Modulo a Prime Number

We have seen that modular arithmetic can both be easier than normal arithmetic (in how powers behave), and more difficult (in that we can't always divide).

But when n is a prime number, then modular arithmetic keeps many of the nice properties we are used to with whole numbers.

(Recall that a prime number is a whole number, greater than or equal to 2, whose only factors are 1 and itself. So $2, 3, 5, 7, 11$ are prime numbers whilst, $6 = 2 \times 3$ and $35 = 5 \times 7$ aren't.)

Inverses, Modulo a Prime

Theorem 1 *When n is a prime number then it is valid to divide by any non-zero number — that is, for each $a \in \{1, 2, ..., n-1\}$ there is one, and only one, number $u \in \{1, 2, ..., n-1\}$ such that*

$$au = 1 \pmod{n}.$$

Then, dividing by a is the same as mulitplying by u, i.e. division by a is given by the rule

$$\frac{b}{a} = bu \pmod{n}.$$

For example, in $\mod 7$, we have

$$\frac{1}{1} = 1, \quad \frac{1}{2} = 4, \quad \frac{1}{3} = 5, \quad \frac{1}{4} = 2, \quad \frac{1}{5} = 3, \quad \frac{1}{6} = 6.$$

Roots of a Polynomial

Theorem 2 *When n is prime number, then a polynomial of degree k, say*

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_k x^k = 0 \pmod{n} \text{ *with* } a_i \in \{0, 1, 2, \ldots, n-1\},$$

has at most k solutions.

So it is impossible, when n is a prime, for a quadratic like $x^2 - 1$ to have more than 2 roots, as we saw it having in $\bmod 8$ arithmetic.

Note that a quadratic, like x^2+x+1 in $\bmod 2$ arithmetic, can have *fewer* than two roots; but this type of behaviour we have seen before as $x^2 + 1 = 0$ has no solutions amongst the real numbers.

Fermat's Little Theorem

Theorem 3 *When n is a prime number, then*

$$a^n = a \pmod{n}.$$

for any a.

This, in fact, tells us more than Theorem 1; this tells us that when $a \neq 0$, then dividing by a is the same as multiplying by a^{n-2}, as

$$\frac{1}{a} = a^{n-2} \pmod{n}.$$

We now have an explicit expression for a's multiplicative inverse.

Problems

1. Find $\frac{1}{1},\ \frac{1}{2},\ \frac{1}{3},\ \frac{1}{4}$ in $\mod 5$ arithmetic.
2. Check, with $n = 7$ that Fermat's Little Theorem holds for each value of $a = 0, 1, 2, 3, 4, 5, 6 \mod 7$.
3. Which numbers is it valid to divide by in $\mod 9$ arithmetic? For these numbers find their inverses.
4. How many solutions has $x^2 + x = 0$ in $\mod 6$ arithmetic? (Try out each of the numbers $0, 1, 2, 3, 4, 5 \mod 6$.)
5. Find two different ways to factorise $x^2 + x$ in $\mod 6$ arithmetic.

Answers

1. In $\bmod 5$ arithmetic, the reciprocals are given by

$$1/1 = 1, \quad 1/2 = 3, \quad 1/3 = 2, \quad 1/4 = 4 \pmod 5.$$

2. Verifying Fermat's Little Theorem, $\bmod 7$, we see:

$$\begin{aligned} 0^7 = 0 \pmod 7; (\pm 1)^7 = \pm 1 \pmod 7; \\ (\pm 2)^7 = \pm 2^7 = \pm 2^4 \times 2^3 = \pm 16 \times 8 = \pm 2 \times 1 = \pm 2 \pmod 7; \\ (\pm 3)^7 = \pm 3^7 = \pm 3^4 \times 3^3 = \pm 81 \times 27 = \pm 4 \times 6 = \pm 24 = \pm 3 \pmod 7. \end{aligned}$$

3. It is valid to divide by $1, 2, 4, 5, 7, 8 \pmod 9$. The reciprocals are:

$$1/1 = 1, \quad 1/2 = 5, \quad 1/4 = 7, \quad 1/5 = 2, \quad 1/7 = 4, \quad 1/8 = 8.$$

4. In $\bmod 6$ arithmetic, $x^2 + x$ takes values $0, 2, 0, 0, 2, 0 \pmod 6$ when $x = 0, 1, 2, 3, 4, 5 \pmod 6$. So $x^2 + x = 0 \pmod 6$ has four roots $0, 2, 3, 5$.

5. $x^2 + x$ factorises in $\bmod 6$ arithmetic as:

$$x^2 + x = x(x + 1) = (x + 4)(x + 3) \pmod 6.$$

Appendix: Modular Arithmetic is Well-defined.

In our earlier definition of addition, subtraction, multiplication and powers in modular arithmetic a subtle check was omitted. Modular arithmetic is about the addition (etc.) of remainders. When we write $1 \times 1 = 1 \pmod 2$, we are saying that multiplying *any* two odd numbers results in an odd number. But we haven't proven this anywhere yet; it's very important that it doesn't matter which odd numbers we choose.

Let's prove this first: the odd numbers are precisely those that can be put in the form $2k+1$ where k is a whole number. So if we have two of these, $2k_1+1$ and $2k_2+1$ say, their product is

$$\begin{aligned}(2k_1+1)(2k_2+1) &= 4k_1k_2 + 2k_1 + 2k_2 + 1 \\ &= 2(2k_1k_2 + k_1 + k_2) + 1\end{aligned}$$

which we can see is another odd number; we finally see a product of odd numbers is odd, irrespective of the odd numbers chosen.

And more generally, we *defined* the operations of modular arithmetic as

$$\begin{aligned}
a \pmod{n} + b \pmod{n} &= \text{remainder when } a+b \text{ is divided by } n;\\
a \pmod{n} - b \pmod{n} &= \text{remainder when } a-b \text{ is divided by } n;\\
a \pmod{n} \times b \pmod{n} &= \text{remainder when } a+b \text{ is divided by } n;\\
(a \pmod{n})^k &= \text{remainder when } a^k \text{ is divided by } n.
\end{aligned}$$

It is important to check that it does not matter which numbers we are choosing from the classes of numbers, $a \pmod{n}$ and $b \pmod{n}$.

Specifically, we really need to check that if

$$a = A \pmod{n}) \text{ and } b = B \pmod{n}$$

then in $\mod n$ arithmetic, we must also have

$$a + b = A + B; \quad a - b = A - B; \quad ab = AB; \quad a^k = A^k.$$

The first two lines are easy checks and the third, multiplication, is very similar to the previous calculation with odd numbers.

To prove that powers are well-defined in modular arithmetic, suppose that $a = A \pmod{n}$. As a and A leave the same remainder, then they must be a multiple of n apart, i.e.

$$a = A + cn$$

for some whole number c. So, factorizing first, we have

$$\begin{aligned} a^k - A^k &= (a - A)\left(a^{k-1} + Aa^{k-2} + A^2a^{k-3} + \cdots + A^{k-1}\right) \\ &= cn\left(a^{k-1} + Aa^{k-2} + A^2a^{k-3} + \cdots + A^{k-1}\right) \\ &= 0 \pmod{n}, \end{aligned}$$

because it is a multiple of n. Hence $a^k = A^k \pmod{n}$, and powers are well-defined in modular arithmetic as well.

Codes

A code is a way of representing words and messages as numbers or symbols. One of the best known examples of a code is the Morse code which represents letters as short sequences of dots and dashes (see below). Note how the code aims to use short codings for commonly used letters like E.

A	· –	G	– – ·	M	– –	S	· · ·	Y	– · – –	4	· · · · –
B	– · · ·	H	· · · ·	N	– ·	T	–	Z	– – · ·	5	· · · · ·
C	– · – ·	I	· ·	O	– – –	U	· · – –	1	· – – – –	6	– · · · ·
D	– · ·	J	· – – –	P	· – – ·	V	· · · –	2	· · – – –	7	– – · · ·
E	·	K	– · –	Q	– – · –	W	· – –	3	· · · – –	8	– – – · ·
F	· · – ·	L	· – · ·	R	· – ·	X	– · · –	4	· · · · –	9	– – – – ·

Another well-known codeis the ASCII code used in computing which assigns a number between 0 and 255 to a wide array of symbols (upper and lower case alphabets, punctuation marks, mathematical symbols etc.)

ISBN codes

The ISBN code (which stands fo *International Standard Book Number* is a 10-digit number which appears on the back of every book published. For example, the ISBN code of David Acheson's *1089 and All That* is

0198516231

It contains information about the language, publisher and title. The last digit of the code is a check digit, which helps combat the problem that the code might be miscommunicated.

A ISBN code abcdefghij is set up in such a way that

$$10a + 9b + 8c + 7d + 6e + 5f + 4g + 3h + 2i + j$$

is divisible by 11. The check digit j can be chosen in the range $0 \leqslant j \leqslant 10$ so that this is possible – in the event that $j = 10$ then the letter X is used instead.

Turning Words into Numbers

The reason we need a code is that the RSA encryption, which we will meet later, is based on clockwork arithmetic; we need to be able to convert any message into numbers first, before we can apply modular arithmetic.

The following important features will be considered in creating any code:

- each coded message can be decoded uniquely;
- the code has limited redundancy – i.e the coded message has optimal length;
- the original message can be recovered in the event of limited transmission errors.

There is a standard way in mathematics, called a Huffman code, for turning a message into numbers, in a systematic and efficient way.

Example of a Huffman Code

Suppose we wish to send a message to a Martian Rover, which understands five commands

Stop, Go North, Go West, Go East, Go South.

Say that these commands are given with respective probabilities

$$\frac{1}{3}, \quad \frac{1}{6}, \quad \frac{1}{6}, \quad \frac{1}{6}, \quad \frac{1}{6}.$$

Then a Huffman code for this might be of the form

Command	Code	Probability	Length
Stop	00	1/3	2
Go North	010	1/6	3
Go West	011	1/6	3
Go East	10	1/6	2
Go South	11	1/6	2

Average code length $= \frac{2}{3} + \frac{1}{2} + \frac{1}{2} + \frac{1}{3} + \frac{1}{3} = 2\frac{1}{3}$ digits per command.

Creating a Huffman Code

Stop	$\frac{1}{3}$	⟶	$\frac{1}{3}$	⟶	$\frac{1}{3}$	↘ 0	$\frac{2}{3}$	↘ 0	1	00
Go North	$\frac{1}{6}$	↘ 0	$\frac{1}{3}$	⟶	$\frac{1}{3}$	↗ 1				010
Go West	$\frac{1}{6}$	↗ 1								011
Go East	$\frac{1}{6}$	⟶	$\frac{1}{6}$	↘ 0	$\frac{1}{3}$	⟶	$\frac{1}{3}$	↗ 1		10
Go South	$\frac{1}{6}$	⟶	$\frac{1}{6}$	↗ 1						11

- Group the two least likely commands together.
- Consider these now a single command, with the two probability combined.
- We have one fewer command now. Repeat this until one command remains.
- Assign a code to each command, working backwards. Call the last two subsets grouped the 0-group and the other the 1-group.
- The first code letter of the 0-group will be “0” and the first letter of the 1-group will be “1”.
- Assign to each command a list of 0s and 1s, depending on whether, at various stages, they came from a 0-group or a 1-group.

Theorem 4 (Huffman 1952) *Huffman codes are:*

- *uniquely decipherable. In fact, Huffman codes are more than this: they are instantaneous. Not only can a coded message be uniquely deciphered, it can be done so by reading the coded message number by number, without knowledge of the whole code.*
- *compact. This means that the average length of a Huffman-coded message is the best possible using the numbers 0 and 1.*

More general Huffman codes, using more than just the numbers 0 and 1, can be easily made.

The Huffman code has no built in redundancy to deal with transmission errors.

Problems

1. Check that the ISBN code 0198516231 is a valid one.
2. I mistakenly wrote down the ISBN code of *Codes and Cryptography* as 0198532773, but with one of the digits out by one. Show that this is an incorrect ISBN code. Which numbers may have been wrong and what should they have been?
3. The following coded messages have all been made using the previous Huffman code. Decode their command sequence.

   $$010011101100, \quad 011011101100, \quad 01001100111100.$$

4. Write down a sequence of 0s and 1s which the Martian Rover couldn't interpret as meaningful commands.
5. By grouping the Rover's commands differently, create a different Huffman code.
6. What is this average code length per command for your new code?
7. Find different decodings of the following Morse Code (with the gaps removed) — which do you think is the right one?

   $$\bullet\bullet\bullet\bullet\bullet\bullet - \bullet\bullet\bullet - \bullet\bullet - \;--$$

Answers

1. The ISBN code 0198516231 gives $0+9+72+56+30+5+24+6+6+1 = 209 = 19 \times 11$.
2. The ISBN code 0198532773 gives $0+9+72+56+30+15+8+21+14+3 = 228 = 8 \pmod{11}$. So either the first 7 should have been an 8 (increasing this sum by 3), or the 9 should have been an 8 (decreasing this sum by 8).
3. 010011101100 = North + West + East + South + Stop;
   011011101100 = West + West + East + South + Stop;
   01001100111100 = North + West + Stop + South + South + Stop.
4. Any odd length string of 0s would be meaningless to the rover.
5. Grouping first N & W, then E & S, then N+W & E+S, then all together, we could get the Huffman code:
   Stop = 0,   North = 100,   West = 101,   East = 110,   South = 111.
6. The new code's average length is $1/3 + 3/6 + 3/6 + 3/6 + 3/6 = 7/3$ as before.
7. The piece of Morse code could be translated as "HEEDRJ", "SSBDO", or more likely "HELLO".