CS 70 Discrete Mathematics and Probability Theory
Summer 2021 **HW 3**

1 Lagrange? More like Lamegrange.

In this problem, we walk you through an alternative to Lagrange interpolation.

(a) Let's say we wanted to interpolate a polynomial through a single point, (x_0, y_0). What would be the polynomial that we would get? (This is not a trick question.)

(b) Call the polynomial from the previous part $f_0(x)$. Now say we wanted to define the polynomial $f_1(x)$ that passes through the points (x_0, y_0) and (x_1, y_1). If we write $f_1(x) = f_0(x) + a_1(x - x_0)$, what value of a_1 causes $f_1(x)$ to pass through the desired points?

(c) Now say we want a polynomial $f_2(x)$ that passes through (x_0, y_0), (x_1, y_1), and (x_2, y_2). If we write $f_2(x) = f_1(x) + a_2(x - x_0)(x - x_1)$, what value of a_2 gives us the desired polynomial?

(d) Suppose we have a polynomial $f_i(x)$ that passes through the points (x_0, y_0), ..., (x_i, y_i) and we want to find a polynomial $f_{i+1}(x)$ that passes through all those points and also (x_{i+1}, y_{i+1}). If we define $f_{i+1}(x) = f_i(x) + a_{i+1}\prod_{j=0}^{i}(x - x_j)$, what value must a_{i+1} take on?

Solution:

(a) We want a degree zero polynomial, which is just a constant function. The only constant function that passes through (x_0, y_0) is $f_0(x) = y_0$.

(b) By defining $f_1(x) = f_0(x) + a_1(x - x_0)$, we get that

$$f_1(x_0) = f_0(x_0) + a_1(x_0 - x_0) = y_0 + 0 = y_0.$$

So now we just need to make sure that $f_1(x_1) = y_1$. This means that we need to choose a_1 such that

$$f_1(x_1) = f_0(x_1) + a_1(x_1 - x_0) = y_1.$$

Solving this for a_1, we get that

$$a_1 = \frac{y_1 - f_0(x_1)}{x_1 - x_0}.$$

(c) We apply similar logic to the previous part. From our definition, we know that

$$f_2(x_0) = f_1(x_0) + a_2(x_0 - x_0)(x_0 - x_1) = y_0 + 0 = y_0.$$

and that

$$f_2(x_1) = f_1(x_1) + a_2(x_1 - x_0)(x_1 - x_1) = y_1 + 0 = y_1.$$

Thus, we just need to choose a_2 such that $f_2(x_2) = y_2$. Putting in our formula for $f_2(x)$, we get that we need a_2 such that

$$f_1(x_2) + a_2(x_2 - x_0)(x_2 - x_1) = y_2.$$

Solving for a_2, we get that

$$a_2 = \frac{y_2 - f_1(x_2)}{(x_2 - x_0)(x_2 - x_1)}.$$

(d) If we try to calculate $f_{i+1}(x_k)$ for $0 \leq k \leq i$, we know one of the $(x - x_j)$ terms (specifically the kth one) will be zero. Thus, we get that

$$f_{i+1}(x_k) = f_i(x_k) + a_{i+1}(0) = y_k + 0 = y_k.$$

So now we just need to pick a_i such that $f_{i+1}(x_{i+1}) = y_{i+1}$. This means that we need to choose a_{i+1} such that

$$f_i(x_{i+1}) + a_{i+1} \prod_{j=0}^{i} (x_{i+1} - x_j) = y_{i+1}.$$

Solving for a_{i+1}, we get that

$$a_{i+1} = \frac{y_{i+1} - f_i(x_{i+1})}{\prod_{j=0}^{i} (x_{i+1} - x_j)}.$$

The method you derived in this question is known as Newtonian interpolation. (The formal definition of Newtonian interpolation uses divided differences, which we don't cover in this class, but it's in effect doing the same thing.) This method has an advantage over Lagrange interpolation in that it is very easy to add in extra points that your polynomial has to go through (as we showed in part (c), whereas Lagrange interpolation would require you to throw out all your previous work and restart. However, if you want to keep the same x values but change the y values, Newtonian interpolation requires you to throw out all your previous work and restart. In contrast, this is fairly easy to do with Lagrange interpolation–since changing the y values doesn't affect the δ_is, you don't have to recalculate those, so you can skip most of the work.

2 Polynomials over Galois Fields

Real numbers, complex numbers, and rational numbers are all examples of *fields*. A field is a set of numbers that has some nice properties over some operations. Galois fields are fields with only a finite number of elements, unlike fields such as the real numbers. Galois fields are denoted by GF(q), where q is the number of elements in the field.

(a) In the field GF(p), where p is a prime, how many roots does $q(x) = x^p - x$ have? Use this fact to express $q(x)$ in terms of degree one polynomials. Justify your answers.

(b) Prove that in GF(p), where p is a prime, whenever $f(x)$ has degree $\geq p$, it is equivalent to some polynomial $\tilde{f}(x)$ with degree $< p$.

(c) Show that if P and Q are polynomials over the reals (or complex numbers, or rationals) and $P(x)Q(x) = 0$ for all x, then either $P(x) = 0$ for all x, $Q(x) = 0$ for all x, or both.

(d) Show that the claim in part (c) is false for finite fields $\mathrm{GF}(p)$, where p is a prime.

Solution:

(a) We can factor $q(x) = x^p - x$ into $x(x^{p-1} - 1)$. By Fermat's Little Theorem, for any $a \in \{1, 2, \dots p-1\}$, $q(a) = a(a^{p-1} - 1) = a(1-1) = 0$. And $q(0) = 0(0^{p-1} - 1) = 0$. So every element of $GF(p)$ is a root. The polynomial $q(x)$ has p roots.

We can write $q(x)$ as a product of its roots:

$$q(x) = \prod_{k=0}^{p-1}(x-k)$$

(b) One proof uses Fermat's Little Theorem. Let $d \geq p$; we'll find a polynomial equivalent to x^p. For any integer, we know

$$\begin{aligned} a^d &= a^{d-p}a^p \\ &\equiv a^{d-p}a \mod p \\ &\equiv a^{d-p+1} \mod p. \end{aligned}$$

In other words x^d is equivalent to the polynomial $x^{d-(p-1)}$. If $d-(p-1) \geq p$, we can show in the same way that x^d is equivalent to $x^{d-2(p-1)}$. Since we subtract $p-1$ every time, the sequence $d, d-(p-1), d-2(p-1), \dots$ must eventually be smaller than p. Now if $f(x)$ is any polynomial with degree $\geq p$, we can apply this same trick to every x^k that appears for which $k \geq p$.

Another proof uses Lagrange interpolation. Let $f(x)$ have degree $\geq p$. By Lagrange interpolation, there is a unique polynomial $\tilde{f}(x)$ of degree at most $p-1$ passing through the points $(0, f(0)), (1, f(1)), (2, f(2)), \dots, (p-1, f(p-1))$, and we designed it exactly so that it would be equivalent to $f(x)$.

(c) First, notice that if r is a root of P such that $P(r) = 0$, then r must also be a root of R, since $P(r)Q(r) = 0 \cdot Q(r) = 0$. The same is true for any roots of Q. Also notice that if some value s is neither a root of P nor Q, such that $P(s) \neq 0$ and $Q(s) \neq 0$, then s cannot be a root of R since $P(s)Q(s) \neq 0$. We therefore conclude that the roots of R are the union of the roots of P and Q.

Now we will show the contrapositive. Suppose that P and Q are both non-zero polynomials of degree d_P and d_Q respectively. Then $P(x) = 0$ for at most d_P values of x and $Q(x) = 0$ for at most d_Q values of x. This implies that R has at most $d_P + d_Q$ roots. Since there are an infinite number of values for x (because we are using complex, real, or rational numbers) we can always find an x, call it $x_{\text{not zero}}$, for which $P(x_{\text{not zero}}) \neq 0$ and $Q(x_{\text{not zero}}) \neq 0$. This gives us $P(x_{\text{not zero}})Q(x_{\text{not zero}}) \neq 0$ so R is non-zero.

(d) In GF(p), $x^{p-1}-1$ and x are both non zero polynomials, but their product, $x^p - x$ is zero for all x by Fermat's Little Theorem.

Examples for a specific p are also acceptable. For example, for GF(2), $P(x) = x$ and $Q(x) = x+1$.

3 How Many Polynomials?

Let $P(x)$ be a polynomial of degree at most 2 over GF(5). As we saw in lecture, we need $d+1$ distinct points to determine a unique d-degree polynomial, so knowing the values for say, $P(0)$, $P(1)$, and $P(2)$ would be enough to recover P. (For this problem, we consider two polynomials to be distinct if they return different values for any input.)

(a) Assume that we know $P(0) = 1$, and $P(1) = 2$. Now consider $P(2)$. How many values can $P(2)$ have? How many distinct possibilities for P do we have?

(b) Now assume that we only know $P(0) = 1$. We consider $P(1)$ and $P(2)$. How many different $(P(1), P(2))$ pairs are there? How many distinct possibilities for P do we have?

(c) Now, let P be a polynomial of degree at most d on GF(p) for some prime p with $p > d$. Assume we only know P evaluated at $k \leq d+1$ different values. How many different possibilities do we have for P?

(d) A polynomial with integer coefficients that cannot be factored into polynomials of lower degree on a finite field, is called an irreducible or prime polynomial.

Show that $P(x) = x^2 + x + 1$ is a prime polynomial on GF(5).

Solution:

(a) 5 polynomials, each for different values of $P(2)$.

(b) Now there are 5^2 different polynomials.

(c) p^{d+1-k} different polynomials. For $k = d+1$, there should only be 1 polynomial.

(d) We can try all possible inputs for x and show that in each case $P(x) \pmod{x} \neq 0$, which means that $P(x)$ does not have any root on the finite field GF(5).

$$\begin{aligned}
x = 0 &\Rightarrow P(0) \equiv 1 \pmod{5} \\
x = 1 &\Rightarrow P(1) \equiv 3 \pmod{5} \\
x = 2 &\Rightarrow P(2) \equiv 2 \pmod{5} \\
x = 3 &\Rightarrow P(3) \equiv 3 \pmod{5} \\
x = 4 &\Rightarrow P(4) \equiv 1 \pmod{5}
\end{aligned}$$

Hence $P(x)$ is a prime polynomial.

4 Counting, Counting, and More Counting

The only way to learn counting is to practice, practice, practice, so here is your chance to do so. Although there are many subparts, each subpart is fairly short, so this problem should not take any longer than a normal CS70 homework problem. You do not need to show work, and **Leave your answers as an expression** (rather than trying to evaluate it to get a specific number).

(a) How many ways are there to arrange n 1s and k 0s into a sequence?

(b) How many 7-digit ternary (0,1,2) bitstrings are there such that no two adjacent digits are equal?

(c) A bridge hand is obtained by selecting 13 cards from a standard 52-card deck. The order of the cards in a bridge hand is irrelevant.
i. How many different 13-card bridge hands are there? ii. How many different 13-card bridge hands are there that contain no aces? iii. How many different 13-card bridge hands are there that contain all four aces? iv. How many different 13-card bridge hands are there that contain exactly 6 spades?

(d) Two identical decks of 52 cards are mixed together, yielding a stack of 104 cards. How many different ways are there to order this stack of 104 cards?

(e) How many 99-bit strings are there that contain more ones than zeros?

(f) An anagram of ALABAMA is any re-ordering of the letters of ALABAMA, i.e., any string made up of the letters A, L, A, B, A, M, and A, in any order. The anagram does not have to be an English word.
i. How many different anagrams of ALABAMA are there? ii. How many different anagrams of MONTANA are there?

(g) How many different anagrams of ABCDEF are there if: (1) C is the left neighbor of E; (2) C is on the left of E (and not necessarily E's neighbor)

(h) We have 9 balls, numbered 1 through 9, and 27 bins. How many different ways are there to distribute these 9 balls among the 27 bins? Assume the bins are distinguishable (e.g., numbered 1 through 27).

(i) How many different ways are there to throw 9 identical balls into 27 bins? Assume the bins are distinguishable (e.g., numbered 1 through 27).

(j) We throw 9 identical balls into 7 bins. How many different ways are there to distribute these 9 balls among the 7 bins such that no bin is empty? Assume the bins are distinguishable (e.g., numbered 1 through 7).

(k) There are exactly 20 students currently enrolled in a class. How many different ways are there to pair up the 20 students, so that each student is paired with one other student? Solve this in at least 2 different ways. **Your final answer must consist of two different expressions.**

(l) How many solutions does $x_0 + x_1 + \cdots + x_k = n$ have, if each x must be a non-negative integer?

(m) How many solutions does $x_0 + x_1 = n$ have, if each x must be a *strictly positive* integer?

(n) How many solutions does $x_0 + x_1 + \cdots + x_k = n$ have, if each x must be a *strictly positive* integer?

Solution:

(a) $\binom{n+k}{k}$

(b) There are 3 options for the first digit. For each of the next digits, they only have 2 options because they cannot be equal to the previous digit. Thus, $3 * 2^6$

(c) We have to choose 13 cards out of 52 cards, so this is just $\binom{52}{13}$.

We now have to choose 13 cards out of 48 non-ace cards. So this is $\binom{48}{13}$.

We now require the four aces to be present. So we have to choose the remaining 9 cards in our hand from the 48 non-ace cards, and this is $\binom{48}{9}$.

We need our hand to contain 6 out of the 13 spade cards, and 7 out of the 39 non-spade cards, and these choices can be made separately. Hence, there are $\binom{13}{6}\binom{39}{7}$ ways to make up the hand.

(d) If we consider the 104! rearrangements of 2 identical decks, since each card appears twice, we would have overcounted each distinct rearrangement. Consider any distinct rearrangement of the 2 identical decks of 52 cards and see how many times this appears among the rearrangement of 104 cards where each card is treated as different. For each identical pair (such as the two Ace of spades), there are two ways they could be permuted among each other (since $2! = 2$). This holds for each of the 52 pairs of identical cards. So the number 104! overcounts the actual number of rearrangements of 2 identical decks by a factor of 2^{52}. Hence, the actual number of rearrangements of 2 identical decks is $104!/2^{52}$.

(e) **Answer 1:** There are $\binom{99}{k}$ 99-bit strings with k ones and $99-k$ zeros. We need $k > 99-k$, i.e. $k \geq 50$. So the total number of such strings is $\sum_{k=50}^{99}\binom{99}{k}$.
This expression can however be simplified. Since $\binom{99}{k} = \binom{99}{99-k}$, we have

$$\sum_{k=50}^{99}\binom{99}{k} = \sum_{k=50}^{99}\binom{99}{99-k} = \sum_{l=0}^{49}\binom{99}{l}$$

by substituting $l = 99-k$. Now $\sum_{k=50}^{99}\binom{99}{k} + \sum_{l=0}^{49}\binom{99}{l} = \sum_{m=0}^{99}\binom{99}{m} = 2^{99}$. Hence, $\sum_{k=50}^{99}\binom{99}{k} = (1/2)\cdot 2^{99} = 2^{98}$.

Answer 2: Symmetry Since the answer from above looked so simple, there must have been a more elegant way to arrive at it. Since 99 is odd, no 99-bit string can have the same number of zeros and ones. Let A be the set of 99-bit strings with more ones than zeros, and B be the set of 99-bit strings with more zeros than ones. Now take any 99-bit string x with more ones than zeros i.e. $x \in A$. If all the bits of x are flipped, then you get a string y with more zeros than ones, and so $y \in B$. This operation of bit flips creates a one-to-one and onto function (called a bijection) between A and B. Hence, it must be that $|A| = |B|$. Every 99-bit string is either in A or in B, and since there are 2^{99} 99-bit strings, we get $|A| = |B| = (1/2) \cdot 2^{99}$. The answer we sought was $|A| = 2^{98}$.

(f) ALABAMA: The number of ways of rearranging 7 distinct letters and is 7!. In this 7 letter word, the letter A is repeated 4 times while the other letters appear once. Hence, the number 7! overcounts the number of different anagrams by a factor of 4! (which is the number of ways of permuting the 4 A's among themselves). Aka, we only want $1/4!$ out of the total rearrangements. Hence, there are $7!/4!$ anagrams.

MONTANA: In this 7 letter word, the letter A and N are each repeated 2 times while the other letters appear once. Hence, the number 7! overcounts the number of different anagrams by a factor of $2! \times 2!$ (one factor of 2! for the number of ways of permuting the 2 A's among themselves and another factor of 2! for the number of ways of permuting the 2 N's among themselves). Hence, there are $7!/(2!)^2$ different anagrams.

(g) (1) We consider CE is a new letter X, then the question becomes counting the rearranging of 5 distinct letters, and is 5!. (2) Symmetry: Let A be the set of all the rearranging of ABCDEF with C on the left side of E, and B be the set of all the rearranging of ABCDEF with C on the right side of E. $|A \bigcup B| = 6!, |A \cap B| = 0$. There is a bijection between A and B by construct a operation of exchange the position of C and E. Thus $|A| = |B| = 6!/2$.

(h) Each ball has a choice of which bin it should go to. So each ball has 27 choices and the 9 balls can make their choices separately. Hence, there are 27^9 ways.

(i) Since there is no restriction on how many balls a bin needs to have, this is just the problem of throwing k identical balls into n distinguishable bins, which can be done in $\binom{n+k-1}{k}$ ways. Here $k = 9$ and $n = 27$, so there are $\binom{35}{9}$ ways.

(j) **Answer 1:** Since each bin is required to be non-empty, let's throw one ball into each bin at the outset. Now we have 2 identical balls left which we want to throw into 7 distinguishable bins. There are 2 cases to consider:
Case 1: The 2 balls land in the same bin. This gives 7 ways.
Case 2: The 2 balls land in different bins. This gives $\binom{7}{2}$ ways of choosing 2 out of the 7 bins for the balls to land in. Note that it is *not* 7×6 since the balls are identical and so there is no order on them.
Summing up the number of ways from both cases, we get $7 + \binom{7}{2}$ ways.

Answer 2: Since each bin is required to be non-empty, let's throw one ball into each bin at the outset. Now we have 2 identical balls left which we want to throw into 7 distinguishable bins. From class (see note 11), we already saw that the number of ways to put k identical balls into n distinguishable bins is $\binom{n+k-1}{k}$. Taking $k = 2$ and $n = 7$, we get $\binom{8}{2}$ ways to do this.
EASY EXERCISE: Can you give an expression for the number of ways to put k identical balls into n distinguishable bins such that no bin is empty?

(k) **Answer 1:** Let's number the students from 1 to 20. Student 1 has 19 choices for her partner. Let i be the smallest index among students who have not yet been assigned partners. Then no matter what the value of i is (in particular, i could be 2 or 3), student i has 17 choices for her partner. The next smallest indexed student who doesn't have a partner now has 15 choices for her partner. Continuing in this way, the number of pairings is $19 \times 17 \times 15 \times \cdots \times 1 = \prod_{i=1}^{10}(2i-1)$.

Answer 2: Arrange the students numbered 1 to 20 in a line. There are 20! such arrangements. We pair up the students at positions $2i-1$ and $2i$ for i ranging from 1 to 10. You should be able to see that the 20! permutations of the students doesn't miss any possible pairing. However, it counts every different pairing multiple times. Fix any particular pairing of students. In this pairing, the first pair had freedom of 10 positions in any permutation that generated it, the second pair had a freedom of 9 positions in any permutation that generated it, and so on. There is also the freedom for the elements within each pair i.e. in any student pair (x,y), student x could have appeared in position $2i-1$ and student y could have appeared in position $2i$ and also vice versa. This gives 2 ways for each of the 10 pairs. Thus, in total, these freedoms cause $10! \times 2^{10}$ of the 20! permutations to give rise to this particular pairing. This holds for each of the different pairings. Hence, 20! overcounts the number of different pairings by a factor of $10! \times 2^{10}$. Hence, there are $20!/(10! \cdot 2^{10})$ pairings.

Answer 3: In the first step, pick a pair of students from the 20 students. There are $\binom{20}{2}$ ways to do this. In the second step, pick a pair of students from the remaining 18 students. There are $\binom{18}{2}$ ways to do this. Keep picking pairs like this, until in the tenth step, you pick a pair of students from the remaining 2 students. There are $\binom{2}{2}$ ways to do this. Multiplying all these, we get $\binom{20}{2}\binom{18}{2}\ldots\binom{2}{2}$. However, in any particular pairing of 20 students, this pairing could have been generated in 10! ways using the above procedure depending on which pairs in the pairing got picked in the first step, second step, ..., tenth step. Hence, we have to divide the above number by 10! to get the number of different pairings. Thus there are $\binom{20}{2}\binom{18}{2}\ldots\binom{2}{2}/10!$ different pairings of 20 students.

You may want to check for yourself that all three methods are producing the same integer, even though they are expressed very differently.

(l) $\binom{n+k}{k}$. This is just n indistinguishable balls into k+1 distinguishable bins (stars and bars). There is a bijection between a sequence of n ones and k plusses and a solution to the equation: x_0 is the number of ones before the first plus, x_1 is the number of ones between the first and second plus, etc. A key idea is that if a bijection exists between two sets they must be the same size, so counting the elements of one tells us how many the other has. Note that this is the exact same answer as part a - make sure you understand why!

(m) $n-1$. It's easy just to enumerate the solutions here. x_0 can take values $1,2,\ldots,n-1$ and this uniquely fixes the value of x_1. So, we have $n-1$ ways to do this. But, this is just an example of the more general question below.

(n) $\binom{(n-(k+1))+k}{k} = \binom{n-1}{k}$. This is just n-(k+1) indistinguishable balls into distinguishable k+1 bins. By subtracting 1 from all $k+1$ variables, and $k+1$ from the total required, we reduce it to problem with the same form as the previous problem. Once we have a solution to that we reverse the process, and adding 1 to all the non-negative variables gives us positive variables.

5 Functional Counting

In each of the following questions, assume that n is a positive integer.

(a) How many strictly increasing functions are there from the set $\{1,2,\ldots,n-1,n\}$ to the set $\{1,2,\ldots,2n-1,2n\}$?

(b) How many surjective functions are there from the set $\{1,2,\ldots,2n-1,2n\}$ to the set $\{1,2,\ldots,n-1,n\}$?

Solution:

(a) The answer is $\binom{2n}{n}$. The reason for this is that every strictly increasing function can be uniquely identified with its range. Since the domain of any such function has n elements, the range of any such function must be an n element subset of $\{1,2,\ldots,2n-1,2n\}$. There are $\binom{2n}{n}$ such subsets, so there are $\binom{2n}{n}$ strictly increasing functions from $\{1,2,\ldots,n-1,n\}$ to $\{1,2,\ldots,2n-1,2n\}$.

(b) The answer is

$$\sum_{k=0}^{n-1}(-1)^k\binom{n}{k}(n-k)^{2n}.$$

To see why, we use the Principle of Inclusion-Exclusion. In order to compute the number of surjective functions, we can equivalently compute the number of functions that aren't surjective. Let A denote the set of functions which aren't surjective. If we let A_j be the set of functions whose ranges' don't contain the number j, we can see that

$$A = A_1 \cup A_2 \cup \cdots \cup A_{n-1} \cup A_n.$$

Now, observe that

$$|A_j| = (n-1)^{2n},$$

as we have $n-1$ choices for the value of the function at each of our $2n$ points in the domain. Moreover, observe that $A_i \cap A_j$ is the set of functions which don't take on either of the values i or j, and similarly with 3 or more sets. In this way, we can see that the size of the intersection of k of these sets is $(n-k)^{2n}$, as we have $n-k$ choices for the values of the function. Now, applying the Principle of Inclusion-Exclusion, we get that

$$|A| = \binom{n}{1}(n-1)^{2n} - \binom{n}{2}(n-2)^{2n} + \cdots + (-1)^n\binom{n}{n-1}(n-(n-1))^{2n}.$$

Now, this is the number of functions that aren't surjective. To get the number of functions which are surjective, we need to subtract this from the total number of functions. The total number of functions is n^{2n} (n choices for the value of each of the $2n$ points in the domain), hence our final answer is

$$\begin{aligned} n^{2n} - |A| &= \binom{n}{0}(n-0)^{2n} - \binom{n}{1}(n-1)^{2n} + \cdots + (-1)^{n-1}\binom{n}{n-1}(n-(n-1))^{2n} \\ &= \sum_{k=0}^{n-1}(-1)^k\binom{n}{k}(n-k)^{2n}. \end{aligned}$$

6 Good Khalil Hunting

As a sidejob, Khalil is also working as a janitor in Berkeley EECS. One day, he notices a problem on the board and decides to solve it. The problem is as follows: **Find all homeomorphically irreducible trees having 10 vertices.** A tree is homeomorphically irreducible if it has no vertices of degree 2. Assume all vertices and edges are indistinguishable from another.

Let's help Khalil solve this using strategic **casework**. We will partition the problem based off the number of the leaves in the tree. For sake of clarity, label the vertices $v_1, \dots, v_{10}$ and their degrees $d_1, \dots, d_{10}$ in decreasing order of degree.

(a) Show that the number of leaves, ℓ, we can have is $6 \leq \ell \leq 9$. (*Hint*: What do you know about the degrees of a leaf? What about a non-leaf in this case?)

For the following parts, drawings are neither necessary nor sufficient for your answer, but are highly encouraged to help you get the answer. Please briefly justify your answers by formulating equations involving the degrees of the vertices, along with short explanations.

(b) How many 10 vertex, homeomorphically irreducible trees of 9 leaves are there? Justify your answer.

(c) How many 10 vertex, homeomorphically irreducible trees of 8 leaves are there? Justify your answer.

(d) How many 10 vertex, homeomorphically irreducible trees of 7 leaves are there? Justify your answer.

(e) How many 10 vertex, homeomorphically irreducible trees of 6 leaves are there? Justify your answer.

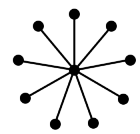
In total, you should have counted 10 trees. Great work!

Solution:

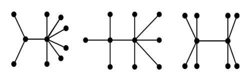
We solve this problem by casework. It is very hard to count upfront the number of 10 vertex homeomorphically irreducible trees. However, counting them with the constraint of how many leaves they contain is much easier. First, we must make sure that our cases are mutually exclusive (no overlap) and collectively exhaustive (covers all the possible cases). They are mutually exclusive because a tree can't have, say, 1 and 2 leaves at the same time. We show in part (a) that as long as we consider the cases of 6,7,8,9 leaves, then that is collectively exhaustive.

(a) We know that the sum of the degrees is 18 as the tree has 9 edges. Note that leaves are degree 1 and non-leaves are at least degree 3 since we aren't considering degree 2 vertices. Thus, if there are k leaves and 10 - k non-leaves, then the sum of the degrees, $18 \leq k + 3 * (10 - k) = 30 - 2k$, thus $k \geq 6$. (the sum of the degrees is lower-bounded when we substitute for all non-leaves $d_i, d_i = 3$). Also, $k \leq 9$ since a 10 vertex tree has at max 9 leaves.

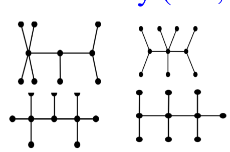
(b) 1. This is a node in the middle with 9 leaves around it.



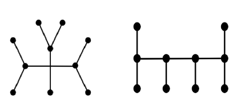
(c) 3. If there are 8 leaves and 2 non-leaves, then the non-leaves, $d_1 + d_2 = 10$ with $d_1, d_2 \geq 3$, so either: $d_1 = 7, d_2 = 3$, or $d_1 = 6, d_2 = 4$, or $d_1 = d_2 = 5$. Each one of these has one tree, because since we are assuming indistinguishability, swapping d_1 and d_2 does not make a difference. This makes up all the cases.



(d) 4. If there are 7 leaves and 3 non-leaves, then the non-leaves, $d_1 + d_2 + d_3 = 11$. Thus the 2 cases are $d_1 = 5, d_2 = 3, d_3 = 3$ or $d_1 = 4, d_2 = 4, d_3 = 3$. This makes up all the cases since any other case will violate $d_i \geq 3$ for non-leaves. Now, for each one of these 2 cases, there are 2 trees. For example, 3,3,5 and 3,5,3 are different trees, however, 5,3,3 is the same as 3,3,5 since they are equivalent by indistinguishability (just flip the tree). Note, there cannot exist any other trees since otherwise we will violate irreducibility (aka, degree 2 nodes will exist).



(e) 2. If there are 6 leaves and 4 non-leaves, then the non-leaves $d_1 + d_2 + d_3 + d_4 = 12$, thus $d_1 = d_2 = d_3 = d_4 = 3$ by pidgeonhole (or just algebra). Sneakily, this case actually has two trees. One of which contains a degree 3 node with no leaves attached to it, while the other does. You can visualize the latter tree containing 4 non-leaves being in a chain as expected, and the former tree containing 4 non-leaves like a perpendicular sign. Note, there cannot exist any other trees since otherwise we will violate irreducibility (aka, degree 2 nodes will exist).



This problem was designed to give extensive practice on casework. If you haven't noticed already, it is based off the problem seen in the movie, Good Will Hunting. The math was heavily derived from this paper: http://math.unideb.hu/media/horvath-gabor/publications/gwh2.pdf The images are derived from this medium article: https://medium.com/cantors-paradise/the-math-problems-from-good-will-hunting-w-solutions-b081895bf379. Congrats, you are now as smart as Matt Damon.

7 Proofs of the Combinatorial Variety

Prove each of the following identities using a combinatorial proof.

(a) For every positive integer $n > 1$,

$$\sum_{k=0}^{n} k \cdot \binom{n}{k} = n \cdot \sum_{k=0}^{n-1} \binom{n-1}{k}.$$

(b) For each positive integer m and each positive integer $n > m$,

$$\sum_{a+b+c=m} \binom{n}{a} \cdot \binom{n}{b} \cdot \binom{n}{c} = \binom{3n}{m}.$$

(Notation: the sum on the left is taken over all triples of nonnegative integers (a,b,c) such that $a+b+c=m$.)

Solution:

(a) Suppose we have n people and want to pick some of them to form a special committee. Moreover, suppose we want to pick a leader from among the committee members - how many ways can we do this?

We can do so by first picking the committee members, and then choosing the leader from among the chosen members. We can pick a committee of size k in $\binom{n}{k}$ ways, and once we have picked the committee, we have k choices for which member becomes the leader. In order to account for all possible committee sizes, we need to sum over all valid values of k, hence we get the expression

$$\sum_{k=0}^{n} k \cdot \binom{n}{k},$$

which is exactly the left hand side of the identity we want to prove.

Now, we can also count this set by first picking the leader for the committee, then choosing the rest of committee. We have n choices for the leader, and then among the remaining $n-1$ people, we can pick any subset to form the rest of the committee. Picking a subset of size k can be done in $\binom{n-1}{k}$ ways, hence summing over k, we get the expression

$$n \cdot \sum_{k=0}^{n-1} \binom{n-1}{k},$$

which is exactly the right hand side of the identity we want to prove.

(b) Suppose we have n distinguishable red pencils, n distinguishable blue pencils, and n distinguishable green pencils ($3n$ pencils total), and want to choose m of these pencils to bring to class. How many ways can be do this?

We can do so by just picking the m pencils without considering color, as they are all distinguishable. There are $\binom{3n}{m}$ ways of doing this, which is exactly the right hand side of the identity we want to prove.

We can also count this set by picking some red pencils, the picking some blue pencils, and then finally picking some green pencils. We can pick a red pencils, b blue pencils, and c green pencils (with the tacit assumption that $a+b+c=m$) in $\binom{n}{a}\cdot\binom{n}{b}\cdot nc$ ways. Finally, in order to account for all possible distributions of pencils, we need to sum over all valid triples (a,b,c), which gives us the expression

$$\sum_{a+b+c=m}\binom{n}{a}\cdot\binom{n}{b}\cdot\binom{n}{c},$$

which is exactly the left hand side of the identity we want to prove.