CS 70 Discrete Mathematics for CS
Spring 2000 Wagner Sample Final

You will have three hours for the final itself. This sample final is similar in style and scope to the actual final; however, it is rather longer than the actual final. The exam is open-book, open-notes. We have not assigned points to the questions (but we will do this for the final itself).

1. **(?? pts.) The longest common subsequence problem**
In this question, we will need an induction principle for *pairs of strings* over an alphabet Σ. We say that a pair of strings (a',b') is "less than" another pair (a,b), written $(a',b') \prec (a,b)$, if $|a'|+|b'| < |a|+|b|$, where $|a|$ denotes the length of string $a \in \Sigma^*$. Our induction principle is the following, where $P(\cdot,\cdot)$ is a proposition that applies to pairs of strings:

   - if $\forall a,b \in \Sigma^*$

     $$[\forall a',b' \in \Sigma^*,\ \ (a',b') \prec (a,b) \Rightarrow P(a',b')] \Rightarrow P(a,b)$$

     then $\forall a,b \in \Sigma^*\ P(a,b)$.

   Now consider the following problem. A *subsequence* of a string a is any sequence that can be obtained by deleting one or more (not necessarily consecutive) letters from a, leaving the order of the remaining letters unchanged. A *longest common subsequence* (written *lcs*) of two strings a and b is a sequence of maximum length that is a subsequence of both a and b. Thus for example a lcs of $a =$ `GACTTACCCAGT` and $b =$ `GTTATGTACA` is `GATTACA`; a lcs of $a =$ `ATT` and $b =$ `GCG` is the empty string λ.

   Suppose we want to compute a lcs of two given strings, a and b. [This problem arises, for example, in DNA analysis and in file comparison utilities.] To make the notation simpler, we'll actually just compute the *length* of a lcs (which is essentially the same problem). Here is a suggested algorithm:

   ```
   algorithm lcs_length(a,b)
     if a = λ or b = λ then return(0)
     else if head(a) = head(b) then return(
       max{1 + lcs_length(tail(a),tail(b)), lcs_length(a,tail(b)), lcs_length(tail(a),b)})
     else return(max{lcs_length(a,tail(b)), lcs_length(tail(a),b)})
   ```

   Your task is to prove that this algorithm is correct, using the induction principle above.

   (a) Write down the proposition $P(a,b)$ that you need to prove for all $a,b \in \Sigma^*$.

   (b) What are the *base cases* for the induction principle in this application? [Hint: There are infinitely many of them!]

   (c) Prove $\forall a,b\ P(a,b)$ using the induction principle given above. Be careful to justify each step in your argument.

2. **(?? pts.) Propositional logic**
Prove each of the following statements:

(a) Given n propositional variables, there are exactly 3^n logically distinct (nonequivalent) 1-CNF expressions. [Reminder: k-CNF is a restriction of CNF in which each clause has exactly k literals.]

(b) If $A \models B$, then $A \models B \vee C$, for any Boolean expressions A, B, C.

(c) If the terms of a DNF expression D_1 are a subset of the terms of a DNF expression D_2, then $D_1 \models D_2$. [Reminder: a *term* is a conjunction of literals.]

(d) Suppose D_1, D_2 are DNF expressions with the property that, for every term T_i of D_1, there is a term T_j in D_2 such that the literals of T_j are a subset of the literals of T_i; then $D_1 \models D_2$.

(e) For every Boolean expression, there is an equivalent Boolean expression that uses only the NOR operator. [Hint: Use induction over Boolean expressions.]

3. (?? pts.) A 300-year-old puzzle

Consider the following game played by our old friends Alice and Bob. Alice rolls six fair dice and wins \$1 if she gets at least one six (and wins nothing otherwise). Bob rolls twelve fair dice and wins \$1 if at least two of them come up six (and wins nothing otherwise).

(a) What is the expected number of sixes rolled by Alice?

(b) What is the expected number of sixes rolled by Bob?

(c) Which of the two players, if any, has a higher chance of winning something? [Hint: You may find it easier to calculate the probability of *not* winning.]

[Note: Apparently the great man of letters, Samuel Pepys, asked Sir Isaac Newton the question in part (c) above way back in 1693. Despite years of effort, Newton was unable to convince Pepys of the correct answer.]

4. (?? pts.) Probability; short answers

Let X and Y be integer-valued random variables on the same probability space, with expectations $\mathrm{E}(X) = \mathrm{E}(Y) = 1$. Which of the following statements is/are always true about X and Y? If you believe the statement to be true, give a *brief* justification referring to results from class; otherwise, give a *simple* counterexample.

(a) $\mathrm{E}(2X + Y) = 3$.

(b) $\mathrm{E}(XY) = 1$.

(c) $\mathrm{Var}(2X) = 4\mathrm{Var}(X)$.

(d) $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$.

(e) $\Pr[X \geq 3] \leq \frac{1}{3}$.

(f) $\Pr[X \geq 3] \leq \frac{1}{4}\mathrm{Var}(X)$.

(g) $\Pr[X \geq 1] > 0$.

(h) $\Pr[X \geq 0 \vee Y \geq 0] \leq \Pr[X \geq 0] + \Pr[Y \geq 0]$.

(i) $\Pr[X \geq 0 \wedge Y \geq 0] \leq \Pr[X \geq 0] \times \Pr[Y \geq 0]$.

(j) If $\mathrm{Var}(X) > \mathrm{Var}(Y)$ then $\Pr[|X| > |Y|] > 0$.

5. (?? pts.) Decomposition of variances

The number of job offers received by graduating seniors at Berkeley is determined largely by their department. Suppose that, among students graduating from department A, the number of job offers received has mean μ_0 and variance σ_0^2. Among students graduating from all other departments, the

number of job offers received has mean μ_1 and variance σ_1^2. Students graduating from department A constitute a fraction p of all graduating seniors at Berkeley, so the mean number of job offers among all graduates is $p\mu_0 + (1-p)\mu_1$. In this problem, we will see how the variance of the number of job offers among the population of all Berkeley graduates can also be found from the above quantities, and that it decomposes as a sum of a "between-department" variance term, and two "within-department" variance terms.

Suppose a graduating senior is chosen uniformly at random. Let the r.v. X denote the number of job offers received by this senior, and define the r.v. Z by

$$Z = \begin{cases} \mu_0 & \text{if the senior is from Department A;} \\ \mu_1 & \text{otherwise.} \end{cases}$$

[Note that $\mathrm{E}(Z) = \mathrm{E}(X)$; you might find it helpful to verify this.]

Show that

$$\mathrm{Var}(X) = \mathrm{Var}(Z) + p\sigma_0^2 + (1-p)\sigma_1^2.$$

The first term on the right-hand side is the "between-department" variance, and the last two terms are the "within-department" variances.

6. (?? pts.) Arithmetic/polynomials; true or false

For each of the following statements, say whether the statement is true or false. You need not provide any justification for your answer; however, you may be awarded partial credit for an incorrect answer if you attempt a justification. (And, of course, an attempt at a justification may help you to find an error in your reasoning.)

(a) For all integers a, b, d, if $ad = bd \bmod n$ then $a = b \bmod n$

(b) For all integers a, b, d, if $ad = bd \bmod n$ then $a = b \bmod \frac{n}{\gcd(n,d)}$.

(c) For all integers a, b, if $a = b \bmod m$ and $a = b \bmod n$ then $a = b \bmod nm$.

(d) There is a known polynomial-time algorithm for finding the prime factors of a given number.

(e) If n is a prime and $a > 0$, then it is always the case that $a^n = a \bmod n$.

(f) If n is not a prime and $a > 0$, then it is always the case that $a^n = a \bmod n$.

(g) If we take all multiples of 3 mod 55, then we get a permutation of the numbers $0, 1, 2, \ldots, 54$.

(h) If we take all multiples of 33 mod 55, then we get a permutation of the numbers $0, 1, 2, \ldots, 54$.

(i) In any field, a polynomial of degree n cannot have more than n roots.

(j) In any field, a polynomial of degree n cannot have fewer than n roots.

7. (?? pts.) RSA

In the RSA encryption scheme, suppose that the two primes are $p = 5, q = 11$, and that the encryption exponent is $e = 3$. Answer the following questions.

(a) What is the decryption exponent d? Show your working.

(b) How would Alice send to Bob the message $m = 6$?

(c) How would Bob sign the message $m = 2$?

(d) Suppose that Alice wants to send to Bob the message $m = 10$. What goes wrong? Is this a problem that limits the applicability of the RSA scheme?