CS 70 Discrete Mathematics and Probability Theory
Summer 2011 Kamil HW 4

Due Monday, July 11, 4:59pm

Note the special due date.

1. (10 pts.) Practice with Lagrange interpolation

This problem will have you practice with Lagrange interpolation. Here, we are looking for a polynomial $p(x)$ of degree at most 2 that passes through the points $(1,2)$, $(2,3)$, and $(3,5)$, working in $GF(7)$. In other words, we want $p(x)$ to satisfy $p(1) \equiv 2 \pmod 7$, $p(2) \equiv 3 \pmod 7$, and $p(3) \equiv 5 \pmod 7$.

(a) Find the three polynomials $\Delta_1(x), \Delta_2(x), \Delta_3(x)$. Simplify them to the form $ax^2+bx+c \pmod 7$ where a,b,c are integers satisfying $0 \le a,b,c < 7$. Circle or box your final answer.

(b) Using your answer to part 1 and Lagrange interpolation, find the polynomial $p(x)$. Simplify it to the form $ax^2+bx+c \pmod 7$ where a,b,c are integers satisfying $0 \le a,b,c < 7$. Circle or box your final answer.

2. (12 pts.) A practice problem with polynomials

Define the sequence of polynomials by $P_0(x) = x+17$, $P_1(x) = x^2-7x+6$, and $P_n(x) = xP_{n-2}(x) - P_{n-1}(x)$. (For instance, $P_2(x) = 24x-6$ and $P_3(x) = x^3-7x^2-18x+6$.)

(a) Show that $P_n(6) \equiv 0 \pmod{23}$ for every $n \in \mathbb{N}$. (i.e., working in the field $GF(23)$)

(b) Show that, for every prime q, if $P_{2011}(x) \not\equiv 0 \pmod q$, then $P_{2011}(x)$ has at most 2011 roots modulo q (i.e., in the field $GF(q)$).

3. (12 pts.) ISBN checksums

An ISBN is a 10-digit number that serves as a serial number for books. The last digit is a checksum, which can be helpful for detecting data entry errors when typing in an ISBN. If the first nine digits are given by $x_1, \ldots, x_9$ (where $0 \le x_i \le 9$), the checksum digit x_{10} is given by

$$x_{10} = \bmod(x_1 + 2x_2 + \cdots + 8x_8 + 9x_9, 11).$$

(The checksum digit is in the range $0 \le x_{10} \le 10$. If the checksum digit is "10", the letter X is substituted when writing out an ISBN.) An equivalent way to describe the ISBN algorithm is like this: the checksum digit x_{10} is chosen so that the following equation is true:

$$10x_1 + 9x_2 + \cdots + 3x_8 + 2x_9 + x_{10} \equiv 0 \pmod{11}.$$

For instance, a sample ISBN is 0201896834; this has a valid checksum, since

$$10\cdot 0 + 9\cdot 2 + 8\cdot 0 + 7\cdot 1 + 6\cdot 8 + 5\cdot 9 + 4\cdot 6 + 3\cdot 8 + 2\cdot 3 + 1\cdot 4 = 176 \equiv 0 \pmod{11}.$$

For each of the following claims about this checksum algorithm, say whether the claim is true or false. Justify your answer.

(a) The ISBN checksum detects all single-digit errors (i.e., all errors where a single digit is entered incorrectly).

(b) The ISBN checksum detects all two-digit errors (i.e., all errors where a pair of digits, not necessarily adjacent, are entered incorrectly).

(c) The ISBN checksum detects all errors where a pair of adjacent digits are transposed (e.g., where we enter 0021896834 instead of 0201896834).

(d) The ISBN checksum detects all errors where any pair of digits (not necessarily adjacent) are transposed (e.g., where we enter 3201896804 instead of 0201896834).

4. (12 pts.) Error-correcting codes: an optimization

In class, we saw an error-correcting code where the n message packets $m_1, \ldots, m_n$ are encoded to the $n+k$ encoded packets $c_1, \ldots, c_{n+k}$ by setting $P(x) = m_n x^{n-1} + \cdots + m_2 x + m_1$, then defining $c_i = P(i)$ (all this is in $GF(q)$, where q is prime and larger than $n+k$, so each packet is a number in the range $0 \ldots q-1$). However, one possible criticism of this error-correcting code is that decoding always requires a Lagrange interpolation step, even if no packets are lost.

(a) In this part, you will develop a scheme that addresses this criticism. Let's preserve the basic approach where $c_i = Q(i)$ (for $i = 1, 2, \ldots, n+k$), for some appropriately chosen polynomial $Q(x)$ which encodes the entire message, and which has degree at most $n-1$. (As before, we'll work in $GF(q)$, where $q > n+k$ and q is prime.) At the same time, let's ensure $c_1 = m_1$, $c_2 = m_2$, …, $c_n = m_n$, so that if no packets are lost, we can just use the first n encoded packets to immediately read off the message. Describe how to choose $Q(x)$ with this desired property, given $m_1, \ldots, m_n$. In other words, describe an efficient algorithm we can use for encoding.

(b) For your scheme from part 1, if some packets are lost, the recipient can use Lagrange interpolation to recover $Q(x)$. Describe how the recipient could recover $m_1, \ldots, m_n$ from $Q(x)$.

5. (10 pts.) Secret Sharing

Give a secret sharing scheme which encodes an n bit secret but works with numbers with at most 32 bits, i.e., choose modular arithmetic q less than $2^{32} - 1$ for your field. (Assume n is much larger than 31 and is divisible by 31.) Your scheme should share the secret among p people and have the property that any k people can recover the entire secret, with $k-1$ people there are at least 2^{31} possible secrets, with $k-i$ people there are at least 2^{31i} possible secrets until $i = n/31$ at which point there are 2^n possibilities for secrets. You may assume that $k \geq n/31$.

6. (32 pts.) 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?

(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.)

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

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

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

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

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

(h) 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?

(i) 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.)

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

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

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

(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).

(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).

(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).

(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?

7. (12 pts.) Grade these proofs

You be the grader. Students have submitted the following proofs. For each, decide whether you think the proof is valid or not, and assign the student answer either an A (valid proof) or an F (invalid proof). If the proof is invalid, explain *clearly and concisely* where the logical error in the proof is, including exactly which step of the reasoning is erroneous. (If you think the proof is correct, you do not need to give any explanation.) Simply saying that the claim (or the induction hypothesis) is false is *not* an acceptable explanation.

$\mathbb{R}^+$ denotes the set of all positive real numbers.

(a) **Claim:** We have $2n \leq n^2 + 1$ for all $n \in \mathbb{N}$.

Proof: We will prove this by simple induction on n. Let $P(n)$ denote the proposition that $2n \leq n^2 + 1$.
Base case: If $n = 0$, then $2n = 0 \leq 0^2 + 1 = n^2 + 1$, so $P(0)$ is true.
Inductive hypothesis: Assume $P(n)$ is true for some $n \in \mathbb{N}$. That is, we assume $2n \leq n^2 + 1$.
Inductive step: We must show that $P(n+1)$ is true. Now

$$2(n+1) = 2n + 2 \leq n^2 + 1 + 2 \leq (n+1)^2 + 1,$$

where we have used the inductive hypothesis as well as the fact that $n^2 + 2 \leq (n+1)^2$. We see that $P(n) \implies P(n+1)$ holds for every $n \in \mathbb{N}$, so by the principle of mathematical induction, $P(n)$ is true for every $n \in \mathbb{N}$, and the claim follows. □

(b) **Claim:** For all $n \in \mathbb{N}$, for all $r \in \mathbb{R}^+$, if $r + \frac{1}{r}$ is an integer, then $r^n + \frac{1}{r^n}$ is an integer.

Proof: We will prove this by strong induction on n. Let $P(n)$ denote the proposition that, for all $r \in \mathbb{R}^+$, if $r + \frac{1}{r}$ is an integer, then $r^n + \frac{1}{r^n}$ is an integer.

Base cases:

- For $n = 0$, $P(0)$ states that for all r, if $r + \frac{1}{r}$ is an integer, then $1 + 1$ is an integer. This is certainly true, since $1 + 1$ is an integer regardless of r.
- For $n = 1$, $P(1)$ states that for all r, if $r + \frac{1}{r}$ is an integer, then $r + \frac{1}{r}$ is an integer. This is also true.

Induction hypothesis: Let k be some natural number with $k \geq 2$. Assume that $P(k-2)$ and $P(k-1)$ are true, i.e., for all r, if $r + \frac{1}{r}$ is an integer, then $r^{k-2} + \frac{1}{r^{k-2}}$ and $r^{k-1} + \frac{1}{r^{k-1}}$ are both integers.

Inductive step: We want to prove $P(k)$. In other words, we want to prove that, for all r, if $r + \frac{1}{r}$ is an integer, then $r^k + \frac{1}{r^k}$ is an integer. So let r be arbitrary and assume $r + \frac{1}{r}$ is an integer. By the inductive hypothesis, both $r^{k-2} + \frac{1}{r^{k-2}}$ and $r^{k-1} + \frac{1}{r^{k-1}}$ are integers.

Notice the following identity (obtained by multiplying out the terms):

$$\left(r + \frac{1}{r}\right)\left(r^{k-1} + \frac{1}{r^{k-1}}\right) = r^k + \frac{1}{r^{k-2}} + r^{k-2} + \frac{1}{r^k}.$$

Re-arranging terms, we find

$$r^k + \frac{1}{r^k} = \underbrace{\left(r + \frac{1}{r}\right)}_{\text{integer}}\underbrace{\left(r^{k-1} + \frac{1}{r^{k-1}}\right)}_{\text{integer (by I.H.)}} - \underbrace{\left(r^{k-2} + \frac{1}{r^{k-2}}\right)}_{\text{integer (by I.H.)}}.$$

We see that the right-hand side must be an integer, so the left-hand side must be an integer too. In other words, we have shown that $r^k + \frac{1}{r^k}$ is an integer. The claim follows by induction. □

(c) **Claim:** Let m be any natural number with $m > 1$ and x, y be integers. If $x^4 + y^4 \equiv 2x^2y^2 \pmod{m}$, then either $x \equiv y \pmod{m}$ or $x \equiv -y \pmod{m}$.

Proof: Suppose x, y form an integer solution to the equation, so we are given

$$x^4 + y^4 \equiv 2x^2y^2 \pmod{m}.$$

Subtracting $2x^2y^2$ from both sides, we find that

$$x^4 - 2x^2y^2 + y^4 \equiv 0 \pmod{m}.$$

We can factor the left-hand side, to get

$$(x^2 - y^2)^2 \equiv 0 \pmod{m}.$$

Taking the square root of both sides, we see that

$$x^2 - y^2 \equiv 0 \pmod{m},$$

or in other words,

$$x^2 \equiv y^2 \pmod{m}.$$

Taking the square root of both sides again, we find that either $x \equiv y \pmod{m}$ or $x \equiv -y \pmod{m}$ (we have to include both possibilities, because the square root on each side could be either positive or negative). This proves the claim. □