CS 70 Spring 2008 | Discrete Mathematics for CS David Wagner | **Final Soln**

Sample solutions for the Final exam

Problem 1. [True or false] (8 points)

Circle TRUE or FALSE. Do not justify your answers on this problem.

(a) TRUE or FALSE: If $x, y \in \mathbb{N}$ and $2^{z-1} < y < x < 2^z$, then we can compute $x \times y$ using at most $O(z)$ bit operations.

(The elementary-school algorithm requires $\Theta(z^2)$ bit operations, since x and y are z-bit numbers.)

(b) TRUE or FALSE: If $x, y \in \mathbb{N}$ and $2^{z-1} < y < x < 2^z$, then we can compute $\gcd(x,y)$ in at most $O(z^4)$ bit operations by using the Euclidean algorithm.

(The Euclidean algorithm only requires $O(z^2)$ bit operations, per the lecture notes, and $O(z^2) \subseteq O(z^4)$.)

(c) TRUE or FALSE: 10 has a multiplicative inverse modulo 14.

(False, since $\gcd(10,14) \neq 1$.)

(d) TRUE or FALSE: There are at most $O(n^2)$ different paths in any undirected graph with n vertices.

(There are infinitely many such paths.)

(e) TRUE or FALSE: If $X_1, \ldots, X_n$ are random variables and $X = X_1 + \cdots + X_n$, then $\mathbf{E}[X] = \mathbf{E}[X_1] + \cdots + \mathbf{E}[X_n]$ is guaranteed to hold, whether or not $X_1, \ldots, X_n$ are mutually independent.

(Linearity of expectation.)

(f) TRUE or FALSE: If $X_1, \ldots, X_n$ are random variables and $X = X_1 + \cdots + X_n$, then $\text{Var}(X) = \text{Var}(X_1) + \cdots + \text{Var}(X_n)$ is guaranteed to hold, whether or not $X_1, \ldots, X_n$ are mutually independent.

(This equation is valid only if $X_1, \ldots, X_n$ are pairwise independent.)

Problem 2. [Grade this proof] (4 points)

Read this proof:

> **Theorem:** For all $n \geq 4$, we have $2^n < n!$.
>
> **Proof**: We will use proof by mathematical induction on n.
> *Base case:* If $n = 4$, then $2^4 = 16 < 24 = 4!$.
> *Inductive hypothesis:* Suppose we have $2^n < n!$ for some n with $n \geq 4$.
> *Inductive step:* We must show that $2^{n+1} < (n+1)!$, so we calculate:
>
> $$2^{n+1} = 2 \times 2^n < 2 \times n! < (n+1) \times n! = (n+1)!,$$
>
> where we used the inductive hypothesis as well as the fact that $2 < n+1$ if $n \geq 4$. Thus, we have shown that $2^n < n! \implies 2^{n+1} < (n+1)!$ for every $n \geq 4$, so by the principle of mathematical induction, we see that $2^n < n!$ holds for every $n \geq 4$. □

Now answer the question(s) below:

(a) TRUE or FALSE: This proof is valid.

Problem 3. [Grade this proof, too] (4 points)

Read this proof:

> **Theorem:** We have $2n \leq n^2 + 1$ for all $n \in \mathbb{N}$.
>
> **Proof**: We will prove this by simple induction on n. Let $P(n)$ denote the proposition that $2n \leq n^2 + 1$.
> *Base case:* If $n = 0$, then $2n = 0 \leq 0^2 + 1 = n^2 + 1$, so $P(0)$ is true.
> *Inductive hypothesis:* Assume $P(n)$ is true for some $n \in \mathbb{N}$. That is, we assume $2n \leq n^2 + 1$.
> *Inductive step:* We must show that $P(n+1)$ is true. Now
>
> $$2(n+1) = 2n + 2 \leq n^2 + 1 + 2 \leq (n+1)^2 + 1,$$
>
> where we have used the inductive hypothesis as well as the fact that $n^2 + 2 \leq (n+1)^2$. We see that $P(n) \implies P(n+1)$ holds for every $n \in \mathbb{N}$, so by the principle of mathematical induction, $P(n)$ is true for every $n \in \mathbb{N}$, and the theorem follows. □

Now answer the question(s) below:

(a) TRUE or **[FALSE]**: This proof is valid.

(b) If you think this proof is invalid (i.e., if you answered False to part (a)), explain where the logical error in the proof lies.

Solution: The error is in the claim that $n^2 + 2 \leq (n+1)^2$. If $n = 0$, this claim is faulty (it's not true that $0 + 2 \leq (0+1)^2$).

Comment: The theorem is in fact true, and the claim $n^2 + 2 \leq (n+1)^2$ *is* true for all $n \geq 1$. So if the proof were modified to add the caveat that $n^2 + 2 \leq (n+1)^2$ is only claimed true for $n \geq 1$, and if we added an additional base case for $n = 1$, then the proof would be valid.

Problem 4. [Short answer] (11 points)

Do not justify your answers to this problem. Do not show your work. Just show us your final answer.

(a) You flip a fair coin twice. Let A denote the event that at least one flip comes up heads and B the event that both flips come up heads. Calculate $\Pr[A]$.

Solution: $\frac{3}{4}$.

(b) With the same notation as part (a), calculate $\Pr[A|B]$:

Solution: 1.

(c) With the same notation as part (a), calculate $\Pr[B|A]$:

Solution: $\frac{1}{3}$.

(d) Flip a fair coin $2n$ times. Find a formula for the probability that we get exactly n heads and n tails.

Solution: $\binom{2n}{n}\frac{1}{2^{2n}}$.

(e) Define the sequence $a(0), a(1), a(2), \ldots$ by $a(0) = 0$, $a(1) = 1$, and

$$a(n) = a(n-1) + a(n-2) + 1 \qquad \text{for } n \geq 2.$$

Remember that the Fibonacci numbers are defined by $F(0) = 0$, $F(1) = 1$, and

$$F(n) = F(n-1) + F(n-2) \qquad \text{for } n \geq 2.$$

Suppose we want to prove that $a(n) \leq F(n+2)$ for all $n \in \mathbb{N}$, using *simple* induction over n with some proposition $P(n)$. Show the definition of the proposition $P(n)$ that we should use. (Don't write down your proof here, just show us the proposition $P(n)$ you would use.)

Solution: $a(n) \leq F(n+2) - 1 \ \wedge \ a(n+1) \leq F(n+3) - 1$.

Alternate solution: $a(n) = F(n+2) - 1 \ \wedge \ a(n+1) = F(n+3) - 1$.

Alternate solution: $a(0) \leq F(2) - 1 \ \wedge \ a(1) \leq F(3) - 1 \ \wedge \ \cdots \ \wedge \ a(n) \leq F(n+2) - 1$.

Alternate solution: $a(0) = F(2) - 1 \ \wedge \ a(1) = F(3) - 1 \ \wedge \ \cdots \ \wedge \ a(n) = F(n+2) - 1$.

Comment: Many other solutions are possible. This problem required both strengthening the hypothesis and converting strong induction to simple induction.

(f) Call an infinite sequence $a_1, a_2, a_3, \ldots$ of integers *good* if only finitely many of the a_is are nonzero. Multiple-choice: which of the following is true? Circle your choice.

(i) The set of all good sequences is finite.
(ii) **[The set of all good sequences is countably infinite.]**
(iii) The set of all good sequences is uncountably infinite.

Problem 5. [A Monty Hall Variant] (12 points)

Tired of hosting the same game year after year, Monty Hall decided to make some changes to his game. There are still three doors, but now one contains 1000 dollars, one contains 500 dollars, and one contains 0 dollars, with the order of the prizes randomly permuted. The contestant first selects a door. Then she has the choice of paying X dollars for Monty to open, among the two unchosen doors, the one that contains the smaller amount of money. If the contestant paid Monty, she then has the choice of switching to the other unopened door.

(a) Suppose the contestant refuses to pay Monty. In this case, what is the expected value of her prize?

Solution: $\frac{0 + 500 + 1000}{3} = \$500$, since all three prize amounts are equally likely.

(b) Suppose that the contestant decides to pay, and then Monty opens a door that contains \$500. Given this, what is the expected value of her prize if she switches?

Solution: \$1000, since the only way that the chosen door can show \$500 is if the other unchosen door contains a larger amount, i.e., \$1000.

Given this, what is the expected value of her prize if she sticks with her original door?

Solution: \$0, since the two unchosen doors must contain the \$500 and \$1000 prizes.

Multiple-choice: which of the following best describes her optimal strategy, in this situation, assuming she wants to maximize her profits? Circle your choice.

(i) She should stick with her initial door—that's strictly better than switching.

(ii) It doesn't matter whether she switches or sticks.

(iii) $\boxed{\text{She should switch doors—that's strictly better than sticking.}}$

(c) Now for a different scenario: Suppose that the contestant pays, and then Monty opens a door that contains $0. Given this, what is the expected value of her prize if she switches?

Solution: $\frac{500+1000}{2} = \$750$. The other unchosen door must contain either the $500 or $1000 prize, and both possibilities are equally likely.

Given this, what is the expected value of her prize if she sticks with her original door?

Solution: $\frac{500+1000}{2} = \$750$. The door she chose must contain either the $500 or $1000 prize, and both possibilities are equally likely.

Multiple-choice: which of the following best describes her optimal strategy, in this situation, assuming she wants to maximize her profits? Circle your choice.

(i) She should stick with her initial door—that's strictly better than switching.

(ii) $\boxed{\text{It doesn't matter whether she switches or sticks.}}$

(iii) She should switch doors—that's strictly better than sticking.

Justification: Her expected prize is the same either way, in this scenario.

(d) Now suppose a second contestant, Bob, decides in advance that he will always pay and always switch to the unopened door (no matter what he sees behind the door that Monty opens). What is the overall expected value of his prize, with this strategy?

Solution: $\frac{1000+1000+500}{3} = \$833.33\ldots$. There are three cases: (i) If Bob initially chooses the $0 door, Monty shows him $500 and after switching he takes home $1000; (ii) If Bob initially chooses the $500 door, Monty shows him $0 and after switching he takes home $1000; (iii) If Bob initially chooses the $1000 door, Monty shows him $0 and after switching he takes home $500. These three cases are equally likely.

(e) What is the most money Monty can charge for opening one of the two unchosen doors and still make it on average profitable for the contestant to pay Monty?

Solution: Monty can charge any amount less than $\$833.33\cdots - \$500 = \$333.33\ldots$ and it will still be profitable for the contestant to pay Monty, since this is the difference between the answers to part (d) and part (a). In other words, Monty can choose any price $X < \$333.33\ldots$.

Problem 6. [Counting] (7 points)

For the purposes of this problem, a telephone number is an arbitrary sequence of 7 decimal digits. Do not justify your answer on this problem. You can leave your answer as an unevaluated expression.

(a) Call a telephone number *non-repetitious* if no pair of adjacent digits are the same. (For instance, 262-3762 is non-repetitious, but 523-3678 is repetitious.)

How many non-repetitious telephone numbers are there?

Solution: 10×9^6.

Justification: The first digit can be anything (10 possibilities). The next digit is constrained not to be the same as the first digit, but can be anything else (9 possibilities, no matter what we choose for the first digit). The same for the third digit (9 possibilities, no matter what we choose for the first two digits), and so on.

(b) Call a telephone *easy to dial* if it consists of only one or two different digits, in any sequence. (For instance, 661-6116 and 222-2222 and 888-8988 are easy to dial, but 121-1333 is not.)

How many easy-to-dial telephone numbers are there?

Solution: $10 + \binom{10}{2}(2^7 - 2)$.

Justification: There are 10 single-digit telephone numbers (000-0000, 111-1111, 222-2222, etc.).

How many two-digit telephone numbers are there? Well, there are $\binom{10}{2}$ ways to choose a pair of two different digits, say x and y. Now there are 2^7 ways to choose a sequence of 7 digits where each digit must be either x or y (e.g., *xxx-xxxx*, *xxx-xxxy*, etc.)—this corresponds to the number of 7-bit strings. However, this overcounts: the strings *xxx-xxxx* and *yyy-yyyy* are actually single-digit telephone numbers, so they should be excluded from our tally. Therefore, there are $2^7 - 2$ two-digit telephone numbers made up out of the pair of digits x, y, and there are $\binom{10}{2}(2^7 - 2)$ two-digit telephone numbers in all.

The answer then follows by the sum rule, since an easy-to-dial telephone number must be either a single-digit number or a two-digit number (but cannot be both).

Problem 7. [Inversions] (6 points)

An *inversion* in a permutation $[a_1, a_2, \ldots, a_n]$ is a pair (a_i, a_j) such that $i < j$ but $a_i > a_j$. For example, in the list $[2,4,3,1]$ there are four inversions: $(2,1), (4,3), (4,1)$, and $(3,1)$.

Write down the numbers $1, 2, \ldots, n$ in a random order, with all $n!$ orders equally likely. Let X denote the number of inversions in the resulting permutation.

(a) Compute $\mathbf{E}[X]$. Your answer should be a simple function of n. Show your work.

Solution: For each pair of indices i, j with $1 \le i < j \le n$, define the indicator variable $X_{i,j}$ to be 1 if (a_i, a_j) is an inversion, and 0 otherwise. Then $\Pr[a_i > a_j] = 1/2$, so $X_{i,j} \sim \text{Bernoulli}(1/2)$ and $\mathbf{E}[X_{i,j}] = 1/2$. Also, we have

$$X = \sum_{i<j} X_{i,j},$$

so by linearity of expectation,

$$\mathbf{E}[X] = \sum_{i<j} \mathbf{E}[X_{i,j}] = \binom{n}{2}\frac{1}{2}.$$

(b) Multiple-choice: With X defined as above, which of the following equations is true? Circle your choice.

(i) $\mathbf{E}[X^2] < \mathbf{E}[X]^2$.

(ii) $\mathbf{E}[X^2] = \mathbf{E}[X]^2$.

(iii) $\boxed{\mathbf{E}[X^2] > \mathbf{E}[X]^2}$.

Justify your answer *concisely*.

Concise solution: $\mathbf{E}[X^2] - \mathbf{E}[X]^2 = \mathrm{Var}(X) > 0$.

In more detail: It's easy to see that $\mathrm{Var}(X) > 0$ (since there's more than one possible value of X). Also $\mathrm{Var}(X) = \mathbf{E}[X^2] - \mathbf{E}[X]^2$ (by Theorem 19.1 from the lecture notes), so we must have $\mathbf{E}[X^2] - \mathbf{E}[X]^2 > 0$. (The concise answer given above would suffice for full credit.)

Problem 8. [A lame sorting routine] (8 points)

Consider the following bad algorithm for sorting an array A of n different integers:

1. Randomly shuffle A.
2. For $i = 0$ to $n-2$, do the following:
3. If $A[i] > A[i+1]$, go back to step 1 and start over again.

(a) Let X be the total number of times the algorithm executes step 1. Compute $\mathbf{E}[X]$. Your answer should be a simple function of n. Show your work.

Solution: $X \sim \mathrm{Geometric}(p)$, where $p =$ the probability that a random shuffle of A leaves A completely sorted. Therefore $\mathbf{E}[X] = \frac{1}{p}$. Since there is only one ordering of A that leaves A perfectly sorted, out of a total of $n!$ possible orderings, we have $p = \frac{1}{n!}$. Therefore $\boxed{\mathbf{E}[X] = n!}$.

(b) Now suppose $n = 3$, and let Y denote the total number of times the algorithm executes the comparison in step 3. Calculate $\mathbf{E}[Y]$. Show your work.

Solution: Let $y = \mathbf{E}[Y]$ denote the desired value. There are three cases for the order of the elements in A immediately after the first shuffle: (i) A is perfectly sorted (happens with prob. $1/6$); (ii) $A[0] < A[1] > A[2]$ (prob. $1/3$); or (iii) $A[0] > A[1]$ (prob. $1/2$). In case (i), we execute 2 comparisons and then terminate. In case (ii), we execute 2 comparisons immediately, and then we start over and execute an average of y more comparisons, for a total of $2+y$ comparisons. In case (iii), we execute 1 comparison immediately, and then we start over and execute an average of y more comparisons, for a total of $1+y$ comparisons. So

$$y = \frac{1}{6}\cdot 2 + \frac{1}{3}(2+y) + \frac{1}{2}(1+y).$$

Performing some algebraic manipulations, we see $y = \dfrac{3}{2} + \dfrac{5}{6}y$, i.e., $\dfrac{1}{6}y = \dfrac{3}{2}$, which yields the solution $y = 9$. Our final answer is $\boxed{\mathbf{E}[Y] = 9}$.

Alternate solution: Define the indicator variable $Y_{k,i}$ by

$$Y_{k,i} = \begin{cases} 1 & \text{if the comparison } A[i] \stackrel{?}{>} A[i+1] \text{ is executed, after the } k\text{th shuffle} \\ 0 & \text{otherwise.} \end{cases}$$

We have $Y_{k,i} \sim \mathrm{Bernoulli}(p_i)$, where

$$p_i = \Pr[A[0] < A[1] < \cdots < A[i]] = \frac{1}{(i+1)!},$$

so $\mathbf{E}[Y_{k,i}] = \frac{1}{(i+1)!}$. Also

$$Y = \sum_{k=1}^{X} \sum_{i=0}^{n-2} Y_{k,i},$$

so by linearity of expectation,

$$\begin{aligned}
\mathbf{E}[Y] &= \mathbf{E}[X] \sum_{i=0}^{n-2} \mathbf{E}[Y_{k,i}] \\
&= \mathbf{E}[X] \sum_{i=0}^{n-2} \frac{1}{(i+1)!} \\
&= n! \sum_{j=1}^{n-1} \frac{1}{j!}.
\end{aligned}$$

For $n = 3$, we get

$$E[Y] = 3! \sum_{j=1}^{2} \frac{1}{j!} = 3! \left(\frac{1}{1!} + \frac{1}{2!} \right) = 6 \left(1 + \frac{1}{2} \right) = 9,$$

so our final answer is $\boxed{\mathbf{E}[Y] = 9}$.

Comment: The latter solution has the advantage of solving the problem for arbitrary n. For large n, we get $\mathbf{E}[Y] \approx n! \cdot (\mathrm{e} - 1) \approx 1.718 \cdot n!$. So, this is a really lousy sorting algorithm.

Have a great summer, everyone! — David

P.S. I wouldn't want anyone to have to suffer through an entire summer without mathematics (heaven forfend!), so here's one last math puzzle for you:

Next semester, in lieu of a final exam, the CS70 instructor secretly plans to line up all 100 registered CS70 students single file in a random order and randomly place either a red or blue hat on each student (e.g., by secretly flipping a fair coin for each student). Each student can see the color of the hats on all the students in front of him/her, but not the students behind him/her. The instructor will start at the back of the line and, moving forward one by one, ask each student to call out aloud his/her best guess at his/her own hat color. If the student guesses correctly, the instructor congratulates the student warmly and awards him/her an A in the course; otherwise, that student is summarily assigned an F and sternly reprimanded. Either way, after each student has guessed, the instructor then moves to the next student in front of him/her. Once the exam is in progress, no other communication among the students is allowed.

Fortunately, a friendly TA has warned the students in advance of the instructor's unorthodox plan. The entire group of students have come to you for advice, asking you to come up with a strategy that will save the maximum number of students from failing. How many students can you help to ace the course?