CS 70 Discrete Mathematics and Probability Theory
Fall 2015 Satish Rao HW 7

Due Wednesday Oct 14 at 10PM

1. **(Error-correcting codes)**

   (a) Recall from class the error-correcting code for erasure errors, which protects against up to k lost packets by sending a total of $n+k$ packets (where n is the number of packets in the original message). Often the number of packets lost is not some fixed number k, but rather a *fraction* of the number of packets sent. Suppose we wish to protect against a fraction α of lost packets (where $0 < \alpha < 1$). What is the total number of packets that we need to send (as a function of n and α)?

   Answer: Suppose we send a total of m packets (where m is to be determined). Since at most a fraction α of these are lost, the number of packets received is at least $(1-\alpha)m$. But in order to reconstruct the polynomial used in transmission, we need at least n packets. Hence it is sufficient to have $(1-\alpha)m \geq n$, which can be rearranged to give $m \geq \frac{1}{1-\alpha}n$.

   (b) Repeat part (a) for the case of general errors.

   Answer: Suppose we send a total of $m = n+2k$ packets, where k is the number of errors we can guard against. The number of corrupted packets is at most αm, so we need $k \geq \alpha m$. Hence $m \geq n+2\alpha m$. Rearranging gives $m \geq \frac{1}{1-2\alpha}n$.
   Note: Recovery in this case is impossible if $\alpha \geq 1/2$.

2. **(Berlekamp–Welch algorithm)**

   In this question we will go through an example of error-correcting codes with general errors. We will send a message (m_0, m_1, m_2) of length $n = 3$. We will use an error-correcting code for $k = 1$ general error, doing arithmetic modulo 5.

   (a) Suppose $(m_0, m_1, m_2) = (4, 3, 2)$. Use Lagrange interpolation to construct a polynomial $P(x)$ of degree 2 (remember all arithmetic is $\bmod 5$) so that $(P(0), P(1), P(2)) = (m_0, m_1, m_2)$. Then extend the message to lengeth $n+2k$ by appending $P(3), P(4)$. What is the polynomial $P(x)$ and what is the message $(c_0, c_1, c_2, c_3, c_4) = (P(0), P(1), P(2), P(3), P(4))$ that is sent?

   Answer: We use Lagrange interpolation to construct the unique quadratic polynomial $P(x)$ such

that $P(0) = m_0 = 4, P(1) = m_1 = 3, P(2) = m_2 = 2$.

$$\begin{aligned}
\Delta_0(x) &= \frac{(x-1)(x-2)}{(0-1)(0-2)} = \frac{x^2-3x+2}{2} \\
\Delta_1(x) &= \frac{(x-0)(x-2)}{(1-0)(1-2)} = \frac{x^2-2x}{-1} \\
\Delta_2(x) &= \frac{(x-0)(x-1)}{(2-0)(2-1)} = \frac{x^2-x}{2} \\
P(x) &= m_0\Delta_0(x) + m_1\Delta_1(x) + m_2\Delta_2(x) \\
&= 4\Delta_0(x) + 3\Delta_1(x) + 2\Delta_2(x) \\
&= -x+4
\end{aligned}$$

[Note that all arithmetic is $\bmod 5$, so for example $2^{-1} \equiv 3 \pmod 5$]. Then we compute $P(3) = 1$ and $P(4) = 0$, so our message is 43210.

(b) Suppose the message is corrupted by changing c_0 to 0. We will locate the error using the Berlekamp–Welsh method. Let $E(x) = x + b_0$ be the error-locator polynomial, and $Q(x) = P(x)E(x) = a_3x^3 + a_2x^2 + a_1x + a_0$ be a polynomial with unknown coefficients. Write down the system of linear equations (involving unknowns a_0, a_1, a_2, a_3, b_0) in the Berlekamp–Welsh method. You need not solve the equations.

Answer: The message received is $(c_0', c_1', c_2', c_3', c_4') = (0,3,2,1,0)$. Let $R(x)$ be the function such $R(i) = c_i'$ for $0 \le i < 5$. Let $E(x) = x + b_0$ be the error-locator polynomial, and $Q(x) = P(x)E(x) = a_3x^3 + a_2x^2 + a_1x + a_0$. Since $Q(i) = P(i)E(i) = R(i)E(i)$ for $1 \le i < 5$, we have the following equalities $\pmod 5$:

$$\begin{aligned}
Q(0) &= 0E(0) \\
Q(1) &= 3E(1) \\
Q(2) &= 2E(2) \\
Q(3) &= 1E(3) \\
Q(4) &= 0E(4)
\end{aligned}$$

They lead to the following system of linear equations:

$$\begin{array}{rcrcrcrcrcl}
 & & & & & & a_0 & & & = & 0 \\
a_3 & + & a_2 & + & a_1 & + & a_0 & - & 3b_0 & = & 3 \\
8a_3 & + & 4a_2 & + & 2a_1 & + & a_0 & - & 2b_0 & = & 4 \\
27a_3 & + & 9a_2 & + & 3a_1 & + & a_0 & - & b_0 & = & 3 \\
64a_3 & + & 16a_2 & + & 4a_1 & + & a_0 & & & = & 0
\end{array}$$

(c) The solution to the equations in part (b) is $b_0 = 0, a_0 = 0, a_1 = 4, a_2 = 4, a_3 = 0$. Show how the recipient can recover the original message (m_0, m_1, m_2).

Answer: From the solution, we know

$$\begin{aligned}
Q(x) &= a_3x^3 + a_2x^2 + a_1x + a_0 = -x^2 + 4x \\
E(x) &= x + b_0 = x
\end{aligned}$$

Since $Q(x) = P(x)E(x)$, the recipient can compute $P(x) = Q(x)/E(x) = -x+4$ [note that this is the same polynomial $P(x)$ from part (a) used by the sender]. The recipient may deduce the

location of the error from $E(x)$ as follows. There is only one error at location e_1, we have $E(x) = (x - e_1) = x$, so $e_1 = 0$ and the error is at position 0. To correct the error we evaluate $P(0) = 4$. Since the other two positions m_1, m_2 of the message are uncorrupted, we recover the original message $(m_0, m_1, m_2) = (4, 3, 2)$.

3. **(Counting, counting, and more counting)**
The only way to learn counting is to practice, practice, practice, so here is your chance to do so. We encourage you to leave your answer as an expression (rather than trying to evaluate it to get a specific number).

   (a) How many 10-bit strings are there that contain exactly 4 ones?

   Answer: This is just the number of ways to choose 4 positions out of 10 positions to place the ones, and so is $\binom{10}{4}$.

   (b) How many different 13-card bridge hands are there? (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.)

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

   (c) How many different 13-card bridge hands are there that contain no aces?

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

   (d) How many different 13-card bridge hands are there that contain all four aces?

   Answer: 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 $\binom{48}{9}$.

   (e) How many different 13-card bridge hands are there that contain exactly 6 spades?

   Answer: 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.

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

   Answer:

   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, $\sum_{k=50}^{99}\binom{99}{k} = \frac{1}{2}2^{99} = 2^{98}$.

   Answer 2: 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 $|A| = |B| = \frac{1}{2}2^{99}$. The answer we sought was $|A| = 2^{98}$.

   (g) How many different anagrams of FLORIDA are there? (An anagram of FLORIDA is any re-ordering of the letters of FLORIDA, i.e., any string made up of the letters F, L, O, R, I, D, and A, in any order. The anagram does not have to be an English word.)

   Answer: This is the number of ways of rearranging 7 distinct letters and is 7!.

(h) How many different anagrams of ALASKA are there?

Answer: In this 6 letter word, the letter A is repeated 3 times while the other letters appear once. Hence, the number 6! overcounts the number of different anagrams by a factor of 3! (which is the number of ways of permuting the 3 A's among themselves). Hence, there are $\frac{6!}{3!}$ different anagrams.

(i) How many different anagrams of ALABAMA are there?

Answer: 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). Hence, there are $\frac{7!}{4!}$ anagrams.

(j) How many different anagrams of MONTANA are there?

Answer: 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 $\frac{7!}{2! \times 2!}$ different anagrams.

(k) If we have a standard 52-card deck, how many ways are there to order these 52 cards?

Answer: The first position of the ordering has 52 choices for the card, the second position has 51 choices, and so on, until the last position where there is only one choice. Thus, there are 52! ways to order the deck.

(l) 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?

Answer: 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 $\frac{104!}{2^{52}}$.

(m) 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).

Answer: 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.

(n) 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).

Answer:

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 notes 10), 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?

(o) 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).

Answer: 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.

(p) 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?

Answer:

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 $\frac{20!}{10! \cdot 2^{10}}$ 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 $\frac{\binom{20}{2}\binom{18}{2}\cdots\binom{2}{2}}{10!}$ 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.

4. **(Algebraic vs. combinatorial proofs)**

Consider the following identity:

$$\binom{2n}{2} = 2\binom{n}{2} + n^2.$$

(a) Prove the identity by algebraic manipulation (using the formula for the binomial coefficients).

Answer:

$$\begin{aligned}\binom{2n}{2} &= \frac{2n(2n-1)}{2} \\ &= n(2n-1) \\ &= n(n-1)+n^2 \\ &= 2\frac{n(n-1)}{2}+n^2 \\ &= 2\binom{n}{2}+n^2.\end{aligned}$$

(b) Prove the identity using a combinatorial argument.

Answer: The left hand side is the number of ways to choose two elements out of $2n$. Counting in another way, we first divide the $2n$ elements (arbitrarily) into two sets of n elements. Then we consider three cases: either we choose both elements out of the first n-element set, both out of the second n-element set, or one element out of each set. The number of ways we can do each of these things is $\binom{n}{2}$, $\binom{n}{2}$, and n^2, respectively. Since these three cases are mutually exclusive and cover all the possibilities, summing them must give the same number as the left hand side. This completes the proof.

Comment: To see why picking one element from each set is n^2, we see that for each choice from the first set, there are n choices from the second set. And since there are n elements in the first set, by the Product Rule, the total number of ways to do this is $(n \cdot n) = n^2$.

5. **(To infinities and beyond)**

Show whether each of the following sets is finite, countably infinite, or uncountable:

(a) $\mathbb{N}$ (the set of all natural numbers)

Answer: Countable and infinite. The identity map is a trivial bijection from $\mathbb{N}$ to $\mathbb{N}$.

(b) $\mathbb{Z}$ (the set of all integers)

Answer: Countable and infinite. Let $f : \mathbb{Z} \to \mathbb{N}$ be given by

$$f(x) = \begin{cases} 2x & x \geq 0 \\ -2x-1 & x < 0 \end{cases}$$

f takes the non-negative integers to the even natural numbers and the negative integers to the odd natural numbers. It is obvious that f is one-to-one. It should also be easy to see that f is a bijection.

(c) $\mathbb{Q}$ (the set of all rational numbers, i.e., numbers that can be expressed in the form a/b, where $a, b \in \mathbb{Z}$ and $b \neq 0$)

Answer: Countable and infinite. See Lecture Note 10.

(d) $\mathbb{R}$ (the set of all real numbers)

Answer: Uncountable. This can be proved using a diagonalization argument, as shown in class. See Lecture Note 10.

(e) $\mathbb{C}$ (the set of all complex numbers)

Answer: Uncountable. $\mathbb{R} \subset \mathbb{C}$, and $\mathbb{R}$ is uncountable, so $\mathbb{C}$ must be uncountable too.

(If $\mathbb{C}$ was countable, there would be an enumeration of $\mathbb{C}$; but then we could cross off the numbers in that enumeration that aren't real numbers and obtain an enumeration of $\mathbb{R}$, which is impossible. In general, any subset of a countable set is countable, and any superset of an uncountable set is uncountable.)

(f) $\{0,1\}^*$ (the set of all finite-length binary strings)

Answer: Countable and infinite. As explained in Lecture Note 10, the set $\{0,1\}^*$ can be enumerated:

$$\{0,1\}^* = \{\varepsilon, 0, 1, 00, 01, 10, 11, 000, 001, 010, 011, 100, 101, 110, 111, 1000, \ldots, \}.$$

(g) $\{0,1,2\}^*$ (the set of all finite-length ternary strings)

Answer: Countable and infinite. Similar reason as in part (f).

(h) The set of all primes.

Answer: It is well known that there are infinitely many primes. For the purposes of this class, **we consider the statements "$|P|$ is infinite" and $|P| \geq |\mathbb{Z}|$ to be equivalent**[1]. Since $P \subset \mathbb{Z}$, we also know that $|P| \leq |\mathbb{Z}|$. To be pedantic, we can say that the "identity" function $f(x) = x$, thought of as a function from primes to the integers, is clearly one-to-one.

(i) The set of all graphs

Answer: It's really enough to notice that graphs have finite descriptions, by just writing out the number of nodes, and listing all of the edges. There are infinitely many of them (for each $n \in \mathbb{Z}$, there's at least a graph with n nodes and no edges), and since each graph has a finite description, the set of all graphs is countable.

Stating exactly how one can store a graph in a bit string so that no information about the graph is lost is a tedious exercise that we do not need to bother with in CS70.

[1] A proof due to Euclid himself uses the fundamental theorem of arithmetic to prove the infinitude of primes. Suppose there are finitely many of them; take the product of all primes, and add 1; the new number is congruent to 1 modulo all of the "known" primes, so its prime factorization must contain a prime that wasn't on your original list. To be truly rigorous in proving that $|\mathbb{Z}| \leq |P|$, we can then define a function $f : \mathbb{N} \to P$ which starts out with, e.g., just $\{2\}$ as the set of "known" primes, and, given n repeats this procedure n times to produce the n'th "new" prime.