Massachusetts Institute of Technology
6.042J/18.062J: Mathematics for Computer Science
Professors Nancy Lynch and Tom Leighton

FINAL EXAM
December 17, 1996

FINAL EXAM

- This exam is closed book. You may use three handwritten crib sheets. **No calculators allowed. You may leave your answers in terms of the ratio of products of numbers and binomial coefficients**.
- Put your name on the top of each page.
- Write your solutions in the space provided. If you need more space, write on the back of the sheet containing the problem.
- Show your work, as partial credit will be given. Full credit will not be given for unjustified answers. **Be neat**.

Problem	Points	Grade
1	10	
2	10	
3	10	
4	15	
5	15	
6	10	
7	10	
8	18	
9	10	
10	12	
11	15	
Total	135	

Name:____

Please circle your tutorial TA's name: Dimitri Gunnar

Please circle your tutorial time: 10 11 12 1 2 3

Problem 1. [10 points]

Recall the definition of Fibonacci numbers:

$$\begin{aligned} F_0 &= 0 \\ F_1 &= 1 \\ F_n &= F_{n-1} + F_{n-2} \quad \text{for } n \geq 2 \end{aligned}$$

(a) Prove by induction that $3|(F_{n+8} - F_n)$, for $n \geq 0$.

(b) Prove that $3|F_n$ if and only if $4|n$, for $n \geq 0$. (*Hint*: Use the result from part (a).)

Problem 2. [10 points]

(a) Six people communicate via radio frequency with one another. Each pair of people choose one of **two** available frequencies with which to communicate. Show that there is a subset of 3 people who are each using the identical frequency with which to communicate with others in the subset.

(b) Now suppose there are seventeen people communicating via radio frequency with one another. Each pair of people choose one of **three** available frequencies with which to communicate. Show that there is a subset of 3 people who are each using the identical frequency with which to communicate with others in the subset.

Problem 3. [10 points]

We consider an algorithm based on a given rooted, undirected tree, oriented with the root on the top. Every node in the tree has an associated *value* variable. Initially, the *value* variables of the leaves contain natural number input values, and the *value* variables of the root and other internal nodes contain a special "undefined" value $\perp$. The algorithm is supposed to end up with the root's *value* variable containing the sum of all the *value*s of all the leaves. This is accomplished by propagating summarized *value* information up the tree structure.

This algorithm is described precisely by the following state machine. This machine assumes a fixed undirected tree $T = (V, E)$, rooted at node r. The state components are:

for each $v \in V$:
$value(v) \in \mathbb{N} \cup \{\perp\}$; initially $value(v) \in \mathbb{N}$ if v is a leaf, and $value(v) = \perp$ if v is an internal node.

The transitions are:

summarize(v), v an internal node
- Precondition:
  - $value(v) = \perp$
  - for every $w \in children(v)$
    - $value(w) \in \mathbb{N}$
- Effect:
  - $value(v) := \sum_{w \in children(v)} value(w)$

The property we are interested in is that $value(r)$ for the root r, if it ever gets defined, is equal to the sum of the *value*s of all the leaves. Formally:

Inv1: If $value(r) \in \mathbb{N}$, where r is the root, then $value(r) = \sum_{l \in leaves} value(l)$.

State and prove a strengthened version of Inv1, which can be proved by induction.

Problem 4. [15 points]

The following three parts are unrelated.

(a) A vampire arrives in Boston at day 0 and starts biting people at day 1. People bitten become vampires themselves. New vampires bite two persons on the next day after they were bitten and four persons every day afterwards. Write an expression for the number of vampires in n days for $n \geq 2$, assuming the first vampire is a newly made vampire. *Do not solve the recurrence relation.*

(b) Obtain Tight (Θ) asymptotic bounds for the recurrence

$$T(n) = 2T(\lceil n/2 \rceil) + \frac{8}{9}T(\lfloor 3n/4 + \sqrt{n} \rfloor) + n^2, \quad \text{for } n \geq 100$$

You may assume that $T(n) = \Theta(1)$ for $n = O(1)$.

(c) Find tight (Θ) asymptotic bounds for the summation

$$\sum_{i=n}^{\infty} \frac{1}{i^2}$$

Problem 5. [15 points]

In a (five-card stud) poker game, two players A and B are dealt 5 cards each at random from the *same* standard deck of 52.

(a) A *straight flush* consists of 5 consecutive cards from the same suit. (Ace, 2, 3, 4, 5 is the lowest and 10, Jack, Queen, King, Ace is the highest, so an Ace may not be in the middle of a straight flush.) What is the probability that player A gets a straight flush?

(b) A *four-of-a-kind* consists of 4 cards of the same rank (for example, four 7s or four Kings). The fifth card can be anything. What is the probability that player B gets a four-of-a-kind?

(c) What is the conditional probability that player B gets a four-of-a-kind *given* that player A gets a straight flush?

(d) What is the probability that player A gets a straight flush *and* player B gets a four-of-a-kind?

(e) What is the probability that player A gets a straight flush *given* player B gets a four-of-a-kind?

Problem 6. [10 points]

(a) At Marilyn's Casino, a new gambling game has been introduced. The game features 3 fair mutually independent 6-sided dice. To play, the contestant picks a number in the interval $[1, 6]$. The contestant wins \$1 if his/her number comes up on one or more of the dice and loses \$1 otherwise. This game has proved to be very popular with the gamblers such as Bobo who is a regular. Your job is to figure out the probability of winning the game.

(b) In an effort to attract even more gamblers to her casino, Marilyn is thinking about a variation of the game in which the bettor also picks a second chance number and the bettor wins if the first pick comes up at least once or the second pick comes up at least twice. Compute the probability of winning this game.

Problem 7. [10 points]

For a certain test for a newly discovered virus, called JABOLA, it is known that 99% of the people carrying the virus will test positive, whereas only 1% of the people not carrying the virus will test positive. Suppose 0.5% of the population carries the JABOLA virus. Given that a person picked at random tests positive, what is the conditional probability that this person carries the JABOLA virus?

Problem 8. [18 points]

What is your expected return in each of the following games? You may assume the events are independent in each part.

(a) You get to roll six ordinary six-sided dice. You win the number of dollars equal to the total number of spots that appear on all the dice.

(b) You roll one ordinary die. You win a number of dollars that is equal to six times the number of spots.

(c) You toss a fair coin six times. For each time it comes up heads, you toss an ordinary six-sided die. You win the number of dollars equal to the total number of spots that appear on all the dice.

(d) You toss a fair coin repeatedly, until it comes up "tails" for the first time. (This could involve infinitely many tosses, if "tails" never comes up.) For each time the coin comes up heads, you toss an ordinary six-sided die. You win the number of dollars equal to the total number of spots that appear on all the dice.

(e) You roll two six-sided dice. You win the product of the number of spots that appear on the two dice.

(f) You roll one ordinary die. You win a number of dollars that is the square of the number of spots.

Problem 9. [10 points]

There is a newly developed test called EQ (*Emotional Quotient*), that is used to decide whether or not you will succeed in life. EQ is expressed as a nonnegative integer. The average EQ over all humans is 1000.

(a) Knowing only the average EQ, what fraction of humans can have EQ$\geq$ 1500?

(b) Assuming that the standard deviation of EQ is 100, what fraction of humans can have EQ$\geq$ 1500?

Problem 10. [12 points]

A pentium chip has 753 components. Past experience with Intel engineers indicates that 225 of these components fail with probability 1/15, 400 fail with probability 1/10, and 128 fail with probability 1/8. Assume that all component failures are independent.

(a) What is the expected number of components that will fail on a pentium chip?

(b) Compute the variance for the number of components that will fail on a pentium chip.

(c) Derive a formula for the exact probability that a pentium chip has no failures.

(d) Derive an upper bound of the form e^{-x}, where x is an integer, on the probability that a pentium chip has no failures.

Problem 11. [15 points]

In an effort to relieve congestion on the Internet, ARPA is considering building SUPERNET. SUPERNET is a heierarchical routing network for which the country is divided into 6 regions: New England (NE), New York (NY), Washington DC (DC), the south (SO), Los Angeles (LA), and the rest of the country (RE). The 6 regions will be connected with 9 ultra-high bandwidth, bidirectional links so that region u is directly connected to region v if and only if $u \in \{\text{NE,NY,DC}\}$, and $v \in \{\text{SO,LA,RE}\}$.

(a) Draw the graph corresponding to the ultra-high bandwidth connections.
Is the graph planar? Is it bipartite?

(b) When a message on the SUPERNET needs to go from one region to another, it travels on a shortest path through the ultra-high bandwidth network. In the event that there is more than one shortest path, then a shortest path is picked uniformly at random from among the shortest paths. (Note that there is always a unique shortest path or exactly three shortest paths.) ARPA engineers believe that there will be 15 million messages per second emanating from each region and that each such message has a random destination (i.e., each message is destined for each of the 5 other regions with equal probability–the message cannot stay within the region). What is the expected message load on each ultra-high bandwidth connection? When counting load, make sure to count messages that are travelling in either direction. (*Hint*: The expected load on each edge is the same.)

(c) Because of budget cutbacks, ARPA engineers have been forced to do without buffers and so if ever the message traffic across a link during a second of time exceeds its capacity, the entire SUPERNET will come to a crashing halt. The president has decreed that the probability of such a crash during the next 4 years should be at most 10^{-9}. ARPA has the option of buying links with capacity $\sqrt{E}, E/2, .99E, E, 1.01E, 2E, 2.718E$, and E^2 (where E is the expected load that you computed in part (b) above), but the cost of a link goes up with its capacity. Your job is to decide the least expensive link that will fit within the President's mandate. You may assume that the destinations of each of the messages are mutually independent. There are at most e^{19} seconds in 4 years.

Which capacity should you choose? Show your work and justify your answer. Compute an upper bound of the form e^x, where x is an integer on the probability that SUPERNET crashes in the next 4 years given your choice of link capacity. (You may use the following facts: $e^{-21} < 10^{-9} < e^{-20}$, $\ln(1.01) + 1/1.01 = 1.00005$, $\ln 2 + 1/2 = 1.19$. These estimates are close enough for government work.)