Discrete Structures (CSCI 246)
Homework 7

Purpose & Goals
The following problems provide practice relating to:

- graph definitions,
- proof by induction,
- tree diagrams for probability,
- expected values, and
- the problem solving process.

Submission Requirements

- **Type or clearly hand-write your solutions** into a pdf format so that they are legible and professional. Submit your pdf to Gradescope. **Illegible, non-pdf, or emailed solutions will not be graded.**
- Each problem should start on a new page of the document. When you submit to Gradescope, associate each page of your submission with the correct problem number. Please post in Discord if you are having any trouble using Gradescope.
- Try to model your formatting off of the proofs from lecture and/or the textbook.
- Submit to Gradescope early and often so that last-minute technical problems don't cause you any issues. Only the latest submission is kept. Per the syllabus, assignments submitted within 24 hours of the due date will receive a 25% penalty and assignments submitted within 48 hours will receive a 50% penalty. After that, no points are possible.

Academic Integrity

- You may work with your peers, but **you must construct your solutions in your own words on your own.**
- Do not search the web for solutions or hints, post the problem set, or otherwise violate the course collaboration policy or the MSU student code of conduct.
- Violations (regardless of intent) will be reported to the Dean of Students, per the MSU student code of conduct, and you will receive a 0 on the assignment.

Tips

- Answer each problem to the best of your ability. Partial credit is your friend!
- Read the hints for where to find relevant examples for each problem.
- Refer to the problem solving and homework tips guide.
- It is not a badge of honor to say that you spent 5 hours on a single problem or 15 hours on a single assignment. Use your time wisely and get help (see "How to Get Help" below).

How to Get Help
When you are stuck and need a little or big push, **please ask for help!**

- Timebox your effort for each problem so that you don't spend your life on the problem sets. (See the problem solving tips guide for how to do this effectively.)
- Post in Discord. If you're following the timebox guide, you can post the exact statement that you produced after spending 20 minutes being stuck.
- Come to office hours or visit the CS Student Success Center.

Problem 1 (11 points)

In class, we saw a proof of the following property of undirected graphs $G = (V, E)$:

$$\sum_{v \in V} deg(v) = 2|E|.$$

In this problem, you will prove the same fact using induction by filling in the missing pieces.

Proof. For any integer $m \geq 0$, let $P(m)$ denote that in any graph $G = (V, E)$ with m edges, $\sum_{v \in V} deg(v) = 2m$. We show that $\forall m \geq 0 : P(m)$ using mathematical induction over m.

base case:

You fill the statement and proof of the base case

Inductive case: We want to show that $\forall m \geq 1 : P(m-1) \Rightarrow P(m)$. For the inductive hypothesis, we assume $P(m-1)$; that is, we assume that for any graph $G = (V, E)$ with $m-1$ edges has $\sum_{v \in V} deg(v) = 2(m-1)$. Now, let $H = (W, F)$ be any graph with m edges. We WTS that $P(m)$ holds; that is, $\sum_{v \in W} deg(v) = 2m$.

Let $\{u, w\}$ be any edge of H. Consider the graph G constructed from H by removing that one edge $\{u, w\}$; that is, $G = (W, F \setminus \{\{u, w\}\})$. Note that we keep the same set of vertices; we only remove a single edge. For any vertex $v \in W$, let $deg_G(v)$ denote the degree of v in graph G, and $deg_H(v)$ denote the degree in graph H, as now these might be different.

You pick it up from here and fill in the rest of the proof of the inductive case

Since we showed $P(0)$ and $P(m-1) \Rightarrow P(m)$, $P(m)$ holds for all $m \geq 0$ by the principle of mathematical induction.

$\square$

Grading Notes. While a detailed rubric cannot be provided in advance as it gives away the solution details, the following is a general idea of how points are distributed for this problem. We give partial credit where we can.

(9) **Correctness.** If your proof is not correct, this is where you'll get docked. You'll need to

- (1) State the base case correctly
- (2) and prove it, and
- (1) for the inductive case, start with the LHS of $P(m)$ and manipulate to RHS (or vice versa). Do not start with LHS=RHS.
- (1) Find a way to get LHS of $P(m-1)$ into your algebra somewhere,
- (1) correctly apply the IH,
- (1) clearly label where the IH is used,
- (2) find a way to get RHS $P(m)$ into your algebra (or LHS if you started with RHS).

(2) **Communication.** We need to see a mix of notation and intuition. If you skip too many steps at once, or we cannot follow your proof, or if your proof is overly wordy or confusing, this is where you'll get docked.

Problem 2 (10 points)

In lecture we showed that every undirected acyclic graph has either a vertex of degree 0 or degree 1. Prove the following directed version of this claim:

Every directed acyclic graph $G = (V, E)$ has a vertex v such that $outdeg(v) = 0$.
Hint: model your proof after the proof of the undirected claim. You may also want to try out some examples: try to draw a directed acyclic graph that *doesn't* have a vertex of outdegree 0 and convince yourself that it cannot be done before attempting the proof.

Grading Notes. While a detailed rubric cannot be provided in advance as it gives away the solution details, the following is a general idea of how points are distributed for this problem. We give partial credit where we can.

(8) **Correctness.** If your proof is not correct, this is where you'll get docked.

(2) **Communication.** We need to see a mix of notation and intuition. If you skip too many steps at once, or we cannot follow your proof, or if your proof is overly wordy or confusing, this is where you'll get docked.

Problem 3 (14 points)

Recall *hash tables* from Homework 4:

Given a set U of possible values and a set $S \subseteq U$ of values from U, a hash table allows us to quickly answer the question "is x in S?" without needing to examine every element of S. A hash table is defined as follows:

- Let $T[1 \ldots n]$ be a table with n cells.
- Let $h : U \to \{1, 2, \ldots, n\}$ be a function (called a *hash function*).
- Each element $x \in S$ is stored in the cell $T[h(x)]$.

In Homework 4, you proved that if we map a universe of 200 elements to a table with only 100 slots, there will be *collisions* (a fancy way of saying that two elements will map to the same slot in the table). In this problem, you will explore two different ways of resolving such collisions.

(a) (5 points) One way to resolve collisions is using *linear probing*. When we insert an element x into the table, we put x in the first unoccupied cell, moving left to right, starting at cell $h(x)$. For example, suppose T is initially empty. We map x to $h(x) = 1$ and store x in $T(1)$. Then, we map y, and it happens that $h(y) = 1$ as well. Since $T(1)$ is occupied, we check the next cell, and since it is unoccupied, we store y in $T(2)$. Note that if we reach the end of the table as we are looking for unoccupied slots, we circle back to the front again.

Suppose we hash 2 elements into a hash table with 5 slots using a hash function h that puts elements into slots of the table with equal probability.

(i) (3 points) Draw a tree diagram to represent all outcomes of the probabilistic process of putting 2 elements into the table.

(ii) (1 point) What is the probability that the two elements end up in adjacent cells (including the first and last)?

(iii) (1 point) What is the probability that the two elements end up in non-adjacent cells (counting the first and last as adjacent)?

(b) (9 points) A similar way of resolving collisions using *quadratic probing*. Here, if $h(x)$ is already occupied, we try $h(x) + 1^2$, then $h(x) + 2^2$, etc.

(i) (5 points) Draw a tree diagram to represent all outcomes of the probabilistic process of putting 2 elements into the table under quadratic probing.

(ii) (2 points) What is the probability that the two elements end up in adjacent cells (including the first and last)?

(iii) (2 points) What is the probability that the two elements end up in non-adjacent cells (counting the first and last as adjacent)?

Problem 4 (10 points)

You are dealt a 5-card hand from a standard deck. Define a *pair* as any two cards with the same rank—so a hand consisting of an ace of clubs, ace of hearts, ace of diamonds, two of diamonds, and three of diamonds has three pairs (ace of clubs and ace of hearts, ace of clubs and ace of diamonds, ace of hearts and ace of diamonds). Let P denote the number of pairs in your hand.

(a) (4 points) Compute $E[P]$ "the hard way" by computing $Pr[P = 0]$, $Pr[P = 1]$, $Pr[P = 2]$, and so forth. (Note that there can be as many as six pairs in your hand.)

(b) (6 points) Compute $E[P]$ "the easy way" by defining an indicator random variable $R_{i,j}$ that's 1 if and only if cards i and j are a pair, computing $E[R_{i,j}]$, and using linearity of expectation.