CS 70 Discrete Mathematics for CS
Fall 2000 Papadimitriou/Russell/Sinclair **MT2 Sol**

Solutions to Midterm 2

1. (20 pts.) Extended GCD

(a) (5 pts) Recall that, on input (a, b) with $a > b$, the extended gcd algorithm returns a triple (d, y, x) such that $d = \gcd(a, b)$ and $bx + ay = d$. Since $\gcd(21, 13) = 1$, this will do the trick for us.

Running extended gcd on input $(21, 13)$ yields the following sequence of recursive calls:

call (a, b)	return (d, y, x)
$(21, 13)$	$(1, 5, -8)$
$(13, 8)$	$(1, -3, 5)$
$(8, 5)$	$(1, 2, -3)$
$(5, 3)$	$(1, -1, 2)$
$(3, 2)$	$(1, 1, -1)$
$(2, 1)$	$(1, 0, 1)$
$(1, 0)$	$(1, 1, 0)$

The returned value in the first line indicates that we may take $x = -8$ and $y = 5$ to satisfy $13x + 21y = 1$.

A remarkable number of people "guessed", or otherwise acquired, a correct solution (perhaps using the fact that 21 and 13 are successive Fibonacci numbers). Since guessing isn't always that easy, you should make sure that you are able to solve such problems systematically via the above algorithm.

(b) (5 pts) The painless way to do this is to note that the left-hand side (lhs) is the same as in part (a), while the right-hand side is multiplied by two. So we just need to multiply the x, y values found in part (a) by two, giving $x = -16$ and $y = 10$. [Note that this approach would work for *any* equation of the form $ax + by = c$: a solution exists iff c is an integer multiple of $d = \gcd(a, b)$, say $c = kd$. Then we can run extended gcd to solve $ax + by = d$, and then multiply the resulting x, y by k to solve the desired equation.]

A surprising number of people decided that there is no solution in this case, quoting as a reason the fact that the rhs is not equal to $\gcd(13, 21)$. *Clearly this is not enough: you would have to show that the rhs is not an integer multiple of the gcd (see part (c)).*

(c) (5 pts) There is no integer solution. This follows from the fact that the rhs, 1, is not an integer multiple of $\gcd(33, 21) = 3$; however, you should justify this conclusion as follows. Plainly, for any integers x, y, $33x + 21y$ is an integer multiple of 3, i.e., it is of the form $3z$ for some integer z. But the equation $3z = 1$ is clearly not satisfied by any integer z.

Almost everybody stated that there is no solution in this case. Many people, though, failed to justify this claim properly; the most common bogus argument was that the rhs is not equal to $\gcd(33, 21)$, *which is not enough as part (b) shows.*

(d) (5 pts) Since $\gcd(33, 21) = 3$, we can either run extended gcd directly on inputs $(33, 21)$, or we can first divide by 3 to get $11x + 7y = 1$ and then run extended gcd on inputs $(11, 7)$. Either way we get the solution $x = 2$, $y = -3$.

Almost everybody got this — though see above comments for part (a).

2. (20+5 pts.) Perfect Squares

(a) (6 pts) By brute force, the squares of the numbers $0, 1, \ldots, 10 \bmod 11$ are 0,1,4,9,5,3,3,5,9,4,1. So the perfect squares mod 11 are $0, 1, 3, 4, 5, 9$.

Pretty much everybody got this right.

(b) (8 pts) Let $0 < y < p$. If $x^2 = y$ for some $x \in \{0, 1, \ldots, p-1\}$ then clearly also $(-x)^2 = y$; and since p is odd and $x \neq 0$, $-x = p - x \neq x$, i.e., x and $-x$ are distinct. So if y has one square root then it has at least two.

Now suppose x, w are two distinct square roots of y. Then $x^2 - w^2 = 0 \bmod p$, or equivalently, $(x + w)(x - w) = 0 \bmod p$. This happens iff $x + w = 0 \bmod p$ or $x - w = 0 \bmod p$, i.e., iff $x = \pm w$. Thus the only possible square roots of y are $\pm x$. Hence y has either zero or two square roots.

Most people observed that, if x is a square root of y, then so is $-x$. This shows that y has at least *two square roots. But many people completely omitted to show that y cannot have* more than *two roots; this is the main issue here.*

(c) (6 pts) Squaring each of the numbers $1, 2, \ldots, p-1$ yields a perfect square, and these are the only possible squares except for 0. By part (b), each of these squares has exactly two roots, i.e., it occurs exactly twice in the above set of $p-1$ squares. So we get a total of $\frac{p-1}{2} + 1 = \frac{p+1}{2}$ perfect squares (where the extra +1 is for 0).

This was generally done OK, even by those who missed part (b).

(d) (5 pts, extra credit) The crucial point here is that no number y in $\{0, 1, \ldots, p-1\}$ can have *more than* three cube roots. This is because any cube root x is a root of the degree-3 polynomial $x^3 - y = 0$. Hence if we cube all the numbers in $\{0, 1, \ldots, p-1\}$, no cube appears more than three times. Hence there are *at least* $p/3$ distinct cubes. [Notice that we can't figure out the *exact* number of cubes, as we did for squares, because the equation $x^3 - y = 0$ may have fewer than three roots.]

Very few people gave a clear answer to this part.

3. (20 pts.) RSA

(a) (4 pts) First we note that since $pq = 33$ that $p = 3$ and $q = 11$ (or vice versa, but it is makes no difference to the rest of the problem). Your public exponent should be 7, as 5 is not relatively prime to $(p-1)(q-1) = 20$ and therefore has no multiplicative inverse modulo 20. We also need to know that 7 has a multiplicative inverse modulo 20 which we see in the next part of this problem.

(b) (4 pts) Your private exponent is the multiplicative inverse of your public exponent modulo $(p-1)(q-1)$. In this case your private exponent should be 3 as $7 \cdot 3 = 21 = 1 \bmod 20$. Your private key can be $(33, 3)$.

(c) (6 pts) To encrypt the message 2 you raise it to the public key exponent 7 modulo $p \cdot q = 33$. This is gives $128 \bmod 33 = 29$ as the encrypted message.

(d) (6 pts) You would sign the message by encrypting it using the private key, this would give $2^3 \bmod 33 = 8$. One might send this by itself, or possibly 2,8 as a pair with the message in plain text followed by the message encrypted by the private key, leaving the message readable immediately, but allowing others to verify that it was in fact signed. Finally one might treat 23 as the new message and encrypt this using the recipient's public key to send a signed encrypted message. Note that in our situation encrypting anything larger than 2 is problematic, in addition we do not know the recipient's public exponent.

4. (20 pts.) Polynomials

(a) A function $f(x)$ on GF_p is defined by its values on all inputs from GF_p: $f(0)$, $f(1)$, ..., $f(p-1)$. For each input x, $f(x)$ can be one of p elements and we have to choose $f(x)$ p times (once for every x). Therefore, there are p^p functions.

(b) A polynomial is defined by its coefficients a_0, ..., a_{p-1}. There are p coefficients a_i and each of them can be one of p elements of GF_p. Therefore, there are p^p polynomials.

(c) Let $q(x) = a_{p-1}x^{p-1} + a_{p-2}x^{p-2} + \ldots + a_0$ and $r(x) = b_{p-1}x^{p-1} + b_{p-2}x^{p-2} + \ldots + b_0$. Then, $q(x) - r(x) = (a_{p-1} - b_{p-1})x^{p-1} + (a_{p-2} - b_{p-2})x^{p-2} + \ldots + (a_0 - b_0)$. Since both $q(x)$ and $r(x)$ are polynomials of degree at most $p-1$, $q(x) - r(x)$ is a polynomial of degree at most $p-1$ as well.

If $q(x)$ and $r(x)$ are apparently distinct, then, for some i, $a_i \neq b_i$ and $a_i - b_i \neq 0$. Thus, $q(x) - r(x)$ is a polynomial of degree at most $p-1$ and it is not identically 0. Therefore, $q(x) - r(x)$ must have at most $p-1$ roots. Since GF_p contains p values, there must be $x \in GF_p$ for which $q(x) - r(x) \neq 0$ and $q(x) \neq r(x)$. This means that $q(x)$ and $r(x)$ are different functions.

A lot of people claimed that, if a coefficient of $q(x) - r(x)$ is not 0 then there must be x for which $q(x) - r(x) \neq 0$. This is true for integers but, in general, not true for GF_p. For example, $x^p - x = 0$ for all x in GF_p. Because of this, it is very important to notice that $q(x) - r(x)$ has degree at most $p-1$ and, therefore cannot be 0 for p values of x unless all of its coefficients are 0.

(d) Every polynomial defines a function and, by (c), no two polynomials define the same function. Therefore, the number of functions that are defined by some polynomial is the same as the number of polynomials. Since there are p^p polynomials (part (a)) and p^p functions (part (b)), this means that every function is defined by some polynomial.

The most common mistake in this part was forgetting to mention part (c). It is not enough to have the same number of functions and polynomials. It could be the case that some function was defined by two (or more) polynomials and some other function by no polynomial and only using part (c) shows that this is impossible.

5. (20+5 pts.) Probability spaces

(a) (6 pts) The *sample space* Ω is the set of possible atomic events that can occur—in this case, the $\binom{100}{20}$ possible committees of 20 senators chosen from 100.

Since the senators are chosen at random, every set of 20 senators is equally likely. Since the probabilities of the sample points must sum to 1, the probability of each sample point is $1/|\Omega| = 1/\binom{100}{20}$.

Several people confused the sample space with its size; the sample space is a set of atomic events, not a number. Many people forgot to give the probability of each sample point,

and later assumed a value of $1/\binom{100}{20}$ *without justification. It is important to realize that* not all probability spaces are uniform*!*

(b) (6 pts) Because the probability space is uniform, $P(CC) = |CC|/|\Omega|$. To compute the number of outcomes in CC, simply note that the two California senators occupy two seats, leaving 18 seats to be filled from 98 senators. Each way of filling those seats corresponds to exactly one outcome in CC, hence

$$P(CC) = \frac{\binom{98}{18}}{\binom{100}{20}} = \frac{98!}{80!\,18!}\frac{80!\,20!}{100!} = \frac{20 \cdot 19}{100 \cdot 99} = \frac{19}{495} \approx 0.0384$$

Almost everyone got this right who did it by counting atomic events in CC. Almost everyone who tried a more "intuitive" argument got it wrong. Some people did not put the answer in reduced form, which makes it difficult to compare probabilities in part (d).—

(c) (6 pts) Let us begin with the definition of the quantity we want:

$$P(CC|W) = \frac{P(CC \cap W)}{P(W)} = \frac{|CC \cap W|}{|\Omega|}\frac{|\Omega|}{|W|} = \frac{|CC \cap W|}{|W|}$$

because the space is uniform. We evaluate the denominator first. W means "at least one senator from Wyoming," which is a disjunctive event. Often, it is easier to work with a conjunctive event: let $\neg W$ be the event that NO senators from Wyoming are chosen—which means the 20 senators are chosen from the remaining 98. Then

$$|W| = |\Omega| - |\neg W| = \binom{100}{20} - \binom{98}{20}$$

Similarly, we have

$$|CC \cap W| = |CC| - |CC \cap \neg W| = \binom{98}{18} - \binom{96}{18}$$

Hence

$$P(CC|W) = \frac{\binom{98}{18} - \binom{96}{18}}{\binom{100}{20} - \binom{98}{20}} = \frac{20 \cdot 19 \cdot (98 \cdot 97 - 80 \cdot 79)}{98 \cdot 97 \cdot (100 \cdot 99 - 80 \cdot 79)} = \frac{30267}{850787} \approx 0.0356$$

We apologize for the horribleness of the arithmetic, which was unintentional. Anyone who got the reasoning right up to that point got full credit; we gave some extra credit for perfect answers.

The majority of answers took the following form: if we have chosen one senator from Wyoming, then there are $\binom{99}{19}$ *ways to complete the committee; similarly, there are* $\binom{97}{17}$ *ways to complete the committee if we have chosen one from Wyoming and two from California. (Some people correctly multiplied each of these by 2 because there are two ways to choose the first Wyoming senator.) This gives* $\binom{97}{17}/\binom{99}{19} = 19/539$*, which is close to the true answer (within 1%) but not correct. The error comes from double-counting the cases where both Wyoming senators are chosen. We gave half credit for this answer.*

(d) (2 pts) CC and W are not independent, since $P(CC) \neq P(CC|W)$. This is easy to tell if the probabilities are in reduced form.

Some people gave a qualitative argument as to why they were dependent, and some gave a qualitative argument as to why they were not. Qualitative arguments for dependence are

unconvincing, because two events can be causally connected in such a way that the effects of the first on the second exactly cancel, leaving independence. Consider the probability of rolling an odd total with two dice, given that the first die is even. The latter event is clearly "relevant" to the former, but the two are mathematically independent.

(e) (5 pts, extra credit) Let MM be the event that at least one state has two members on the committee. As before, it's easier to deal with the negation—$\neg MM$ means that no states have two members on the committee. In this case, we can make the committee by first selecting 20 states in one of $\binom{50}{20}$ ways, then selecting one senator from each in one of 2^{20} ways:

$$P(MM) = 1 - \frac{\binom{50}{20} \cdot 2^{20}}{\binom{100}{20}} \approx 0.908$$

A variety of other answers were obtained, ranging from 10^{-8} *to 190. It's always a good idea to check your answer for numerical plausibility!*