Class 20: Number Theory

Schedule

Problem Set 8 is due **Friday at 6:29pm**.

Tuesday's class will include a review for Exam 2. Before 6:59pm Monday, send to uvacs2102staff@gmail.com topics you would like to review. We will have a similar rule to Exam 1:

- Fewer than 10 total requests: it means, class is not being sufficiently challenged and we should be doing more difficult material (except for the 10 requestors who get exam exemptions).
- Exactly 10 total requests: all requestors get automatic exam exemptions, review requested topics.
- More than 10 requests implies we should spend Thursday's class on reviewing requested topics, no exam exemptions.

Exam 2

Exam 1 will be in class on Thursday, 9 Nov. You can get a good idea for what to expect on Exam 2 based on Exam 1, and also by looking at the Practice Exam 2 (from last year's class). We strongly encourage you to try to problems on your own, before looking at the posted solutions. The main difference is that this year, we did not cover the topic of "stable marriage", but the our exam 2 will also cover the topic of infinities (i.e. infinite sets, countable and uncountable sets, etc.) as well.

Resources. Similarly to Exam 1, you will be permitted to use a *single paper page* of notes that you prepare and bring to the exam. It is fine to collaborate with others to prepare your notes. The page should be no larger than a US Letter size page (8.5×11 inches), and you may write (or print) on both sides of the page. You may not use any special devices (e.g., magnifying glasses) to read your page. No other resources, other than your own brain, body, and writing instrument, are permitted during the exam.

Content. The problems on the exam will cover material from Classes 1–19, Problem Sets 1–8 (including the provided solutions), and the relevant material from MCS Chapters 1–8. However, only one problem will be about the material covered by Exam 1, and the rest will be on the material we covered since Exam 1 (namely, classes 13-19, problem sets 6–8, and chapters 6,7,8). Everything on the exam will be something you have seen in at least two of these (Classes, Problem Sets, and MCS Book), and most of the exam will be things you have seen in all three. If you understand the problems on the problem sets and questions on the class notes well enough to be able to answer similar questions, you should do well on the exam.

Simiarly to Exam 1, for most students, we believe the best way to prepare for the exam will be to (1) go over the problem sets and their solutions, and make sure you understand well any of the problems you did not get before; (2) go through the provided practice exam and try to solve all the problems on your own before reading the solutions; (3) go through the questions in the class notes and convince yourself you can answer them well; (4) re-read chapters of the book, solving the associated practice problems, especially for any sections on topics where you had difficulty on the problem sets. If you do #1 and #2 and understand well the problems on the practice exam, you should be confident you'll do well on the exam; if you struggled on the problem sets, you would benefit from doing #3 and #4 as well.

Divisibility

Number theory is a study of integers $\mathbb{Z} = \{0, -1, 1, -2, 2, \dots\}$. Adding, multiplying, and negating, are 3 operators on integers that we can define in a straight-forward way. Dividing is more tricky. For that, we first define the notation $a \mid b$ to denote the simpler situation where $b = k \cdot a$ for some $k \in \mathbb{Z}$. In this case, we say that a is a *divisor* of b and that b is a *multiple* of a.

Show that if we assume $a \mid b$ and $a \mid c$ (for any integers a, b, c), then for all integers x, y it will hold that $a \mid xa + yb$.

When b is *not* a multiple of a, we can still have some form of division, stated in the following theorem.

Theorem [Division Theorem] For all integers n, d, there are integers q (called the quotiont) and r (called the remainder, denoted by $rem(m, d)$) such that

$$n = q \cdot d + r \text{ and } 0 \leq r < |d|.$$

Note that the remainder, as defined here, is always non-negative. For example when dividing -3 by 5 we get $q = -1$ and $r = 2$.

Definition A *prime* p, is an integer $p > 1$ such that 1 and p are the only positive divisors of p. Other $n > 1$ numbers (that are not prime) are called composite.

The following well-known theorem can be proved by strong induction on n.

Theorem [Fundamental theorem of arithmetic]. Any $n > 1$ has a unique represention in the following form:

$$n = p_1^{a_1} \dots p_k^{a_k}$$

where $p_i < p_{i+1}$ and p_i is prime for all i.

Hint for proof: Use strong induction on n, and try to show that any two different representations for n will have the same 'smallest' prime p_1 and the same power a_1 for p_1.

Easy-to-state, but hard-to-solve problems

In number theory, we can easily state problems that are super hard to solve. These are some examples.

Goldbach's conjecture: Any even integer $n > 2$ can be written as $n = p + q$ where p, q are both primes. The conjecture is still open, though we know that it is correct for all $n \leq 10^{18}$.

Twin prime's conjecture: There infinite prime numbers p where $p + 2$ is also prime. Note that we cannot 'check the conjecture up to n' anymore! And it is still open...

Fermat's last theorem: For any integer $n > 2$ there are no integers a, b, c where $a^n + b^n = c^n$. Note that for $n = 2$ there are integer solutions like $3^2 + 4^2 = 5^2$. Fermat claimed to have a proof, but nobody knows if he really did or not. Andrew Wiles eventually proved this in 1994, hence we call it a *theorem* and not a conjecture anymore.

Greatest Common Divisor

Definition For natural numbers $m, n > 0$ we call $d > 0$ their *greatest common divisor*, denoted by $d = gcd(m, n)$ if d is the largest number such that $d \mid m$ and $d \mid n$. (Make sure to verify why we can always say that such d exists.)

We can use the fundamental theorem of arithmetic, to *compute* the gcd of any given pairs of integers. But how efficient is it? The problem is that factoring numbers m, n into their prime factors is not something we know how to do efficiently. In particular, no known algorithm is guaranteed to factor all numbers with $1000,000$ digits in less than an (estimated) thousand years! The conjecture is that no such 'efficient' algorithm exists, and many cryptographic algorithms *assume* this is the case!

Euclid's Algorithm

Despite not knowing how to factor integers efficiently, Euclid showed how to find gcd quite efficiently. This shows, even if the most "natural" algorithm fails for doing something efficiently, we should always consider the possibility that an alternative way could still exist.

```
def Euclid_gcd(m, n):
   while n > 0:
      r = rem(m,n)
      m = n
      n = r
   return m
```

What is a good state machine modeling the behavior of Euclid's algorithm?

Suppose d is any integer (in particular, it could be the gcd of the original state (m, n) of the state machine for Eclid's algorithm). Show that $gcd(a, b) = d$ is a preserved invariant for the transition graph of this state machine. Namely, if we go from one state to another state, the gcd does not change.

If the Euclid's algorithm ends, at the very last step we will have $n = 0$ which means m is indeed the right gcd for that particular (final) state. Together with the preserved invariant (that you proved above), argue that the program has *partial correctness*. Namely: if it ends, it outputs correctly.

Now prove the termination. Namely, show that the Euclid's gcd algorithm will always end. Hint, look at $a + b$ for any state (a, b) that we are in, and show that it always decreases.

Amazing thing about Euclid's algorithm is that it is indeed quite efficient. A better analysis shows that it will always terminate in $\log(m+n)$ steps. (Hint: show that after two steps, m will be at most *half* of what it was before..)