cycle can be represented by many sequences. For example, the 4-cycle represented by 1234 is as the same cycle represented by 2341 or 3412 because a cycle does not have to start at any particular vertex. It is also represented by 4321, because the cycle is undirected.

(a) (3 points) How many different sequences of vertices represent a given simple k-cycle? ________

(b) (2 points) In a complete graph on n vertices (i.e. all possible edges are present), how many simple k-cycles are there? ________

Suppose we construct a simple n-node graph, $G = (V, E)$, randomly as follows: For every set of two distinct vertices $\{v, v'\}$, toss a biased coin whose probability of coming up heads is p. The undirected edge between v and v' is included in E iff the coin comes up heads. Assume that all coin tosses are mutually independent.

(c) (5 points) What is the expected number of simple k-cycles in G? (Note: If you were not able to solve part (b), you may let b denote the answer to part (b) and express your answer in terms of b, k, and p). ________

Problem 9 (10 points). Let R be a positive integer valued random variable such that

$$f_R(n) = \frac{1}{cn^3},$$

where

$$c ::= \sum_{n=1}^{\infty} \frac{1}{n^3}.$$

(a) (5 points) Prove that $\mathrm{E}[R]$ is finite.

(b) (5 points) Prove that $\mathrm{Var}[R]$ is infinite.

Problem 10 (20 points). Suppose that you are given two biased coins C_p and C_r. Coin C_p flips a head with probability p and coin C_r flips a head with probability r.

The game is to flip C_p and then C_r and keep alternating between coins until you get $2l$ heads in a row, where l is a given positive integer. However, as soon as you get a tail, then you start over again by first flipping C_p and then alternating between coins[1]. Assume that coin flips are mutually independent of each other.

(a) (10 points) For $l = 3$, write an expression in p and r for the the expected number of coin flips to see $2l$ heads in a row.

Hint: It may be more convenient to use the special expectation formula for natural number valued variables (in the Appendix).

[1] For example, if $l = 2$ and the sequence of flips was $HTTHHTHHHH$, then the game took 10 flips and the sequence of coins used was $C_pC_rC_pC_pC_rC_pC_pC_rC_pC_r$

(b) (10 points) What is the expected number of coin flips to see $2l$ heads in a row?

Note: You may express your answer using summations determined by l, but if you do, you should briefly indicate how results from the course imply that there are closed forms for your summations.

Problem 11 (10 points). **(a) (5 points)** Let R be an indicator variable for getting a head on a flip of a fair coin. Calculate the bound on $\Pr\{R \geq 1\}$ using

(i) Markov Bound ____________

(ii) Chebyshev Bound ____________

(iii) Chernoff Bound ____________

Note: $e^{\ln 2} \approx 2$

(b) (5 points) Prove that $\text{Var}[X+Y] = \text{Var}[X] + \text{Var}[Y]$, if X, Y are independent. Be sure to justify each step in your proof.

Problem 12 (10 points). A test tube with a bacterial culture is delivered to a Lab technician at a hospital. Her task is to estimate the percentage of bacteria in the culture that are resistant to penicillin. She will do this by sampling a certain number of bacteria and testing them. Assume her sampling technique is good, *i.e.*, the bacteria are sampled randomly and independently with replacement. Let the percentage of antibiotic-resistant bacteria be b.

(a) (6 points) Write a formula for a number, n, of samples she could take in order to be C% confident that her estimation was no more than d away from the actual value b. The formula for n must not, of course, involve the unknown quantity, b.

(b) (4 points) The lab technician is supposed to perform her task daily. She is instructed to warn the medical staff whenever her tests show with 99% confidence that the percentage of penicillin-resistant bacteria in that day's tube exceeds a specified danger threshold.

In early April, the staff receives its first warning of the year from her, and a new resident on the staff orders the nurses to review the medication of all patients taking penicillin, because he realizes there is a very high probability of a dangerous level of penicillin-resistance bacteria in the hospital. When a more senior physician arrives, he overrules the residents' order.

Briefly indicate how you would justify the senior physician's decision to the resident.