DISCRETE MATHEMATICS AND STRUCTURES

A Concise Introduction

Written by

Jonathan Lin

University of Maryland
College Park

Last updated August 8, 2020
UMD Bois 2 Publishing inc.

Contents

1 Propositions and Formal Logic **1**
- 1.1 Untitled 1 1
- 1.2 Recursive definition of formal expressions 3

2 Circuits and Adders **5**
- 2.1 Circuit Basics 5
- 2.2 Half and Full Adders 8

3 Set Theory, Functions, Relations **11**
- 3.1 Sets 11
- 3.2 Functions and Relations 13
- 3.3 Functions 13
- 3.4 Relations 14
- 3.5 Exercises 15

4 How to Prove it: Direct and Indirect Proof Techniques **17**
- 4.1 Proof Writing Guidelines 17
- 4.2 Indirect Proof 18

5 How to Prove it: The Well Ordering Principle and Proof by Induction **19**
- 5.1 Weak Induction 19
- 5.2 Summation notation, product notation 19
- 5.3 The Principle of Mathematical Induction 20
- 5.4 Exercises 21
- 5.5 Strong Induction 23
- 5.6 Principle of Strong Induction 23
- 5.7 Exercises 24
- 5.8 Proving the Induction Principle 25
- 5.9 Exercises 27

6 Number Theory **29**
- 6.1 Modular Arithmetic 29
- 6.2 random 250 comment 32
- 6.3 Preliminary Definitions and Results 32
- 6.4 Bezout's Lemma, and the $p \mid ab$ Lemma 34
- 6.5 Immediate applications to irrationality proofs 35
- 6.6 Exercises 36

7 Counting and Probability 39
7.1 The Principle of Multiplication . . . 39
7.2 Permutations . . . 40
7.3 Combinations and the Principle of Division . . . 41
7.4 Probability in terms of Combinatorics . . . 41

8 The Pigeonhole Principle and its Generalizations 43
8.1 The Pigeonhole Principle . . . 43

9 Linear Equations and their Applications 45
9.1 Linear Equations Reintroduced . . . 45
9.2 Row Reduction . . . 46
9.3 Determinant . . . 46
9.4 Cramer's Rule . . . 46

Bibliography 51

Chapter 1

Propositions and Formal Logic

1.1 Untitled 1

Important observation I wanted to highlight today:

> In math, we work with *functions* a lot. Recall the basic property of a function: it will take a value (a number, an animal, anything) and output another value (a letter of the alphabet, the number of the animal's legs, anything). Later on we will learn a more rigorous definition than that described above.

Then a boolean expression is really a function. Each variable takes in either true or false, at which point we can *evaluate* the expression and we get some boolean value (either true or false). Consider the expression

$$(p \wedge q) \vee r.$$

Then we can feed in any combination of trues or falses. For example, if $p \equiv T$, $q \equiv F$, and $r \equiv T$, then the expression above will evaluate these values to

$$(T \wedge F) \vee T \equiv F \vee T \equiv T,$$

or true.

A **truth table** then is really just a way of recording the output values of this function.

1. Create truth tables for the following expressions.
    - $p \wedge (\sim p)$ (This formula is called *unsatisfiable*. Why is this?)
    - $p \wedge T$ (Here, T represents true. Similarly, F represents false).
    - $p \wedge F$
    - $p \vee T$ and $p \vee F$.

> Two logical expressions are called **logically equivalent** if they have the same truth value given a set of truth values for each variable.
> Equivalently, two logical expressions are logically equivalent if they have the same truth table.

2. Let's consider the logical expressions in problem 1 again. Figure out simpler expressions which are equivalent to these.

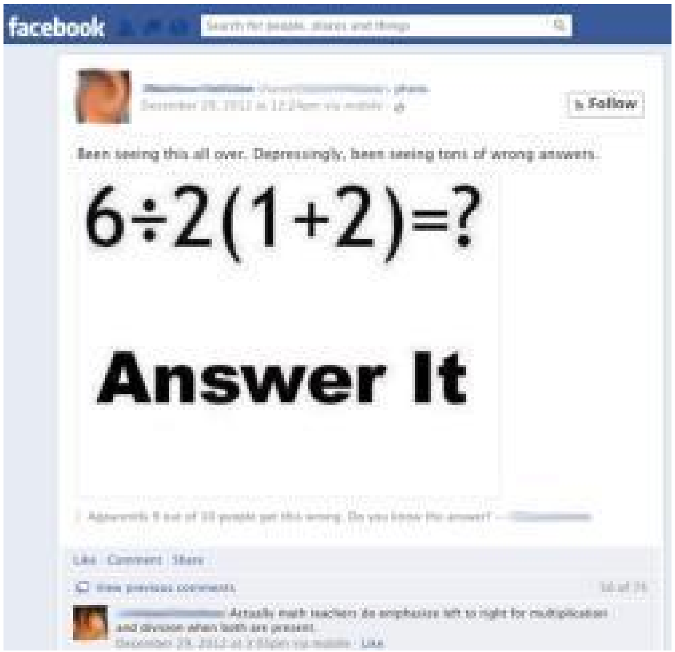
3. Consider the following chunk of code. Here, `var1`, `var2` are boolean variables (ie, they are true or false).

   ```
   if (var1 && (var1 || var2)) { printf("She loves me!"); }
   else { printf("She loves me not..."); }
   ```

   Simplify the code.

4. In logical expressions, parentheses matter!

   Build the truth tables for $(\sim a) \wedge b$ and $\sim (a \wedge b)$. Compare it to the viral facebook math problem below:



Been seeing this all over. Depressingly, been seeing tons of wrong answers.

6÷2(1+2)=?

Answer It

Apparently 9 out of 10 people get this wrong. Do you know the answer? — [redacted]

[redacted] Actually math teachers do emphasize left to right for multiplication and division when both are present.

Figure 1.1: Viral Facebook Math Meme

5. This problem mostly deals with the implies ($\implies$) logical operator. The motivation behind this operator is that eventually we want to do **proofs**. That is, given a set of initial hypotheses, we want to make deductions and logically deduce that a conclusion is true. For example, eventually all of you will be able to prove, roughly stated:

   $$(n \text{ is even}) \implies (n+1 \text{ is odd}).$$

   Construct the truth tables for all the logical expressions below.

   - $p \implies (p \wedge q)$ and $p \implies (p \vee q)$.

- $(p \wedge q) \implies (p \vee q)$. This formula is an example of a *tautology*.
- $(p \vee q) \implies (p \wedge q)$.
- $(p \implies q) \implies (p \iff q)$.
- $(p \iff q) \implies (p \implies q)$.
- $p \implies q \implies r$ and $p \implies r$.

1.2 Recursive definition of formal expressions

In this worksheet, we will briefly discuss a property of boolean formulas called **satisfiability**. This was briefly discussed on the last worksheet, with respect to the boolean formula $p \wedge (\neg p)$.

A boolean formula is **satisfiable** if there exists an assignment of true or false to each of its variables which makes the formula evaluate to true. Note that we only need *one* such assignment of variables. A boolean formula is **unsatisfiable** if it isn't satisfiable. In particular, all assignments of variables will evaluate to false.

For example, the formula $p \wedge q \wedge r$ is satisfiable (with satisfying assignment $p \equiv q \equiv r \equiv T$. And this is the ONLY satisfying assignment).

1. Determine whether or not the following Boolean formulae are satisfiable:
    - $x_1 \wedge x_2 \wedge \cdots \wedge x_n$, and $x_1 \vee x_2 \vee \cdots \vee x_n$, where n is a natural number.
    - $(x_1 \wedge (\neg x_2)) \vee (x_2 \wedge (\neg x_3)) \vee \cdots \vee (x_{n-1} \wedge (\neg x_n))$.
    - $(x_1 \vee \neg x_2) \wedge (\neg x_1 \vee x_2) \wedge (x_1 \vee x_2) \wedge (\neg x_1 \vee \neg x_2)$.
    - $(x_1 \vee (\neg x_2) \vee x_3) \wedge (x_1 \vee (\neg x_2) \vee (\neg x_3)) \wedge ((\neg x_1) \vee x_2 \vee (\neg x_3))$.
2. Explain why a boolean formula with n variables has 2^n different assignments.
    a) Suppose I was to tell you that a certain boolean formula was satisfiable. What would I need to tell/demonstrate to you in order to convince you of this fact? How much time would it take?
    b) Suppose on the other hand I was to tell you that a certain boolean formula was *unsatisfiable*. What would I need to demonstrate to you in order to convince you of this fact? How much time would it take?
3. Sometimes, despite the computational time it takes to find a satisfying assignment, some unfortunate circumstance may require that you do so. This problem will outline one or two heuristics in order to possibly simplify your job a little further.

    Consider the boolean formula

    $$(x_1 \wedge x_5 \wedge x_2) \vee (x_3 \wedge (\neg x_3) \wedge x_5) \vee (x_2 \wedge (\neg x_7) \wedge x_6 \wedge x_1 \wedge (\neg x_4)).$$

    a) Notice that this formula is in *disjunctive normal form*. That is, it can be represented as an OR of ANDS. Explain how this might simplify our problem a little bit.
    b) Simplify the formula by considering the variable x_3.
    c) Discuss (with your neighbors) the theoretical or practical use of such simplifications.

Chapter 2

Circuits and Adders

2.1 Circuit Basics

This worksheet primarily deals with logical circuits and their relationship with logical expressions. Below is a helpful chart of commonly used logical gates and their associated logical expressions. Also, in terms of logical notation, $\oplus$ and $\iff$ will represent XOR and XNOR, respectively. NAND and NOR will be represented using $\uparrow$ and $\downarrow$, respectively.

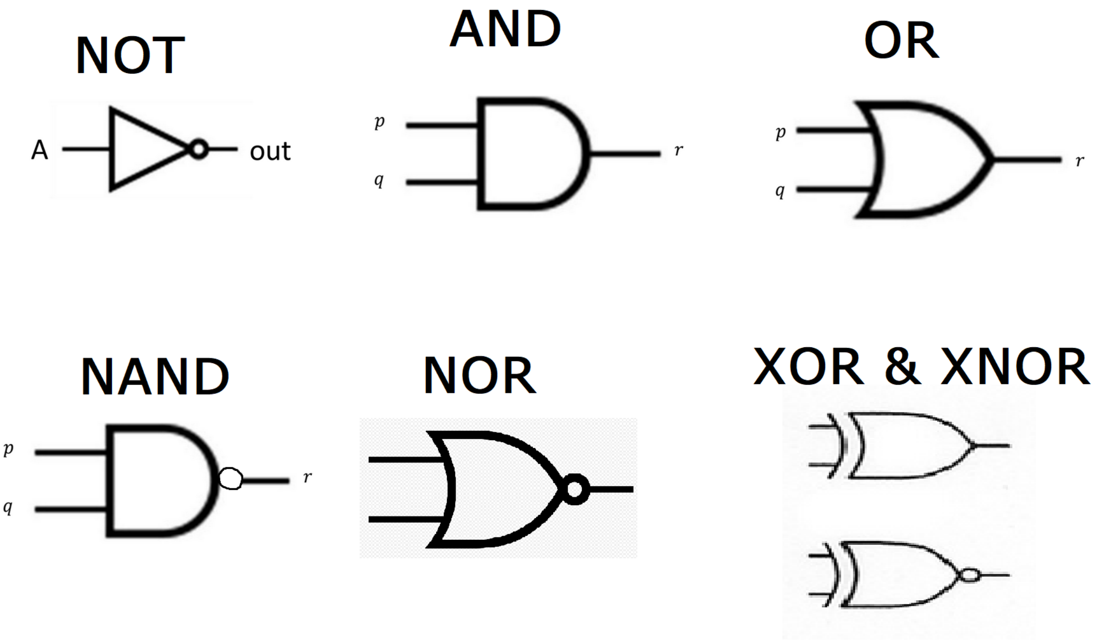


Figure 2.1: Circuit diagram gates

1. The following question is a review from Thursday's lecture.

a) Write $x \oplus y$ (XOR) using only the and ($\wedge$), or ($\vee$), and not ($\sim$) operations.

b) Construct a circuit representing the boolean formula you wrote in part a).

2. For this problem and the next one assume that logic gates can take more than 2 inputs for inputs where they are **associative**. The reason why we require this is that some operators, such as NAND and NOR, are NOT associative and thus expressions like

$$p \uparrow q \uparrow r$$

are ambiguous.

Simplifying boolean expressions isn't all that useless! It seems to have "applications" in digital logic via the simplification of circuits, as discussed on circuit lecture. This problem is mostly a bunch of practice problems about this topic.

a) Consider the logical expression $(r \vee ((s \vee q) \wedge r)) \wedge (\sim(\sim(s \wedge q))$.

   i. Draw the circuit associated with this diagram.

   ii. Simplify the logical expression above and draw an updated circuit associated with this diagram.

b) Do the same thing, but for the logical expression $(a \wedge (\sim b)) \vee (a \wedge b)$.

c) i. Okay, instead of converting a boolean formula into a circuit, convert this boolean circuit I made in mspaint to a boolean formula.

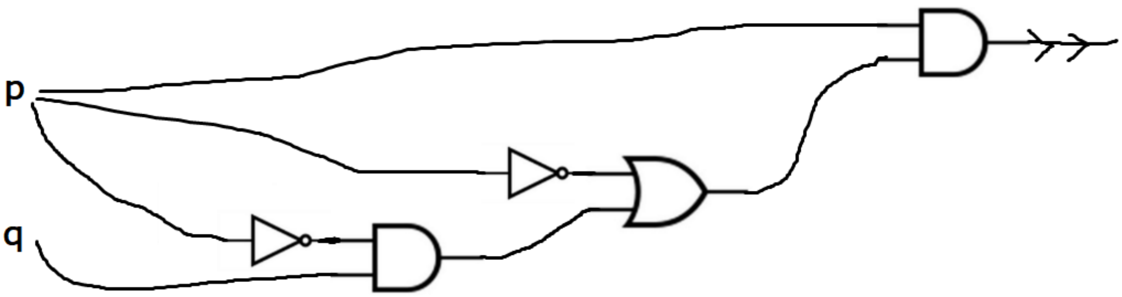


Figure 2.2: Caption

ii. Simplify this formula. After that, go ask around to see if there are any EE or CE majors in the room, because if you are not one of these people it is highly likely that you will not know how to draw the circuit for the simplified formula.

3. Assume for this problem that logic gates have a **cost** (just like in real life!). The cost list is below.

Gate	**Cost**
NAND	2
NOR	4
NOT	3
AND	7
OR	3
XOR	6
XNOR	11

Table 2.1: Some fictional gate costs.

Suppose (for the sake of imagination) that this is the cost list for the fictional Alpha Beta Gamma computer hardware company. You are a consultant who is brought in to improve business pricing and advertising (although in this problem we will only focus on the pricing). Assume that people won't buy a gate if there is a cheaper alternative available.

a) The NAND gate is really, really cheap. (Actually, in real life, the NAND gate turns out to be the cheapest logic gate to construct, it requires the least transistors). Construct a NOT gate using the least number of NAND gates possible. Conclude if the current price for NOT is appropriate or not.

b) Try to do the same thing for all the other logical gates and conclude whether or not their pricing is appropriate or not for each gate.

c) (Challenge Problem, will probably not be discussed). Show that given any circuit, you can create an equivalent circuit using only the NAND gate. In other words, the NAND operator will generate any truth table. (Hint: what does the conjunctive normal form show?)

2.2 Half and Full Adders

This worksheet mostly has to do with the half and full-adder constructions discussed in lecture. I will review the half adder and we will build up to the full adder and n-bit adders in the problems below. I'll also review addition in binary, just to make sure everyone has it down.

1. In this problem, assume that all numbers are in binary.

   Add 10101 and 11011 together. After that, Add 11011 and 10011 together.

2. Assuming that we have a correct half adder circuit, we will construct a full adder circuit. The following box below describes how to add 3 one bit numbers p, q, and r.

   If you are a little confused by this description, try to keep in mind that we are really just doing normal addition, just in binary.

> We want to add the three bits p, q, and r. First we will add the two bits p and q. We will get something like
>
> $$p + q = c_1 s_1$$
>
> where c_1 and s_1 are bits in binary. The half adder we made will compute these two bits. After this, we add r to s_1 to get $c_2 s_2$. To finish off, we add the two carry bits c_1 and c_2 to get a value c. Then the final result of the full adder is cs_2. Note that adding 3 one bit numbers will not get you a number longer than 2 bits.

a) Explain why in this process, the two carry bits c_1 and c_2 cannot be both 1. This is why when combining c_1 and c_2 at the end it suffices to use an OR gate instead of an XOR gate.

b) Based on the information in the box and the considerations of part i), construct the full adder circuit having "black-boxed" the half-adder circuit, as in figure 2.3

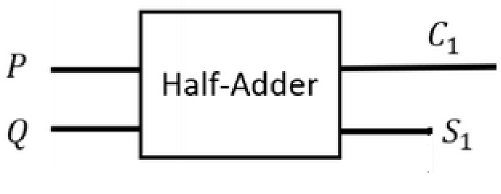


Figure 2.3: Black boxed half-adder

c) How many AND, OR, and NOT gates does this construction use? (The half adder we will be considering uses 2 AND gates, one OR gate, and one NOT gate.)

3. Now we will construct a 2 bit adder, bit by bit. Consider the process in the two bit box below.

> Suppose that we want to add two 2-bit binary numbers ab and cd. The process is as follows, modeling regular vertical addition. First we add the end bits b and d to get some number $c_1 s_1$. Then we add c_1, a and c to get some number $c_2 s_2$. Our binary answer is then $c_2 s_2 s_1$ concatenated.

a) Convince yourself that this process for adding two 2-bit numbers works. I will try to illustrate this on the board for you when we are at this problem.

b) Construct the 2-bit adder circuit. (Hint: I have described two bit addition in a way which nicely uses the operations we've discussed previously.)

c) How many AND, OR, and NOT gates are in a 2-bit adder?

d) Outline in a diagram how an n-bit adder is constructed, based on the idea of the 2-bit adder.

e) How many AND, OR, and NOT gates are in an n-bit adder?

f) This will be an instructive exercise for people who've learned proof by induction already.

   i. First, suppose we have the $(n-1)$-bit adder black boxed. Show that we can make an n-bit adder using this $(n-1)$-bit adder and a full adder. (You should be able to do this problem even if you do not know proof by induction).

   ii. Use this characterization of the n-bit adder as a combination of the $(n-1)$-bit adder and the full adder to verify the number of AND, OR, and NOT gates for an n-bit adder.

Chapter 3

Set Theory, Functions, Relations

3.1 Sets

Hopefully with this worksheet we'll cover the set theory required for all of you to finish the homework good and proper.

1. It's important to realize the limitations of what kinds of sets we can define in our usual notation. The following example, independently discovered by Russel and Zermelo in the early 20th century illustrates that there cannot be a set of all sets.

   Suppose there was a universal set U, the set of all sets. Then U contains itself (since U is also a set). Consider the subset of U defined by

   $$N = \{S \in U \mid S \notin S\}.$$

   That is, N contains all the sets S which do not contain themselves.

   a) Show that N is non-empty.

   b) Suppose $N \in N$. Deduce that $N \notin N$. This is a contradiction, since formally we get the logical expression

   $$(N \in N) \wedge \neg(N \in N) \equiv F.$$

   So then $N \notin N$?

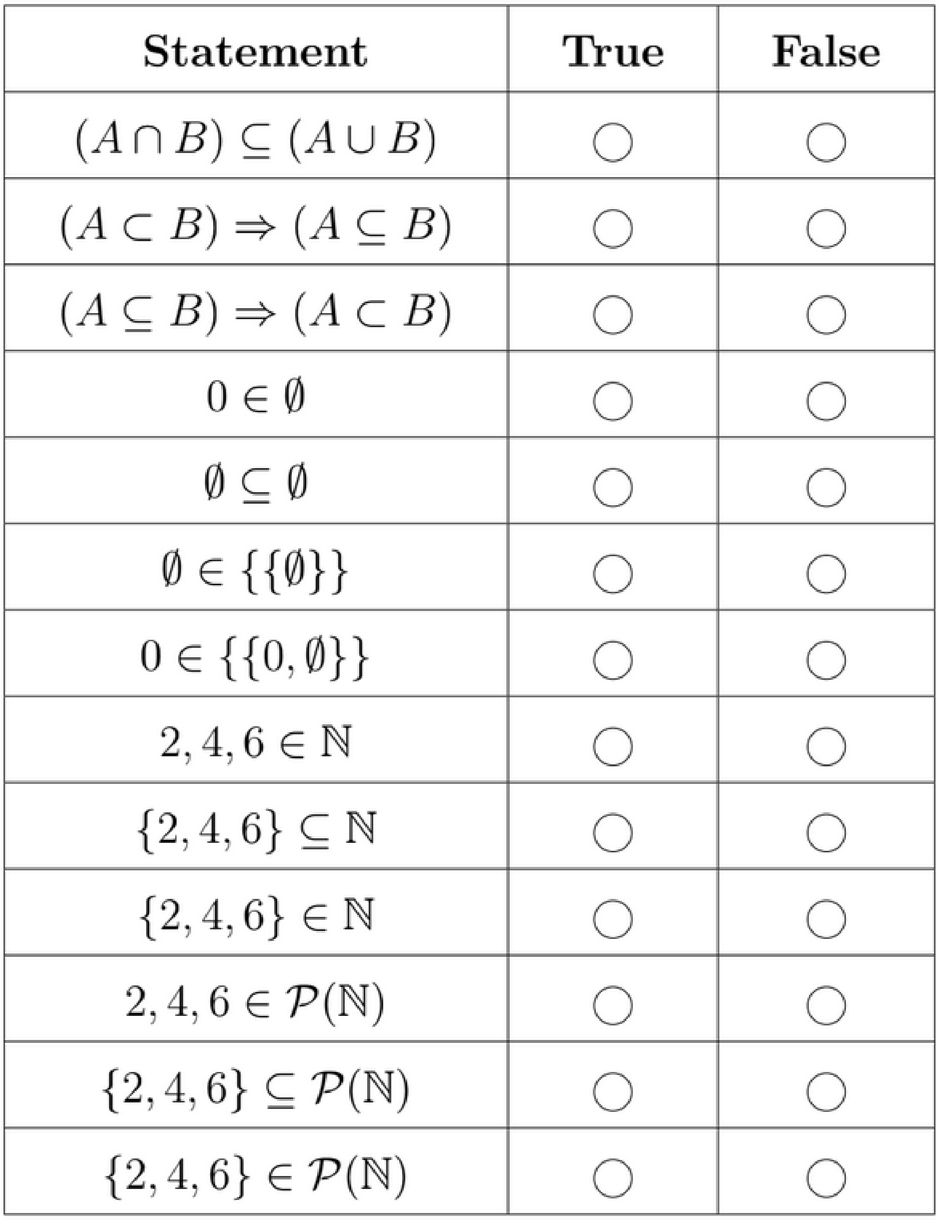
   c) But now suppose $N \notin N$. Reason that $N \in N$.

   So the subset N cannot exist. The conclusion is that the set U is not well-defined (ie, does not exist).

2. Fill in the table on page 2.

   Here are some definitions we will go over. A and B denote any sets.

   - $x \in A$ means an object x is in the set A.
   - $A \cup B = \{x \mid (x \in A) \vee (x \in B)\}$.
   - $A \cap B = \{x \mid (x \in A) \wedge (x \in B)\}$.
   - $A \subseteq B \iff ((x \in A) \implies (x \in B))$.
   - $A \subset B \iff (A \neq B) \wedge (A \subseteq B)$.



Statement	**True**	**False**
$(A \cap B) \subseteq (A \cup B)$	○	○
$(A \subset B) \Rightarrow (A \subseteq B)$	○	○
$(A \subseteq B) \Rightarrow (A \subset B)$	○	○
$0 \in \emptyset$	○	○
$\emptyset \subseteq \emptyset$	○	○
$\emptyset \in \{\{\emptyset\}\}$	○	○
$0 \in \{\{0, \emptyset\}\}$	○	○
$2, 4, 6 \in \mathbb{N}$	○	○
$\{2, 4, 6\} \subseteq \mathbb{N}$	○	○
$\{2, 4, 6\} \in \mathbb{N}$	○	○
$2, 4, 6 \in \mathcal{P}(\mathbb{N})$	○	○
$\{2, 4, 6\} \subseteq \mathcal{P}(\mathbb{N})$	○	○
$\{2, 4, 6\} \in \mathcal{P}(\mathbb{N})$	○	○

- $\mathcal{P}(A) = \{U \mid U \subseteq A\}$. This is the **powerset** of A.
- $|A|$ simply denotes the number of elements in the set A. For example, $|\{1, 2, 3\}| = 3$.

3. This problem hopefully illustrates that the set cardinality operation is a "depth 1" operation. All that matters when considering cardinality is the number of elements/objects in the first level of the set.

   Determine the numbers for the sets below.

- a) $|\{3,6,9\}|$
- b) $|\{\{3\},6,9\}|$
- c) $|\{\{3\},6,9,\{9\}\}|$
- d) $|\{\{3\},6,9,\{\{9\},9\}\}|$
- e) $|\{\mathbb{R}\}|$

4. Just to get you used to the powerset operation. Compute $\mathcal{P}(\{a\})$, $\mathcal{P}(\{a,b\})$, and $\mathcal{P}(\{a,b,c\})$. Compute $\mathcal{P}(\mathcal{P}(\{a,b\}))$.

5. Last question, not really about set theory, but rather about quantifiers. Consider the propositional statement

$$(\forall x \in \varnothing)[x = 5].$$

Is this true? False? Discuss with your tablemates. After this, consider the proposition

$$(\forall x \in \varnothing)[x \in A]$$

where A is any set. (Hopefully, from this we see that $\varnothing \subseteq A$ for all sets A!)

3.2 Functions and Relations

3.3 Functions

Definition and Terminology

In this section, we will answer two questions. Namely, what is a function, and what do functions do?

From prior experience in precalculus and calculus courses most people have an intuitive picture of what functions are. For most purposes, they are “rules” that sends values to other values. In other words:

> A function is a rule that assigns to certain elements in a set, certain elements to another set.

The word rule is not very clear or rigorous. For example, we might consider the real valued functions $f(x) = x^3$ and $g(x) = x^3 + 3x - 3(x + 3) + 9$. These have the same value everywhere, but the rule for evaluating them is slightly different.

It turns out that when defining functions rigorously, the “rule” is a means to an end. All we care about when talking about any specific function is the output given the input.

Definition 1. A function $f : A \to B$ (read as “f from A to B”) is defined as a set of ordered pairs (a, b) where $a \in A$ and $b \in B$ (more specifically, a subset of the set of ordered pairs $A \times B$). f has the following properties:

1. For all $a \in A$, there exists some $b \in B$ such that $(a, b) \in f$.

2. For all $a \in A$, $b, c \in B$, if $(a, b) \in f$ and $(a, c) \in f$, then $b = c$.

In other words, every value $a \in A$ will have some value b such that (a, b) is in f. Moreover, this value is **unique**. We define $f(a)$ as the unique value in B such that $(a, f(a)) \in f$.

Here is some more terminology.

Definition 2. Let $f : A \to B$ be a function. We say A is the **domain** of the function f and B is the **codomain** of f. We say that a is mapped to b or f maps a to b if $(a, b) \in f$. You can also write $a \mapsto b$ as long as it is clear what f is.

Using these definitions, we can say the following: Given any function $f : A \to B$, every value in the domain will be mapped to exactly one value in the codomain.

Injectivity, Surjectivity, Bijectivity

Here are the relevant definitions:

- A function $f : A \to B$ is injective (or one to one) if $f(a) = f(b) \implies a = b$. So two different values in A do not map to the same value in B.
- A function $f : A \to B$ is surjective (or onto) if for all $b \in B$ we can find a value in A such that $f(a) = b$.
- A function f is bijective if it is both injective and surjective.

Here is an interesting observation: if $f : A \to B$ is a bijection, then there is an inverse function $g : B \to A$. For each $b \in B$, we can define $g(b)$ as the unique value $a \in A$ such that $f(a) = b$. We know such a value exists because f is surjective, and we know that this value is unique because f is injective.

3.4 Relations

Let A be a set. Traditionally, we have lots of different notation for when we might want to *compare* two objects in A. For example, the $=$ notation "compares" for equality, and the $\leq$ symbol compares two objects for magnitude. Relations generalize these comparison operators.

Definition 3. Let A and B be sets. A relation R on A and B is a subset of $A \times B$. That is, R is any subset of ordered pairs (a, b) where $a \in A$ and $b \in B$.

For example, a function f is a relation.

Most of the time we will consider relations in the special case where $A = B$. This is, as we might observe, the use case for the equality and inequality operators mentioned at the beginning of this section.

Definition 4. Let A be a set and R be a relation on $A \times A$. A relation is

- *reflexive* if $(a, a) \in R$ for every $a \in A$,
- *symmetric*, if for all $a, b \in A$, $(a, b) \in R \implies (b, a) \in R$.
- *transitive* if for all $a, b, c \in A$, $(a, b) \in R, (b, c) \in R \implies (a, c) \in R$.

We write aRb if $(a, b) \in R$.

A relation is called an **equivalence relation** if it is reflexive, symmetric, and transitive.

3.5 Exercises

1. Let X be any non-empty set, and consider the relation $R = X \times X \subseteq X \times X$. Verify that this relation is an equivalence relation.
2. Give an example of a relation R on a set A that is
   - reflexive and symmetric but not transitive,
   - symmetric and transitive but not reflexive,
   - reflexive and transitive but not symmetric.
3. Suppose $f : A \to B$ and $g : B \to A$ are functions such that
   $$g(f(a)) = a$$
   for all $a \in A$. Show that f is injective and that g is surjective.
4. Let f and g be functions as in the last problem. Suppose also that $f(g(b)) = b$ for all $b \in B$. Show that g is the only function with these properties, that is if h has these properties then $h = g$. (Notice that this equality is technically realized as an equality of **sets**.)
5. (The classification of symmetric, transitive relations.) Let A be a non-empty set, and suppose that R is a non-empty relation which is symmetric and transitive. Show there is a non-empty set $B \subseteq A$ for which R is an equivalence relation restricted to $B \times B$. Explicitly describe the set B. (It might help to do the second part of problem 2 first.)
6. (Equivalence relations as surjective functions) Let X be a set, and let $\sim \;\subset X \times X$ be an equivalence relation. Given any $a \in X$, define
   $$X_a = \{b \in X \mid a \sim b\}$$
   We call this the equivalence class of a.
   - a) Given any $a, b \in X$, show that either $X_a \cap X_b = \varnothing$ or $X_a = X_b$. So distinct equivalence classes are disjoint.
   - b) Let $C = \{X_a \mid a \in A\}$. That is, C is the set of all the equivalence classes (which are sets, specifically subsets of X). Consider the function $f : X \to C$ given by $f(a) = X_a$. First, justify that f is actually a function (that is, it satisfies all the properties that functions satisfy). Then prove that f is a surjective function (this is pretty easy).
   - c) Now consider any surjective function
     $$f : X \to C,$$
     where C is any set. Consider the sets
     $$M_z = \{x \in X \mid f(x) = z\}.$$
     Show that the relation $\sim \;\subset X \times X$ defined by $x \sim y$ if and only if $f(x) = f(y)$ is an equivalence relation. Show that the equivalence classes of $\sim$ are precisely the sets M_z for $z \in C$.
   - d) Let D be the set of equivalence classes of $\sim$ in the previous part. Now define $\tilde{f} : D \to C$ by $\tilde{f}(X_a) = f(a)$. Show that $\tilde{f}$ with this rule defines a valid function and that $\tilde{f}$ is in fact bijective.

   Remark: for the more mathematically inclined the reason why this construction works boils down to fact that the inverse image of a function is well behaved under set union and intersection; see if you can figure out why this fact basically makes the problem work.

Chapter 4

How to Prove it: Direct and Indirect Proof Techniques

4.1 Proof Writing Guidelines

In this worksheet we have several different exercises to get you used to the CMSC250 styled format of step-by-step proofs.

Here are some guidelines you should take when writing proofs in the 250 style:

- When starting the proof, be sure to give all the objects you are working with variable names.
- It helps to number your steps, because you might want to refer to a previous step. If this reminds you too much of geometry two column proofs you can label specific steps with a marker like $(*)$ or $(**)$, for example

  $$\text{we have the equation } e^{\pi i} = -1 \ (*).$$

  $$\text{then by } (*) \text{ and algebra, } e^{\pi i} + 1 = 0.$$
- If you are in doubt whether or not to skip a step, I recommend that you do not skip a step.

Here are some other guidelines:

- Suppose I have a number $n = 2r + 1$. Have I shown that n is odd yet? No! I still need to check that the expression r is an integer. This justification should be one of its own lines in the proof. The lesson here is make sure you have verified **all parts** of a definition *in writing* before claiming something is of that specific definition type.
- What does it mean for an integer n to not be divisible by 3? It means that there is some integer q where either

  $$n = 3q + 1 \text{ or } n = 3q + 2.$$

  This follows from something called the **remainder theorem** for integers. What if we replace 3 with some other number?
- In discussion I will show how to structure a proof by cases.

1. Suppose $x \in \mathbb{N}^{\leq 4}$ (That is, the set of natural numbers less than or equal to 4). Then, prove that $n^3 \leq 3^n$.

2. Prove (by cases) that 100 is not the perfect cube of a natural number. (Hint: there are two cases to consider. The first problem might help.)
3. Answer each question, and prove your answer is correct.
   a) For any $m, n \in \mathbb{Z}$ is $6m + 8n$ even?
   b) For any $m, n \in \mathbb{Z}$ is $10mn + 7$ odd?
   c) For any $c \in \mathbb{Z}$ what can we say about $(c+1)^2 - (c-1)^2$?
4. Suppose $a \in \mathbb{Z}$ is an odd number. Prove that $a^2 = 8m + 1$ for some $m \in \mathbb{Z}$ (that is, a^2 is n more than a multiple of 8). There are many ways to do this problem.
5. Suppose m is an even integer. Prove that $m(m+2)$ is a multiple of 8.
6. (Challenge Problem) Suppose p is a prime number greater than 3. (A prime number p is a positive integer which has no positive integer divisors other than 1 and itself. For example, 2, 3, and 101 are prime, but 15 is not (since 3 divides 15). Prove that p^2 is 1 more than a multiple of 24, that is there exists some integer k such that $p^2 = 24k + 1$. (Hint: at some point you will have to use the result of the previous problem.)

4.2 Indirect Proof

In this document we will outline some typical examples of where indirect proofs will be used.

1. This problem gives you some more practice with the divides $|$ notation.
   a) Prove by use of the contrapositive method that if $a \nmid b$ and $a \mid c$, then $a \nmid (b + c)$.
   b) Prove the result in the first part, but using contradiction instead. (Try doing this part if you're having trouble with the first part)
2. Prove that if $2 \mid a^2$, then $2 \mid a$. Which proof method should we use, contradiction or contrapositive?
3. Prove by contradiction that there are no integers a and b such that $a^2 - 4b = 2$. (You will need to use the previous problem at some point in your argument.)
4. Prove that u is an irrational number if and only if $1 + 3u$ is an irrational number. If you are feeling particularly saucy today, prove one direction using contradiction and another direction using contrapositive. (Reminder that proving an if and only if implication $P \iff Q$ requires proving both $P \implies Q$ and $Q \implies P$.)
5. Prove that for all positive real numbers r and s that $\sqrt{r+s} \neq \sqrt{r} + \sqrt{s}$. So the non-identity called the "freshman dream" is not true.
6. a) Suppose a, b, and c are odd integers. Prove that the polynomial $p(x) = ax^2 + bx + c$ only has irrational roots.
   b) After you have proved this, look up the Rational Root Theorem on your own time. Try to find a proof of this statement. Is the proof direct or indirect? Does it depend on facts which you don't know? Is the approach the same as this exercise (This is a judgement which depends on your taste).
   c) Think about this part after tomorrow's lecture. How similar is this proof to the proof that $\sqrt{2}$ is irrational? In some sense this part is even more subjective than the previous part.

Chapter 5

How to Prove it: The Well Ordering Principle and Proof by Induction

5.1 Weak Induction

5.2 Summation notation, product notation

One should read

$$\sum_{k=m}^{n} a_m$$

as the value of `result` after executing this code:

```
result = 0;
for (k = m; k <= n; k = k + 1) {
    result = result + a_m.
}
```

Ideally, most of the time you should see the above sum as just a compactified way to write

$$a_m + a_{m+1} + \cdots + a_n$$

Notice that this code does not make sense if $m > n$, because the index k is **always** incremented. In this case we always take the sum to be 0 by default. So a sum like

$$\sum_{k=1}^{0} k$$

is equal to 0, not 1.

Similarly for products. one should read

$$\prod_{k=m}^{n} a_m$$

as the value of `result` after executing this code:

```
result = 1;
for (k = m; k <= n; k = k + 1) {
    result = result * a_m.
}
```

Again this code does not make sense if $m > n$. In this case we always take the product to be 1 by default.

The following equations are some useful properties related to summations.

$$\sum_{i=m}^{n}(a_i + b_i) = \sum_{i=m}^{n} a_i + \sum_{i=m}^{n} b_i$$

The above property is often colloquially referred to as "splitting a sum".

$$\sum_{i=m}^{n} ca_i = c\sum_{i=1}^{m} a_i$$

if c is a constant that does not depend on the index i.

$$\sum_{i=m}^{n} a_i = a_m + \sum_{i=m+1}^{n} a_i$$

$$\sum_{i=m}^{n} a_i = a_n + \sum_{i=n}^{n-1} a_i$$

Note that the above two equations make sense if $n \leq m$. (What happens when $n = m$?)

By yourself come up with some analogous formulas for products.

5.3 The Principle of Mathematical Induction

Definition 5. The proposition $P(n)$ is a logical statement involving some value n.

It is very helpful to consider an example or two. Define $P(n)$ to be the logical statement that $n = 1$. In other words, $P(n) \equiv (n = 1)$. Then $P(n) \equiv T$ if and only if $n = 1$, which is true but is silly to say. Another example: define $P(n) \equiv (2 \mid n)$. Then

$$P(n) \equiv \begin{cases} T, & 2 \mid n \\ F, & 2 \nmid n \end{cases}$$

But as of now we aren't really interested in examples like this. We'd like to know if there is a way to show that $P(n) \equiv T$ for $n \geq k$ for some integer k, for example.

For example, we'd like to figure out whether or not

$$\sum_{i=1}^{n}(2i - 1) = n^2$$

for all n. It seems true! But how might we prove it? This leads us to the principle of mathematical induction, which is the basis for the majority of proofs we'll do for the rest of the semester.

Theorem 1. *Let a be any natural number. And let $P(k)$ be a proposition about $k \in \mathbb{N}$. Assume the following claims are true:*

- $P(a) \equiv T$
- *Assuming $P(k)$ is true for any natural number $k \geq a$, we can prove that $P(k+1)$ is true.*

Then $P(k)$ is true for all $k \geq a$.

Proof. Omitted. To convince yourself it is true you might draw a number line. □

To actually prove a proposition $P(x)$ is true for all $x \geq a$ using induction, it is a two step process:

1. First show that $P(a) \equiv T$.
2. Next, assuming that $P(k)$ is true for $k \geq a$, prove that $P(k+1)$ is true.

Once you show the above two steps by the principle of mathematical induction you have shown the claim for all natural numbers $x \geq a$.

In 250, here is how we format this proof, which is a 3 step process. For convenience, I will take $a = 0$. How to prove $P(n)$ for all $n \geq 0$:

1. Inductive Base: Show that $P(0) \equiv T$.
2. Inductive hypothesis: State the following: For all $n = k \geq 0$ assume that $P(k)$ is true.
3. Inductive Step: Assuming that $P(k)$ is true, prove that $P(k+1)$ is true.

We will do several examples of how to do this in lecture and discussion, and you have chance to practice with the below exercises (which are rather hard apart for most of problem 1).

5.4 Exercises

1. Prove by induction the following equalities for appropriate integers. You, the reader, will have to determine the appropriate base case.

   a)
   $$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}.$$
   (For a solution, peep Jason's lecture!)

   b)
   $$\sum_{k=1}^{n} k(k+1) = \frac{n(n+1)(n+2)}{3}.$$

   c)
   $$\sum_{k=1}^{n} k(k+1)(k+2) = \frac{n(n+1)(n+2)(n+3)}{4}.$$

d) i.

$$\sum_{k=0}^{n} 2^k = 2^{k+1} - 1.$$

ii.

$$\sum_{k=0}^{n} 3^k = \frac{3^{k+1} - 1}{2}$$

iii.

$$\sum_{k=0}^{n} 4^k = \frac{4^{k+1} - 1}{3}.$$

Also explain alternatively how we can use the example in the Monday discussion to prove this result.

iv. State the pattern you see, and prove your general result using induction.

v. Using the so called “high-school” algebra identity

$$(x-1)(x^n + x^{n-1} + \cdots + x + 1) = x^{n+1} - 1,$$

give another proof of your general result. (This will be a non-inductive proof).

2. Consider the pseudocode for the following recursive function, which takes a list `ls` $= (a_0, \ldots, a_n)$ and a function f and returns a list $(f(a_0), f(a_1), \ldots, f(a_n))$ for any non-negative integer n.

   In the code:

   - `length` is a function returning the length of a list,
   - `ls[0]` accesses the first element of the list.
   - `pop(ls)` returns the original list but without the first element. For example, `pop`$(2, 5, 3) = (5, 3)$.
   - `::` represents a push operator. For example, 2`::`$(5, 3)$ will be the list $(2, 5, 3)$.

```
function map(ls, f) =
    if (length(ls) = 0) {
        return () /* the empty list. */
    }
    else {
        return f(ls[0])::map(pop(ls), f);
    }
```

   Prove that the code above is **correct**. That is, given any list $(a_0, \ldots, a_n)$ prove that `map` will return $(f(a_0), f(a_1), \ldots, f(a_n))$. (Hint: you'll need to use induction. What are you inducting on? Once you have this in mind this problem turns out to be very simple.)

3. Take a couple of minutes and try to prove by induction that for any integer $n \geq 1$, we have

$$\sum_{k=1}^{n} \frac{1}{k^2} < 2.$$

   (You won't be able to do it!)

Let's prove something seemingly "stronger": Prove by induction that for any integer $n \geq 1$ we have

$$\sum_{k=1}^{n} \frac{1}{k^2} < 2 - \frac{1}{n},$$

which implies the above result.

4. Define a function $f(x, y)$, where x and y are allowed to be **natural numbers.** (which include 0). Suppose f has the following properties:

   - $f(n, n) = f(n, 0) = 1$ for all $n \in \mathbb{N}$.
   - $f(n+1, k+1) = f(n, k) + f(n, k+1)$ for all natural numbers n and k.
   - $f(n, m) = 0$ if $m > n$.

   Prove by induction (on $n \geq 0$) that

   $$\sum_{i=0}^{n} f(n, i) = 2^n.$$

   (Hint: when proving the inductive step, you will have to split the sum and reindex.)

5. A *triomino* is an "L" shaped tile formed by 3 adjacent squares of a chessboard. We say that an arrangement of triominos is a *tiling* of a chessboard if every square of the chessboard is covered without any triominos overlapping.

   Prove by induction that for all integers $n \geq 1$ the $2^n \times 2^n$ sized chessboard missing one square (in ANY location) can be tiled. The fact that $3 \mid (2^n 2^n - 1 = 4^n - 1)$ is a consequence of problem 1diii.

6. (Challenge Problem): Prove Euler's Formula: For any convex polyhedron, the formula

   $$v - e + f = 2$$

   holds, where v is the number of vertices, e is the number of edges, and f is the number of faces. You will want to induct on the number of vertices for $v \geq 4$ (because the notion of "face") doesn't make sense for $v < 4$. Your proof will probably not be completely rigorous, and that's okay. (Hints: For the base case, a tetrahedron is the only You cannot solve this problem by just chopping off a vertex, as in general the vertices which are adjacent to the vertex you are attempting to chop off are not generally going to be all in a shared plane. Look up "edge contraction" if you are stuck without ideas about the inductive step.)

5.5 Strong Induction

5.6 Principle of Strong Induction

Here is the Principle of Strong Induction:

Theorem 2. *Suppose $P(n)$ is a proposition, and that the following statements hold:*

- *$P(0)$ is true,*
- *If $P(i)$ is true for all $i \leq n$, then we can prove that $P(n+1)$ is true.*

Then $P(n)$ holds for all $n \in \mathbb{N}$.

This is very powerful.

5.7 Exercises

1. Determine whether or not $P(n)$ is true for all $n \geq 0$ given that P satisfies the rules in part (a), part (b), etc. (No proof is needed)

   a)
      - $P(0)$, $P(1)$, $P(2)$ are true.
      - If $P(n)$ is true, then $P(n+3)$ is true.

   b)
      - $P(0)$ and $P(1)$ are true,
      - If $P(n)$ is true, then $P(2n)$ is true,
      - If $P(n)$ is true, then $P(n-1)$ is true.

   c) Challenge Problem:
      - $P(0)$, $P(1)$ are true.
      - If n is odd, then $P(n)$ is true if and only if $P(3n+1)$ is.
      - If n is even, then $P(n)$ is true if and only if $P(n/2)$ is.

2. In discussion, I presented the following notion of strong induction (which we called Weak Induction++):

   Theorem 3. *Let $P(n)$ be a proposition, and suppose the following statements hold:*

   - *$P(0)$ and $P(1)$ is true.*
   - *If $P(n-1)$ and $P(n)$ are true for $n \geq 1$, then $P(n+1)$ is true.*

   The conclusion is that $P(n)$ is true for all $n \in \mathbb{N}$.

   Generalize this theorem to multiple base cases.

3. Recall during discussion when we proved that $a_n \leq (7/4)^{n+1}$, we stated an inductive hypothesis which was similar to that in Theorem 1. Explain why we couldn't state and prove the inductive hypothesis with only one base case $n = 0$, mimicking the format of the above theorem. (Recall that for $n \geq 2$, we had $a_n = a_{n-1} - a_{n-2}$.)

4. The following problem outlines a proof of the so called **arithmetic-geometric mean inequality**: for all non-negative reals $a_1, \ldots, a_n$ we have

$$\frac{1}{n}\sum_{i=1}^{n} a_i \geq \sqrt[n]{\prod_{i=1}^{n} a_i}.$$

   a) When $n = 1$ this inequality is trivially true. For $n = 2$: Note that for all $a > 0$ there is $b > 0$ such that $b^2 = a$. Use this fact and the fact that $(x-y)^2 \geq 0$ for all (non-negative) x and y to show directly that the inequality is true when $n = 2$.

   b) Suppose this inequality is true for $n = k$. Using the $n = 2$ case, show that this inequality is true for $n = 2k$.

   c) Supposing this inequality is true for $n = k$, show that this inequality is true for $n = k - 1$.

   d) Conclude that we are done. (This method of proof is called **Cauchy Induction**.)

5. Suppose that there are n people that board a flight. They seat themselves one at a time. The first person has forgotten which seat he is assigned, and picks a random seat to sit in. Each subsequent person either sits in their designated seat, or if it is already occupied, they sit in a remaining seat randomly (as in uniformly random).
   a) When $n = 2$, what is the probability that the last person (person 2) sits in the seat that he is assigned?
   b) What about when $n = 3$? $n = 4$?
   c) Make a conjecture for general n and prove your conjecture using strong induction.

5.8 Proving the Induction Principle

The Well Ordering Principle

We have (?) stated the principle of Mathematical Induction. In the following sections, we attempt to demonstrate that the principle of Mathematical Induction is a statement that requires "proof". The main technical detail needed to prove this principle is the well ordering principle of the natural numbers. First we indicate what it means for a set to be well ordered.

Definition 6. Let S is a set with an order relation $\leq$ on $S \times S$. Then S (along with the order relation $<$) is called well ordered if every subset of S has a least element.

In other words, for any non-empty subset $T \subset S$ there exists $t_0 \in T$ such that $t_0 \leq t$ for all $t \in T$.

Example 1. The set $\mathbb{Z}$ is not well ordered, for $\mathbb{Z}$ has no least element. In that case, consider the set $Y = \mathbb{Z} \cup \{-\infty\}$ with the usual order relation plus the order $-\infty < a$ for all $a \in \mathbb{Z}$. Then Y has a least element (namely, $-\infty$) but is not well ordered, for the subset $\mathbb{Z}$ of Y has no least element.

The above example leads us to a first elementary observation: any subset of a well ordered set is well ordered. This simply follows from the fact that a subset of a subset is a subset of the original set.

With these preliminary details we can state the well ordering principle of the natural numbers. Usually, this is taken as an **axiom** of the natural numbers[1]. This means that we take it to be the

> The set of natural numbers $\mathbb{N}$, along with the usual ordering $<$, is a well ordered set.

This makes it possible to come up with various interesting examples of well ordered sets.

Example 2. Consider the set $\{1, 2\} \times \mathbb{N}$ in the *dictionary order*: we say that $(a, b) < (c, d)$ if $a < c$ or if $a = c$ then $b < d$. Then $\mathbb{N} \times \mathbb{N}$ is well ordered. To see this, let S be a non-empty subset of $\{1, 2\} \times \mathbb{N}$. Then consider all the elements of S with 1 as the first component. If this subset is non-empty then it follows by the well ordering principle on $\mathbb{N}$ that there is a least element of the form $(1, k)$. Otherwise, we can apply the same reasoning to conclude that there must be a least element of the form $(2, k)$.

The well ordering principle is powerful. We will first use it to prove the quotient remainder theorem (which was first introduced in (atm nowhere as of now)) here, and we will use it later in Chapter 7 to prove some fundamental results about divisibility.

---

[1] If we formally define the natural numbers, then the well ordering principle can actually be proven. But this is beyond the scope of this book (for now).

Theorem 4. *Suppose that* a *is any integer and* b *is any positive integer. Then there exist unique integers* q *and* r *such that* $0 \leq r < b$ *and we have*

$$a = qb + r.$$

The numbers q *and* r *are called the* ***quotient*** *and* ***remainder*** *of the division of* a *by* b*.*

The proof idea is very simple. In the theorem statement, the *remainder* r is intuitively in a sense the "smallest" positive value of remainder we can get by subtracting integer multiples of b from a. We make this reasoning precise by invoking the well ordering principle to explicitly obtain r.

Proof. Let

$$S = \{n \in \mathbb{Z} \mid n = a - kb, k \in \mathbb{Z}, n \geq 0\}.$$

That is, S is all the non-negative values $a - kb$ where k ranges across integer values. Then S is a subset of $\mathbb{N} \cup \{0\}$, which is well ordered because $\mathbb{N}$ is well ordered (by the well ordering principle). It follows that S has a least element r. Let q be the value such that $a - qb = r$ (which is possible precisely because r is contained in S). Then $a = qb + r$.

Now all that is left is to show that this representation $a = qb + r$ is unique. To show this suppose that there is another pair q' and r' of integers with $0 \leq r' < b$ we also have $a = q'b + r'$. Subtracting both equations we have that

$$0 = (q - q')b + (r - r').$$

Since $0 \leq r, r' < b$ it follows that $|r - r'| < b$ so we get hte equation

$$|q - q'|b < b \implies |q - q'| < 1.$$

Since q and q' are integers it follows that $q = q'$ and hence $r = r'$, showing that this representation $a = qb + r$ is unique. □

Equivalence of the Well Ordering Principle and the Principle of Mathematical Induction

Now we prove from the Well Ordering principle the Principle of Mathematical Induction. The idea will be as follows: we will consider the *smallest* element such that a proposition is false, and then we will use the assumptions of the Priniciple of Mathematical Induction to derive contradictions. Hence the Principle of Mathematical Induction is true as stated.

In fact, if we assume the Principle of Mathematical Induction as a given statement, we can use it to show the well ordering principle of the natural numbers. But to give a rigorous proof we need the Principle of Strong Induction, which we will prove from the Principle of Weak Induction later.

Theorem 5. *Suppose that* $P : \mathbb{N} \to \{T, F\}$ *is a boolean function with the following properties:*

- $P(1) \equiv T$
- $P(n) \equiv T \implies P(n+1) \equiv T$ *for all* $n \in \mathbb{N}$*.*

Then $P(n) \equiv T$ *for all* $n \in \mathbb{N}$*.*

Proof. Assume for the sake of contradiction that there is some $s \in \mathbb{N}$ such that $P(s) \equiv F$. Let S be the set of all s where $P(s) \equiv F$. By assumption, $S \neq \varnothing$, and hence by the well ordering principle there is a minimal element $r \in S$. We see that $r \neq 1$ because $P(1) \equiv T$ by assumption. So $r - 1 \in \mathbb{N}$. But since $r \in S$ is minimal, we must have that $P(r-1) \equiv T$ (or else our assumption about r being the minimal element in S would be wrong). But since $P(r-1) \equiv T$, this implies that $P(r) \equiv T$ (by modus ponens). This is a contradiction and the conclusion follows. □

To prove the other direction (namely, that the Principle of Mathematical Induction implies the Well Ordering Principle), we need the stronger principle of Strong Induction. So what we will do is use the Principle of Mathematical Induction to prove the Principle of Strong Induction first.

Theorem 6. *Suppose that the Principle of Mathematical Induction is true. Suppose that* $P : \mathbb{N} \to \{T, F\}$ *is a boolean function with the following properties:*

- $P(1) \equiv T$
- *If* $P(k) \equiv T$ *for all* $1 \leq k \leq n$ *for* $n \geq 1$*, then* $P(n+1) \equiv T$.

Then $P(n) \equiv T$ *for all* $n \in \mathbb{N}$.

Proof. Define an auxiliary function $Q : \mathbb{N} \to \{T, F\}$ by the following:

$$Q(n) \equiv \begin{cases} T & P(k) \equiv T \text{ for } 1 \leq k \leq n \\ F & \text{otherwise.} \end{cases}$$

Then by assumption, $Q(1) \equiv T$ and $Q(k) \equiv T$ implies that $P(n+1)$ is true, which together with $Q(n)$ implies that $Q(n+1) \equiv T$. Hence by the Principle of Mathematical Induction we have that $Q(n)$ is true for all $n \in \mathbb{N}$ which implies that $P(n)$ is true for all $n \in \mathbb{N}$, as desired. □

The converse statement is also true, this will be left as an exercise. Now that we have the principle of strong induction, we can use it now to prove the well ordering principle. This will show that the Well Ordering Principle and the Principle of Mathematical Induction are equivalent.

Theorem 7. *The Strong Induction Principle implies the Well Ordering Principle. Hence the Well Ordering Principle and the Principle of Mathematical Induction are equivalent.*

Proof. Suppose that $S \subset \mathbb{N}$ is a non-empty set with no least element. Let P be the following proposition:

$$P(n) = \begin{cases} T & n \notin S \\ F & n \in S \end{cases}.$$

So the proposition $P(n)$ reflects whether or not n is in the set S. Since 1 is the least element in $\mathbb{N}$, we know that $1 \notin S$ or else S would have a least element. So $P(1) \equiv T$. Now suppose for any $n \in \mathbb{N}$ that $P(k) \equiv T$ for $1 \leq k \leq n$. Then by the definition of P no natural number less than or equal to n is in S. It follows that $n+1 \notin S$ or else $n+1$ would be the least element in S. It follows that $P(n+1) \equiv T$.

By the Principle of Strong Induction it follows that $P(n) \equiv T$ for all $n \in \mathbb{N}$. This implies that no element $n \in \mathbb{N}$ is in S. So $S = \varnothing$. The Well Ordering Principle follows, as desired. □

5.9 Exercises

- Show that the following sets are well-ordered under the ordering described. (It will be helpful to use the fact that $\mathbb{N}$ is well-ordered.
  - $\mathbb{N} \cup \{\omega\}$, where $\omega > n$ for any $n \in \mathbb{N}$ and where $\mathbb{N}$ has the usual ordering.
  - $\mathbb{N} \times \mathbb{N}$ under the dictionary order (which was defined in Example 2).
  - $\mathbb{Z}$ under the following ordering R: aRb if $|a| < |b|$, and if $|a| = |b|$ then $a < b$ if a is negative and b is positive.

- Show that the following sets are not well-ordered:
    - The set of rationals $\mathbb{Q}$ under the usual ordering.
    - The set $\{2^n \mid n \in \mathbb{Z}\}$ under the usual ordering on $\mathbb{Q}$.
    - The cube $[0,1] \times [0,1]$ under the dictionary ordering.
- Show that for any well-ordered set S with order relation $<$ and any element $r \in S$, either r is a maximal element ($r \geq s$ for all $s \in S$) or r has an immediate successor r^+, that is, a smallest element greater than r. Is every non-minimal element r an immediate successor of some other element?
- Suppose that a is any integer and b is any positive integer. State and prove when there exist integers q and r such that $r < |b/2|$ and we have

  $$a = qb + r.$$

  When can we ensure that q and r are unique?
- Prove that the Principle of Strong Induction implies the Principle of Mathematical Induction.
- Prove the Principle of Strong Induction using the Well Ordering Principle.
- Suppose that $P(n)$ is a proposition. In each part assume $P(n)$ satisfies the given list of properties and use the Well Ordering Principle to prove that $P(n)$ is true for all $n \geq 0$. After this, generalize.
    1. - $P(0)$, $P(1)$, and $P(2)$ are true,
       - $P(n) \equiv T \implies P(n+3) \equiv T$.
    2. - $P(0)$ and $P(1)$ are true,
       - If $P(n)$ is true, then $P(2n)$ is true,
       - IF $P(n)$ is true, then $P(n-1)$ is true.

Chapter 6

Number Theory

6.1 Modular Arithmetic

Mods are a tricky concept to get right because they rely on internalizing concepts that you've learned earlier very well. Once you've internalized those concepts you'll see that modular arithmetic is really just a recasting of divisibility. In this handout, I'd like to highlight the correspondence between divisibility, mods, and the quotient remainder theorem.

Quotient Remainder Theorem: Suppose that n is an integer and m is a positive integer. Then there exist *unique* integers q and r with $0 \leq r < m$ where

$$n = qm + r.$$

Let a and b be any integers. We say that a divides b (and we write $a \mid b$) if there exists an integer k such that

$$a \cdot k = b.$$

We are going to define mods in terms of divisibility. To motivate the definition a little bit, when we take mods we want to have that two numbers are equivalent modulo n if they have the same remainder (as in the quotient remainder theorem) when they are divided out by n. So for example, $15 \equiv 3 \pmod{4}$ because $15 = 4(3) + 3$. Here is the definition:

Let a and b be integers, and let n be a positive integer. We say that a is congruent to b modulo n (and we write $a \equiv b \pmod{n}$) if $n \mid (a - b)$.

Let's pick apart this definition a little bit. In the above example, we have $15 \equiv 3 \pmod{4}$. Is that true? It is, since we see that $15 - 3 = 12 = 4(3)$, so $4 \mid (15 - 3)$, as we wanted. (If you are confused, I am taking $a = 15$, $b = 3$ in this example).

To get you used to the proof approaches using modular arithmetic I will do a simple proof involving mods below.

$a \equiv b \pmod{n}$ if and only if $a - b \equiv 0 \pmod{n}$.

Proof:

1. Suppose that $a \equiv b \pmod{n}$.
2. By definition, $n \mid (a - b)$.
3. By algebra, $a - b = (a - b) - 0$.
4. By substituting (3) into (2),
$$n \mid (a - b) - 0$$
.
5. By definition $a - b \equiv 0 \pmod{n}$, as desired.
6. The converse is left as an exercise to the reader.

There are three important properties of mods which you will use in modular arithmetic problems. I provide proofs for two of these below.

If $a \equiv b \pmod{n}$ and $c \equiv d \pmod{n}$, then $a + c \equiv b + d \pmod{n}$.

Proof:

1. Suppose the above.
2. By definition, $n \mid (a - b)$ and $n \mid (c - d)$.
3. By definition of divisibility, there exist integers k_1 and k_2 such that $a - b = nk_1$ and $c - d = nk_2$.
4. By (3) and algebra, we add the two equations together to get
$$(a - b) + (c - d) = n(k_1 + k_2).$$
5. Rearranging the LHS we get
$$(a + c) - (b + d) = n(k_1 + k_2).$$
6. $k_1, k_2 \in \mathbb{Z}$ by (3). So by closure $k_1 + k_2 \in \mathbb{Z}$.
7. By (6), (5), definition of mods, it follows that $a + c \equiv b + d \pmod{n}$, as desired.

If $a \equiv b \pmod{n}$ and c is any integer then $ca \equiv cb \pmod{n}$.

Proof: left as exercise for the reader.

If $a \equiv b \pmod{n}$ and $c \equiv d \pmod{n}$ then $ac \equiv bd \pmod{n}$.

1. Given the above:
2. By the first thing proved in this handout, $a - b \equiv 0 \pmod{n}$ and $c - d \equiv 0 \pmod{n}$.
3. By one of the above facts, we have $c(a-b) \equiv 0 \pmod{n}$ and $b(c-d) \equiv 0 \pmod{n}$ (because $c \times 0 = 0$ and similarly with b).
4. By addition of mods, we get
$$c(a-b) + b(c-d) \equiv 0 \pmod{n}$$
5. By algebra on the left hand side we get
$$ac - bd \equiv 0 \pmod{n}$$
6. This gives us $ac \equiv bd \pmod{n}$, as desired.

Finally, we'll recast the remainder theorem in terms of mods. From the equation $n = qm + r$, we get $n - r = qm$, so we have $m \mid n - r \iff n \equiv r \pmod{m}$. With this in mind, here is the reformulation of the quotient remainder theorem:

QRT in terms of modular arithmetic: Suppose that n is any integer and m is a positive integer. Then there exists a unique integer r where $0 \leq r \leq m$ and
$$n \equiv r \pmod{m}.$$

Below are the problems. Everyone should attempt problem 2 (and 3a). Everything else is more challenging and should be attempted at your own discretion.

1. Write out a proof of the equivalence of definitions of the very last box. That is, suppose that the quotient remainder theorem is true. Then, using the definition of mods as given, show that the assertion made in the last box holds as a result. Conversely, assume that the statement in the last box holds, and use it to give a proof of the quotient remainder theorem.
2. Suppose that a is any integer. Then show that either $a^2 \equiv 0 \pmod{4}$ or $a^2 \equiv 1 \pmod{4}$. (In light of the quotient remainder theorem it follows that both cannot happen at the same time.)
3. a) Suppose that $S = \{1, 2, 3, 4\}$. For any $k \in S$, prove that there exists $\ell \in S$ such that $k \cdot \ell \equiv 1 \pmod{5}$. (Hint: this is a finite domain problem.)

b) (**Challenge Problem***) Suppose that $S = \{1, 2, \dots, p-1\}$ for p a prime number. Then show that for all $k \in S$ there exists $\ell \in S$ such that $k \cdot \ell \equiv 1 \pmod{p}$.

c) (**Challenge Problem*****) Prove Wilson's Theorem: For any prime p,

$$(p-1)! \equiv -1 \pmod{p}.$$

Here $n! = n \times (n-1) \times \cdots \times 2 \times 1$ for any positive integer n.

4. (**Challenge Problem**) Suppose $ab \equiv ac \pmod{d}$. Then prove that $b \equiv c \pmod{\frac{d}{\gcd(a,d)}}$. Here $\gcd(a, d)$ is the *greatest common divisor* of a and d, that is, the greatest integer k such that $k \mid a$ and $k \mid d$.

5. Prove that there is no positive integer n such that $4n^2 - 4 \equiv 0 \pmod{19}$.

6.2 random 250 comment

In CMSC250, there are only about two ways in which you might prove a number $\sqrt[n]{q}$ irrational for some $n \in \mathbb{Z}^{>0}$ and some $q \in \mathbb{Z}^{>1}$:

- A Euclidean argument. That is, you prove some lemma of the form "if $q \mid a^n$ then $q \mid a$".
- Use the unique prime factorization theorem and compare the exponents on prime factors modulo n.

There are two main problems with both approaches:

- The Euclidean argument scales poorly to large values of n and large values of q, especially when proving a statement like this by contrapositive, where the naive thing to do is to prove separately each remainder case out of the quotient remainder theorem.
- Proofs by UPFT are hard to write correctly, are a mess of indices, and are difficult to grade and give feedback. Especially if you have an equation of the form

  $$2p_1^{2e_1} \cdots p_n^{2e_n} = q_1^{2f_1} \cdots q_k^{2f_k},$$

  it is tricky to talk about the exponent of the 2 in the unique prime factorization of the number on the left hand side.

Here in this handout I outline better tools (Bezout's Lemma, and the $p \mid ab$ lemma) which completely resolve these two issues. Proofs using these two methods both scale well to large numbers and are not too hard to write correctly.

6.3 Preliminary Definitions and Results

Definition 7. Given integers a and b, we say that a divides b (and we write $a \mid b$) if there exists an integer k such that $a \cdot k = b$.

Definition 8. Given integers a and b and positive integer m we say that a is congruent to b modulo m (and we write $a \equiv b \pmod{m}$) if $m \mid (a - b)$.

Here in this document $\mathbb{Z}$ denotes the integers and $\mathbb{N}$ denotes the integers greater than 0. **Note that this is different from standard CMSC250 convention!** I'll also abuse the following fact a lot, which is often taken as a given (axiomatic) property of the natural numbers:

Let $S \subseteq \mathbb{N}$. Then S has a least element. That is, there is an $s \in S$ such that for all $t \in S$, $s \leq t$.

This property is called the **well ordering property** of $\mathbb{N}$. As an aside, note that this property that T is equivalent to the principle of mathematical induction, as proven in the box below.

To fill in later. It's not too important.

Definition 9. Let a and b be **non-zero** integers. Then the greatest common denominator of a and b (we write $\gcd(a, b)$) is the largest number g which divides a and b.

Example 3. Here are some simple examples illustrating the gcd.

The greatest common denominator of 12 and 8 is 4.
For any integer n, the greatest common denominator of n and $n+1$ is 1. How do we show this? Suppose $a \geq 2$, $a \mid n$ (that is, $ak = n$ for some $k \in \mathbb{Z}$). Then $ak + 1 = n + 1$, and by the quotient remainder theorem it follows that a does not divide $n+1$. We've deduced that no divisor of a greater than 2 can also divide $n+1$. Hence the *greatest* common divisor of n and $n+1$ is 1.

Definition 10. A prime number p is a natural number greater than 1 with the property that the only natural numbers which divide p are 1 and p.

The following lemma is the key to most irrationality applications. It says that if p is prime and $p \nmid a$, then p and a have no commmon factors other than 1. Intuitively and formally, this is clear because p only has 2 factors, 1 and p.

Lemma 1. *Suppose p is a prime and a is an integer with $p \nmid a$. Then* $\gcd(p, a) = 1$.

Proof. The only divisors of p are 1 and p. The condition $p \nmid a$ shows that p is not a divisor of a. Hence 1 must be the *greatest* common divisor of p and a, as desired. □

We also use the following lemma a lot, which is just a way to bound numbers by divisors.

Lemma 2. *Suppose a and b are integers, b non-zero, with $a \mid b$. Then $|a| \leq |b|$. In particular, if a and b are positive, then $a \leq b$.*

Proof. $a \mid b$ implies that there is an integer k such that $a \cdot k = b$. Since $b \neq 0$ it follows that $k \neq 0$. So $|k| \geq 1$. It follows that

$$|a| = |a| \cdot 1 \leq |a||k| = |ak| = |b|,$$

as desired. □

6.4 Bezout's Lemma, and the $p \mid ab$ Lemma

Definition 11. Suppose m and n are integers. An **integer linear combination** of m and n is any integer of the form

$$mx + ny$$

where $x, y \in \mathbb{Z}$.

Example 4. 1 is an integer linear combination of 13 and 9, since

$$1 = 91 - 90 = 13(7) - 9(10).$$

Given non-zero integers m and n, let

$$S = \{mx + ny \mid x, y \in \mathbb{Z}, mx + ny > 0\}.$$

In English, S is all the positive integer linear combinations. Since $S \subseteq \mathbb{N}$, it has a least element ℓ. Bezout's Lemma states that $\ell = \gcd(m, n)$, a surprising result to take in at first.

Lemma 3. *Let m and n be non-zero integers, and let $g = \gcd(m, n)$. Then g is the smallest positive integer linear combination of m and n, that is, there exist some integers x_0, y_0 where*

$$g = mx_0 + ny_0,$$

and moreover g is the smallest number where such x_0 and y_0 exist.

Proof. The proof strategy is as follows: define ℓ as earlier, that is, as the least positive integer linear combination of m and n (that is, the smallest natural number where $mx_0 + ny_0 = \ell$ for some integer x_0 and y_0). We prove two intermediate results:

1. $\ell \mid n$ and $\ell \mid n$.
2. $g \mid \ell$

These two things imply $\ell = g$, for the first step tells us ℓ is a common divisor of m and n, and the second step tells us that the *greatest* common divisor divides ℓ, hence is less than or equal to ℓ. But g is the greatest common divisor, so $g = \ell$.

Let's prove these two steps. Suppose (by contradiction), $\ell \nmid m$. Then by the quotient remainder theorem there exist integers q and r where $0 < r < \ell$ such that $m = q\ell + r$. Then we have

$$r = m - \ell q = m - (mx_0 + ny_0)\ell = m(1 - x_0\ell) + n(b_0\ell).$$

Both expressions in the parentheses in the right hand side are integers by closure under addition and multiplication. So we have expressed $r > 0$, strictly less than ℓ, as an integer linear combination of m and n. But ℓ was defined as the **least** integer linear combination! This is a contradiction, so $\ell \mid m$. Similarly, $\ell \mid n$.

Let $g = \gcd(m, n)$. There exist $k_1, k_2 \in \mathbb{Z}$ such that $gk_1 = m$ and $gk_2 = n$. Then using the x_0 and y_0 as above we get

$$\ell = mx_0 + ny_0 = g(k_1x_0 + k_2y_0),$$

which shows that $g \mid \ell$, as desired. ☐

What immediately follows is a very nice lemma about a prime dividing a product.

Lemma 4. *Let p be a prime and a and b be integers. If $p \mid ab$ then $p \mid a$ or $p \mid b$.*

Proof. To prove this statement we need to show that the statement

$$(p \mid a) \vee (p \mid b).$$

is always true. To prove this or statement, there are two cases. Either $(p \mid a)$, which means the or statement is true. So there is nothing to do here. So suppose $p \nmid a$. We need to show that $(p \mid b)$ to show that the or statement is true.

If $p \nmid a$ then $\gcd(a, p) = 1$. So there exist integers x, y such that $ax + py = 1$. Multiplying this equation by b, we get $abx + pby = b$. But since $p \mid ab$ there exists an integer k such that $pk = ab$. So substituting this into the above equation we get

$$pkx + pby = p(kx + by) = b,$$

and since $kx + by$ is an integer it follows that $p \mid b$, as desired. ☐

6.5 Immediate applications to irrationality proofs

Using the $p \mid ab$ lemma we can immediately deduce the Euclidean lemma for any prime and any exponent, as highlighted below.

Lemma 5. *Suppose p is a prime and a is an integer with $p \mid a^2$. Then $p \mid a$.*

I'll give two different proofs of this fact, one using Bezout's Lemma and one using the $p \mid ab$ lemma.

Proof. By contrapositive. Suppose $p \nmid a$. Then $\gcd(p, a) = 1$, so there exist integers x and y such that $px + ay = 1$. Squaring this equation we get

$$p^2x^2 + 2pxay + a^2y^2 = p(px^2 + 2pxy) + a^2(y^2) = 1.$$

This shows that 1 is an integer linear combination of p and a^2. Since it is the smallest possible integer linear combination we can conclude that $\gcd(a^2, p) = 1$. That is, $p \nmid a^2$, as desired. ☐

Isn't that neat? No nightmare of indices. No dealing with $p - 1$ cases (especially as p gets very large). Just square the equation $ax + py = 1$ to get what you want. The second proof is even easier.

Proof. Suppose $p \mid a^2$. Then $p \mid a \cdot a$. So $p \mid a$ or $p \mid a$, which means that $p \mid a$, as desired. ☐

In fact, this proof is so nice that it leads the way to an easy proof by induction.

Lemma 6. *Suppose that p is a prime and a is an integer with $p \mid a^n$. Then $p \mid a$.*

Proof. We'll prove this by induction on n. For the case $n = 1$, clearly $p \mid a$ implies $p \mid a$. Now assume that for $n \geq 1$ that $p \mid a^n$ implies $p \mid a$. Suppose $p \mid a^{n+1}$. Then $p \mid a$ or $p \mid a^n$. Clearly if $p \mid a$ we would be done. Even if we suppose not, then $p \mid a^n$ and our inductive step gives us $p \mid a$ anyway. So the inductive step is complete and we are done. ☐

The next lemma shows that mod a prime, any non-zero integer has an inverse modulo p.

Lemma 7. *Suppose p is any prime, and k is not a multiple of p. Then there exists an integer ℓ such that*

$$k\ell \equiv 1 \pmod{p}.$$

Proof. Since k is not a multiple of p, it follows that $\gcd(k,p) = 1$. So there exist integers x and y such that $kx - py = 1$, or $kx - 1 = py$. This implies that $p \mid (kx-1)$, or $kx \equiv 1 \pmod{p}$, as desired. ☐

Example 5. We'll find all the solutions n to the equation $5n \equiv 3 \pmod{7}$.

By Lemma 7 there exists some k such that $5k \equiv 1 \pmod{7}$. After some searching we see that $k = 3$ is an integer such that $5k \equiv 1 \pmod{7}$. Hence

$$5kn \equiv 3k \pmod{7} \implies n \equiv 3(3) \equiv 2 \pmod{7}.$$

So all such n must be of the form $7k+2$ for some integer k. Conversely, if $n = 7k+2$, then

$$5n = 35k + 10 = 7(5k+1) + 3,$$

or $5n \equiv 3 \pmod{7}$. This shows that all solutions n to the equation are the integers $7k+2$ where k is **any** integer.

6.6 Exercises

1. Suppose m and n are integers. Then consider the set S of all k such that k is an integer linear combination of x and y. Let $g = \gcd(m,n)$.
    a) Suppose $g \mid k$. Prove that there are $x, y \in \mathbb{Z}$ such that $mx + ny = k$.
    b) Suppose $g \nmid k$. Prove (by contradiction) that there are no $x, y \in \mathbb{Z}$ such that $mx + ny = k$. Hence the only integer linear combinations of m and n are multiples of g.

2. Suppose that a and b are non-zero integers such that $\gcd(a,b) = 1$. Prove that there exists an integer k such that
$$ak \equiv 1 \pmod{b}.$$
(Hint: imitate the proof of Lemma 7.)

3. Suppose a, b, c, d are integers such that $ab \equiv ac \pmod{d}$. Then prove that $b \equiv c \pmod{\frac{d}{\gcd(a,d)}}$. Is the converse statement true?

4. Let $S \in \mathbb{Z}$ be a non-empty (this is important!) set such that
    - $(\forall x, y \in S)[x + y \in S]$ (Closure under addition)
    - $(\forall x \in S)[-x \in S]$. (Existence of additive inverse)

    Show that there exists $g \in S$ such that $g \mid k$ for all $k \in S$. Describe S in terms of g. (Hint: consider the smallest element ℓ of $S \cap \mathbb{N}$). Why did I specify that S be non-empty?

5. For each congruence equation, find/characterize all the n which make the equation true.
    a) $5x \equiv 4 \pmod{9}$
    b) $6x \equiv 4 \pmod{10}$
    c) $10n \equiv 0 \pmod{15}$
    d) $4n \equiv 6 \pmod{8}$

e) $57n \equiv 87 \pmod{105}$.

6. In this problem we outline a proof of the **Fundamental Theorem of Arithmetic**: the statement that every positive integer greater than 1 factors uniquely into primes.

   a) Use induction to show that for a prime p and integers $a_1, \ldots, a_n$ that if $p \mid a_1 \cdot \cdots \cdot a_n$, then $p \mid a_i$ for some $1 \leq i \leq n$.

   b) This fact implies unique factorization in the following sense: Suppose that n is an integer with two different factorizations. Take a prime p in one factorization. Then prove that p must occur in the other factorization. (For an elegant proof, use the well ordering principle to derive a contradiction.)

Chapter 7

Counting and Probability

> 7. I went to the Math Olympiad Summer Program (MOP) after my tenth grade year. The first classes started with Professor Rousseau writing 'Counting' on the board. *Good*, I thought. *I can count . . .* Within ten minutes I was thoroughly lost.
>
> — Richard Rusczyk

Combinatorics is the mathematics of enumeration. It focuses on counting things. In a typical combinatorics problem, there will be a setup of objects, and you are asked to count a new "special" object with certain properties.

The biggest underlying principle behind combinatorics is the **principle of multiplication**. To motivate this we will go through the basic toy examples below.

7.1 The Principle of Multiplication

Example 6. Suppose I have a diner menu. I have 3 choices for drink (milk, tea, coffee), 3 choices of toast (wheat, white, rye). How many ways can I choose a drink and toast combo?

Solution. For each choice of drink, there are 3 ways to choose a toast, and these 3 ways are unrestricted. there are 3 drinks, so it follows there must be $3+3+3=3\times 3=9$ ways to choose a drink and toast combo. ☐

This was the most elementary application of what is called the principle of multiplication. Here is another less elementary application:

Example 7. Suppose I want to bet on a horse race. There are 10 horses. How many ways can we fill the podium (that is, the first 3 places)? Assume there are no ties.

Solution. There are 10 horses which can come in first place. Then, whichever horse comes in first, there are only 9 horses which can come in second place (as the horse that won cannot have been second place at the same time). So there are $10\times 9=90$ ways in which the first two places can be filled. Similarly, after second place has been chosen there are only 8 horses that can come in third place. It follows that there are $90\times 8=720$ ways to fill the first 3 spots in the race. ☐

Let's illustrate what was different between the first and second examples. In the first example, the choice of the drink did not affect the number of toasts we could choose from after we had chosen the drink. However, in the horse race example, the first choice (choosing first place) limited the number of horses we had to choose from in the second place. However, the key idea is that the **number** of horses we are allowed to choose from in the second choice is always the same. So in the drink and toast example, there are three ways we could make the first choice (the drink), and no matter the first choice there are three ways we could make the second choice (the toast). In the horse race example, there are always 10 ways we could make the first choice, and no matter the first choice there are always 9 ways we could make the second choice (and so on). This observation makes the basis for the principle of multiplication:

Suppose we have to make s successive choices, and that we know that at step j there are n_j ways to make a choice. Then there are

$$\prod_{j=1}^{s} n_j$$

total ways to make all s choices.

7.2 Permutations

A general class of examples which follow from the principle of multiplication are the class of permutation problems, from which the horse race problem is an example. Recall that there were $10 \times 9 \times 8$ ways to choose a race podium for 10 horses, with no ties. We can rewrite this as follows:

$$10 \times 9 \times 8 = \frac{10 \times 9 \times 8 \times (7 \times \cdots \times 1)}{(7 \times \cdots \times 1)} = \frac{10!}{7!}.$$

This division of factorials will form the general formula for the general class of permutation problems. First we will define the permutation problem. Suppose that we have n distinct objects, and we need to choose s of these objects in an *ordered* manner (so which object we choose first/second/third matters). How many ways are there of doing this? We call such a selection of s objects a *permutation* of s objects from n.

By the logic of the horse race problem, there are n ways of choosing the first object, and then there are $(n-1)$ ways of choosing the second object, and so on. We do this s times, and on the sth choice there are $(n-s+1)$ ways of choosing the sth object. It follows by the principle of multiplication that there are

$$n(n-1)\dots(n-s+1) = \frac{n(n-1)\dots(n-s+1)(n-s)(n-s-1)(\dots)2\cdot 1}{(n-s)(n-s-1)(\dots)2\cdot 1} = \frac{n!}{(n-s)!}$$

ways of choosing s objects from n objects in an ordered manner.

This formula is important enough that we give it some convenient notation. We represent $P(n,k) = \frac{n!}{(n-k)!}$. It follows that $P(n,k)$ is the number of ways to choose k objects in an ordered manner from n objects.

When $k = n$, then the number $P(n,k)$ amounts to the number of ways to order n objects in a line. In this case, we have $P(n,n) = n!$. So there are $n!$ ways to order n objects in a line.

7.3 Combinations and the Principle of Division

Let's modify the permutation problem, so that we do not care in which order the s objects are chosen from the n objects. This amounts to the following problem: How many ways are there to choose a size s subset of objects from a size n set of objects? (Remember that order does not matter when listing a subset.)

In order to indicate the general solution we will solve a more specific example.

Example 8. Suppose I have 6 different colors of paint, and I must choose 4 of them to mix together (so that the order in which I choose the paint does not matter). How many ways are there for me to make such a choice?

solution. From the permutation problem, we have seen that there are $P(6,4) = 6!/2! = 360$ ways to choose 4 colors in order. However, this is not the answer to the problem, because we count certain combinations as different when they are really the same. For example, the ordered selection (red, blue, green, yellow) will be counted differently from the ordered selection (blue, green, red, yellow). This means for each distinct combination, we count it $4! = 24$ times (since this is the number of ways to arrange $n!$ in a line.

This means that the number of permutations is the number of combinations times 4!. So there are $\frac{1}{4!}P(6,4) = 15$ ways to choose 4 colors without regard to order. ☐

We can directly generalize this problem to solve the general combination problem. We call a collection of r objects chosen from a collection of n objects, where the order does not matter, a *combination* of r objects from n. From the permutation problem there are $P(n,r) = n!/(n-r)!$ ways to choose r objects in a certain order. For each *combination* of r objects there are $r!$ different orderings associated with it. Since each distinct combination gives rise to $r!$ different permutations, and each permutation can be realized as a combination, we conclude that there must

$$\frac{1}{r!}P(n,r) = \frac{n!}{r!(n-r)!}$$

distinct combinations of r objects from n objects. This number is important enough that we give it a special notation. We define the "binomial coefficient" $\binom{n}{r}$ (read "n choose r") to be the value $\frac{n!}{r!(n-r)!}$.

The method of solving the combination problem from the permutation problem is an important (often only implicitly discussed!) technique, which we call **the principle of division**.

Suppose I want to count the number a certain collection of objects S, and that I know the number of objects to a certain other collection of objects T with the following properties:

- For each element $s \in S$ we can associate it with exactly r objects in T.
- Each object in T can only be associated with a unique object in S.

Then it follows that the number of objects in the collection S is $|T|/r$, or the number of objects in T divided by r.

7.4 Probability in terms of Combinatorics

With knowledge of combinatorics, it is possible to give the answer to some basic questions about probability. Here is the basic idea:

Suppose that I have a *finite* (this is important!) collection S of objects, and I want to know the probability that I choose some object with a certain property from S. Let T denote the subcollection of objects with this property. Then the desired probability is $|T|/|S|$, or the size of T divided by the size of S.

Remark that this collection is finite is important, because implicitly this probability was attained under the uniform distribution, which does not make sense for infinite collections.

Chapter 8

The Pigeonhole Principle and its Generalizations

In this chapter we explore some consequences of a basic mathematical fact called the **Pigeonhole Principle**. Although the statement itself is very elementary, the various consequences and results that one can obtain as a consequence of this basic fact are very deep and surprising. In the first section we will explore the pigeonhole principle's application in the finite case and the infinite case.

8.1 The Pigeonhole Principle

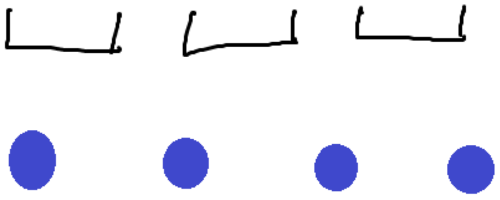
The most basic statement of the pigeonhole principle seems so obvious that it may not be apparent that it actually requires proof.

Proposition 1. *Suppose I have n pigeons and m holes, with* $n > m$*. No matter how I put the pigeons into the holes, I will have a hole that contains at least 2 pigeons.*

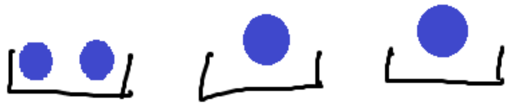
Proof. Suppose this is not true. Then there is a way to place the m pigeons such that each hole has less than or equal to one pigeon. Then the total number of pigeons would be less than or equal to $m \cdot 1 = m$ pigeons. But there are $n > m$ pigeons. This is a contradiction and the statement follows. ☐

We can more formally state the pigeonhole principle as follows: suppose $n > m$. Then for any function $f : \{1, \ldots, n\} \to \{1, \ldots, m\}$ there exists a k between 1 and m inclusive such that $|f^{-1}(k)| > 1$.

Example 9. Here is an illustration of the pigeonhole principle when $n = 4$ and $m = 3$. In the following diagram, the "holes" are the boxes and the "pigeons" are the blue balls.



Here is an example of balls being put into the boxes.



With this principle, it is possible to prove some facts that don't seem like they follow from this simple proposition.

Example 10. Suppose that we are at a party with n people. At this party, people shake hands with one another. Then no matter how they shake hands, at least two people shook hands with the same number of people. The reasoning is as follows. Consider the possible values for hands shaken for each individual. Normally, this ranges from 0 (no hands shaken) to $n-1$ (hands shaken with everyone else). If we consider these the holes, the pigeonhole principle does not immediately guarantee that 2 people will have shaken the same number of hands. However, we observe that if someone has shaken hands with $n-1$ people, they've shaken hands with everyone else, so that no one will have shaken hands with 0 people. The same logic applies once someone has shaken 0 hands. Hence we can use the pigeonhole principle to conclude that two people have shaken the same number of hands.

Chapter 9

Linear Equations and their Applications

Note: Chapter number is not fixed. Ramsey should come after Probability.

9.1 Linear Equations Reintroduced

Most readers who attended the equivalent of American middle or high school will recall having to solve systems of equations. A general form would be of two by two systems, which look like this:

$$\begin{aligned} ax + by &= z_1 \\ cx + dy &= z_2. \end{aligned}$$

In general, we will be dealing with much larger systems of linear equations where the number of variables to be solved for and the number of equations can vary.

Definition 12. An $m \times n$ system of linear equations is a system of equations of the form

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= y_1 \\ &\cdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= y_m. \end{aligned}$$

where the a_{ij} and the y_k are predetermined coefficients in any field (refer to ??). The variables $x_1, \ldots, x_n$ are referred to as **unknowns** or **unknown variables**.

In this chapter, what we will be most focused on are methods that that can determine solutions to these linear equations. That is, we would like to find values $(b_1, \ldots, b_n)$ such that all the equations in the definition above are true if we replace x_i with b_i for each i. For example, if we have the linear equation

$$\begin{aligned} ax + by &= c \\ dx + ey &= f \end{aligned}$$

then $x = g$, $y = h$ is a solution to this system of linear equations (In fact, it's the only solution).

We might first observe the following. Suppose that $(b_1, \ldots, b_n)$ is a solution to the $m \times n$ system of linear equations

$$\begin{gathered} a_{11}x_1 + \cdots + a_{1n}x_n = y_1 \\ \cdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = y_m. \end{gathered}$$

Then $(b_1, \ldots, b_n)$ is also a solution to the equation

$$\sum_{i=1}^{m} c_i(a_{i1}x_1 + \cdots + a_{in}x_n) = \sum_{i=1}^{m} c_i y_i$$

where the c_i are arbitrary coefficients. This is an example of a **linear combination** of the equations in our $m \times n$ system of linear equations. So if we had a second system of linear equations where each equation in the second system was a linear combination of the equations in the first system, then every solution of the first system would be a solution in the second system. This observation is important for proving the following, which is the basis for our methods to solve linear systems of equations:

Theorem 8. *Suppose A and B are two systems of linear equations which have the property that each equation in one system is a linear combination of equations in the other system. Then A and B have exactly the same solutions.*

9.2 Row Reduction

9.3 Determinant

9.4 Cramer's Rule

Motivation

Consider a system of linear equations in two variables, such as

$$\begin{aligned} y_1 &= ax_1 + bx_2 \\ y_2 &= cx_1 + dx_2. \end{aligned}$$

We would like to solve for x_1 and x_2 in terms of y_1 and y_2. Presumably, these solutions will have expressions involving a, b, c, and d. To solve this equation for x_1, we first eliminate the second variable by multiplying the first and second equations by appropriate constants, respectively:

$$\begin{aligned} dy_1 &= adx_1 + bdx_2 \\ by_2 &= bcx_1 + bdx_2. \end{aligned}$$

Subtracting the second equation from the first, and solving for x_1, we get

$$x_1 = \frac{dy_1 - by_2}{(ad - bc)}.$$

In a similar manner we can solve for x_2 as well.

Now consider an alternative perspective. We can rewrite the system of linear equations that we are considering as a matrix vector equation

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Then in the language of determinants, we find that

$$x_1 = \frac{\det \begin{bmatrix} y_1 & b \\ y_2 & d \end{bmatrix}}{\det \begin{bmatrix} a & b \\ c & d \end{bmatrix}}$$

and

$$x_2 = \frac{\det \begin{bmatrix} a & y_1 \\ c & y_2 \end{bmatrix}}{\det \begin{bmatrix} a & b \\ c & d \end{bmatrix}}.$$

This is the 2 dimensional case of *Cramer's Rule.*

Sums and Sequences

Bibliography

[1] Kenneth Hoffman and Ray Kunze. *Linear Algebra.* Pearson, 2nd edition, 1971.

[2] Iven Niven, Herbert Zuckerman, and Hugh Montgomery. *An Introduction to the Theory of Numbers.* John Wiley & Sons, Inc., 5th edition, 1991.

[3] Kenneth Rosen. *Discrete Mathematics and its Applications.* McGraw-Hill, 7th edition, 2012.

[4] Alan Slomson and R.B.J.T. Allenby. *How to Count: An Introduction to Combinatorics.* Chapman and Hall/CRC, 2nd edition, 2010.

[5] Michael Spivak. *Calculus.* Publish or Perish, 4th edition, 2008.

[6] Paul Zeitz. *The Art and Craft of Problem Solving.* John Wiley & Sons, Inc., 2nd edition, 2007.