Staff Solutions to Final Exam

Problem 1 (Basics) (40 points).

(a) i. The power set of the set \( \{a, b, \emptyset\} \) is:

Solution.
\[
\emptyset, \{a\}, \{b\}, \{\emptyset\}, \{a, b\}, \{\emptyset, a\}, \{\emptyset, b\}, \{a, b, \emptyset\}
\]

ii. What is the number of elements in the power set of \( \{2, 4, \ldots, 14\} \)?

Solution.
\[2^7 = 128.\]

iii. Circle the condition(s) under which the size of the intersection, \( A \cap B \), of two finite sets \( A \) and \( B \) is the sum of their sizes.

Always \( A \cup B = \emptyset \) \( \quad \) \( A \subseteq B \) \( \quad \) \( A \cap B \neq \emptyset \) \( \quad \) Never

Solution.
\[A \cup B = \emptyset.\]

(b) There are twenty families in a small village. Four families each have exactly 1 child, seven families each have exactly 2 children and among these five families have children of different ages and two families have twins, three families each have exactly 3 children, and a single family has exactly 4 children. The remaining families have no children.

A family was chosen uniformly at random. What is the probability of the family having:

i. no children at all?

Solution.
\[
5 \quad 1
\]
\[
20 \; 4
\]

ii. exactly 2 or 3 children?

Solution.
iii. fewer than 3 children?

Solution.

\[ \frac{16}{20} = \frac{4}{5} \]

(c) Explain the following concepts.

i. Walk relation in a digraph:

Solution. The walk relation, \( G^* \), of a digraph \( G \) is the binary relation on the vertices of the digraph such that \( v G^* w \) if and only if there is a directed walk from vertex \( v \) to \( w \).

ii. Mutual independence:

Solution. Informally, events \( A_1, A_2, \ldots, A_n \) are mutually independent when the probability of each is unchanged given which of the other ones occur. Formally, they are mutually independent iff for all \( i \in [1, n] \) and \( J \subseteq [1, n] - \{i\} \)

\[
\Pr[A_i] = \Pr \left[ A_i \left| \bigcap_{j \in J} \Pr[A_j] \right. \right].
\]

An equivalent definition is that for every nonempty\(^1\) subset \( J \subseteq [1, n] \)

\[
\Pr \left[ \bigcap_{j \in J} A_j \right] = \prod_{j \in J} \Pr[A_j].
\]

(d) What propositional connective has the following truth table?

<table>
<thead>
<tr>
<th>( P )</th>
<th>( Q )</th>
<th>( P \oplus Q )</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>F</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
</tbody>
</table>

Solution. XOR.

(e) For each of the following formulas, state whether it is Valid or Not by writing V or N.

\(^1\)Equality is guaranteed to hold when \( J = \emptyset \) by the conventions for empty products and intersections.
i. NOT(∀x ∈ A. P(x)) IFF ∃x ∈ A. NOT(P(x)).
ii. P(x) AND NOT(Q(x)) IFF NOT(P(x) IMPLIES Q(x)).
iii. ∀x ∃y. (P(x) OR Q(y)) IFF ∃x ∀y NOT(P(x)) AND NOT(Q(y)).
iv. ∀x ∃y. x^2 + y < 17 where the domain of discourse is {1, 2, 3, 4}.

Solution. V, V, N, N.
The third formula is not valid because it is false if P and Q are true of all elements on the domain of discourse.
The fourth formula is false because if x = 4, so x^2 = 16, there is no y ∈ [1, 4] such that 16 + y < 17.

(f) The probability that a move in a certain game will succeed is 2/3. Three such moves are done independently. What is the probability that:

i. Not all three moves will succeed?

Solution. 
\[
1 - (2/3)^3 = 19/27.
\]

ii. At most one move will succeed?

Solution. 
\[
(1/3)^3 + 3(2/3)(1/3)^2 = 7/27.
\]

(g) Let p(n) ::= n^2 - 16. What is the largest prime factor of p(51)?

Solution. 47.

\[ p(n) = (n - 4)(n + 4) \text{ So } p(51) = 47 \cdot 55 = 47 \cdot 11 \cdot 5. \]

(h) What is the value of

\[
\sum_{n=0}^{\infty} \frac{2}{3^n} + \frac{2^n}{5^n+1}.
\]

Solution. 
\[
\frac{2}{1 - (1/3)} + \frac{1}{5} \cdot \frac{1}{1 - (2/5)} = 3 + \frac{1}{3} = \frac{10}{3}.
\]

(i) Circle all the expressions below that equal the number of size-5 subsets of the set \{1, 2, 3, 4, 5, 6, 7\}.

\[ 7!/(2!5!) \quad 3 \cdot 7 \quad 5 \cdot 5! \quad 7!/(5! \cdot 5!) \quad 7!/(3! \cdot 4!) \]

Solution.

\[ 7!/(2!5!), 3 \cdot 7. \]
(j) A shooter shoots two bullets at a target. The probability that he will hit the target in the first shot is 0.70. If he hits the target in his first shot, then the probability that he will hit the target in the second shot is 0.80. If he misses the target in his first shot, then the probability that he will hit the target in the second shot is \( p \). Given that the probability that he will hit the target in the second shot is 0.62, what is \( p \)?

Solution. 0.20.

Let \( T_1 \) be the event that the shooter hits the target the first time, and let \( T_2 \) be the event that he hits the target the second time. We apply the rule of total probability.

\[
\Pr[T_2] = \Pr[T_2 \mid T_1] \Pr[T_1] + \Pr[T_2 \mid \overline{T_1}] \Pr[\overline{T_1}]
\]

\[0.62 = 0.8 \times 0.7 + p \times 0.3\]

We conclude that \( p = (0.62 - 0.8 \times 0.7)/0.3 = 0.20\).

(k) A jar contains an equal number of green and red balls, indistinguishable except for color. George is blindfolded and asked to pick a ball from a jar and guess its color. As soon as he picks, his two friends, Anna and Sam, tell him what to answer. Anna is color blind, so her answer is random. Sam can see the ball George draws and answers correctly 3/4 of the time. If Sam and Anna do not agree, George returns the ball to the jar without answering and then picks again. If they do agree, George will use their answer.

i. What is the probability that George will give the correct answer when he responds?

Solution. 3/4.

George responds when Anna and Sam agree, so the probability that George is correct when he responds is the probability that both Anna and Sam are correct, given that Anna and Sam agree, namely,

\[
\Pr[\text{both correct} \mid \text{both agree}] = \frac{\Pr[\text{Anna correct}] \Pr[\text{Sam correct}]}{\Pr[\text{both agree}]}
\]

\[
= \frac{(1/2)(3/4)}{(1/2)(3/4) + (1/2)(1/4)} = \frac{3}{4}.
\]

ii. Suppose that there is only a single red ball in the jar. Will your previous answer change? Explain:

Solution. Doesn’t matter. The previous reasoning that led to the 3/4 probability did not depend on how many balls of either color were in the jar.

Problem 2 (GCD & Invariants) (25 points).
The following Binary GCD state machine computes the GCD of positive integers \( a \) and \( b \):

\[
\text{states ::= } \mathbb{N}^3
\]

\[
\text{start state ::= } (a, b, 1)
\]

\[
\text{transitions ::= } \text{if } \min(x, y) > 0, \text{ then}
\]
The predicate
\[ \text{gcd}(a, b) = e \text{gcd}(x, y) \] (7)
is claimed to be a preserved invariant of this state machine.

(a) Verify that equation (7) is a preserved invariant under transition rules (1), and (2). Explicitly state any basic GCD properties you need, but you do not have to prove them.

Solution. To verify preserved invariance, we assume that the invariant holds for state \((x, y, e)\) and show that if \((x, y, e) \rightarrow (x', y', e')\), then \(\text{gcd}(a, b) = e' \text{gcd}(x', y')\). We do so by proof by cases according to which kind of transition occurs.

Case (1): \((2 \mid x \text{ and } 2 \mid y)\). In this case, \((x', y', e') = (x/2, y/2, 2e)\).

We use the easily verified fact
\[ \text{gcd}(au, av) = a \text{gcd}(u, v). \] (8)

Now,
\[
\text{gcd}(a, b) = e \text{gcd}(x, y) \\
= e2 \text{gcd}(x/2, y/2) \quad \text{(by the invariant for \((x, y, e)\))} \\
= e' \text{gcd}(x', y') \quad \text{(by (8))}
\]
which shows that the invariant holds for \((x', y', e')\).

Case (2): \((2 \mid x \text{ and } 2 \notmid y)\). In this case, \((x', y', e') = (x/2, y, e)\).

We use the easily verified fact
\[ \text{gcd}(au, v) = \text{gcd}(u, v) \] (9)
for all \(a\)'s relatively prime to \(v\).

Now,
\[
\text{gcd}(a, b) = e \text{gcd}(x, y) \\
= e \text{gcd}(x/2, y) \quad \text{(by the invariant for \((x, y, e)\))} \\
= e' \text{gcd}(x', y').
\]
which shows that the invariant holds for \((x', y', e')\).

(b) Prove that rule (1) is never executed after any of the other rules is executed.

Hint: A preserved invariant about the parities of \(x\) and \(y\).
Solution. We claim that another preserved invariant is

\[ \text{NOT}(2 \mid x \ \land \ 2 \mid y), \tag{10} \]

that is, at least one of \( x \) and \( y \) is odd.

To verify this, suppose a state \((x, y, e)\) satisfies (10). Then the first rule (1) will not be executed.

Suppose the second rule (2) gets executed, resulting in state \((x/2, y, e)\). Then \( x \) must be even and \( y \) must be odd; so this state satisfies (10) since \( y \) is odd.

Suppose the fourth rule gets executed, resulting in state \((x - y, y, e)\). Then if \( y \) is odd this state satisfies (10), and if \( y \) is even, then \( x \) must be odd, so this state satisfies (10) because \( x - y \) must be odd.

The verification that (10) is preserved by the other rules follows by similar routine reasoning.

Now, if the first rule is not executed for some state \((x, y, e)\), then (10) must hold, and since (10) is preserved, we can conclude that the first rule will never be executed for any subsequent state. \( \blacksquare \)

Problem 3 (Propositional formulas, Relations, Partial Orders, Matchings) (20 points).

Let \( \hat{R} \) be the “implies” binary relation on propositional formulas defined by the rule that

\[ F \hat{R} G \quad \text{iff} \quad [(F \implies G) \text{ is a valid formula}]. \tag{11} \]

For example, \((P \ \land \ Q) \hat{R} P\), because the formula \((P \ \land \ Q) \implies P\) is valid. Also, it is not true that \((P \ \lor \ Q) \hat{R} P\) since \((P \ \lor \ Q) \implies P\) is not valid.

(a) Let \( A \) and \( B \) be the sets of formulas listed below. Explain why \( \hat{R} \) is not a weak partial order on the set \( A \cup B \).

Solution. \( \hat{R} \) is not a weak partial order since the condition of antisymmetry is violated; for example, \( \text{NOT}(P \ \land \ Q) \) implies and is implied by \( \overline{P} \lor \overline{Q} \), but the two formulas are not equal. \( \blacksquare \)

(b) Fill in the \( \hat{R} \) arrows from \( A \) to \( B \).

\[
\begin{array}{ccc}
A & \text{arrows} & B \\

P \ \lor \ P & \overline{P} \lor \overline{Q} \\

P \ \land \ Q & \overline{P} \lor \overline{Q} \lor (\overline{P} \ \land \ \overline{Q}) \\

\text{NOT}(P \ \land \ Q) & P \\
\end{array}
\]
Solution. Arrows for $\vec{R}$:

- $P \text{ XOR } Q$ implies $\overline{P}$ OR $\overline{Q}$
- $P \text{ XOR } Q$ implies $\overline{P}$ OR $\overline{Q}$ OR $(\overline{P}$ AND $\overline{Q})$
- $P \text{ AND } Q$ implies $\overline{Q}$
- $P \text{ AND } Q$ implies $P$
- $\neg(P \text{ AND } Q)$ iff $\overline{P}$ OR $\overline{Q}$
- $\neg(P \text{ AND } Q)$ iff $\overline{P}$ OR $\overline{Q}$ OR $(\overline{P}$ AND $\overline{Q})$

(e) The diagram in part (b) defines a bipartite graph $G$ with $L(G) = A$, $R(G) = B$ and an edge between $F$ and $G$ iff $F \vec{R} G$. Exhibit a subset $S$ of $A$ such that both $S$ and $A - S$ are nonempty, and the set $N(S)$ of neighbors of $S$ is the same size as $S$, that is, $|N(S)| = |S|$.

Solution. $S := \{P \text{ XOR } Q, \neg(P \text{ AND } Q)\}$.

Now $N(S) = \{\overline{P}$ OR $\overline{Q}, \overline{P}$ OR $\overline{Q}$ OR $(\overline{P}$ AND $\overline{Q})\}$, so $|S| = |N(S)| = 2$.

(d) Let $G$ be an arbitrary, finite, bipartite graph. For any subset $S \subseteq L(G)$, let $\overline{S} := L(G) - S$, and likewise for any $M \subseteq R(G)$, let $\overline{M} := R(G) - M$. Suppose $S$ is a subset of $L(G)$ such that $|N(S)| = |S|$, and both $S$ and $\overline{S}$ are nonempty. Circle the formula that correctly completes the following statement:

There is a matching from $L(G)$ to $R(G)$ if and only if there is both a matching from $S$ to its neighbors, $N(S)$, and also a matching from $\overline{S}$ to $N(\overline{S})$.

Hint: The proof of Hall’s Bottleneck Theorem.

Solution. $N(\overline{S})$

Problem 4 ($\mathbb{Z}_n$) (25 points).

(a) Prove that $k^m = 1 \ (\mathbb{Z}_n)$ IMPLIES $\text{ord}(k, n) \ | \ m$.

Hint: Take the remainder of $m$ divided by the order. Reminder: The order of $k \in \mathbb{Z}_n$ is the smallest positive $m$ such that $k^m = 1 \ (\mathbb{Z}_n)$.

Solution. Let $d$ be the order of $k$, and $q, r$ be the quotient and remainder of $m$ divided by $d$, so $m =qd + r$

where $r \in [0, k)$. Now,

$1 = k^m = k^{qk+r} \ (\mathbb{Z}_n)$

$= (k^d)^q \cdot k^r$

$= 1^q \cdot k^r \ (\mathbb{Z}_n)$

$= k^r$. 

$^2$Reminder: $N(S)$ is the set of vertices that are adjacent to at least one vertex in $S$. 

But \( r \in [0, d) \), and since \( d \) is the smallest positive power of \( k \) equal to 1 in \( \mathbb{Z}_n \), we must have \( r = 0 \), that is, \( d \mid m \).

Now suppose \( p > 2 \) is a prime of the form \( 2^s + 1 \). For example, \( 2^1 + 1, 2^2 + 1, 2^4 + 1 \) are such primes.

(b) Conclude from part (a) that if \( 0 < k < p \), then \( \text{ord}(k, p) \) is a power of 2.

\[
\text{Solution.} \quad p - 1 \text{ is a power of two, and therefore so are all its factors. But by Fermat’s Little Theorem, } \\
k^{p-1} = 1 \quad (\mathbb{Z}_p),
\]

and we conclude from part (a) that \( \text{ord}(k, p) \) is a factor of \( p - 1 \).

Problem 5 (Trees, Coloring) (20 points).
Prove by induction that, using a fixed set of \( n > 1 \) colors, there are exactly \( n \cdot (n - 1)^{m-1} \) different colorings\(^3\) of any tree with \( m \) vertices.

\[
\text{Solution. Proof.} \quad \text{By induction on the number of vertices, } m.
\]

\textbf{Induction hypothesis: } \( P(m) : = \) For all \( m \)-vertex trees, \( T \), there are \( n \cdot (n - 1)^{m-1} \) different colorings of \( T \).

\textbf{Base case: } (m = 1). There are \( n = n(n - 1)^{1-1} \) ways to color one vertex.

\textbf{Induction step: } Let \( T \) be a tree with \( m + 1 \) vertices for some \( m \geq 1 \). Let \( v \) be a leaf of \( T \). Then removing \( v \) and its incident edge, we obtain a tree \( T - v \) with \( m \) vertices. We may assume by induction that there are \( n \cdot (n - 1)^{m-1} \) colorings of \( T - v \). For each such coloring of \( T - v \), there are \( n - 1 \) ways to assign a color to \( v \) to obtain an coloring of \( T \), so there are \( n \cdot (n - 1)^{m-1} \cdot (n - 1) = n \cdot (n - 1)^m \) colorings of \( T \), which proves \( P(m + 1) \).

\textbf{Conclusion: } Thus, there are exactly \( n \cdot (n - 1)^{m-1} \) different colorings of any tree with \( m \) vertices.

Problem 6 (Binary sequences) (25 points).

(a) Explain why the union of two countable sets is countable.

\textbf{Solution. Proof.} \quad \text{Countable and nonempty means can be listed, possibly with repeats. Suppose a list of all the elements of } A \text{ is } a_0, a_1, \ldots \text{ and a list of } B \text{ is } b_0, b_1, \ldots \text{ Then a list of all the elements in } A \cup B \text{ is just } \\
a_0, b_0, a_1, b_1, \ldots, a_n, b_n, \ldots.
\]

Let \( \{0, 1\}^* \) be the set of finite binary sequences, \( \{0, 1\}^\omega \) be the set of infinite binary sequences, and \( F \) be the set of sequences in \( \{0, 1\}^\omega \) that contain only a finite number of occurrences of 1's.

(b) Describe a simple surjective function from \( \{0, 1\}^* \) to \( F \).

\textbf{Solution.} \quad \text{Add an infinite sequence of 0's to the end of a finite sequence to make it into an infinite sequence with only finitely many 1's. That is, } x \in \{0, 1\}^* \text{ maps to } x0^\omega \in F.

\(^3\)That is, an assignment of colors to vertices so that no two adjacent vertices are assigned the same color.
(c) The set $F := \{0, 1\}^\omega - F$ consists of all the infinite binary sequences with infinitely many 1’s. Use the previous problem parts to prove that $F$ is uncountable.

*Hint:* We know that $\{0, 1\}^*$ is countable and $\{0, 1\}^\omega$ is not.

**Solution.** By part (b), we have $\{0, 1\}^*$ surj $F$, and so $F$ is countable (Lemma 7.1.9). Now if $F$ was also countable, then by part (a), $F \cup F = \{0, 1\}^\omega$ would be countable, a contradiction.

There is also a simple diagonal argument that proves that $F$ is uncountable using a $30^\circ$ diagonal instead of the basic $45^\circ$ diagonal. Supposing there was a list of the elements of $F$, define a “diagonal” sequence, $s$, that differs from the $n$th sequence in the list at position $2n$. This still ensures that $s$ is not in the list, but it leaves all the odd numbered bits of $s$ unspecified. So we can define all the odd numbered bits of $s$ to be 1, guaranteeing that $s$ has infinitely many 1’s.

**STAFF NOTE:** Consider why the usual $45^\circ$ diagonal argument does not work here. If $s'$ is the diagonal sequence that differs from the $n$th sequence in the list at position $n$, for all $n \in \mathbb{N}$, then $s'$ is not in the list, but there is no guarantee that $s'$ has infinitely many 1’s. So the diagonalization may not yield a “missing” sequence, and the proof breaks down.

---

**Problem 7 (Bijections, Inclusion-Exclusion) (20 points).**

(a) Let $r$ be the number of length $n$ binary strings in which $011$ occurs starting at the 4th position. Write a formula for $r$ in terms of $n$.

**Solution.** $r = 2^{n-3}$

This is the number of patterns of the remaining $n - 3$ bits besides the substring $011$ occupying positions 4–6.

(b) Let $A_i$ be the set of length $n$ binary strings in which $011$ occurs starting at the $i$th position. (So $A_i$ is empty for $i > n - 2$.) If $i \neq j$, the intersection $A_i \cap A_j$ is either empty or of size $s$. Write a formula for $s$ in terms of $n$.

**Solution.** $s = 2^{n-6}$.

To be nonempty, the copies of $011$ at $i$ and $j$ use up 6 positions, leaving $n - 6$ positions that can contain any pattern of bits. So $|A_i \cap A_j| = 2^{n-6}$.

(c) Let $t$ be the number of intersections $A_i \cap A_j$ that are nonempty, where $i < j$. Write a binomial coefficient for $t$ in terms of $n$.

**Solution.**

$$t = \binom{n-4}{2}.$$
This is the same as asking how many ways there are to place two copies of 011 in a length $n$ binary sequence. Since the copies can’t overlap, this is the same as the number of sequences of $n - 6$ 0’s and two 1 where the 0’s indicate positions not occupied by the two copies and the 1’s indicate where the copies are placed. By the Bookkeeper Principle, this is

$$\binom{(n - 6) + 2}{2}.$$ 

\[ \textbf{(d)} \] How many length 9 binary strings are there that contain the substring 011? You should express your answer as an integer or as a simple expression which may include the constants, $r$, $s$ and $t$ above.

*Hint:* Inclusion-exclusion for $|\bigcup_1^7 A_i|$.

\[ \text{Solution.} \]

$$\left| \bigcup_{i=1}^{9} A_i \right| = 7 \cdot r - t \cdot s + 1 = 369. \quad (12)$$

By Inclusion-exclusion

$$\left| \bigcup_{i=1}^{9} A_i \right| = \sum_{i=1}^{9} |A_i| - \sum_{i \neq j} |A_i \cap A_j| + \sum_{i \neq j \neq k} |A_i \cap A_j \cap A_k|. \quad (13)$$

Since $A_8 = A_9 = \emptyset$, there are 7 terms in the first sum in (13), and each term is $r$.

There are $t$ terms in the second sum in (13), each of size $s$.

Finally, among the terms in the third sum,

$$A_1 \cap A_4 \cap A_7 = \{011011011\}$$

and all the other intersections are empty, so the third term is 1. This leads to equation 12.

\[ \textbf{Problem 8 (Generating Functions) (20 points).} \]

Every day in the life of Dangerous Dan is a disaster filled with mishaps.

- Dan may or may not spill his breakfast cereal on his computer keyboard.

- Dan blurts something foolish an even number of times.

- Dan breaks a sequence of two or more dinnerware items (where each item is either a bowl or a plate or a cup.)

Let $T_n$ be the number of different combinations of $n$ mishaps that Dan can suffer in one day (where we regard different sequences of breaks as different combinations). For example, $T_0 = T_1 = 0$, since there are always two or more breaks. On the other hand, $T_3 = 36$; the reasoning is that there can be three breaks (which can happen in $3^3 = 27$ different combinations), or a spill and two breaks (which can happen in $3^2 = 9$ different combinations).
(a) Express the generating function
\[ T(x) := T_0 + T_1 x + T_2 x^2 + T_3 x^3 + \cdots \]
as a quotient of polynomials and explain your derivation.

Solution.

\[
T(x) = \frac{(1 + x)}{\text{spills}} \cdot \frac{(1 + x^2 + x^4 + \cdots)}{\text{blurts}} \cdot \frac{(3^2 x^2 + 3^3 x^3 + 3^4 x^4 + \cdots)}{\text{dinnerware breaks}}
\]
\[
= (1 + x) \cdot \frac{1}{1-x^2} \cdot \frac{9x^2}{1-3x}
\]
\[
= (1 + x) \cdot \frac{1}{(1-x)(1+x)} \cdot \frac{9x^2}{1-3x}
\]
\[
= \frac{9x^2}{(1-x)(1-3x)}.
\]

\[ \blacksquare \]

(b) \( T(x) \) can be written in the form
\[
T(x) = \frac{9x}{2} \left( \frac{1}{1-3x} - \frac{1}{1-x} \right).
\]
Using this fact, show that for \( n \geq 2 \),
\[
T_n = \frac{3^{n+1} - 9}{2}.
\]

Solution. Therefore, for \( n \geq 2 \),
\[
T_n = \frac{9}{2} \left( [x^{n-1}] \frac{1}{1-3x} - [x^{n-1}] \frac{1}{1-x} \right)
\]
\[
= \frac{9}{2} (3^{n-1} - 1)
\]
\[
= \frac{3^{n+1} - 9}{2}.
\]

\[ \blacksquare \]

Problem 9 (Simple Graphs, Asymptotics) (20 points).
Let \( \mathbb{G} \) be the set of all finite connected simple graphs, and let \( f, g : \mathbb{G} \to \mathbb{R}^+ \). We will extend the \( O() \) notation to such graph functions as follows:
\[
[f = O(g)] \text{ IFF } \exists c \in \mathbb{R}^+ \exists n_0 \in \mathbb{N} \forall n > n_0 \forall n\text{-vertex } G \in \mathbb{G}. f(G) \leq cg(G).
\]
For each of the following assertions, state whether it is true or false, and briefly explain your answer. You are not expected to offer a careful proof or detailed counterexample.

Reminder: \( V(G) \) is the set of vertices and \( E(G) \) is the set of edges of \( G \).
(a) \(|V(G)| = O(|E(G)|)\).

Solution. TRUE.
Since \(G\) is connected, the number of edges, \(|E(G)|\), is at most one less than the number of vertices. That is,

\[|V(G)| \leq |E(G)| + 1 = O(|E(G)|).\]

(b) \(|E(G)| = O(|V(G)|)\).

Solution. FALSE.

The complete graph \(K_n\) with \(n\) vertices has \(n(n - 1)/2\) edges.

(c) \(|V(G)| = O(\chi(G))\), where \(\chi(G)\) is the chromatic number of \(G\).

Solution. FALSE. For example, every tree with more than one vertex has a chromatic number 2.

(d) \(\chi(G) = O(|V(G)|)\).

Solution. TRUE.
Assigning a different color to each vertex gives a valid coloring, so \(\chi(G) \leq |V(G)|\).

Problem 10 (Combinatorics) (25 points).

The complete graph, \(K_n\), is a simple graph with \(n\) vertices and an edge between every pair of vertices. We’ll refer to length-3 cycles as triangles.

(a) Show that there are exactly twenty different triangles in \(K_6\).

Solution. Every set of three vertices in \(K_n\) determines a triangle, so there are

\[\binom{6}{3} = 20\]

triangles.

(b) Now suppose that every edge in \(K_6\) is colored either red or blue. A 2-set is a set of two edges of different colors that have a vertex in common; this vertex is called the center of the 2-set. Show that at most six 2-sets can have the same center.

Hint: A vertex incident to \(r\) red edges and \(b\) blue edges is the center of \(r \cdot b\) 2-sets.

Solution. If vertex \(v\) is incident to \(r\) red edges and \(b\) blue edges, then it is the center of exactly \(r \cdot b\) 2-sets. But the degree of \(v\) is five, so \(r + b = 5\). Under this constraint, it’s easy to see that the maximum value of \(r \cdot b\) is six.

(c) A triangle is multicolored if it has two edges of different colors. Describe a 2-to-1 mapping between the 2-sets and the set of multicolored triangles. Conclude that there are at most eighteen multicolored triangles.
Solution. The edges in a two 2-set are edges of a unique multicolor triangle, so map the set to the triangle. Since the edge set of a multicolor triangle contains exactly two 2-sets, this mapping is 2-to-1.

This implies that the number of multicolored triangles is exactly half the number of 2-sets.

Now by part (b), each vertex is the center of at most six 2-sets, and since there are six vertices, there are at most $6 \cdot 6 = 36$ 2-sets and hence at most $(6 \cdot 6)/2 = 18$ multicolored triangles.

(d) If two people are not friends, they are called strangers. If every pair of people in a group are friends, or if every pair are strangers, the group is called uniform. Explain why the following assertion follows directly from the facts established in parts (a)–(c):

Every set of six people includes two uniform three-person groups.

Solution. Label the vertices of $K_6$ with the names of the 6 people. Color the edge between two people red if they are friends and blue if they are strangers. Then, the statement is equivalent to the claim that there are at least two monochromatic triangles, that is, triangles that are not multicolored.

But we know there are twenty triangles, and of these at most eighteen are multicolored, so there must be at least $20 - 18 = 2$ monochromatic triangles.

Problem 11 (Independence) (20 points).

Events $A, B, C$ are said to be conditionally independent if

\[ \Pr[A \cap B \mid C] = \Pr[A \mid C] \cdot \Pr[B \mid C]. \]

Describe events $A, B, C, D$ and explain why they disprove the following:

False Claim. If $A$ and $B$ are conditionally independent given $C$, and are also conditionally independent given $D$, then $A$ and $B$ are conditionally independent given $C \cap D$.

Hint: Choose $A, B, C, D$ 3-way but not 4-way independent.

Solution. Choose $A, B, C, D$ 3-way but not 4-way independent. Then $A$ and $B$ are conditionally independent, given $C$ or given $D$ by 3-way independence. But

\[
\Pr[A \cap B \mid C \cap D] := \frac{\Pr[A \cap B \cap C \cap D]}{\Pr[C \cap D]} \\
\neq \frac{\Pr[A] \cdot \Pr[B] \cdot \Pr[C] \cdot \Pr[D]}{\Pr[C \cap D]} \quad \text{(not 4-way independent)} \\
= \frac{\Pr[A] \cdot \Pr[B] \cdot \Pr[C] \cdot \Pr[D]}{\Pr[C] \cdot \Pr[D]} \quad \text{(2-way independence of $C, D$)} \\
= \Pr[A] \cdot \Pr[B] \\
= \Pr[A \mid C \cap D] \cdot \Pr[B \mid C \cap D] \quad \text{(3-way independence)},
\]

so $A$ and $B$ are not independent given $C \cap D$.

An explicit example would be $A, B$ and $C$ being the events that the first, second and third tosses, respectively, came up Heads in three independent tosses of a fair coin, with $D$ being the event that an even number of Heads were tossed (Problem 17.23).
Problem 12 (Conditional Expectation, Theta()) (20 points).
You have a random process for generating a positive integer, $K$. The behavior of your process each time you use it is (mutually) independent of all its other uses. You use your process to generate an integer, and then use your procedure repeatedly until you generate an integer as big as your first one. Let $R$ be the number of additional integers you have to generate.

(a) State and briefly explain a simple closed formula for $\text{Ex}[R \mid K = k]$ in terms of $\text{Pr}[K \geq k]$.

Solution.

$$\text{Ex}[R \mid K = k] = \frac{1}{\text{Pr}[K \geq k]}$$

If $K = k$, then we can think of generating a number $\geq k$ as a failure, so the expected number of repeats is the mean time to failure, which is the reciprocal of probability of failure on a given try.

Suppose $\text{Pr}[K = k] = \Theta(k^{-4})$.

(b) Show that $\text{Pr}[K \geq k] = \Theta(k^{-3})$.

Solution. It follows from the definition of $\Theta()$ that

$$\text{If } f = \Theta(g), \text{ then } \sum_{n \in S} f(n) = \Theta \left( \sum_{n \in S} g(n) \right)$$  \hspace{1cm} (14)

for any countable set $S$. Therefore, with a slight abuse of $\Theta()$ notation, we have

$$\text{Pr}[K \geq k] = \sum_{m \geq k} \text{Pr}[K = m]$$

$$= \Theta \left( \sum_{m \geq k} k^{-4} \right) \hspace{1cm} \text{by (14)}$$

$$= \Theta \left( \int_{k}^{\infty} x^{-4} \, dx \right) \hspace{1cm} \text{(the Integral Method)}$$

$$= \Theta \left( k^{-3} \right)$$

(c) Show that $\text{Ex}[R]$ is infinite.

Solution. It follows from the definition of $\Theta()$ that

$$\text{If } f_1 = \Theta(g_1) \text{ and } f_2 = \Theta(g_2), \text{ then } f_1 \cdot f_2 = \Theta(g_1 \cdot g_2).$$  \hspace{1cm} (15)
Now we have By Total Expectation:

\[ \text{Ex}[R] = \sum_{k \in \mathbb{Z}^+} \text{Ex}[R \mid K = k] \cdot \Pr[K = k] \]

\[ = \sum_{k \in \mathbb{Z}^+} \frac{1}{\Pr[K \geq k]} \cdot \Pr[K = k] \quad \text{by part (a)} \]

\[ = \Theta \left( \sum_{k \in \mathbb{Z}^+} k^3 \cdot k^{-4} \right) \quad \text{(by part (b), and equations (15) and (14))} \]

\[ = \Theta \left( \sum_{k \in \mathbb{Z}^+} k^{-1} \right) = \infty. \]

Problem 13 (Deviation, Estimation) (20 points).

If \( A \) is a finite set of real numbers, then the collection-variance \( \text{CVar}(A) \) of \( A \) is defined as \( A \)'s average square deviation from its mean:

\[ \text{CVar}(A) := \frac{\sum_{a \in A} (a - \mu)^2}{|A|}, \]

where \( \mu \) is the average value of the numbers in \( A \).

There is a herd of cows whose average body temperature turns out to be 100 degrees, while the collection-variance of all the body temperatures is 20. Our thermometer produces such sensitive readings that no two cows have exactly the same body temperature.

The herd is stricken by an outbreak of \textit{wacky cow disease}, which will eventually kill any cow whose body temperature differs from the average by 10 degrees or more.

(a) Apply the Chebyshev bound to the temperature \( T \) of a random cow to show that at most 20\% of the cows will be killed by this disease outbreak.

\textbf{Solution.} Let \( A \) be the set of body temperatures of the herd. Let \( T \) be the temperature of a random cow. Then

\[ \Pr[|T - 100| \geq 10] \leq \frac{\text{Var}[T]}{10^2} \quad \text{(Chebyshev’s bound)} \]

\[ = \frac{20}{100}. \]

So at most 20\% of the herd can have a temperature that differs from the average by as much as 10 degrees.

(b) The conclusion of part (a) is a bound on a certain fraction of the herd and was derived by bounding the deviation of a random variable. Justify this approach by explaining how to define a random variable, \( T \), for the temperature of a cow. Carefully specify the probability space on which \( T \) is defined: what are the outcomes? what are their probabilities? Explain the precise connection between the characteristics of \( T \) and the characteristics of the actual herd that justify the application of the Chebyshev bound to reach the conclusion about the herd.

\textbf{Solution.} The sample space for \( T \) is the set of cows in the herd, that is, each cow is an outcome. (Alternatively, the outcomes could be chosen to be the temperatures, since our sensitive thermometer ensures there
is a bijection between cows and temperatures.) The probabilities are defined to be uniform—the probability of any outcome, \(c\), is \(1/n\) where \(n\) is the size of the herd—and \(T(c)\) is the temperature of cow \(c\). Since the probabilities are uniform, it is easy to prove that:

- The average temperature of the herd equals \(\text{Ex}[T]\).
- The collection variance of the herd equals \(\text{Var}[T]\).
- The fraction of cows whose temperatures have any given property, \(P\), is the probability that \(T\) has property \(P\).

So the fact that \(\Pr[|T - 100| \geq 10] \leq 0.2\) implies that at most 20% of the herd will die from the outbreak.