CS22 Discrete Structures and Probability Practice Final Exam

May 1, 2025

Please read these directions carefully before you begin!

- This is based on an actual exam from CS22 last year.
- The practice final is not a replacement for reviewing this year's lectures, recitations, and homeworks.
- 2 Hours were given originally to complete this exam, without the use of technology. The only aid provided was the Final Exam Reference Sheet, which is on the website, and you will also get during the exam.
- Good luck and happy proving!

1. Translate the sentences below into formulas of first-order logic using the following symbols.

   The 22 TAs are having their end of year party! Phones are in hand, food is ready to beaten, and laughter fills the air.

   Note: the sets listed here can be used as domains of quantification. That is, you could write $\forall x : T, \ldots$ to quantify over all 22 TAs. You should not use any other domains of quantification.

   - Sets:
     - T: the set of all 22 TAs.
   - Predicates:
     - $N(x)$: "x has an allergy to nuts"
     - $C(x)$: "x eats the 22 cake!"
     - $P(x, y)$: "x gets a selfie with y"
   - Constants:
     - t: A constant, the 22 HTA Tyler.

   (a) (6 points) There is a TA who got a selfie with every other TA, but didn't get a selfie with Tyler.

   Answer: $\exists x : T, \neg P(x,t) \wedge \forall y : T, (\neg x = y \wedge \neg t = y) \to P(x,y)$

   (b) (6 points) If a TA has a nut allergy, they won't eat the 22 cake.

   Answer: $\forall x : T, N(x) \to \neg C(x)$

2. I have a large set B of CS22 problems that I plan to use to create a final exam. There are 1000 problems total in B: 450 are logic problems, 300 are probability problems, and 250 are number theory problems. Each problem has a difficulty score represented by an integer between 1 and 5 (inclusive).

   I will assemble this year's final exam by drawing 10 questions at random from B with uniform probability. The questions will be ordered on the exam from easiest to hardest. Before I do so, I will compute the expected number of problems of each kind and the expected total difficulty of these problems.

   For each question below, select exactly one option.

   (a) (2 points) What would be a reasonable choice for a *sample space* to model this problem?
   - ◯ {logic, probability, number theory}
   - ◯ $\{450, 300, 250\}$
   - ◯ The set of 10 problems I pick
   - ◯ B
   - √ $\{A \in \mathcal{P}(B) \mid |A| = 10\}$

   (b) (2 points) Exactly 10 problems have a difficulty score of 1. What is the *event* that I pick a collection of questions that all have a score of 1?
   - ◯ $\frac{1}{1000}$
   - ◯ The set of problems with difficulty score 1
   - √ **The set containing only the set of problems with difficulty score 1**
   - ◯ $\{\frac{1}{100}\}$
   - ◯ $\{1, 1, \ldots, 1\}$ (10 times)

   (c) (2 points) Which of the following options best corresponds to a random variable in this model?
   - ◯ The average total difficulty across all possible exams
   - √ **The total difficulty of an exam**
   - ◯ The variance of the total difficulty across all possible exams
   - ◯ The probability of picking an exam with a total difficulty of 10

   (d) (2 points) What is the expected number of probability questions on the exam?
   - ◯ 300
   - ◯ 30
   - √ **3**
   - ◯ $\frac{3}{10}$

3. Prove that the sum of any three consecutive integers is divisible by 3.

Solution:

We can represent three consecutive integers, starting at some arbitrary integer n, as n, $n+1$, and $n+2$.

Recall that a number x is divisible by 3 if we can write it as $x = 3k$ (where k is some integer).

$$n + (n + 1) + (n + 2) = 3n + 3$$

Which, we can rewrite as $3(n + 1)$, where $n + 1 \in \mathbb{Z}$.

Thus, using the same division rule, $3 \mid n + (n + 1) + (n + 2)$ for any $n \in \mathbb{Z}$; the sum of any three consecutive integers is divisible by 3, as desired.

4. Consider the set $A = \{a, b, c, d, e\}$. In the following questions, you may leave your answers in unsimplified form if you wish: e.g. 10^2, $200!$, $\binom{10}{2}$.

(a) (3 points) How many subsets of A are there? Remember that $\emptyset \subseteq A$.

$2^5 = 32$

(b) (3 points) How many three-element subsets of A contain the element c?

$\binom{4}{2} = 6$

(c) (2 points) Now consider permutations of the elements of A. In how many permutations does a occur before e? For instance: in the permutation (a, c, e, d, b), a occurs before e. In the permutation (c, e, b, d, a), e occurs before a.

(Note: there's an easy way and a hard way to do this!)

$5!/2 = 60 = 4! + 3 \cdot 3! + 3 \cdot 2 \cdot 2! = 4! + 5 \cdot 3!$

(d) (4 points) Briefly explain your answer to part c.

Solution:

a occurs before e in exactly half, since for any permutation in which a comes first, there is a corresponding one where a and e have been flipped.

Hard way:

There are $4!$ permutations satisfying this requirement where a is the first letter.

If a is not the first letter: 3 choices for the first letter (since it can't be b). Then, if a is the second letter: $3!$ remaining. $4! + 3 \cdot 3!$ so far.

If a is not the first or second: $3 \cdot 2$ choices for those. Then, if a is the third: $2!$ remaining. $4! + 3 \cdot 3! + 3 \cdot 2 \cdot 2!$

Otherwise, a must be 4th and b 5th. So $3!$ ways to place the first 3.

Total: $4! + 3 \cdot 3! + 3 \cdot 2 \cdot 2! + 3! = 24 + 18 + 12 + 6 = 60$.

5. We say that a natural number n is *even* if $2 \mid n$.

   I am a very smart mathematician and I have written the following proof.

   > **Theorem:** all natural numbers are even.
   >
   > **Proof:** I proceed by *strong* induction on n, using the predicate "n is even."
   >
   > *Base case*: I show that 0 is even. $2 \mid 0$, since $2 \cdot 0 = 0$.
   >
   > *Inductive case*: Let n be an arbitrary natural number. My induction hypothesis is that k is even for all $0 \leq k \leq n$. I show that $n+1$ is even. By the induction hypothesis, since $n-1 \leq n$, $n-1$ is even. So $2 \mid n-1$. But $2 \mid 2$, so by the linear combination theorem, $2 \mid (n-1)+2$. This means $2 \mid n+1$, so $n+1$ is even.

   You are also a very smart mathematician. Clearly this theorem shouldn't be true.

   (a) (6 points) Where did I go wrong? Identify the mistaken step in my proof, and explain in a sentence or two why it does not work.

   > **Solution:**
   > My argument in the induction step does not work for $n = 0$, when I try to show that 1 is even. I argue that $n-1$ is even, but -1 is not a natural number and my induction hypothesis does not apply, since it requires $0 \leq k$.

   (b) (6 points) I want to try again. This time, I proceed by regular induction on n with the same predicate. My base case is the same.

   Set up the inductive case for this new induction argument. That is: clearly state the induction hypothesis and the proposition that must be proved in order to complete the argument. You do *not* need to provide this proof! (It shouldn't be possible to prove it.)

   > **Solution:**
   > Induction hypothesis: suppose n is even. Goal: I need to show $n+1$ is even.

6. (8 points) There were $n \geq 2$ students who took the CS22 midterm this year. As they left the exam, each student high-fived some of their classmates—but no two students high-fived each other more than once. Discrete math is exciting, but not that exciting! Each student high-fived at least one classmate.

   (Note: we consider high-fiving to be symmetric. If Tyler high-fives Josh, then Josh has also high-fived Tyler. Furthermore, a student cannot high-five themself.)

   Prove that there are two students who high-fived the same number of their classmates.

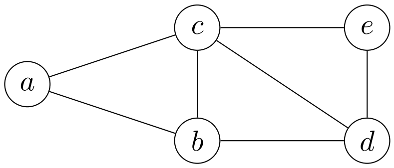
Solution:

We clarified during the exam that every student high-fives at least one other.

Since no student high-fives themselves, the maximum number of people a student can high-five is $n - 1$.

So each student high-fived between 1 and $n - 1$ classmates. There are $n - 1$ numbers in this range and n students, so by the pigeonhole principle, at least two students high-fived the same number.

7. Let $V = \{a, b, c, d, e\}$. Let R be a binary relation on V such that $R(x, y)$ holds if and only if there is a line connecting x and y in the diagram below.

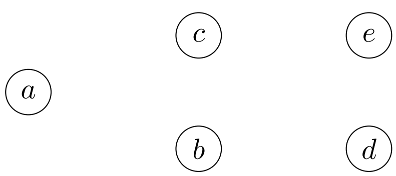


(a) (2 points) Is R reflexive? ○ yes √ **no**

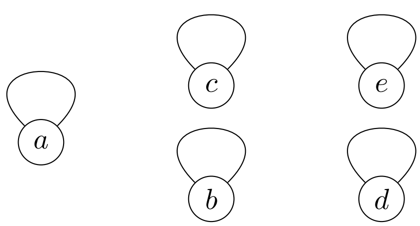
(b) (2 points) Is R symmetric? √ **yes** ○ no

(c) (2 points) Is R transitive? ○ yes √ **no**

(d) (4 points) Draw a new relation on vertices a-e below that has the opposite properties from your answer above. That is: if you answered that the relation above is reflexive, the relation below should *not* be reflexive. If one of these criteria is impossible, briefly explain why.



Solution:



It is impossible for R not to be symmetric because connecting x to y also means connecting y to x.

8. The official CS22 card game assigns point values to playing cards in the following way. Numeric cards (2-10) are worth 5 points each; face cards (J, Q, K) are worth 10 points each; aces (A) are worth 15 points each. A traditional 52-card deck has 4 of each card.

   (a) (3 points) I draw a card at random from a deck. What is its expected point value?

   ___ $90/13 = 360/52$ ___

   (b) (3 points) The point value of a 2-card hand is the sum of the point values of each card in the hand. What is the expected point value of a randomly chosen hand? (The cards are chosen without replacement.)

   ___ $180/13 = 720/52$ ___

9. As a dedicated CS22 student, I am interested in constructing formulas of propositional logic at random. (Logic and probability!)

   Let $C = \{\wedge, \vee, \rightarrow\}$ be my set of binary connectives and $P = \{p, q, r, s\}$ be my set of propositional letters. I construct a formula with the following procedure:

   1. Pick a letter l uniformly at random from P.
   2. Choose with uniform probability whether or not to negate l.
   3. Pick a connective c uniformly at random from C.
   4. Pick a letter r uniformly at random from P. (r could be the same as l.)
   5. Choose with uniform probability whether or not to negate r.
   6. Assemble these choices into the formula $l\ c\ r$.

   For example, I could create the formulas $p \vee \neg q$, $r \wedge r$, and $\neg s \rightarrow \neg p$ with this procedure.

   In the following questions, you may leave your answers in unsimplified form if you wish: e.g. 10^2, $200!$, $\binom{10}{2}$.

   (a) (3 points) What is the probability that I generate a formula with the same letter on both sides of the connective? (Ignore negations: $r \wedge \neg r$ counts! Remember I am only using the letters p, q, r, s.)

   1/4

   (b) (3 points) How many different formulas can I construct with this procedure? I consider the formulas $p \wedge q$ and $q \wedge p$ to be different formulas, even though they are logically equivalent.

   $8 * 3 * 8 = 192$

(c) (2 points) Recall that a formula is *valid* if it is true under every truth assignment. What is the probability that I generate a valid formula?

1/12

(d) (4 points) Briefly explain your answer to part c.

Solution:

There are only two "shapes" of formulas that we can generate that are valid: $x \lor \neg x$ (symmetrically $\neg x \lor x$) and $x \to x$ (symmetrically $\neg x \to \neg x$). There is no way for a $\land$ formula to be valid with our restrictions.

Our choices are all independent, so we just need to multiply probabilities. There is a 2/3 chance we choose $\lor$ or $\to$ for the connective c. The choice of l doesn't matter, but r must match l, which happens with probability 1/4. Similarly, it doesn't matter whether l is negated, but that choice fixes whether r must be negated, and we hit that with probability 1/2. So $\frac{2}{3} \cdot \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{12}$.

10. Let R be a random variable on a probability space Ω such that $\mathbb{E}(R) = 100$ and $\text{Var}(R) = 17$. Define a new random variable G on the same probability space by $G(\omega) = R(\omega) + 20$.

(a) (2 points) What is $\mathbb{E}(G)$?

120

(b) (2 points) What is $\text{Var}(G)$?

17

(c) (2 points) What is $\text{Var}(G - R)$?

0

11. (10 points) Last year's version of this exam had 5 questions, worth 100 points total. I gave myself the following restriction while designing the exam: question 1 needed to be worth at least 1 point; question 2 needed to be worth at least 2 points; and so on.

For example, I could have allocated points so that question 1 was worth 28 points, question 2 was worth 2 points, question 3 was worth 20 points, question 4 was worth 10 points, and question 5 was worth 40 points, summing to 100 total. For another example, I could have made each question worth 20 points.

How many different ways could I allocate points to these five questions while meeting this restriction? Justify your answer.

You may leave your answers in unsimplified form if you wish: e.g. 10^2, 200!, $\binom{10}{2}$.

Solution:

15 points are immediately allocated, so I need to allocate 85 points into 5 boxes. It's stars and bars. That's 4 bars, giving us $\binom{89}{4}$ possibilities.