Name ______________________ **PID** ____________________

Seat Number ______________ **Discussion** _______________

Midterm Exam 2 - CSE 21

Do not turn the page until you are instructed to do so.

- You may not use any electronic devices, written notes, or any other form of assistance during this exam.
- If you have a question, please remain seated. Raise your hand and wait for assistance.
- Show your work. To receive full credit, your answers must be neatly written and logically organized.

#	Points	Score
1	7	
2	6	
3	6	
4	6	
Total	25	

Academic integrity is expected of all students at all times, whether in the presence or absence of members of the faculty. Understanding this, I declare I shall not give, use, or receive unauthorized aid in this examination.

X ____

1. **Counting**

   It is okay to leave your answer in terms of binomial coefficients, factorials, and exponents. You do not need to simplify. Give a short explanation of your answer. (10 points each, 5 for answer , 5 for explanation)

   (a) How many base 10 numbers have exactly one of each numeral, if we allow a leading 0? For example, 0237456891 is one such number.

   Since we need to have each digit appear in some order, the numbers correspond to permutations of the 10 digits. So there are 10! such numbers.

   (b) How many base 10 numbers have exactly one of each numeral, if we disallow 0 as the most significant digit?

   In the count above, if 0 is the most significant digit, the other 9 appear afterwards in any order. So there are 9! numbers we counted with 0 as the most significant digit, leaving $10! - 9!$ if we disallow this case.

   (c) How many n digit base 10 numbers have at least one even digit and at least one odd digit? Allow the number to start with 0.

   There are 10^n n digit numbers total. We need to subtract off those with no odd digits and those with no even digits. No odd digits leave 5 even digits, and 5^n numbers total with only even digits, and the same for odd digits. Since these categories are mutually exclusive, the sets have an empty intersection, so the amount we add back in using inclusion-exclusion is just 0. Thus the total left is $10^n - 2 * (5^n)$.

   (d) *Bonus question for practice only, skip if timed*: How many base 10 numbers have exactly two of each digit,and start with a 1? One example is 12345678912345678900

   Once we start with a 1, there will be one 1 and 2 of every other kind of digit. We can shuffle these 19 digits any way, for 19! shuffles, but switching the two occurences of each digit other than 1 keeps the numbers the same. So by the quotient rule, there are $19!/2^9$ such numbers.

2. **Graphs** In the game of geography, the players say a sequence of place names, where the next in the sequence must start with the last letter in the previous place. For example, one game might involve "BOSTON", "NEWARK", "KENTUCKY", "YORK", "KANSAS". Describe how, given a list of place names, you could create a graph whose paths would represent possible games of geography using those places. (In actual geography, you are not allowed to repeat names, but you don't have to enforce this explicitly in your construction.)

   (10 points each part, 5 points answer, 5 points short explanation)

   (a) What are the vertices of your graph?

   There are two ways to construct the graph. In the first way, we have one vertex per word.

   In the second way, we can have one vertex per letter that starts or ends a word.

   (b) Is your graph directed or undirected? Is it simple?

   Either way the graph will be directed, because going from YORK to KANSAS does not mean we can go from KANSAS to YORK. In the first way, it might have self-loops buut no parallel edges. The second way, we will introduce parallel edges and possibly self-loops. So in neither way is it guaranteed to be a simple graph.

   (c) What are the edges ?

   If the vertices are words, we put an edge from word 1 to word 2 if the last character of word 1 is the first in word 2. Then the vertices in a path in this graph are words that can be listed consecutively in the word puzzle. Technically, only paths that don't repeat vertices would be legal for the game.

   If the vertices are characters, for each word, we put in an edge from the first character in that word to the last character, with parallel edges if there are duplicates, and self-loops if there are words that begin and end with the same character. The edges along a path form words that can be listed consecutively in the game. Technically, only paths that don't repeat edges would be legal for the game.

   (d) *Bonus questions for practice only, skip if timing* Give the adjacency list representation for this graph for the list: ARIZONA, COLORADO, HAWAII, IDAHO, IOWA, NEVADA, OREGON, UTAH.

The vertices are ARIZONA, COLORADO, HAWAII, IDAHO, IOWA, NEVADA, OREGON, UTAH.

In the first construction, we would have vertices : ARIZONA, COLORADO, HAWAII, IDAHO, IOWA, NEVADA, OREGON, UTAH.

- ARIZONA: ARIZONA, eol.
- COLORADO: OREGON, eol.
- HAWAII: IDAHO, IOWA, eol.
- IDAHO: OREGON, eol.
- IOWA: ARIZONA, eol.
- NEVADA: ARIZONA, eol
- OREGON: NEVADA, eol.
- UTAH: HAWAII, eol.

For the second construction, the vertices are A, C, H, I, N, O, U. I'll put the corresponding place names in parenthesis after the link in the list.

- A: A(ARIZONA), eol.
- C: O(COLARADO), eol.
- H: I(HAWAII), eol.
- I: A(IOWA), O(IDAHO), eol.
- N: A(NEVADA) eol
- O: N(OREGON), eol.
- U: H(UTAH), eol.

3. Draw the graph you'd create for the list:

   I'll leave this to you, because it is hard to word process.

4. *Bonus questions for practice only, skip if timing* Give the adjacency matrix for this graph

   I'll leave this to you, because it is hard to word process.

5. **Encoding**
   A *run* in a binary sequence is a consecutive subsequence that has all identical bit values, and can't be extended further in either direction

with that value. Every binary sequence can be decomposed into runs. For example, if the sequence is 011110011111000000, it can be decomposed as 0 1111 00 11111 000000. (10 points each part)

(a) How many binary strings of length n have exactly k runs? You can leave your expression in terms of binomial coefficients and factorials.

The first run begins at position 1, and there are $k-1$ other positions where a run begins. We know that within each run, the bits are all the opposite bit from the previous run. So if we know the first bit, b_1, and we know the $k-1$ positions where the later runs start, we know the complete string. Conversely, every set of $k-1$ positions between 2 and n and each starting bit determine a different string. So there are exactly $2 * \binom{n-1}{k-1}$ such runs.

(b) What is the minimum number of bits we need to code a binary sequence of length n There are exactly $2 * \binom{n-1}{k-1}$ so we need $\lceil log_2 2 * \binom{n-1}{k-1} \rceil = 1 + \lceil log_2 \binom{n-1}{k-1} \rceil$ bits.

(c) Describe an encoding algorithm that takes such a sequence and encodes it using the number of bits above. You can use algorithms from class (such as Ranking/Unranking) without further explanation as subroutines.

We can use the same approach we used to count. First, compute the first bit and then to encode the string from positions 2 to n, encode it with a 1 if that position has a different value than the one before it, i.e., if a run starts at that position and a 0 otherwise. The second part is a string of length $n-1$ with exactly $k-1$ ones, so we can further encode it using *Ranking* with parameters $n-1$ and $k-1$. The first bit contributes the 1 bit, and *Ranking* has length exactly $\lceil log_2 \binom{n-1}{k-1} \rceil$ bits.

(d) *bonus question, skip if timing* Describe the corresponding decoding algorithm.