CS 70 Discrete Mathematics and Probability Theory
Spring 2015 Vazirani Midterm #2 Solution

PRINT your name: _____, _____
(last) (first)

SIGN your name: _____

PRINT your student ID: _____

CIRCLE your exam room: 3106 Etcheverry 3108 Etcheverry 180 Tan

241 Cory 247 Cory 1 Pimentel

Name of the person sitting to your left: _____

Name of the person sitting to your right: _____

Please write your name and student ID on every page.

Please write your answers in the spaces provided in the test. We will not grade anything on the back of an exam page or outside the space provided for a question unless we are clearly told on the front of the page in the space provided for the question to look there.

You may consult one double-sided sheet of handwritten notes. Apart from that, you may not look at books, notes, etc. Calculators and computers are not permitted.

You have 110 minutes. There are 5 questions worth a total of 100 points. Use the number of points as a rough guide for the amount of time to allocate to that question. Note that many of the points are for proofs and justifications for your answers. Please make sure you spend the time to write clear, correct and concise justifications. Good luck!

Do not turn this page until your instructor tells you to do so.

Q.1 / 40	Q.2 / 10	Q.3 / 10	Q.4 / 20	Q.5 / 20	Total / 100

Name: _____ SID: _____

1. **Short Answers (40 points)**

   Provide brief justifications of your answers. In parts (d)–(f) you can leave your answers as factorials, n choose k, etc., but do explain your calculations clearly.

   (a) Are there integers x, y such that $21x + 55y = 3$?

   Answer: **Yes.** Since $\gcd(21, 55) = 1$, we know there exist integers a, b such that $21a + 55b = 1$ (we can find a and b via the extended-gcd algorithm). Setting $x = 3a$ and $y = 3b$ gives us $21x + 55y = 3$, as required.

   (b) Is the set of all C programs countable?

   Answer: **Yes.** Recall that a C program (and any program in general) can be represented as a finite-length binary string, and the set of all finite-length binary strings $\{0,1\}^*$ is countably infinite.

   (c) Consider a function f that takes as input a program P, and outputs:

$$f(P) = \begin{cases} 1 & \text{if program } P \text{ on input } P \text{ does not halt within the first 1000 steps,} \\ 0 & \text{otherwise.} \end{cases}$$

   Is f computable?

   Answer: **Yes.** We can compute f by writing a program that takes as input a program P, runs the program P on input P, and waits for 1000 steps. If the execution $P(P)$ continues for the 1001-st step, then the program f outputs 1, otherwise it outputs 0.

   (d) How many seven-card hands are there with three pairs? That is, there are two cards each of three different ranks, and one card of a different rank. For example, $(2\clubsuit, 2\heartsuit, 5\diamondsuit, 5\spadesuit, 6\clubsuit, 10\diamondsuit, 10\spadesuit)$ is one such hand with three pairs, but $(2\clubsuit, 2\heartsuit, 5\diamondsuit, 5\spadesuit, 5\clubsuit, 10\diamondsuit, 10\spadesuit)$ is not. Here the ordering does not matter.

   Answer:

$$\binom{13}{3}\binom{10}{1}\binom{4}{2}\binom{4}{2}\binom{4}{2}\binom{4}{1}$$

   There are $\binom{13}{3}$ ways to choose 3 distinct ranks for the three pairs, and $\binom{10}{1}$ ways to choose 1 rank for the lone card, which must be different from the three pairs. Once we have fixed the ranks, there are $\binom{4}{2}$ ways to choose which suits for each pair, and $\binom{4}{1}$ ways to choose the suit for the lone card.

(e) How many different ways are there to rearrange the letters of DIAGONALIZATION without the two N's being adjacent?

Answer:

$$\frac{15!}{3!3!2!2!} - \frac{14!}{3!3!2!} = \frac{13 \cdot 14!}{3!3!2!2!}$$

The word DIAGONALIZATION has 15 letters with 3 A's, 3 I's, 2 N's, and 2 O's, so there are $\frac{15!}{3!3!2!2!}$ ways to rearrange the letters in total. The number of rearrangements where the two N's are adjacent is $\frac{14!}{3!3!2!}$, where we have considered "NN" as a single character. The difference $\frac{15!}{3!3!2!2!} - \frac{14!}{3!3!2!}$ is then equal to the number of rearrangements without the two N's being adjacent.

(f) How many non-decreasing sequences of k numbers from $\{1,\ldots,n\}$ are there? For example, for $n = 12$ and $k = 7$, $(2,3,3,6,9,9,12)$ is a non-decreasing sequence, but $(2,3,3,9,9,6,12)$ is not.

Answer:

$$\binom{n+k-1}{k} = \binom{n+k-1}{n-1} = \frac{(n+k-1)!}{(n-1)!k!}$$

Each non-decreasing sequence is specified by how many times each element $i \in \{1,\ldots,n\}$ appears in the sequence. Therefore, the number of non-decreasing sequences of length k is equal to the number of solutions to the equation

$$x_1 + x_2 + \cdots + x_n = k$$

where $x_i \geq 0$ is the number of times i appears in the sequence. Each solution can be represented as a binary string of length $n+k-1$ with exactly k 1's, where the 0's represent the plus signs and the consecutive 1's represent the x_i's. Therefore, the number of such non-decreasing sequences is $\binom{n+k-1}{k} = \binom{n+k-1}{n-1}$.

For the remaining two parts: Three people each independently choose a random number between 1 and 50. Let $A_{i,j}$ be the event that persons i and j choose the same number. Let $A_{1,2,3}$ be the event that all three people choose the same number.

(g) Are $A_{1,2}$ and $A_{2,3}$ independent?

Answer: **Yes.** We can compute explicitly $\Pr[A_{1,2}] = \Pr[A_{2,3}] = \frac{50}{50^2} = \frac{1}{50}$, and

$$\Pr[A_{1,2} \cap A_{2,3}] = \Pr[A_{1,2,3}] = \frac{50}{50^3} = \frac{1}{50^2} = \Pr[A_{1,2}] \cdot \Pr[A_{2,3}]$$

Alternatively, we can argue that conditioned on the event $A_{2,3}$ (i.e., persons 2 and 3 choose the same number), person 1 still makes his choice independently, so he has $\frac{1}{50}$ chance of picking the same number as person 2. Therefore,

$$\Pr[A_{1,2} \mid A_{2,3}] = \frac{1}{50} = \Pr[A_{1,2}]$$

Note that merely saying "event $A_{1,2}$ has no bearing on/does not affect the probability of event $A_{2,3}$" is not enough; this is only restating the definition of what it means for $A_{1,2}$ and $A_{2,3}$ to be independent, but we want you to explain why it holds in this specific problem.

A common mistake students make is in calculating $\Pr[A_{1,2}]$ or $\Pr[A_{1,2,3}]$. Note that the probability that persons 1 and 2 choose the same *specific* number x is equal to $\frac{1}{50} \cdot \frac{1}{50} = \frac{1}{50^2}$. However, to calculate $\Pr[A_{1,2}]$ we need to sum over all possible values $x \in \{1, \ldots, 50\}$, giving us $\Pr[A_{1,2}] = \frac{50}{50^2} = \frac{1}{50}$. Similarly, $\Pr[A_{1,2,3}] = \frac{1}{50^2}$, not $\frac{1}{50^3}$.

(h) Are $A_{1,2}$ and $A_{1,2,3}$ independent?

Answer: **No.** Observe that the event $A_{1,2,3}$ is a subset of the event $A_{1,2}$, so

$$\Pr[A_{1,2,3} \cap A_{1,2}] = \Pr[A_{1,2,3}] \neq \Pr[A_{1,2}] \cdot \Pr[A_{1,2,3}]$$

since $\Pr[A_{1,2}] \neq 1$.

Alternatively, we can argue that if $A_{1,2,3}$ occurs, then $A_{1,2}$ must also occur (this is the same as stating that $A_{1,2,3}$ is a subset of $A_{1,2}$). More precisely,

$$\Pr[A_{1,2} \mid A_{1,2,3}] = 1 \neq \frac{1}{50} = \Pr[A_{1,2}]$$

2. **Las Vegas (10 points)**

A certain dice game in Las Vegas involves the dealer rolling either two or three standard (six-sided) dice with 50-50 probability and then reporting the total of all rolls. Suppose that after such a roll the reported total is 3. What is the probability that the dealer rolled two dice? Express your answer as a rational number in its lowest terms. Make sure you explain your calculation clearly.

Answer: Let A denote the event that two dice were rolled, so that $\bar{A}$ denotes the event that three dice were rolled. Let S denote the event that the total is 3.

The problem tells us that the prior probabilities are $\Pr[A] = \Pr[\bar{A}] = \frac{1}{2}$. We can calculate $\Pr[S \mid A]$ and $\Pr[S \mid \bar{A}]$ as follows. If we roll two dice, then there are 36 possible outcomes, 2 of which get a total of 3 — (1, 2) and (2, 1) — for $\Pr[S \mid A] = \frac{2}{36}$. If we roll three dice, there are 216 outcomes, and only (1, 1, 1) gets a total of 3, so $\Pr[S \mid \bar{A}] = \frac{1}{216}$.

To compute $\Pr[A \mid S]$, we apply Bayes' Rule:

$$\Pr[A \mid S] = \frac{\Pr[S \mid A] \cdot \Pr[A]}{\Pr[S]} = \frac{\Pr[S \mid A] \cdot \Pr[A]}{\Pr[S \mid A] \cdot \Pr[A] + \Pr[S \mid \bar{A}] \cdot \Pr[\bar{A}]} = \frac{\frac{2}{36} \cdot \frac{1}{2}}{\frac{2}{36} \cdot \frac{1}{2} + \frac{1}{216} \cdot \frac{1}{2}} = \frac{12}{13}$$

Common mistakes:

(a) Assuming all the outcomes have equal likelihood (either the 3 outcomes which sum up to 3, or all the possible sums of values of rolling 2 or 3 dice). To see why this is incorrect, consider the probability of one particular outcome rolling 2 dice, which is $\frac{1}{36}$, and the probability of rolling 3 dice, which is $\frac{1}{216}$. Since the probability of rolling 2 dice is the same as the probability of rolling 3 dice, each outcome with 3 dice is far less likely.

For the second case, consider the probability of getting sums of 2 and 7 while rolling 2 dice. Clearly, the probability of getting 7 is much higher, since there are 6 pairs of numbers which will sum to 7, while there is only one pair which will sum to 2.

(b) Treating the conditional probabilities $\Pr[S \mid A]$ and $\Pr[S \mid \bar{A}]$ as $\Pr[S \cap A]$ and $\Pr[S \cap \bar{A}]$. (This actually did not affect the result for this particular problem, since both the numerator and denominator are divided by $\frac{1}{2}$, but is still incorrect.)

(c) Forgetting to multiply by $\Pr[A]$ and/or $\Pr[\bar{A}]$ while applying Bayes' Rule. (If you did this in both the numerator and denominator, it also didn't affect the result, but still conceptually incorrect.)

3. **Infinity (10 points)**

Find the precise error in the following proof:

Theorem: The set of rationals between 0 and 1 is uncountable.

Proof: Suppose towards a contradiction that there is a bijection $f : \mathbb{N} \to \mathbb{Q}[0,1]$, where $\mathbb{Q}[0,1]$ denotes the rationals between 0 and 1. This allows us to list all the rationals between 0 and 1, with the j-th element of the list being $f(j)$. Now consider the number d along the diagonal, whose j-th digit d_j is the j-th digit of $f(j)$. We define a new number e, whose j-th digit e_j is equal to $(d_j + 5)$ mod 10. We claim that e does not occur in our list of rationals between 0 and 1. This is because e cannot be the j-th number on the list for any j, since it differs from the j-th number on the j-th digit. Contradiction. Therefore the set of rationals between 0 and 1 is uncountable.

j	$\leftrightarrow$	$f(j)$						
0	$\leftrightarrow$	0.	**5**	2	1	4	9	…
1	$\leftrightarrow$	0.	1	**4**	1	6	2	…
2	$\leftrightarrow$	0.	9	4	**7**	8	2	…
3	$\leftrightarrow$	0.	5	3	0	**9**	8	…
4	$\leftrightarrow$	0.	6	2	5	7	**2**	…
⋮					⋮			

$\Longrightarrow$ $d = 0.54792\ldots$ $\downarrow$ $e = 0.09247\ldots$

Answer: The problem is that e is irrational and therefore does not belong in our enumerated list of rational numbers $\mathbb{Q}[0,1]$ in the first place. Recall that rational numbers have decimal expansions that are either finite or periodic. We have no guarantee that the diagonal number d is rational, and since e is constructed by modifying every digit in d separately, the number e need not be rational either. Therefore, it is not a contradiction if e, an irrational, does not appear in our list of rationals. This is not a problem with our proof involving real numbers because the modified diagonal is guaranteed to also be a real number.

Common Mistakes:

(a) The initial assumption that there is a bijection and we can enumerate all rationals $\mathbb{Q}[0,1]$ is fine. Here we are trying to prove by contradiction that rationals are uncountable, so our proof is trying to find a contradiction when we start with this statement.

(b) The fact that some rationals have terminating decimal expansions is not a problem because we can always pad with zeros at the end. For example, we can represent 0.5 as $0.5000\ldots$, and thus we can still construct the number d from the diagonal.

(c) Modifying our diagonal element by having $e_j = (d_j + 5)$ mod 10 is not a problem. Recall that in our proof of the uncountability of the reals we could not add 1 to each digit because of the ambiguity $0.999\ldots = 1.000\ldots$. We avoided that problem by adding 2 to each digit, but we could also have added any other number. Similarly, here we can add 5 to each digit. In general, we can perform any procedure to the diagonal to construct a new number that is guaranteed to be different from every element in the list.

(d) It is not sufficient to say that our diagonal element, d, is not rational. If d can be modified and converted to a rational e, then we still need to be able to find e in our list of rationals. We require students to explicitly state that e is not rational to receive full credit.

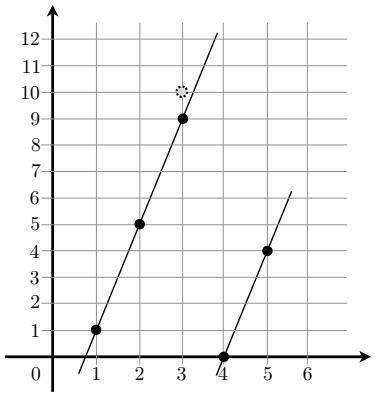
4. **Error Correction (20 points)**

Recall that a polynomial error correcting code sends a message $m_1, \ldots, m_{d+1}$ consisting of $d+1$ characters, each modulo a prime q, by constructing a polynomial $P(x)$ of degree (at most) d such that $P(i) = m_i$ for $i = 1, \ldots, d+1$. We transmit the message along with several additional points on $P(x)$ to guard against errors.

The main idea behind error correction is to try to recover the degree d polynomial that was used to create the encoded message. The way this is done is by finding a polynomial of degree d that goes through the maximum number of received points.

(a) Consider the following example where we are working modulo 13 and the message is $m_1 = 1$ and $m_2 = 5$. We encode the message using a polynomial $P(x) = 4x - 3$ of degree 1 over $GF(13)$. The transmitted points $P(1) = 1$, $P(2) = 5$, $P(3) = 9$, $P(4) = 0$, $P(5) = 4$ lie on a line (shown as black points below; the line "wraps around" because we are working mod 13).

Suppose that the received values are $r_1 = 1$, $r_2 = 5$, $r_3 = 10$, $r_4 = *$, and $r_5 = 4$, with errors in positions 3 (shown as grey dotted point below) and 4. Specify a value for r_4 such that the recipient cannot uniquely decode the message. Justify your answer.



Answer: If $r_4 \not\equiv 0 \bmod 13$, then there will be exactly 3 received values that lie on $P(x)$ (namely r_1, r_2, r_5). Therefore, if we can choose r_4 such that there is another degree 1 polynomial $Q(x)$ that also passes through 3 of the received values, then the recipient will be confused because she does not know whether the correct polynomial is $P(x)$ or $Q(x)$.

In order to do so, we simply have to interpolate a line through r_3 and any other received value r_i ($i = 1$, 2, or 5) to find $Q(x)$, and choose r_4 to be $Q(4)$ such that $Q(x)$ passes through r_i, r_3, r_4.

There are three possible answers:

i. $r_4 = 2$. The received points $r_2 = 5$, $r_3 = 10$, $r_4 = 2$ lie on the line $Q(x) = 5x + 8 \bmod 13$.

ii. $r_4 = 8$. The received points $r_1 = 1$, $r_3 = 10$, $r_4 = 8$ lie on the line $Q(x) = 11x + 3 \bmod 13$.

iii. $r_4 = 7$. The received points $r_3 = 10$, $r_4 = 7$, $r_5 = 4$ lie on the line $Q(x) = 10x + 6 \bmod 13$.

Note that merely stating $r_4 \not\equiv 0 \bmod 13$ is incorrect. The Berlekamp-Welch algorithm states that if we send 6 points, then we can recover from 2 arbitrary errors. However, it does *not* imply that if we send fewer than 6 points, then we cannot recover from 2 errors, because it might happen that the errors are not bad enough that we can still recover the message. For example, if $r_4 = 1$, then we have 2 errors, but the original polynomial $P(x)$ is still the unique polynomial passing through 3 of the received points, so in that case we can still recover the original message.

(b) In part (a) we showed that transmitting 3 extra points does not protect against two errors. In this problem you will prove that transmitting 4 extra points is sufficient to correct two errors.

Specifically, consider the same setting as in part (a), but suppose we transmit 6 points instead of 5: $P(1)$, $P(2)$, $P(3)$, $P(4)$, $P(5)$, and $P(6)$. The received values are r_1, r_2, r_3, r_4, r_5, r_6, with two errors, say in positions 1 and 2: i.e. $r_1 \neq P(1)$ and $r_2 \neq P(2)$. Clearly the original polynomial $P(x)$ agrees with four of the received values. Prove that no other degree 1 polynomial can agree with more than 3 of the received values. This means the recipient can uniquely reconstruct $P(x)$, and therefore the original message.

(Note: You must prove this from first principles relying on simple properties of polynomials. You are not allowed to rely on any results about the Berlekamp-Welch algorithm.)

Answer: The main idea is that if there is another degree 1 polynomial $Q(x)$ that agrees with at least 4 of the received points, then $Q(x)$ must be equal to the original polynomial $P(x)$ because they share at least 2 points. Equivalently, you can argue that any other degree 1 polynomial different from $P(x)$ can only use at most 1 out of the 4 correct points on P, and because there are only 2 other points remaining, this other polynomial can only agree with at most 3 of the received points.

Here are some sample proofs based on this idea:

Proof 1: Assume for the sake of contradiction that there exists a degree 1 polynomial $Q(x)$ which agrees with at least 4 of the received points. Because there are only two errors, at least two of the points on $Q(x)$ must be the correct values. However, we know that these two correct values also lie on the original degree 1 polynomial $P(x)$. Since these two points uniquely determine a degree 1 polynomial, we conclude that $P(x)$ must be equal to $Q(x)$. Contradiction. □

Proof 2: Assume for the sake of contradiction that there exists another degree 1 polynomial $Q(x)$ which agrees with at least 4 of the received values. Because there are only 6 received values in total, $P(x)$ and $Q(x)$ must share at least 2 points. Consider $S(x) = P(x) - Q(x)$. Note that $S(x)$ is nonzero because we assume $P(x)$ is distinct from $Q(x)$, and $S(x)$ has degree at most 1 because both $P(x)$ and $Q(x)$ have degree 1. Therefore, $S(x)$ has at most one zero. However, because $P(x)$ and $Q(x)$ share 2 values, $S(x)$ is zero on at least two points. Contradiction. □

5. **Random RSA (20 points for parts (a)–(c), extra credit for part (d))**

Let $n = pq$ be the product of two distinct primes, and let $E(a) = a^e \bmod n$ be the RSA encryption function, where as usual, the exponent e is relatively prime to $(p-1)(q-1)$. In this problem we will explore the probabilistic aspect of RSA.

(a) Show that if r is chosen uniformly at random mod n, then $E(r)$ is also a uniformly random number mod n. That is, prove that $\Pr[E(r) = i] = 1/n$ for every $i \in \{0, 1, \ldots, n-1\}$.

Answer: The key idea is that we know that the RSA encryption function $E(x)$ is a bijection from $\{0, \ldots, n-1\}$ to $\{0, \ldots, n-1\}$. This means that there exists an inverse function (the decryption function) $D(y) = E^{-1}(y)$ for every $y \in \{0, \ldots, n-1\}$. For every $i \in \{0, \ldots, n-1\}$:

$$\Pr[E(r) = i] \;=\; \Pr[r = D(i)] \;=\; \frac{1}{n}.$$

Thus, $E(r)$ is also a uniformly random number.

Common mistakes:

It is incorrect to just say that the range of $E(x)$ is $\{0, \ldots, n-1\}$, and no points are given if the only explanation is this. For example, the function $E(x) = 0$ has that range, but it is clearly non-uniform. For the probability to be uniform, it is critical that $E(x)$ is a bijection so that each element in the range is hit exactly once.

(b) Show that $E(ab \bmod n) = E(a)E(b) \bmod n$ for any messages a and b.

Answer: The main idea is that we can reduce numbers modulo n when performing multiplication:

$$xy \bmod n \equiv (x \bmod n) \cdot (y \bmod n) \bmod n$$

We use this property to prove the problem:

$$\begin{aligned} E(ab \bmod n) &\equiv (ab \bmod n)^e \bmod n \\ &\equiv (ab)^e \bmod n \\ &\equiv a^e \cdot b^e \bmod n \\ &\equiv (a^e \bmod n) \cdot (b^e \bmod n) \bmod n \\ &\equiv E(a) \cdot E(b) \bmod n \end{aligned}$$

Common mistakes:

i. Cannot just state that $E(ab \bmod n) = E((a \bmod n) \cdot (b \bmod n)) = E(a)E(b) \bmod n$; this is what we are trying to prove in the first place.

ii. Trying to use the Chinese Remainder Theorem or Fermat's Little Theorem.

iii. Incorrectly using the decryption function on both sides. Some answers just assumed that $D(E(a)E(b) \bmod n) = D(E(a)) \cdot D(E(b))$. However, this is essentially restating the problem, so you need to prove this first before you can use it in your general proof.

iv. Saying that $E(ab \bmod n) = a \cdot b^e \bmod n$ instead of $(a \cdot b)^e \bmod n$.

(c) Let a be relatively prime to n. Show that if r is chosen uniformly at random mod n, then $E(a)E(r) \bmod n$ is also a uniformly random number mod n. Make sure you explain where you use the assumption that $\gcd(a,n) = 1$.

Answer: The main idea is to show that the function that maps r to $E(a)E(r) \bmod n$ is a bijection. Because r is uniformly random mod n, this implies $E(a)E(r) \bmod n$ is also uniformly random.

There are several different ways to proceed. Here are two examples:

1) Because $\gcd(a,n) = 1$, the multiplicative inverse $a^{-1} \bmod n$ exists and is unique. We claim this implies that the mapping $g\colon \{0,1,\ldots,n-1\} \to \{0,1,\ldots,n-1\}$ given by $g(r) = ar \bmod n$ is a bijection. This is because if we have r_1, r_2 with $ar_1 = ar_2 \bmod n$, then we can multiply both sides by a^{-1} to conclude $r_1 = r_2$. This shows that g is one-to-one, and since the domain and codomain of g have equal cardinality, this also shows that g is a bijection. Because r is uniformly random mod n, this means $g(r) = ar \bmod n$ is also uniformly random mod n.

   Now because the RSA encryption function E is a bijection (or appealing to the result of part (a)), this implies $E(ar) \bmod n$ is uniformly random. Finally, from part (b) we conclude that $E(a)E(r) = E(ar) \bmod n$ is also uniformly random.

2) Since r is uniformly random and the RSA encryption function E is a bijection (or appealing to the result of part (a)), we get that $E(r)$ is also uniformly random mod n. To conclude that $E(a)E(r) \bmod n$ is uniformly random, we need to show that multiplication by $E(a) = a^e \bmod n$ is a bijection. This is because $\gcd(a,n) = 1$, so the multiplicative inverse a^{-1} exists, and therefore $a^{-e} = (a^{-1})^e \bmod n$ also exists. Thus, if for some y_1, y_2 we have $E(a)y_1 = E(a)y_2 \bmod n$, then we can multiply both sides by a^{-e} to conclude that $y_1 = y_2$. This shows that multiplication by $E(a)$ is injective, and since the domain and codomain have the same cardinality, this mapping is also bijective. This completes the proof.

Note that in a complete proof you need to argue explicitly that the assumption $\gcd(a,n) = 1$ implies that multiplication by a (or multiplication by $E(a)$) is a bijection.

Now let us evaluate some sample solutions.

Sample solution 1: $E(a)E(r) = E(ar) = (ar)^e \bmod n$ by part (b). Let $\Phi(r) = E(ar) \bmod n$. $\Phi^{-1}(r) = (a^{-1})r^d \bmod n$, note a^{-1} exists since $\gcd(a,n) = 1$. We show that Φ^{-1} is the actual inverse of Φ:

$$\begin{aligned}\Phi^{-1}(\Phi(r)) &\equiv \Phi^{-1}((ar)^e \bmod n)\\ &\equiv ((ar)^e)^d a^{-1} \bmod n\\ &\equiv ar \cdot a^{-1} \bmod n\\ &\equiv r \bmod n\end{aligned}$$

So again, the same reasoning as (a) shows that $E(a)E(r) \bmod n$ is uniformly distributed.

Evaluation: This is a valid proof. Note that instead of showing that Φ is onto and one-to-one, the proof instead finds *and* verifies the inverse function for Φ, which would only exist if Φ is a bijection.

Sample solution 2: From part (b), we know that $E(a)E(r) \equiv E(ar \bmod n)$. Since a and n are relatively prime ($\gcd(a,n) = 1$), there exists a^{-1}. Since a^{-1} exists, there exists a bijection from $E(ar \bmod n)$ to $E(r \bmod n)$ which has a bijection to r. Therefore, $E(a)E(r)$ is also a uniformly random number modulo n.

Evaluation: This proof almost gets it completely right. The existence of a^{-1} allows us to conclude that $ar \bmod n$ to $r \bmod n$ is a bijection (or equivalently, $r \bmod n$ to $ar \bmod n$), not $E(ar \bmod n)$ to $E(r \bmod n)$. We also know from part (a) that $ar \bmod n$ to $E(ar) \bmod n$ is a bijection.

Sample solution 3: $E(a)E(r) \bmod n = E(ar) \bmod n = a^e r^e \bmod n$. We know r^e is uniformly random, then because $\gcd(a,n) = 1$, multiplying r^e by $a^e \bmod n$ will not change the uniform random distribution, so $E(a)E(r) \bmod n$ is also uniformly random.

Evaluation: While it is true that the given precondition $\gcd(a,n) = 1$ ensures that $a^e r^e \bmod n$ is uniformly random, the result does not directly follow from the precondition. $\gcd(a,n) = 1$ ensures that a^{-1} exists which ensures that $h'(E(r)) = E(a)E(r) \bmod n$ is a bijection, which in turn ensures the result. This proof therefore gets half the points.

Sample solution 4: Because $\gcd(a,n) = 1$, we know $E(a)$ is uniformly random. We know from part (a) that $E(r)$ is uniformly random. So their product $E(a)E(r) \bmod n$ is also a uniformly random number.

Evaluation: It is not true that $E(a)$ is uniformly random. $\gcd(a,n) = 1$ implies that a is not a multiple of p or q. This ensures that $E(a)$ is not a multiple of p or q either, which means it cannot be uniformly random. Furthermore, note that in this problem a is a given, fixed number, not random. The rest of the proof depends on this false result and so is not correct.

Sample solution 5: Since $\gcd(a,n) = 1$, $f(r) = ar \bmod n$ is a bijection by Fermat's Little Theorem. From part (a), we know that $E(x) = x^e \bmod n$ is a bijection, which means that $E(a)E(r) = E(ar) \bmod n$ is bijection.

Evaluation: The proof of bijection for $f(r)$ is flawed because Fermat's Little Theorem by itself does not make any such claims. Moreover, recall that Fermat's Little Theorem only applies when the modulus is a prime, but here the modulus is $n = pq$. It is true that in the course of proving Fermat's Little Theorem (in Note 7) we did use a similar result: "$f' : S \to S$ *such that* $f'(x) \equiv ax \bmod p$ *is a bijection where* $S = \{1, \ldots, p-1\}$." However, the setting for this problem is different (because the modulus is $n = pq$ instead of a prime), so if you want to use this result, you have to reproduce the proof, which amounts to proving explicitly that $f(r) = ar \bmod n$ is a bijection.

(d) **(Extra credit)** Alice suspects that Bob might have misplaced his RSA private key. So she asks him to decrypt a cypher text to prove to her that he still has it. She assures him this is completely safe because the cypher text that she has chosen is uniformly random and nonzero (mod n). After all, as she explains to Bob, she has the ability to decrypt a random nonzero cypher text herself by just picking a uniformly random r from $\{1,\ldots,n-1\}$ and encrypting it using Bob's public key to get $c = E(r)$. Then c is a uniformly random nonzero cypher text and Alice knows its decryption.

Bob suspects there is something fishy about Alice's claim, but he believes her anyway, and is willing to decrypt a random nonzero cypher text for Alice. It turns out Alice has an ulterior motive: She has intercepted a cypher text $E(a)$ that Eve sent Bob and is dying to decrypt it to recover the message a.

Show how Alice can give Bob a uniformly random nonzero cypher text to decrypt and use his answer to recover the message a.

Answer: The main idea is as follows: Alice picks a uniformly random r from $\{1,\ldots,n-1\}$, and computes $E(r)$ using Bob's public key. She then computes $E(a)E(r) \mod n$, and gives it to Bob. Note that by part (b), $E(a)E(r) \mod n = E(ar \mod n)$, so when Bob decrypts this message, he sends back $ar \mod n$ to Alice, who can now compute $a = ar \cdot r^{-1} \mod n$.

Note that we require that Alice give Bob a uniformly random nonzero cypher text to decrypt, so we can only apply the scheme above when $E(a)E(r) \mod n$ uniformly random and nonzero. Moreover, the last step requires that the multiplicative inverse $r^{-1} \mod n$ exist. The complete solution must take these points into account:

- Observe that 0 always encrypts to 0, so if $E(a) = 0$, then Alice knows the message a must be 0. Now assume $E(a) \neq 0 \mod n$. If $E(a)$ is not relatively prime to n, then Alice can compute $\gcd(E(a), n)$. Since $n = pq$ is a product of two prime factors and $0 < E(a) < n$, we see that $\gcd(E(a), n)$ must be one of the prime factors of n; this allows Alice to factor n and break the RSA encryption, allowing her to recover the message a directly.
- Now assume $E(a)$ is relatively prime to n (which is the same as a being relatively prime to n). By the same argument as in part (c), if we choose r uniformly random over $\{1,\ldots,n-1\}$, then $E(a)E(r) \mod n$ is uniformly random and nonzero. Then Alice can apply the scheme above to obtain $ar \mod n$ from Bob. If r is relatively prime to n, then Alice can compute $a = ar \cdot r^{-1} \mod n$. If r is not relatively prime to n, then the multiplicative inverse $r^{-1} \mod n$ does not exist, but in this case Alice can compute $\gcd(r, n)$ to factor n and thereby break the RSA encryption.