CS174 **Midterm 2** **Spring 99**
J. Canny **April 6**

This is a closed-book exam with 4 questions. You are allowed to use the 4 sides of notes that you brought with you. The marks for each question are shown in parentheses, and the total is 80 points. Make sure you allocate enough time to attempt all the questions. Write all your answers in this booklet. Good Luck!

NAME____

SID Number____

1. (20 points) Give brief answers to the following:

   (a) (4 points) In the byzantine agreement algorithm, do all good processors set their output bits permanently on the same round?

   (b) (4 points) The data-punctuated token trees from class don't work on a string of the same character repeated many times. How would you fix that?

   (c) (4 points) In the fingerprint algorithms from class, the sequence of characters is converted to a polynomial, evaluated, and then reduced modulo a prime. Why is a prime used rather than a non-prime modulus?

   (d) (4 points) What is the minimum number of edges in a spanning forest on n vertices?

   (e) (4 points) In choice coordination with two processors, do all processors halt in the same round?

2. (20 points) Suppose we run the randomized routing algorithm from class on a hypercube with $N = 2^n$ processors. We use bit-fixing, so that at the i^{th} step, the i^{th} bit of a packet's address $a_1 a_2 \cdots a_n$ is set to the destination address.

   (a) (4 points) What is the expected number of collisions during the second time step? (note: there are no collisions during the first time step, but more than one packet may arrive at the same processor in the first time step)

   (b) (4 points) What is the probability that two randomly chosen routes share at least one edge? Give an upper bound.

   (c) (8 points) Let X_{ij} be a random variable that counts the number of edges shared by routes of packets v_i and v_j *given* that these two routes share an edge, so $X_{ij} > 0$. What is the distribution of X_{ij}? You can assume the destination addresses for v_i and v_j are random bit strings.

   (d) (4 points) What is $E[X_{ij}]$ for X_{ij} as defined in the last question?

3. (20 points) Consider the caching problem where the cache holds k items. Assume memory items are identified with positive integers. Let the following sequence of N requests occur: $(1, 2, 3, \ldots, 2k, 1, 2, 3, \ldots, 2k, \ldots)$, that is, the sequence $(1, \ldots, 2k)$ repeated until there are N requests (assume $N >> k$).

   (a) (15 points) How many misses will the optimal offline algorithm MIN make on this sequence as a function of N?

   (b) (5 points) Is there a sequence of requests for $2k$ distinct memory items that would produce more misses for MIN?

4. (20 points) In the algorithm for computing minimum cuts from class, we used contraction to simplify the graph. We showed that the probability that a randomly-selected edge lies in the minimum cut is $\leq 2/n$.

   (a) (8 points) Suppose we pick $n/2$ edges independently and at random without contracting any of them. What is the probability that none of them is in the min-cut? Assume n is very large, and simplify your answer.

   (b) (6 points) Suppose we now contract all of those edges. Give lower and upper bounds on the number of vertices in the contracted graph.

   (c) (6 points) What is the expected number of vertices in the contracted graph after (b) above? Hint: the MIN-CUT algorithm from lecture samples edges randomly from a contracted graph. That's the same as sampling as in part (a) above, and discarding redundant edges, i.e. edges that join vertices that have already been contracted. So you can assume that edges were sampled from the contracted graph. Then use the formula derived in class (or derive it yourself) for the probability of avoiding the min cut as a function of the number of vertices t remaining in the contracted graph.