CSM CS70 Fall 2018 Mock Midterm 2 Solutions

October 28 2018

1. **True/False (2 points each)**

   (a) X is picked randomly variable over the set $\{0, 1, 2, 3, 4, 5, 6\}$, and A is a uniform random variable over the set $\{1, 2, 3, 4, 5, 6\}$. Define $Y = AX \pmod 7$. Then X and Y are identically distributed.
   True. Every value for A has a unique multiplicative inverse mod 7, so Y is also uniformly distributed on $\{0, 1, 2, 3, 4, 5, 6\}$.

   (b) In GF(p), any polynomial of degree greater than p has an equal polynomial representation with degree less than p.
   True. There are at max p points that you can use in GF(p), and we can always construct a degree $p-1$ degree polynomial through these points..

   (c) $(x_1 + x_2 + ... + x_n)^p \equiv x_1^p + x_2^p + ... + x_n^p \pmod p$ for a prime p True, by FLT:

   $$(x_1 + \ldots + x_n)^p = x_1 + \ldots + x_n \mod p$$

   $$x_1^p + \ldots + x_n^p = x_1 + \ldots + x_n \mod p$$

   (d) The set of all finite length strings consisting of characters from the English alphabet is uncountably infinite. False. Treat the set of strings as numbers in base 26, where A = 0, B = 1, ... Z = 25. Then, since we know we can write any natural number in base 26, there is a bijection between the natural numbers and the set of alphabetical strings.

   (e) Say that you are working in GF(p) have a channel that will corrupt k packets, where $k > \frac{p}{2}$. It is still possible to communicate some length message over this channel.
   False. you need to send at least $2k$ extra packets, and $2k > p$. (We only have p unique points to use

   (f) We can construct a program $\mathcal{P}$ that, given another program $\mathcal{Q}$ and input x as inputs, can determine if $\mathcal{Q}(x)$ halts in $3^{|x|}$ steps. True, we can run any program and input for a finite number of steps and check if it halts by that stage. The halting problem only becomes an issue when we deal with infinite loops. (Also, $|x|$ is the size of input x)

   (g) The number of ways to rearrange n distinct letters is greater than the number of ways to choose $\frac{n}{2}$ letters from the first half with replacement. True. The number of ways to rearrange n distinct letters is $n!$. The number of ways to choose a letter from $\frac{n}{2}$ possibilities is $\frac{n}{2}$. Since we replace after each draw, the total number of possibilities becomes $(\frac{n}{2})^{\frac{n}{2}}$, which is less than $n!$.

   (h) $\sum_{i=1}^{\frac{n}{2}} \binom{n}{i} = \sum_{i=\frac{n}{2}+1}^{n} \binom{n}{i}$, where $n \geq 2$ is even. False. Even though $\binom{n}{k} = \binom{n}{n-k}$, the left hand side does not include the $i = 0$ case.

   (i) $\sum_{i=0}^{n} \binom{n}{i} = n^2$ False. The sum actually evaluates to 2^n. This is one of the fundamental properties of Pascal's Triangle, and can be proved using a combinatorial argument. Both sides count the number of possible subsets of a set of size n.

   (j) Given a sample space Ω and event A, A and Ω are always independent. True. Notice that the intersection of A and Ω is just A. Then: $P(A|\Omega) = \frac{P(A \cap \Omega)}{P(\Omega)} = \frac{P(A)}{1} = P(A)$, therefore A and Ω are independent.

(k) Take a deck of n cards where each card has a unique number in $1, ..., n$. You draw cards one by one without replacement. Let X_i be the number on the ith card you pick up. Then X_1 and X_2 are identically distributed. True. They are identical, but not independent.

(l) For disjoint events A and B, the $Pr(A \cap B) = Pr(A) \cdot Pr(B)$
False. $Pr(A \cap B) = 0$ for disjoint events. (NOTE: independent events are not the same as disjoint)

(m) If Pr(A) > 0 and Pr(B) > 0, and A and B are disjoint then A and B are not independent.
True. Disjoint events means that Pr(A|B) = 0. For independent events, Pr(A|B) = Pr(A). Since Pr(A|B) = 0 and Pr(A) $\neq$ 0, then Pr(A|B) $\neq$ Pr(A), so they are not independent.

(n) Given two events A and B with Pr(A) > 0 and Pr(B) > 0, if Pr(A|B) > Pr(A), then Pr(B|A) < Pr(B).
False. Plug into Bayes' rule. Pr(B|A) = $\frac{Pr(B \cap A)}{Pr(A)}$ = $\frac{Pr(A|B)Pr(B)}{Pr(A)}$ > $\frac{Pr(A)Pr(B)}{Pr(A)}$ = Pr(B) $\rightarrow$ $Pr(B|A) > Pr(B)$.

(o) If two events A and B are independent, then A^c and B^c are dependent. False. If A and B are independent, so are A^c and B^c. $P(A^c)P(B^c) = (1 - P(A))(1 - P(B)) = 1 - P(B) - P(A) + P(A)P(B) = 1 - (P(A) + P(B) - P(A \cap B)) = 1 - P(A \cup B)$ (by inclusion-exclusion). The only events that are not in the union of A and B are the events that are both not in A and not in B, which is precisely $A^c \cap B^c$, meaning $1 - P(A \cup B) = P(A^c \cap B^c)$, meaning A^c and B^c are independent.

2. **Short Answer: RSA (10 points)**
LeBron wants to send a RSA-encrypted message M to his friends Kevin and Chris. Kevin and Chris use public keys (N, e_1) and (N, e_2). Notice they use the same modulus, but different exponents. LeBron sends $C_1 = M^{e_1} \mod N$ to Kevin and $C_2 = M^{e_2} \mod N$ to Chris. You eavesdrop on LeBrons transmissions and learn C_1 and C_2.

Show how you can recover M. You may assume that you know e_1, e_2 and N and that e_1 and e_2 are relatively prime.

This is taken from the Spring 2005 iteration of this course. Since $gcd(e_1, e_2) = 1$, e_1 has an inverse in mod e_2, and vice versa. That means there must exist some linear combination of e_1, e_2 that adds up to 1, i.e. there exist integers a, b such that $ae_1 + be_2 = 1$. By the extended Euclid's GCD algorithm, we can find these values. Then, we can compute M as follows: $(C_1)^a(C_2)^b \equiv (M^{e_1})^a(M^{e_2})^b \equiv M^{ae_1+be_2} \equiv M^1 \equiv M \pmod N$

3. **Short Answer: Polynomials and Error Correction (3 points each)**

(a) A polynomial has 4 roots. What is the minimum degree? 4. Degree n polynomials have at most n roots, so with 4 roots the polynomial is at least degree 4.

(b) Suppose $P(x)$ and $Q(x)$ are two distinct polynomials (of degree d_1 and d_2 respectively) which intersect in exactly 5 points. If the lowest degree polynomial that contains those five points has degree 3, what is the minimum value of $d_1 + d_2$?
This is very similar to a problem from Fall 17's midterm 2. Since $P(x) = Q(x)$ at exactly 5 points, the polynomial $P(x) - Q(x)$ must be of at least degree 5. We know from the second part of the problem that any polynomial passing through those points must have degree at least three, therefore we can let $P(x)$ have degree 5 and $Q(x)$ have degree 3. Note that subtracting a polynomial of degree 3 from one of degree 5 still results in a polynomial of degree 5. Then, our answer is 5 + 3 = 8.

(c) Prove or disprove:

i. The set of all polynomials of degree 3 that interpolate $(1, 1), (2, 4)$ and $(3, 10)$ is countably infinite.
False. We know there are infinitely many degree 3 polynomials that pass through these three points. Determining a degree 3 polynomial that passes through them is equivalent to fixing some fourth point and interpolating that polynomial. However, since we're dealing with the set of real numbers, since that set is uncountable, there are uncountably many points we could choose to be our fourth point, and therefore the set of degree 3 polynomials that passes through those three points is uncountably infinite.

ii. Now, re-answer the problem above, but suppose we are working in $GF(p)$ for some prime $p \geq 11$.
False. This set is actually finite! To determine the degree 3 polynomial, we simply need to fix one other point, just as before. However, we know now our coordinates must be integers, and must be less than or equal to $p-1$. There are finitely many integers less than $p-1$, therefore this set is countable.

(d) Suppose each point represents one character of a message, and that the ith letter in the alphabet is represented by the number i (so A = 1, B = 2, ...). Suppose we want to send the message outlined in part a) (that is, ADJ), but we know that 1 character of our message is going to be corrupted. Determine the number of extra points we need to send, and find those points. Use $GF(13)$ and the correct degree polynomial.
This is determined by Lagrange Interpolation. We find that $P(x) = 8x^2 - 8x + 1$. Since $n = 3$ and $k = 1$, we need to send 2 extra points. We send $P(4) = 6$ and $P(5) = 5$, i.e. $(4, 6)$ and $(5, 5)$.

(e) Consider an erasure channel through which you want to send a message of length n.

i. If $\frac{1}{4}$ of your packets are going to be dropped, how many total packets must you send?
Suppose we send k extra. Then (fraction not dropped) * (total number sent) should be equal to n. We then have $\frac{3}{4} * (n + k) = n$, which yields $\frac{3}{4}k = \frac{1}{4}n$, meaning we should send at least $k = \frac{1}{3}n$ extra packets and at least $\frac{4}{3}n$ total.

ii. Now suppose $\frac{1}{4}$ of your packets are going to be corrupted instead of dropped. How many total packets must you send to combat this error?
Let m be the number of extra packets we send. We know from Berlekamp-Welch that (original message) + 2 * (number dropped) = (number received). Since a quarter of our packets are dropped, we know number dropped $= \frac{1}{4}(n+m)$. We also must subtract that amount from our original message, so original message $= n - \frac{1}{4}(n+m)$. Lastly, we have the number received, which is $\frac{3}{4}(n+m)$. Solving the equation $\frac{3}{4}(n+m) = n - \frac{1}{4}(n+m) + 2 \cdot \frac{1}{4}(n+m)$ yields $m = n$, meaning we should send at least $2n$ packets total. *Alternatively, we can say that (fraction not changed) * (# sent) = (# received), with # sent $= n + 2k$ and # received $= n + k$. This yields $k = \frac{n}{2}$ meaning we should send $n + 2\frac{n}{2} = 2n$ packets total.*

4. **Short Answer: Countability**

(a) Let A and B be two countable sets. Define $A \times B = \{(a, b) : a \in A, b \in B\}$. Show that $A \times B$ is countable. This follows from the proof in lecture that $\mathbb{N} \times \mathbb{N}$ is countable.

(b) Show that the set $\mathbb{N}^k = \{(a_1, a_2, ..., a_k) : a_i \in \mathbb{N} \ \forall i = 1, ..., k\}$ is countable. This follows from inductively applying the property in part (a). Base case: $\mathbb{N} \times \mathbb{N}$ is countable, by part (a). Inductive step: Assume $\mathbb{N}^k$ is countable. $\mathbb{N}^{k+1} \cong \mathbb{N}^k \times \mathbb{N}$ is countable, once again by part (a).

(c) Let P_d be the set of integer co-efficient polynomials of degree at most d. Show that P_d is countable for some fixed d. To specify an integer co-efficient polynomial $p(x) = \sum_{i=0}^{d} a_i x^i$, it suffices to specify what the coefficients a_i are. In particular, every sequence of coefficients $(a_0, a_1, ..., a_d)$ uniquely specifies a polynomial in this set, so the size of this set is the same as the size of $\mathbb{N}^{d+1}$, and hence this set if countable.

(d) Show that the set P of all integer coefficient polynomials of finite degree is countable. (Hint: Can you write P in terms of the sets P_d?)
$P = \cup_{d=0}^{\infty} P_d$. i.e. P is simply the union of all the P_ds. This is a countable union of countable sets, and is hence countable.

(e) An algebraic number is a real number that can be written as the root of an integer coefficient polynomial. Show that the set of algebraic numbers is countable.
We saw above that the set of all integer polynomials is countable. Each polynomial has finitely many real roots (at most the degree). The set of algebraic numbers is the union of the sets of roots of all polynomials in P, so the set of algebraic numbers is written as a countable union of finite sets, and is hence countable.

5. **Short Answer: Secret-Sharing (5 points)**

   (a) In a secret-sharing scheme in $GF(p)$ where k is the minimum number of people required to recover the secret, is successful secret recovery more probable when $k-1$ people collaborate compared to random guessing? Explain your answer. k points uniquely determine a polynomial of degree $(k-1)$. Given $(k-1)$ points, there exist p possible polynomials since the kth point can take any value between 0 and $(p-1)$. Probability of successfully guessing the secret is $\frac{1}{p}$, which is the same as the case of random guessing. Thus, having $k-1$ points does not give us any advantage over having 0 points.

   (b) In a secret-sharing scheme in $GF(p)$ where the secret is s. Can we pick any p? Explain your answer. No. p has to be larger than s so that our secret is a valid value inside $GF(p)$

6. **Short Answer: Self-reference/Uncomputability (10 points)**

   (a) Suppose you are given a program $\mathcal{P}$ which, if run forever, eventually prints every input to another program $\mathcal{Q}$ that halts in finite time. Does this exist? If so, describe the program. If not, expain why. It can exist. Create a grid of inputs x to $\mathcal{Q}$ and timesteps 1, 2, ...: simulate each pair in a diagonal snaking pattern from the $|\mathbb{N}| = |\mathbb{N} \times \mathbb{N}|$ proof.

7. **Short Answer: Counting (3 points each)**

   (a) Suppose we have a set of ordered items: $a_1 < a_2 < \ldots < a_n$.

      i. How many sets have a contiguous set of items: a set where every element is either the minimum item, the maximum item, or for a given a_i, both a_{i-1} and a_{i+1} are in the set? $\binom{n}{2} + n$. There are $\binom{n}{2}$ possibilities for sets of size 2 or larger, and there are n possibilities for size 1 sets.

      ii. Now suppose that an item $a_i, i \in \{1, \ldots, n\}$ must be included in each set. How many sets are there? $i \cdot (n-i+1)$. You have to choose a minimum element up to a_i, and a maximum element up to a_n. You can try this on a small set like $n = 3$ to see this.

   (b) How many permutations exist of the word "BERKELEY" with "BRK" appearing together, but not necessarily in that order?
   Block BRK together. Then, there are 6 "blocks", with 3 Es, yielding $\frac{6!}{3!}$ ways to arrange that. However, we need to multiply by the movement of BRK; there are 3! ways to arrange it. Therefore, our final answer is $\frac{6!3!}{3!} = 6!$.

   (c) Determine:

      i. The number of simple undirected (and possibly disconnected) graphs on n vertices, with the vertices labelled $1, 2, 3, ...n$. Justify your answer.
      For ever pair of vertices $(u, v) \in G$, there are two options: Either the two are connected by an edge, or not. There are $\binom{n}{2}$ such pairs (since direction does not matter). Then, the total number of ways of building such a graph is $2^{\binom{n}{2}}$.

      ii. The number of simple directed (and possibly disconnected) graph on n vertices, with the vertices labelled as above. Justify your answer.
      For ever pair of vertices $(u, v) \in G$, there are 3 options: Either there is no edge between them, or there is an edge from u to v, or there is one from v to u. There are $\binom{n}{2}$ such pairs. So the total number of ways of building such a graph is $3^{\binom{n}{2}}$. This is assuming there cannot be two directed edges between a pair of vertices u and v (and no self loops).

8. **Short Answer: Probability (3 points each)**

   (a) Suppose your local McChipotle's ice cream machine is broken on any given day with probability p, independent of other days. What is the probability that April 5th is the second day in April that the machine works? We need the machine to work one day of the first four. There are $\binom{4}{1}$ ways to pick which day of the first four the machine works on. For each of those 4 ways, the probability of the machine working once and being broken three times is $p^3(1-p)$. Multiplying that with the probability of working on the fifth day yields $4p^3(1-p)^2$ to be the final answer.

(b) You have a coin that turns up heads with p probability and one that does so with q probability. You flip each one n times and you mark each pair of flips as a success if one of the two coins was heads. What is the expected number of successes? We can model each event as a success with probability $1-(1-p)(1-q) = p+q-pq$. Then the expected number of timesteps is $\text{Binom}(p+q-pq) = n(p+q-pq)$

(c) Jane is flipping coins!

i. First, Jane flips 9 fair coins. What is the probability that she gets an even number of heads? Justify your answer. *Hint: use a counting argument* 1/2. Because these are fair coins, all outcomes of the 9 flips are equally likely, with probability 2^{-9}. We also know that $\binom{n}{k} = \binom{n}{n-k}$, meaning $P(H=k) = P(H=n-k)$. Because there are an odd number of coins, if k is odd, then $n-k$ is even, and vice-versa- the probability of getting 0 heads is the same as the probability of getting 9 heads, 1 is the same as 8, etc.

ii. Now Jane flips 10 coins. What is the probability that she gets an even number of heads? What about if she flips n coins? Justify your answers. Suppose Jane flips 9 coins first- she'll get even number of heads half of the time, and odd number of heads half of the time. Now when she flips the 10th coin, it'll either be heads or tails with probability 1/2 each. You get an even number of heads when flipping 10 coins by either getting an odd number of heads with 9 coins, and then flipping a head, or by getting an even number of heads with 9 coins, then flipping a tail. These combine to give you probability 1/2. This generalizes to any odd and even number of coins n.

iii. Suppose Jane flips n **unfair** coins (where $P(H) \neq 1/2$), along with a single fair coin. What is the probability that she gets an even number of heads after flipping these $n+1$ coins? Justify your answer.
Note that because outcomes no longer occur with uniform probability, we cannot use the counting argument from the first part anymore. Instead, let the probability of flipping an even number of heads among just the n unfair coins be p. Similar to our argument for 10 coins, we note that flipping an even number of heads altogether is equal to flipping an even number of heads among the unfair coins then flipping a tail on the fair coin, or flipping an odd number of heads among the unfair coins then flipping a head on the fair coin. This occurs with probability $p(1/2) + (1-p)(1/2) = 1/2$.

(d) Pick a random integer n in the range from 0 to 999,999 each with equal probability.

i. What is the probability that the decimal digits of n add up to 8? $\binom{13}{5} * 10^{-6}$. To calculate the number of 6-digit integers with decimal digits that sum up to 8, count the number of ways we can order 8 stars and 5 bars (ex: $**||***|*|*|*$ represents 203,111). The number of possible 6-digit integers is 10^6.

ii. What is the probability that the decimal digits of n add up to 10? $(\binom{15}{5} - 6) * 10^{-6}$. To calculate the number of 6-digit integers with decimal digits that sum up to 10, count the number of ways we can order 10 stars and 5 bars (ex: $**|*|****|*|*|*$ represents 214,111). The difference here is if all 10 stars fall between 2 bars, we can't represent that with a single digit, so we need to subtract out the 6 ways in which this could happen (ex: $|||**********||$)

(e) A particle sits on the real number line, starting at the origin (0). At each timestep, we flip a fair coin and move the particle as follows:

- If we see heads, we move the particle one unit to the left
- If we see tails, we move the particle one unit to the right

Let X_n be the position of the particle at time-step n, and assume $X_0 = 0$.

i. Find $P(X_{1000} = 0)$.
We want to find the probability that the particle returns to its position after 1000 timesteps. At each timestep, the particle either moves forward by one (+1) or backward by one (-1). In order for the particle to return where it started, it must have moved forward and backward exactly the same number of times. So, we want to find the probability that the 1000 timesteps

included exactly 500 +1s. This is equivalent to flipping a coin 1000 times and seeing 500 heads and 500 tails. There are $\binom{1000}{500}$ ways to choose which flips to be heads, so this probability is $\binom{1000}{500}(1/2)^{1000}$.

ii. What is the most likely position for the particle at time $t = 2k$, where k is a non-negative integer? *Hint: Think about the symmetry of Pascal's triangle.*
It turns out that the result above is the most likely one, by symmetry (that is, that the particle ends back at the origin). For any combination of k heads and $1000 - k$ tails, our probability will include a factor of $(1/2)^{1000}$ due to the fact that this is a fair coin. Then, we just need to maximize the combinatorial term, which is maximized in the middle, that is $\binom{1000}{500}$. An equivalent argument is that the most likely outcome when flipping 1000 coins is 500 heads and 500 tails.

iii. In general, what is $P(X_n = 0|X_0 = 0)$ in terms of n?
From the given information, we see that it is impossible to return to the starting position if n is odd. So we have $P(X_n = 0|X_0 = 0) = 0$ for odd n, and $\binom{n}{n/2}(1/2)^n$ for even n.