Discrete Mathematics: An Open Introduction

Oscar Levin, Ph.D.

Fall 2015

Contents

Chapter 0

Introduction and Preliminaries

Welcome to Discrete Mathematics. If this is your first time encountering the subject, you will probably find discrete mathematics quite different from other math subjects. You might not even know what discrete math is! Hopefully this short introduction will shed some light on what the subject is about and what you can expect as you move forward in your study of it.

0.1 What is Discrete Mathematics?

> **dis·crete / dis'krēt**.
> *Adjective*: Individually separate and distinct.
> *Synonyms*: separate - detached - distinct - abstract.

Defining *discrete mathematics* is hard because defining *mathematics* is hard. What is mathematics? The study of numbers? In part, yes. But you also study functions and lines and triangles and parallelepipeds and vectors and .... Or perhaps you want to say that mathematics is a collection of tools that allow you to solve problems. What sort of problems? Okay, those that involve numbers, functions, lines, triangles,.... Whatever your conception of what mathematics is, try applying the concept of "discrete" to it, as defined above. Some math fundamentally deals with ... *stuff* ... that is individually separate and distinct.

In an algebra or calculus class, you might have found a particular set of numbers (maybe the set of number in the range of a function). You would represent this set as an interval: $[0, \infty)$ is the range of $f(x) = x^2$ since the set of outputs of the function are all real numbers 0 and greater. This set of numbers is **NOT** discrete. The numbers in the set are not separated by much at all - in fact, take any two numbers in the set and there are infinitely many more between them which are also in the set. Discrete math could still ask about the range of a function, but the set would not be an interval. Consider the function which gives the number of children each person reading this has. What is the range? I'm guessing it is something like $\{0, 1, 2, 3\}$. Maybe 4 is in there too. But certainly there is nobody reading this that has 1.32419 children. This set *is* discrete because the elements are separate. Also notice that the inputs to the function are a discrete set - each input is an individual person - you would not consider fractional inputs (there is nothing 2/3 between Bob and Carl we care about).

One way to get a feel for the subject is to consider the types of problems you solve in discrete math. Here are a few simple examples:

1. The most popular mathematician in the world is throwing a party for all of his friends. As a way to kick things off, they decide that everyone should shake hands. Assuming all 10 people at the party each shake hand with every other person (but not themselves, obviously) exactly once, how many handshakes take place?

2. At the warm-up event for Oscar's All Star Hot Dog Eating Contest, Al ate one hot dog. Bob then showed him up by eating three hot dogs. Not to be outdone, Carl ate five. This continued with each contestant eating two more hot dogs than the previous contestant. How

many hot dogs did Zeno (the 26th and final contestant) eat? How many hot dogs were eaten all together?

3. While walking through a fictional forest, you encounter three trolls. Each is either a *knight*, who always tells the truth, or a *knave*, who always lies. The trolls will not let you pass until you correctly identify each as either a knight or a knave. Each troll makes a single statement:

   Troll 1: If I am a knave then there are exactly two knights here.

   Troll 2: Troll 1 is lying.

   Troll 3: Either we are all knaves or at least one of us is a knight.

   Which troll is which?

4. Back in the days of yore, five small towns decided they wanted to build roads directly connecting each pair of towns. While the towns had plenty of money to build roads as long and as winding as they wished, it was very important that the roads not intersect with each other (as stop signs had not yet been invented). Also, tunnels and bridges were not allowed. Is it possible for each of these town to build a road to each of the four other towns without creating any intersections?

One reason it is difficult to define discrete math is that it is a very broad description which encapsulates a large number of subjects. In this course we will study four main topics: *combinatorics* (the theory of ways things *combine*, in particular, how to count these ways), *sequences*, *logic*, and *graph theory*. However, there are other topics that belong under the discrete umbrella, including computer science, abstract algebra, number theory, game theory, probability, and geometry (some of these, particularly the last two, have both discrete and non-discrete variants).

Ultimately the best way to learn what discrete math is about is to *do* it. Let's get started! Before we can begin answering more complicated (and fun) problems, we must lay down some foundation. We start by reviewing sets and functions in the framework of discrete mathematics.

0.2 Sets

The most fundamental objects we will use in our studies (and really in all of math) are *sets*. Much of what follows might be review, but it is very important that you are fluent in the language of set theory. Most of the notation we use below is standard, although some might be a little different than what you have seen before.

For us, a set will simply be an unordered collection of objects. For example, we could consider the set of all students enrolled at UNC this semester. Or the set of natural numbers between 1 and 10 inclusive. In the first case, each student here is a element (or member) of the set, while Barack Obama, among many others, is not an element of the set. Also, the two example are of different sets. Two sets are equal exactly if they contain the exact same elements.

0.2.1 Notation

Because we will want to consider many examples, we should have some notation to make talking about sets easier. Consider,

$$A = \{1, 2, 3\}.$$

This is read, "A is the set containing the elements 1, 2 and 3." We use curly braces "{, }" to enclose elements of a set. Some more notation:

$$a \in \{a, b, c\}.$$

The symbol "$\in$" is read "is in" or "is an element of." Thus the above means that a is an element of the set containing the letters a, b, and c. Note that this is a true statement. It would also be true to say that d is not in that set:

$$d \notin \{a, b, c\}.$$

Be warned: we say "$x \in A$" when we wish to express that "one of the elements of the set A is x." For example, consider the set,

$$A = \{1, b, \{x, y, z\}, \emptyset\}$$

This is a strange set, to be sure. It contains four elements: the number 1, the letter b, the set $\{x, y, z\}$ and the empty set ($\emptyset = \{\}$, the set containing no elements). Is x in A? The answer is no. None of the four elements in A are the letter x, so we must conclude that $x \notin A$. Similarly, if we considered the set $B = \{1, b\}$, then again $B \notin A$ - even though the elements of B are also elements of A, we cannot say that the thing B is one of the things in the collection A.

If a set is *finite*, then we can describe it by simply listing the elements. Infinite sets exists though, so we need to be able to describe them as well. For instance, if we want A to be the set of all even natural numbers, would could write,

$$A = \{0, 2, 4, 6, \ldots\}$$

but this is a little imprecise. Better would be

$$A = \{x \in \mathbb{N} : \exists n \in \mathbb{N}(x = 2n)\}.$$

Breaking that down: $x \in \mathbb{N}$ means x is in the set $\mathbb{N}$ (the set of natural numbers, starting with 0), : is read "such that" and $\exists n(x = 2n)$ is read "there exists an n in the natural numbers for which x is two times n" (in other words, x is even). Slightly easier might be,

$$A = \{x : \ x \text{ is even } \}.$$

Note: sometime people use | or ϶ for the "such that" symbol instead of :.

Defining a set using this sort of notation is very useful, although it takes some practice to read them correctly. It is a way to describe the set of all things that satisfy some condition (the condition is the logical statement after the ":" symbol). Here are some more examples. We use the logical symbols $\wedge$ for "and" and $\vee$ for "or" (which always includes the "or both" for us).

> **Example**: Describe each of the following sets both in words and by listing out enough elements to see the pattern.
>
> 1. $\{x : x + 3 \in \mathbb{N}\}$
> 2. $\{x \in \mathbb{N} : x + 3 \in \mathbb{N}\}$
> 3. $\{x : x \in \mathbb{N} \vee -x \in \mathbb{N}\}$
> 4. $\{x : x \in \mathbb{N} \wedge -x \in \mathbb{N}\}$
>
> *Solution*:

1. This is the set of all number which are 3 less than a natural number (i.e., that if you add 3 to them, you get a natural number). The set could also be written as $\{-3, -2, -1, 0, 1, 2, \ldots\}$ (note that 0 is a natural number, so -3 is in this set because $-3 + 3 = 0$).
2. This is the set of all natural numbers which are 3 less than a natural number. So here we just have $\{0, 1, 2, 3 \ldots\}$. In the previous set, we never specified where x could come from. That was okay since the condition forced our hand.
3. This is the set of all integers (positive or negative whole numbers, written $\mathbb{Z}$). In other words, $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$.
4. Now we want all numbers x such that x and $-x$ are natural numbers. There is only one: 0. So we have the set $\{0\}$.

We already have a lot of notation, and there is more yet. Below is a handy chart of symbols. Some of these will be discussed in greater detail as we move forward.

Special sets

$\emptyset$	The *empty set* is the set which contains no elements.
$\mathcal{U}$	The *universe set* is the set of all elements.
$\mathbb{N}$	The set of natural numbers. That is, $\mathbb{N} = \{0, 1, 2, 3 \ldots\}$
$\mathbb{Z}$	The set of integers. $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, 3, \ldots\}$
$\mathbb{Q}$	The set of rational numbers.
$\mathbb{R}$	The set of real numbers.
$\mathcal{P}(A)$	The *power set* of any set A is the set of all subsets of A.

Set Theory Notation

Symbol:	Read:	Example:
$\{, \}$	braces	$\{1, 2, 3\}$. The braces enclose the elements of a set. This is the set which contains the numbers 1, 2 and 3.
$:$	such that	$\{x : x > 2\}$ is the set of all x such that x is greater than 2.
$\in$	is an element of	$2 \in \{1, 2, 3\}$ asserts that 2 is one of the elements in the set $\{1, 2, 3\}$. However, $4 \notin \{1, 2, 3\}$.
$\subseteq$	is a subset of	$A \subseteq B$ asserts that every element of A is also an element of B.
$\subset$	is a proper subset of	$A \subset B$ asserts that every element of A is also an element of B, but $A \neq B$.
$\cap$	intersection	$A \cap B$ is the *set* of all elements which are elements of both A and B.
$\cup$	union	$A \cup B$ is the *set* of all elements which are elements of A or B or both.
$\times$	cross, or Cartesian product	$A \times B$ is the set of all ordered pairs (a, b) with $a \in A$ and $b \in B$.
$\setminus$	set difference	$A \setminus B$ is the *set* of all elements of A which are not elements of B.
$\overline{A}$	compliment (of A)	$\overline{A}$ is the set of everything which is not an element of A. The A can be any set here.
$\lvert A \rvert$	cardinality (of A)	$\lvert\{4, 5, 6\}\rvert = 3$ because there are 3 elements in the set. Sometimes we say $\lvert A \rvert$ is the *size* of A.
Logic symbols:		
$\wedge$	and	$x \in A \wedge x \notin B$ means x is both in the set A **and** also not in B.
$\vee$	or	$x \in A \vee x \notin B$ asserts that x is an element of A **or** not an element of B, or both.
$\neg$	not	Another way to write $x \notin A$ is $\neg x \in A$.
$\forall$	for all	$\forall x(x \geq 0)$ claims that for every number is greater than 0.
$\exists$	there exists	$\exists x(x < 0)$ claims that there are negative numbers (there exists a number such that it is less than 0).

0.2.2 Relationships between sets

We have already said what it means for two sets to be equal: they have exactly the same elements. Thus, for example,

$$\{1, 2, 3\} = \{2, 1, 3\}.$$

(Remember, the order the elements are written down in does not matter.) Also,

$$\{1, 2, 3\} = \{I, II, III\}.$$

Now what about the sets $A = \{1, 2, 3\}$ and $B = \{1, 2, 3, 4\}$? Clearly $A \neq B$. However, we can notice that every element of A is also an element of B. Because of this, we say that A is a subset of B, or in symbols $A \subset B$ or $A \subseteq B$. (Both symbols are read "is a subset of." The difference is that sometimes we want to say that A is either equal to or a subset of B, in which case we use $\subseteq$. Compare the difference between $<$ and $\leq$.)

Example: Let $A = \{1, 2, 3, 4, 5, 6\}$, $B = \{2, 4, 6\}$, $C = \{1, 2, 3\}$ and $D = \{7, 8, 9\}$. Determine which of the following are true, false, or meaningless.

1. $A \subset B$
2. $B \subset A$
3. $B \in C$
4. $\emptyset \in A$
5. $\emptyset \subset A$
6. $A < D$
7. $3 \in C$
8. $3 \subset C$.
9. $\{3\} \subset C$

Solution:

1. False.
2. True: every element in B is an element in A.
3. False: the elements in C are 1, 2, and 3. The *set* B is not equal to 1, 2, or 3.
4. False: A has exactly 6 elements, and none of them are the empty set.
5. True: Everything in the empty set (nothing) is also an element of A. Notice that the empty set is a subset of every set.
6. Meaningless. A set cannot be less that another set.
7. True.
8. Meaningless. 3 is not a set, so it cannot be a subset of another set.
9. True. 3 is the only element of the set $\{3\}$, and is an element of C, so every element in $\{3\}$ is an element of C.

In the example above, B is a subset of A. You might wonder what other sets are subsets of A. If you collect all these subsets of A, they themselves for a set - a set of sets. We call the set of all subsets of A the *power set* of A, and write it $\mathcal{P}(A)$.

Example: Let $A = \{1, 2, 3\}$. Find $\mathcal{P}(A)$.

Solution: $\mathcal{P}(A)$ is a set of sets - all of which are subsets of A. So

$$\mathcal{P}(A) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$$

Notice that while $2 \in A$, it is wrong to write $2 \in \mathcal{P}(A)$ - none of the elements in $\mathcal{P}(A)$ are numbers!. On the other hand we do have $\{2\} \in \mathcal{P}(A)$ because $\{2\} \subseteq A$.

What does a subset of $\mathcal{P}(A)$ look like? Notice that $\{2\} \not\subseteq \mathcal{P}(A)$ because not everything in $\{2\}$ is in $\mathcal{P}(A)$. But we do have $\{\{2\}\} \subseteq \mathcal{P}(A)$. The only element of $\{\{2\}\}$ is the set $\{2\}$ which is also an element of $\mathcal{P}(A)$. We could take the collection of all subsets of $\mathcal{P}(A)$ and call that $\mathcal{P}(\mathcal{P}(A))$. Or even the power set of that set of sets of sets.

Another way to compare sets is by their size. Notice that in the example above, A has 6 elements, B, C, and D all have 3 elements. The size of a set is called the set's cardinality. We would write $|A| = 6$, $|B| = 3$ and so on. For sets that have a finite number of elements, the cardinality of the set is simply the number of elements in the set. Note that the cardinality of $\{1, 2, 3, 2, 1\}$ is 3 – we do not count repeats (in fact, $\{1, 2, 3, 2, 1\}$ is exactly the same set as $\{1, 2, 3\}$). There are sets with infinite cardinality, such as $\mathbb{N}$, the set of rational numbers (written $\mathbb{Q}$), the set of even natural numbers, the set of real number ($\mathbb{R}$). It is possible to distinguish between different infinite cardinalities, but that is beyond the scope of these notes. For us, a set will either be infinite, or finite, and if it is finite, we can determine it's cardinality by counting elements.

Example:

1. Find the cardinality of $A = \{23, 24, \ldots, 37, 38\}$.
2. Find the cardinality of $B = \{1, \{2, 3, 4\}, \emptyset\}$.
3. If $C = \{1, 2, 3\}$, what is the cardinality of $\mathcal{P}(C)$?

Solution:

1. Since $38 - 23 = 15$, we can conclude that the cardinality of the set is $|A| = 16$ (you need to add one since 23 is included).
2. Here $|B| = 3$. The three elements are the number 1, the set $\{2, 3, 4\}$, and the empty set.
3. We wrote out the elements of the power set $\mathcal{P}(C)$ above, and there are 8 elements (each of which is a set). So $|\mathcal{P}(C)| = 8$.[1]

0.2.3 Operations on sets

Is it possible to add two sets? Not really, however there is something similar. If we want to combine two sets – to get the collection of objects that are in either set, then we can take the *union* of the two sets. Symbolically,

$$C = A \cup B$$

means C is the union of A and B. Every element of C is either an element of A or an element of B (or an element of both). For example, if $A = \{1, 2, 3\}$ and $B = \{2, 3, 4\}$, then $A \cup B = \{1, 2, 3, 4\}$.

The other common operation on sets is *intersection*. We write,

$$C = A \cap B$$

to mean that C is the intersection of A and B; everything in C is in both A and in B. So if $A = \{1, 2, 3\}$ and $B = \{2, 3, 4\}$, then $A \cap B = \{2, 3\}$.

Often when dealing with sets, we will have some understanding as to what "everything" is. Perhaps we are only concerned with natural numbers. We would say that our *universe* is $\mathbb{N}$.

[1] You might wonder if there is a relationship between $|A|$ and $|\mathcal{P}(A)|$ for all sets A. This is a good question which we will come back to soon.

Sometimes we call denote this universe by $\mathcal{U}$. Given this context, we might wish to speak of all the elements which are *not* in a particular set. We call this the *compliment* of the set, and write,

$$B = \overline{A}$$

when B contains every element not contained in A. So if our universe is $\{1, 2, \ldots, 9, 10\}$, and $A = \{2, 3, 5, 7\}$, then $\overline{A} = \{1, 4, 6, 8, 9, 10\}$.

Of course we can perform more than one operation at a time. Fore example, consider

$$A \cap \overline{B}$$

This is the set of all element which are both elements of A and not elements of B. What have we done? We've started with A and removed all of the elements which were in B. Another way to write this is the *set difference*:

$$A \cap \overline{B} = A \setminus B$$

It is important to remember that these operations (union, intersection, compliment and difference) on sets produce other sets. Don't confuse these with the symbols from the previous section (element of and subset of). $A \cap B$ is a set, while $A \subseteq B$ is true or false. This is the same difference as between $3 + 2$ (which is a number) and $3 \leq 2$ (which is in this case false).

Example: Let $A = \{1, 2, 3, 4, 5, 6\}$, $B = \{2, 4, 6\}$, $C = \{1, 2, 3\}$ and $D = \{7, 8, 9\}$. The universe is $\mathcal{U} = \{1, 2, \ldots, 10\}$. Find:

1. $A \cup B$
2. $A \cap B$
3. $B \cap C$
4. $A \cap D$
5. $\overline{B \cup C}$
6. $A \cap \overline{B}$
7. $(D \cap \overline{C}) \cup \overline{A \cap B}$
8. $\emptyset \cup C$
9. $\emptyset \cap C$

Solution:

1. $A \cup B = \{1, 2, 3, 4, 5, 6\} = A$ since everything in B is already in A.
2. $A \cap B = \{2, 4, 6\} = B$ since everything in B is in A.
3. $B \cap C = \{2\}$ - the only element of both B and C is 2.
4. $A \cap D = \emptyset$ since A and D have no common elements
5. $\overline{B \cup C} = \{5, 7, 8, 9, 10\}$. First we find that $B \cup C = \{1, 2, 3, 4, 6\}$, then we take everything not in that set.
6. $A \cap \overline{B} = \{1, 3, 5\}$. Everything that is in A which is not in B. This is the same as $A \setminus B$.
7. $(D \cap \overline{C}) \cup \overline{A \cap B} = \{1, 3, 5, 7, 8, 9\}$. The set contains all elements that are either in D but not in C or not in both A and B.
8. $\emptyset \cup C = C$ - nothing is added by the emptyset.
9. $\emptyset \cap C = \emptyset$ - nothing can be both in a set and in the emptyset.

You might notice that the symbols for union and intersection slightly resemble the logic symbols for "or" and "and." This is no accident. What does it mean for x to be an element of $A \cup B$? It means that x is an element of A or x is an element of B (or both). That is,

$$x \in A \cup B \qquad \Leftrightarrow \qquad x \in A \vee x \in B.$$

Similarly,

$$x \in A \cap B \qquad \Leftrightarrow \qquad x \in A \wedge x \in B.$$

Also,

$$x \in \overline{A} \qquad \Leftrightarrow \qquad \neg(x \in A)$$

which says x is an element of the compliment of A if x is not an element of A.

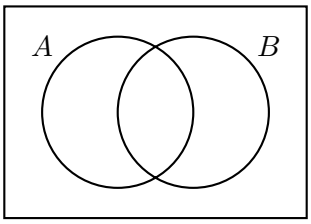
There is one more way to combine sets which will be useful for us: the Cartesian product. This sounds fancy but is nothing you haven't seen before. When you graph a function in calculus, you graph it in the Cartesian plane. This is the set of all ordered pairs of real numbers (x, y). We can do this for *any* pair of sets, not just the real numbers with themselves.

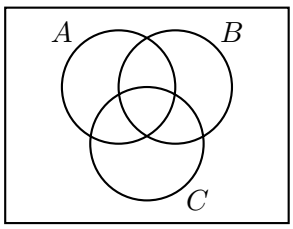
Put another way, $A \times B = \{(a, b) : a \in A \wedge b \in B\}$. The first coordinate comes from the first set, the second coordinate comes from the second set. Sometimes we will want to take the Cartesian product of a set with itself, and this is fine: $A \times A = \{(a, b) : a, b \in A\}$ (we might also write A^2 for this set). Notice that in $A \times A$, we still want *all* ordered pairs, not just the ones where the first and second coordinate are the same. We can also take products of 3 or more sets, getting ordered triples, or quadruples, and so on.

> **Example**: Let $A = \{1, 2\}$ and $B = \{3, 4, 5\}$. Find $A \times B$ and $A \times A$. How many elements do you expect to be in $B \times B$?
>
> *Solution*: $A \times B = \{(1, 3), (1, 4), (1, 5), (2, 3), (2, 4), (2, 5)\}$.
> $A \times A = A^2 = \{(1, 1), (1, 2), (2, 1), (2, 2)\}$.
> $|B \times B| = 9$. There will be 3 pairs with first coordinate 3, three more with first coordinate 4 and a final three with first coordinate 5.

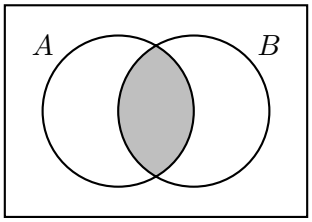
0.2.4 Venn Diagrams

Union, intersection, compliment, set difference - operations can get complicated (see part 7 in the above example). Luckily, there is a very nice visual tool we can use to clarify things. Venn diagrams represent sets as intersecting circles. We can shade the region we are talking about when we carry out an operation. We can also represent cardinality of a particular set by putting the number in the corresponding region.

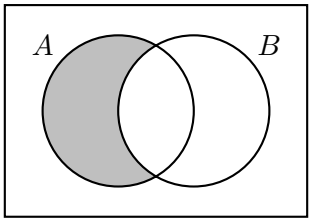




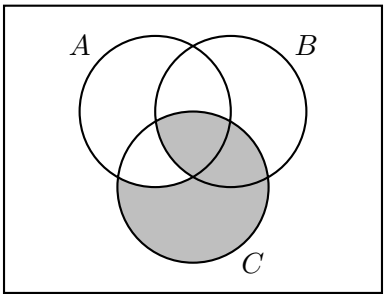
Each circle represents a set. The rectangle containing the circles represents the universe. To represent combinations of these sets, we shade the corresponding region. For example, we could draw $A \cap B$ as:



Here is a representation of $A \cap \overline{B}$, or equivalently $A \setminus B$:



A more complicated example is $(B \cap C) \cup (C \cap \overline{A})$, as seen below.



Notice that the shaded regions above could also be arrived at in another way. We could have started with all of C, then excluded the region where C and A overlap (without B). That region is $(A \cap B) \cap \overline{B}$. So the above Venn diagram also represents $C \cap \left(\overline{(A \cap B) \cap \overline{B}}\right)$. So using just the picture, we have determined that

$$(B \cap C) \cup (C \cap \overline{A}) = C \cap \left(\overline{(A \cap B) \cap \overline{B}}\right).$$

Exercises

1. Let $A = \{1, 2, 3, 4, 5\}$, $B = \{3, 4, 5, 6, 7\}$ and $C = \{2, 3, 5\}$.
   (a) Find $A \cap B$.
   (b) Find $A \cup B$.
   (c) Find $A \setminus B$.
   (d) Find $A \times C$.
   (e) Is $C \subseteq A$?
   (f) Is $C \subseteq B$?
2. Let $A = \{x \in \mathbb{N} : 3 \leq x \leq 13\}$, $B = \{x \in \mathbb{N} : x \text{ is even}\}$, and $C = \{x \in \mathbb{N} : x \text{ is odd}\}$.
   (a) Find $A \cap B$.
   (b) Find $A \cup B$.
   (c) Find $B \cap C$.
   (d) Find $B \cup C$.
3. Find an example of sets A and B such that $A \cap B = \{3, 5\}$ and $A \cup B = \{2, 3, 5, 7, 8\}$.

4. Find an example of sets A and B such that $A \subseteq B$ and $A \in B$.

5. Recall $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ (the integers). Let $\mathbb{Z}^+$ be the positive integers. Let $2\mathbb{Z}$ be the even integers, $3\mathbb{Z}$ be the multiples of 3, and so on.

   (a) Is $\mathbb{Z}^+ \subseteq 2\mathbb{Z}$?

   (b) Is $2\mathbb{Z} \subseteq \mathbb{Z}^+$?

   (c) Find $2\mathbb{Z} \cap 3\mathbb{Z}$. Describe the set in words, and also in symbols (using a : symbol).

   (d) Express $\{x \in \mathbb{Z} : \exists y \in \mathbb{Z}(x = 2y \vee x = 3y)\}$ as a union or intersection of two sets above.

6. Let A_2 be the set of all multiples of 2 except for 2. Let A_3 be the set of all multiples of 3 except for 3. And so on, so that A_n is the set of all multiple of n except for n, for any $n \geq 2$. Describe (in words) the set $\overline{A_2 \cup A_3 \cup A_4 \cup \cdots}$

7. Draw a Venn diagram to represent each of the following:

   (a) $A \cup \overline{B}$

   (b) $\overline{(A \cup B)}$

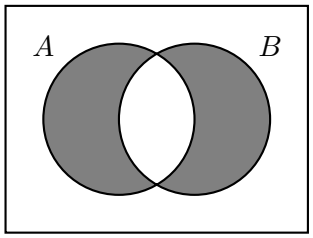
   (c) $A \cap (B \cup C)$

   (d) $(A \cap B) \cup C$

   (e) $\overline{A} \cap B \cap \overline{C}$

   (f) $(A \cup B) \setminus C$

8. Describe a set in terms of A and B which has the following Venn diagram:



9. Find the cardinalities:

   (a) Find $|A|$ when $A = \{4, 5, 6, \ldots, 37\}$

   (b) Find $|A|$ when $A = \{x \in \mathbb{Z} : -2 \leq x \leq 100\}$

   (c) Find $|A \cap B|$ when $A = \{x \in \mathbb{N} : x \leq 20\}$ and $B = \{x \in \mathbb{N} : x \text{ is prime}\}$

10. Let $A = \{a, b, c\}$. Find $\mathcal{P}(A)$.

11. Let $A = \{1, 2, \ldots, 10\}$. How many subsets of A contain exactly one element (i.e., how many *singleton* subsets are there). How many *doubleton* (containing exactly two elements) are there?

12. Let $A = \{1, 2, 3, 4, 5, 6\}$. Find all sets $B \in \mathcal{P}(A)$ which have the property $\{2, 3, 5\} \subseteq B$.

13. Find an example of sets A and B such that $|A| = 4$, $|B| = 5$ and $|A \cup B| = 9$.
14. Find an example of sets A and B such that $|A| = 3$, $|B| = 4$ and $|A \cup B| = 5$.
15. Are there sets A and B such that $|A| = |B|$, $|A \cup B| = 10$ and $|A \cap B| = 5$? Explain.
16. In a regular deck of playing cards there are 26 red cards and 12 face cards. Explain in terms of sets why there are only 32 cards which are either red or a face card.

0.3 Functions

A function is simply a rule that assigns each input exactly one output. The set of all inputs for a function is called the *domain*. The set of all allowable outputs is called the *codomain*. For example, a function might assign each natural number a natural number from 1 to 5. In that case, the domain is the natural numbers and the codomain is the set of natural numbers from 1 to 5. Now it could be that this particular function we are thinking about assigns each even natural number to the number 2 and each odd natural number to the number 1. In this case, not all of the codomain is actually used. We would say that the set $\{1, 2\}$ is the *range* of the function - these are the element in the codomain (allowable outputs) which are actually outputs for some input.

The key thing that makes a rule assigning inputs to outputs a *function* is that there is *only one* output for an input. In other words, it is important that the rule be a good rule. What output do we assign to the input 7? There can only be one answer for a particular function - otherwise where does 7 go to?

To specify the name of the function, as well as the domain and codomain, we write $f : X \to Y$. The function is called f, the domain is the set X and the codomain is the set Y. This however does not describe the rule. To do that, we say something like this:

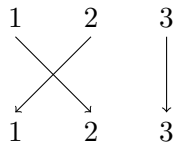
The function $f : X \to Y$ is defined by $f(x) = x^2 + 3$.

This function takes an input x and computes the output by squaring x and then adding 3. In this case, you better hope that X is a set of numbers and Y is a set of number which can be 3 more than squares of numbers from X. It would not work for Y to be negative numbers here - that would not be a valid function.

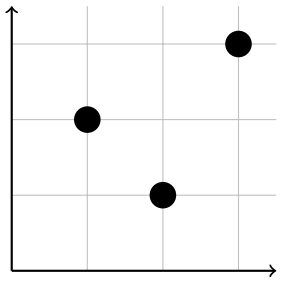
The description of the rule can vary greatly. If X is a finite set, we might just give a list of each output for each input. You could also describe the function with a table or a graph or in words.

Example: The following are all examples of functions:

1. $f : \mathbb{Z} \to \mathbb{Z}$ defined by $f(n) = 3n$. The domain and codomain are both the set of integers. However, the range is only the set of integer multiples of 3.
2. $f : \{1, 2, 3\} \to \{a, b, c\}$ defined by $f(1) = c$, $f(2) = a$ and $f(3) = a$. The domain is the set $\{1, 2, 3\}$, the codomain is the set $\{a, b, c\}$ and the range is the set $\{a, c\}$. Note that $f(2)$ and $f(3)$ are the same element of codomain. This is not a problem - each element in the domain still has only one output (although each output does not have a unique input).
3. $f : \{1, 2, 3\} \to \{1, 2, 3\}$ defined as follows:



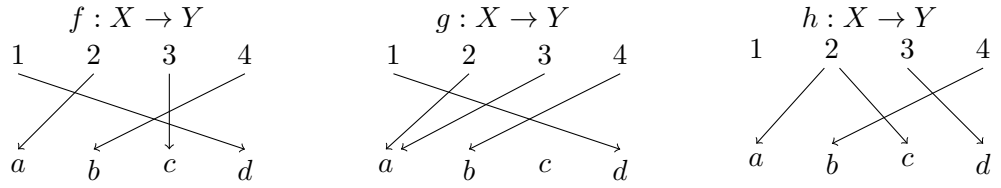
The arrow diagram used to define the function above can be very helpful in visualizing functions. We will often be working with functions on finite sets so this kind of picture is often more useful than a traditional graph of a function. A graph of the function in example 3 above would look like this:



It would absolutely be WRONG to connect the dots or try to fit them to some curve. There are only three elements in the domain. A curve suggests that the domain contains an entire interval of real numbers. Remember, we are not in calculus any more!

It is important to know how to recognize a function from a rule which is not a function. The arrow diagram can help.

Example: Which of the following diagrams represent a function. Let $X = \{1, 2, 3, 4\}$ and $Y = \{a, b, c, d\}$



Solution: f is a function. So is g. There is no problem with an element of the codomain not being the output for any input, and there is no problem with a from the codomain being the output of both 2 and 3 from the input.

However, h is **not** a function - in fact, for two reasons. First, the element 1 from the domain has not been mapped to any element from the codomain. Second, the element 2 from the domain has been mapped to more than one element from the codomain (a and c). Note that either one of these problems is enough to make a rule not a function. Neither of these mappings are functions:

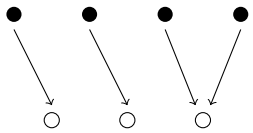
Not functions.

0.3.1 Surjections, Injections, and Bijections

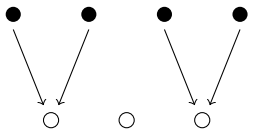
We now turn to investigating special properties functions might or might not possess.

In the examples above, you may have noticed that sometimes there are elements of the codomain which are not in the range - which are not actual outputs for any input. When this sort of the thing *does not* happen, (that is, when everything in the codomain is in the range) we say the function is *onto* or that the function maps the domain *onto* the codomain. This terminology should make sense: the function puts the domain (entirely) on top the codomain. The fancy math term for an onto function is a *surjection*, and we say that an onto function is a *surjective* function.

In pictures:



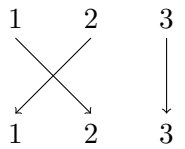
Surjective



Not surjective

Example: Which functions are surjective (i.e., onto)?

1. $f : \mathbb{Z} \to \mathbb{Z}$ defined by $f(n) = 3n$.
2. $g : \{1, 2, 3\} \to \{a, b, c\}$ defined by $g(1) = c$, $g(2) = a$ and $g(3) = a$.
3. $h : \{1, 2, 3\} \to \{1, 2, 3\}$ defined as follows:

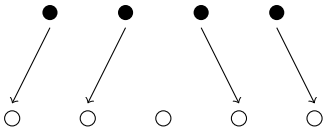


Solution:

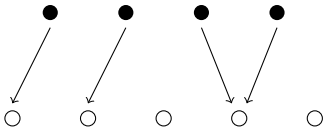
1. f is not surjective. There are elements in the codomain which are not in the range. For example, no $n \in \mathbb{Z}$ gets mapped to the number 1 (the rule would say that $\frac{1}{3}$ would be sent to 1, but $\frac{1}{3}$ is not in the domain). In fact, the range of the function is $3\mathbb{Z}$ (the integer multiples of 3), which is not equal to $\mathbb{Z}$.
2. g is not surjective. There is no $x \in \{1, 2, 3\}$ (the domain) for which $g(x) = b$. So b, which is in the codomain, is not in the range, so once again the function is not onto.
3. h is surjective. Every element of the codomain is also in the range. Nothing is missed.

To be a function, a map cannot assign a single element of the domain to two or more different elements of the codomain. However, we have seen that the reverse is permissible. That is, a function might assign the same element of the codomain to two or more different elements of the domain. When this *does not* occur - that is, when each element of the codomain is assigned to at most one element of the domain - then we say the function is *one-to-one*. Again, this terminology makes sense: we are sending at most one element from the domain to one element from the codomain. One input to one output. The fancy math term for a one-to-one function is a *injection*. We call one-to-one functions *injective* functions.

In pictures:



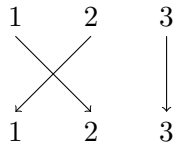
Injective



Not injective

Example: Which functions are injective (i.e., one-to-one)?

1. $f : \mathbb{Z} \to \mathbb{Z}$ defined by $f(n) = 3n$.
2. $g : \{1, 2, 3\} \to \{a, b, c\}$ defined by $g(1) = c$, $g(2) = a$ and $g(3) = a$.
3. $h : \{1, 2, 3\} \to \{1, 2, 3\}$ defined as follows:



Solution:

1. f is injective. Each element in the codomain is assigned to at *most* one element from the domain - if x is a multiple of three, then only $x/3$ is mapped to x. If x is not a multiple of 3, then there is no input corresponding to the output x.
2. g is not injective. Both inputs 2 and 3 are assigned the output a.
3. h is injective. Each output is only an output once.

From the examples above, it should be clear that there are functions which are surjective, injective, both or neither. In the case when a function is both one-to-one and onto (a injection and surjection) we say the function is a *bijection*, or that the function is a *bijective* function.

0.3.2 Inverse Image

When discussing functions, we have notation for talking about an element of the domain (say x) and its corresponding element in the codomain (we write $f(x)$). It would also be nice to start with some element of the codomain (say y) and talk about which element or elements (if any) from the domain get sent to it. We could write "those x in the domain such that $f(x) = y$," but this is a lot of writing. So here is some notation to make our lives easier.

Suppose $f : X \to Y$ is a function. For $y \in Y$ (an element of the codomain), we write $f^{-1}(y)$ to represent the *set* of all elements in the domain X which get sent to y. That is, $f^{-1}(y) = \{x \in X : f(x) = y\}$.

WARNING: $f^{-1}(y)$ is not an inverse function!!!! Inverse functions only exist for bijections, but $f^{-1}(y)$ is defined for any function f. The point: $f^{-1}(y)$ is a <u>set</u>, not an element of the domain.

Example: Consider the function $f : \{1, 2, 3, 4, 5, 6\} \to \{a, b, c, d\}$ given by $f(1) = a$, $f(2) = a$, $f(3) = b$, $f(4) = c$, $f(5) = c$ and $f(6) = c$. Find the complete inverse image of each element in the codomain.

Solution: Remember, we are looking for sets. $f^{-1}(a) = \{1, 2\}$
$f^{-1}(b) = \{3\}$
$f^{-1}(c) = \{4, 5, 6\}$
$f^{-1}(d) = \emptyset$.

Example: Consider the function $g : \mathbb{Z} \to \mathbb{Z}$ defined by $g(n) = n^2 + 1$. Find $g^{-1}(1)$, $g^{-1}(2)$, $g^{-1}(3)$ and $g^{-1}(10)$.

Solution: To find $g^{-1}(1)$, we need to find all integers n such that $n^2 + 1 = 1$. Clearly only 0 works, so $g^{-1}(1) = \{0\}$ (note that even though there is only one element, we still write it as a set with one element in it).

To find $g^{-1}(2)$, we need to find all n such that $n^2+1 = 2$. We see $g^{-1}(2) = \{-1, 1\}$.

If $n^2 + 1 = 3$, then we are looking for an n such that $n^2 = 2$. There are no such integers so $g^{-1}(3) = \emptyset$.

Finally, $g^{-1}(10) = \{-3, 3\}$ because $g(-3) = 10$ and $g(3) = 10$.

Since $f^{-1}(y)$ is a set, it makes sense to ask for $|f^{-1}(y)|$ - the number of elements in the domain which map to y.

Example: Find a function $f : \{1, 2, 3, 4, 5\} \to \mathbb{N}$ such that $|f^{-1}(7)| = 5$.

Solution: There is only one such function. We need five elements of the domain to map to the number $7 \in \mathbb{N}$. Since there are only five elements in the domain, all of them must map to 7. So $f(1) = 7$, $f(2) = 7$, $f(3) = 7$, $f(4) = 7$ and $f(5) = 7$.

Function Definitions

- A *function* is a rule that assigns each element of a set, called the *domain*, to exactly one element of a second set, called the *codomain*.
- Notation: $f : X \to Y$ is our way of saying that the function is called f, the domain is the set X and the codomain is the set Y.
- $f(x) = y$ means the element x of the domain (input) is assigned to the element y of the codomain. We say y is an output. Alternatively, we call y the *image of x under f*.
- The *range* is a subset of the codomain. It is the set of all elements which are assigned to at least one element of the domain by the function. That is, the range is the set of all outputs.
- A function is *injective* (an *injection* or *one-to-one*) if every element of the codomain is the output for **at most** one element from the domain.
- A function is *surjective* (a *surjection* or *onto*) if every element of the codomain is the output of **at least** one element of the domain.
- A *bijection* is a function which is both an injection and surjection. In other words, if every element of the codomain is the output of **exactly one** element of the domain.

- The *complete inverse image* of an element in the codomain, written $f^{-1}(y)$ is the **set** of all element in the domain which are assigned to y by the function.

Exercises

1. Write out all functions $f : \{1, 2, 3\}$ to $\{a, b\}$. How many are there? How many are injective? How many are surjective? How many are both?

2. Write out all functions $f : \{1, 2\}$ to $\{a, b, c\}$. How many are there? How many are injective? How many are surjective? How many are both?

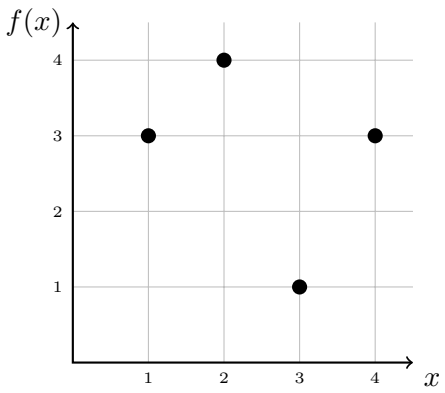
3. Consider the function $f : \{1, 2, 3, 4, 5\} \to \{1, 2, 3, 4\}$ given by the table below:

x	1	2	3	4	5
$f(x)$	3	2	4	1	2

   (a) Is f injective? Explain.

   (b) Is f surjective? Explain.

4. Consider the function $f : \{1, 2, 3, 4\} \to \{1, 2, 3, 4\}$ given by the graph below.



   (a) Is f injective? Explain.

   (b) Is f surjective? Explain.

5. For each function given below, determine whether or not the function is injective and whether or not the function is surjective.

   (a) $f : \mathbb{N} \to \mathbb{N}$ given by $f(n) = n + 4$.

   (b) $f : \mathbb{Z} \to \mathbb{Z}$ given by $f(n) = n + 4$.

   (c) $f : \mathbb{Z} \to \mathbb{Z}$ given by $f(n) = 5n - 8$.

   (d) $f : \mathbb{Z} \to \mathbb{Z}$ given by $f(n) = \begin{cases} n/2 & \text{if } n \text{ is even} \\ (n+1)/2 & \text{if } n \text{ is odd.} \end{cases}$

6. Let $A = \{1, 2, 3, \ldots, 10\}$. Consider the function $f : \mathcal{P}(A) \to \mathbb{N}$ given by $f(B) = |B|$. So f takes a subset of A as an input and outputs the cardinality of that set.

   (a) Is f injective? Prove your answer.

(b) Is f surjective? Prove your answer.

(c) Find $f^{-1}(1)$.

(d) Find $f^{-1}(0)$.

(e) Find $f^{-1}(12)$.

7. Let $A = \{n \in \mathbb{N} : 0 \leq n \leq 999\}$ be the set of all numbers with three or fewer digits. Define the function $f : A \to \mathbb{N}$ by $f(abc) = a + b + c$, where a, b, and c are the digits of the number in A. For example, $f(253) = 2 + 5 + 3 = 10$.

(a) Find $f^{-1}(3)$.

(b) Find $f^{-1}(28)$.

(c) Use one of the parts above to prove that f is not injective.

(d) Use one of the parts above to prove that f is not surjective.

8. Let $f : X \to Y$ be some function. Suppose $3 \in Y$. What can you say about $f^{-1}(3)$ if you know,

(a) f is injective? Explain.

(b) f is surjective? Explain.

(c) f is bijective? Explain.

9. Find a set X and a function $f : X \to \mathbb{N}$ so that $f^{-1}(0) \cup f^{-1}(1) = X$.

10. What can you deduce about the sets X and Y if you know,

(a) there is a injective function $f : X \to Y$? Explain.

(b) there is a surjective function $f : X \to Y$? Explain.

(c) there is a bijection $f : X \to Y$? Explain.

11. Suppose $f : X \to Y$ is a function. Which of the following are possible? Explain.

(a) f is injective but not surjective.

(b) f is surjective but not injective.

(c) $|X| = |Y|$ and f is injective but not surjective.

(d) $|X| = |Y|$ and f is surjective but not injective.

(e) $|X| = |Y|$, X and Y are finite, and f is injective but not surjective.

(f) $|X| = |Y|$, X and Y are finite, and f is surjective but not injective.

12. Consider the function $f : \mathbb{Z} \to \mathbb{Z}$ given by $f(n) = \begin{cases} n+1 & \text{if } n \text{ is even} \\ n-3 & \text{if } n \text{ is odd.} \end{cases}$

(a) Is f injective? Prove your answer.

(b) Is f surjective? Prove your answer.

13. At the end of the semester a teacher assigns letter grades to each of her students. Is this a function? If so, what sets make up the domain and codomain, and is the functions be injective, surjective, bijective, or neither?

14. In the game of *Hearts*, four players are each dealt 13 cards from a deck of 52. Is this a function? If so, what sets make up the domain and codomain, and is the function be injective, surjective, bijective, or neither?

15. Suppose 7 players are playing 5-card stud. Each player initially receives 5 cards from a deck of 52. Is this a function? If so, what sets make up the domain and codomain and is the function be injective, surjective, bijective, or neither?

Chapter 1

Counting

One of the first things you learn in mathematics is how to count. Now we are going to learn that all over again, and you will find that counting is a lot harder than you remember. The problem is that we want to count large collections of things quickly and precisely. For example:

- How many different lotto tickets are possible?
- In a group of 10 people, if everyone shakes hands with everyone else exactly once, how many handshakes took place?[1]
- How many ways can you distribute 10 girl scout cookies to 7 boy scouts?
- How many 10 digit numbers contain exactly 4 prime digits?
- How many functions $f : \{1, 2, 3, 4\} \to \{1, 2, 3\}$ are surjective?
- How many subsets of $\{1, 2, 3, \ldots, 10\}$ have cardinality 7?

Before tackling these difficult questions, let's look at the basics of counting.

1.1 Additive and Multiplicative Principles

Consider this rather simple counting problem: at Red Dogs and Donuts, there are 14 varieties of donuts, and 16 types of hot-dogs. If you want either a donut or a dog, how many options do you have? This is an easy question - you just add 14 and 16. Will that always work? What is important here?

Additive Principle
The *additive principle* states that if event A can occur in m ways, and event B can occur in n *disjoint* ways, then the event "A or B" can occur in $m + n$ ways.

It is important that the events be disjoint. For example, a standard deck of 52 cards contains 26 red cards and 12 face cards. However, the number of ways to select a card which is either red or a face card is not $26 + 12 = 38$. This is because there are 6 cards which are both red and face cards.

The additive principle works with more than two events. Say you would also consider eating one of 15 waffles? How many choices do you have now? You would have $14 + 16 + 15 = 45$ options.

> **Example**: How many two letter "words" start with either A or B? How many start with one of the 5 vowels? (A word is just a strings of letters - they don't have to be English words, or even pronounceable).

[1] We have solved this problem already. Soon we will see how this problem relates to more complicated counting problems.

Solution: First, how many two letter words start with A? We just need to select the second letter, which can be accomplished in 26 ways. So there are 26 words starting with A. There are also 26 words that start with B. So to select a word which starts with either A or B, we can pick the word from the first 26 or the second 26, for a total of 52 words. The additive principle is at work here.

Now what about all the two letter words starting with a vowel? Well there are 26 starting with A, another 26 starting with E, and so on. We will have 5 groups of 26. So we add 26 to itself 5 times. Of course it would be easier to just multiply $5 \cdot 26$ - we are really using the additive principle again, just using multiplication as a shortcut.

Example: Suppose you are going for some FroYo - you can pick one of 6 yogurt choices, and one of 4 toppings. How many choices do you have?

Solution: Break your choices up into disjoint events: A are the choices with the first topping, B the choices featuring the second topping, and so on. So we have events. Each can occur in 6 ways (one for each yogurt flavor). The events are disjoint, so the total number of choices is $6 + 6 + 6 + 6$.

Note that in both of the previous examples, when using the additive principle on a bunch of sets all the same size, it is quicker to multiply. This really is the same - not just because $6+6+6+6 = 4 \cdot 6$. We can first select the topping in 4 ways (that is we first select which of the disjoint events we will take). For each of those first 4 choices, we now have 6 choices of yogurt. We have:

Multiplicative Principle
The *multiplicative principle* states that if event A can occur in m ways, and each possibility for A allows for exactly n ways for event B, then the event "A and B" can occur in $m \cdot n$ ways.

The multiplicative principle generalizes to more than two events as well.

Example: How many license plates can you make out of three letters followed by three numerical digits?

Solution: Here we have six events: the first letter, the second letter, the third letter, the first digit, the second digit and the third digit. The first three events can each happen in 26 ways, the last three can each happen in 10 ways. So the total number of license plates will be $26 \cdot 26 \cdot 26 \cdot 10 \cdot 10 \cdot 10$, using the multiplicative principle.

Does this make sense? Think about how we would pick a license plate - how many choices we would have. First, we need to pick the first letter. There are 26 choices. Now for each of those, there are 26 choices for the second letter. So 26 second letters with first letter A, 26 second letters with first letter B, and so on. So we add 26 to itself 26 times. Or quicker: there are $26 \cdot 26$ choices for the first two letters.

Now for each choice of the first two letters, we have 26 choices for the third letter. That is, 26 third letters for the first two letters AA, 26 choices for the third letter after starting AB, and so on. There are $26 \cdot 26$ of these 26 third letter choices, for a total of $(26 \cdot 26) \cdot 26$ choices for the first three letters. And for each of these $26 \cdot 26 \cdot 26$ choices of letters, we have a bunch of choices for the remaining digits.

In fact, there are going to be exactly 1000 choices for the numbers. We can see this because there are 1000 three-digit numbers (000 through 999). This is 10 choices

for the first digit, 10 for the second, and 10 for the third. The multiplicative principle says we multiply: $10 \cdot 10 \cdot 10 = 1000$.

So there were 26^3 choices for the three letters, and 10^3 choices for the numbers, so to we have a total of $26^3 \cdot 10^3$ choices of license plates.

Careful: "and" doesn't mean "times." For example, how many playing cards are both red and a face card? Not $26 \cdot 12$! The answer is 6, and we needed to know something about cards to answer that question.

Another caution: how many ways can you select two cards, so that the first one is a red card and the second one is a face card? This looks more like the multiplicative principle (you are counting two separate events) but the answer is not $26 \cdot 12$ here either. The problem is that while there are 26 ways for the first card to be selected, it is not the case that *for each* of those there are 12 ways to select the second card. If the first card was both red and a face card then there would be only 11 choices for the second card. The moral of this story? The multiplicative principle only works if the events are independent.[2]

1.1.1 Counting with sets

Do you believe the additive and multiplicative principles? How would you convince someone they are correct? This is surprisingly difficult. They seem so simple, so obvious. But why do they work?

To make things clearer, and more mathematically rigorous, we will use sets. Do not skip this section! It might seem like we are just trying to give a proof of these principles, but we are doing a lot more. If we understand the additive and multiplicative principles rigorously, we will be better at applying them, and knowing when and when not to apply them at all.

We will look at the previous problems in a slightly different way. Instead of thinking about event A and event B, we want to think of a set A and a set B. The sets will contain all the different ways the event can happen. (It will be helpful to be able to switch back and forth between these two models when checking that we have counted correctly.) Here's what we mean:

Example: Suppose you own 9 shirts and 5 pairs of pants.

1. How many outfits can you make?
2. If today is half-naked-day, and you will wear only a shirt or only a pair of pants, how many choices do you have?

Solution: By now you should agree that the answer to the first question is $9 \cdot 5 = 45$ and the answer to the second question is $9 + 5 = 14$. These are the multiplicative and additive principles. There are two events: picking a shirt and picking a pair of pants. The first event can happen in 9 ways and the second event can happen in 5 ways. To get both a shirt and a pair of pants, you multiply. To get just one article of clothing, you add.

Now let's look at this using sets. There are two sets, call them S and P The set S contains all 9 shirts so $|S| = 9$ while $|P| = 5$, since there are 5 elements in the set P (namely your 5 pairs of pants). What are we asking in terms of these sets? Well in question 2, we really want $|S \cup P|$; how many elements are there in the union of shirts and pants. Clearly this is just $|S| + |P|$ (since there is no overlap - in other words,

[2] To solve this problem, you could break into two cases - first count how many ways there are to select the two cards when the first card is a red non-face card and second count how many ways when the first card is a red face card. Doing so makes the events in each separate case independent, so the multiplicative principle can be applied.

since $|S \cap P| = 0$). Question 1 is a little more complicated. Your first guess might be to find $|S \cap P|$ - but this is not right (there is nothing in the intersection). We are not asking for how many clothing items are both a shirt and a pair of pants. Instead, we want one of each. We could think of this as asking how many pairs (x, y) there are, where x is a shirt and y is a pair of pants. As we will see, to find this number, we would take $|S| \cdot |P|$.

From this example we can see right away how to rephrase our additive principle in terms of sets:

Additive Principle (with sets)
Given two sets A and B, if $A \cap B = \emptyset$ (that is, if there is no element in common to both A and B), then

$$|A \cup B| = |A| + |B|$$

This hardly needs a proof - to find $A \cup B$ you take everything in A and throw in everything in B. Since there is no element in both sets already, you will have $|A|$ things and add $|B|$ new things to it. This is what adding does! Of course, we can easily extend this to any number of (disjoint) sets.

From the example above, we see that in order to investigate the multiplicative principle carefully, we need to consider ordered pairs. We should define this carefully.

Definition 1. Given sets A and B, we can form the *set* $A \times B = \{(x, y) : x \in A \wedge y \in B\}$ to be the set of all ordered pairs (x, y) where x is an element of A and y is an element of B. We call $A \times B$ the *Cartesian product* of A and B.

The question is, what is $|A \times B|$ - the cardinality of the Cartesian product? To figure this out, let's write out $A \times B$.

Let $A = \{a_1, a_2, a_3, \ldots, a_m\}$ and $B = \{b_1, b_2, b_3, \ldots, b_n\}$ (so $|A| = m$ and $|B| = n$). The set $A \times B$ contains all pairs with the first half of the pair being a_i for some i and the second being b_j for some j. In other words:

$$\begin{aligned} A \times B = \{&(a_1, b_1), (a_1, b_2), (a_1, b_3), \ldots (a_1, b_n), \\ &(a_2, b_1), (a_2, b_2), (a_2, b_3), \ldots, (a_2, b_n), \\ &(a_3, b_1), (a_3, b_2), (a_3, b_3), \ldots, (a_3, b_n), \\ &\vdots \\ &(a_m, b_1), (a_m, b_2), (a_m, b_3), \ldots, (a_m, b_n)\} \end{aligned}$$

Notice what we have done here: we made m rows of n pairs - so that is a total of $m \cdot n$ pairs.

Each row above is really $\{a_i\} \times B$ for some $a_i \in A$. That is, we fixed the A-element. Clearly we have

$$A \times B = \{a_1\} \times B \cup \{a_2\} \times B \cup \{a_3\} \times B \cup \cdots \cup \{a_m\} \times B$$

So $A \times B$ is really the union of m disjoint sets. Each of those sets has n elements in them. So the total (using the additive principle) is $n + n + n + \cdots + n = m \cdot n$.

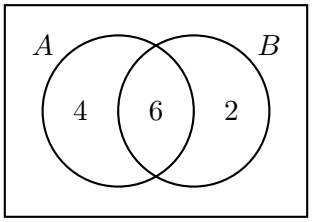
To summarize:

Multiplicative Principle (with sets)
Given two sets A and B, we have $|A \times B| = |A| \cdot |B|$.

Again, we can easily extend this to any number of sets.

1.1.2 Principle of Inclusion/Exclusion

While we are thinking about sets, let's consider what happens to the additive principle when the sets are NOT disjoint. Suppose we know that $|A| = 10$ and $|B| = 8$ (there are 10 elements in A and 8 elements in B). This is not enough information though. We do not know how many of the 8 elements in B are also element of A. But if we do know that $|A \cap B| = 6$, then we can say exactly how many elements are in A, and of those how many are in B and how many are not (6 of the 10 elements are in B, so 4 are in A but not in B). We would fill in a Venn diagram as follows:



This says there are 6 elements in $A \cap B$, 4 elements in $A \setminus B$ and 2 elements in $B \setminus A$. Now these three sets *are* disjoint, so we can use the additive principle to find the number of elements in $A \cup B$. It is $6 + 4 + 2 = 12$.

This will always work, but drawing a Venn diagram is more than we need to do. In fact, it would be nice to relate this problem to the case where A and B are disjoint. Is there one rule we can make that works in either case?

Here is another way to get the answer to the problem above. We start by just adding $|A| + |B|$. This is $10 + 8 = 18$, which would be the answer if $|A \cap B| = 0$. And we see that we are off by exactly 6, which just so happens to be $|A \cap B|$. So perhaps we guess

$$|A \cap B| = |A| + |B| - |A \cap B|.$$

This works for this one example. Will it always work? Think about what we are doing here. We want to know how many things are either in A or B (or both). We can throw in everything in A, and everything in B. This would give $|A| + |B|$ many elements. But of course when you actually take the union, you do not repeat elements that are in both. So far we have counted every element in $A \cap B$ exactly twice - once when we put in the elements from A and once when we included the elements from B. We correct by subtracting out the number of elements we have counted twice. So we added them in twice, subtracted once, leaving them counted only one time.

In other words, we have:

Cardinality of a union (2 sets)
For any finite sets A and B,

$$|A \cup B| = |A| + |B| - |A \cap B|$$

We can do something similar with three sets. First here is an example of how you could use Venn diagrams.

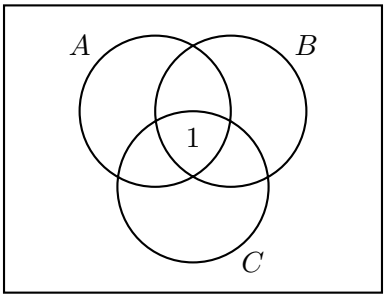
Example: An examination in three subjects, Algebra, Biology, and Chemistry, was taken by 41 students. The following table shows how many students failed in each single subject and in their various combinations.

Subject:	A	B	C	AB	AC	BC	ABC
Failed:	12	5	8	2	6	3	1

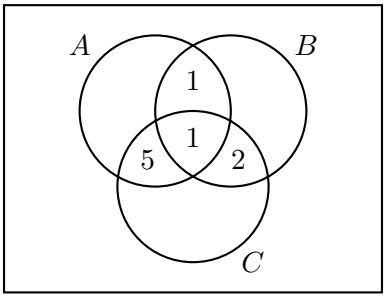
How many students failed at least one subject?

Solution:

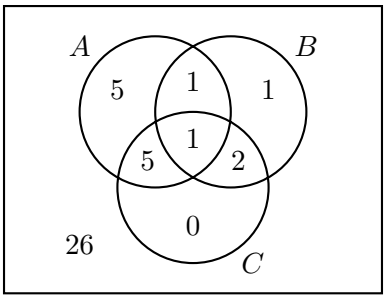
The answer is not 37, even though the sum of the numbers above is 37. The reason is that while 12 students failed algebra, 2 of those students also failed biology and 1 failed chemistry as well. In fact, that 1 student who failed all three subjects is counted a total of 7 times in the total 37. To clarify things, let us think of the students who failed algebra as the elements of the set A, and similarly for sets B and C. The one student who failed all three subjects is the lone element of the set $A \cap B \cap C$. Thus in Venn diagrams:



Now let's fill in the other intersections. We know $A \cap B$ contains 2 elements, but one element has already been counted. So we should put a one in the region where A and B intersect (but C does not). Similarly, we calculate the cardinality of $(A \cap C) \cap \overline{B}$, and $(B \cap C) \cap \overline{A}$:



Next we determine the numbers which should go in the remaining regions, including outside of all three circles. This last number is the number of students who did not fail any subject – the number we were asked to find:



We found that 5 so go in the A only region because the entire circle for A needed to have a total of 12, and 7 were already accounted for. Thus the number of students who passed all three classes is 26. The number who failed at least one class is 15.

Note that we can also answer other questions. For example, now many students failed just chemistry? None. How many passed biology but failed both algebra and chemistry? 5.

Can we solve the problem in an algebraic way? Note that while the additive principle generalizes to any number of sets, when we add a third set here, we must be careful. With two sets, we needed to know the cardinalities of A, B, and $A \cap B$ in order to find the cardinality of $A \cup B$. With three sets we need more information - there are more ways the sets can combine. Not surprisingly then, the formula for cardinality of the union of three non-disjoint sets is more complicated:

Cardinality of a union (3 sets)
For any finite sets A, B, and C,

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|$$

To determine how many elements are in at least one of A, B, or C we add up all the elements in each of those sets. However, when we do that, any element in both A and B is counted twice. Also each element in both A and C is counted twice, as are elements in B and C. So we take each of those out of our sum once. But now what about the elements which are in $A \cap B \cap C$ (in all three sets). We added them in three times, but also removed them three times. So they have not yet been counted. Thus we add those elements back in at the end.

Returning to our example above, we have $|A| = 12$, $|B| = 5$, $|C| = 8$. We also have $|A \cap B| = 2$, $|A \cap C| = 6$, $|B \cap C| = 3$ and $|A \cap B \cap C| = 1$. So:

$$|A \cup B \cup C| = 12 + 5 + 8 - 2 - 6 - 3 + 1 = 15$$

This is what we got when we solved the problem using Venn diagrams.

This process of adding in, then taking out, then adding back in, and so on is called the *Principle of Inclusion/Exclusion*, or simply PIE. We will return to this counting technique later to solve for more complicated problems (involving many more sets).

Exercises

1. Your wardrobe consists of 5 shirts, 3 pairs of pants, and 17 bow ties. How many different outfits can you make?

2. For your college interview, you must wear a tie. You own 3 regular (boring) ties and 5 (cool) bow ties. How many choices do you have for your neck-wear?

3. You realize that the interview is for clown-college, so you should probably wear both a regular tie and a bow tie. How many choices do you have now?

4. You realize that it would also be okay to wear more than two ties.

   (a) You must select some of your ties to wear - everything is okay, from no ties up to all ties. How many choices do you have?

   (b) If you want to wear at least one regular tie and one bow tie, but are willing to wear up to all your ties, how many choices do you have for which ties to wear?

   (c) How many choices do you have if you wear exactly 2 of the 3 regular ties and 3 of the 5 bow ties?

   (d) Once you have selected 2 regular and 3 bow ties, in how many orders could you put the ties on, assuming you must have one of the three bow ties on top?

5. Your Blu-ray collection consists of 9 comedies and 7 horror movies. Give an example of a question for which the answer is:

   (a) 16.

   (b) 63.

6. If $|A| = 10$ and $|B| = 15$, what is the largest possible value for $|A \cap B|$? What is the smallest? What are the possible values for $|A \cup B|$?

7. If $|A| = 8$ and $|B| = 5$, what is $|A \cup B| + |A \cap B|$?

8. A group of college students were asked about their TV watching habits. Of those surveyed, 28 students watch *Elementary*, 19 watch *Castle* and 24 watch of *The Mentalist*. Additionally, 16 watch *Elementary* and *Castle*, 14 watch *elementary* and *The Mentalist* and 10 watch *Castle* and *The Mentalist*. There are 8 students who watch all three shows. How many students surveyed watched at least one of the shows?

9. Find $|(A \cup C) \cap \overline{B}|$ provided $|A| = 50$, $|B| = 45$, $|C| = 40$, $|A \cap B| = 20$, $|A \cap C| = 15$, $|B \cap C| = 23$ and $|A \cap B \cap C| = 12$.

10. Using the same data as the previous question, describe a set with cardinality 26.

11. Consider all 5 letter "words" made from the letters a through h.

    (a) How many of these words are there total?

    (b) How many of these words contain no repeated letters?

    (c) How many of these words (repetitions allowed) start with the sub-word "aha"?

    (d) How many of these words (repetitions allowed) either start with "aha" or end with "bah" or both?

    (e) How many of the words containing no repeats also do not contain the sub-word "bad" (in consecutive letters)?

1.2 Four things to count

Here are some apparently different discrete objects we can count: subsets, bit strings, lattice paths, and binary coefficients. We will give an example of each type of counting problem (and say what these things even are). As we will see, these counting problems are surprisingly similar.

Subsets

Subsets should be familiar - if not, read over the notes on set theory again. Our example problem is this: Let $A = \{1, 2, 3, 4, 5\}$. How many subsets of A contain exactly 3 elements?

First, a simpler question. How many subsets of A are there total? In other words, what is $|\mathcal{P}(A)|$ (the cardinality of the power set of A)? Think about how we would build a subset - we need to decide, for each of the elements of A, whether or not to include the element in our subset. So we need to decide "yes" or "no" for the element 1. And for each choice we make, we need to decide "yes" or "no" for the element 2. And so on. For each of the 5 elements, we have 2 choices. Therefore the number of subsets is simply 2^5 (by the multiplicative principle).

Of those 32 subsets, how many have 3 elements? This is a tricky question. Note that we cannot just use the multiplicative principle. Maybe we want to say we have 2 choices for the first element, 2 choices for the second, 2 choices for the third, and then only 1 choice for the other two. But what if we said "no" to one of the first three elements? Then we would have two choices for the 4th element - this is not working.

Another (bad) idea: we need to pick three elements to be in our subset. There are 5 elements to choose from. So there are 5 choices for the first element, and for each of those 4 choices for the second, and then 3 for the third (last) element. The multiplicative principle would say then that there are a total of $5 \cdot 4 \cdot 3 = 60$ ways to select the 3 element subset. But this is wrong! One of the outcomes we would get from these choices would be the set $\{3, 2, 5\}$ - choosing the element 3 first, then the element 2, then the element 5. Another outcome would be $\{5, 2, 3\}$ - choose the element 5 first, then the element 2 then the element 3. But these are the same set! We can correct this by dividing the supposed 60 outcomes by the number of different outcomes which count as the same for each three elements - there happen to be 6 of these. So we expect there to be 10 3-element subsets of A.

Is this right? Well, we could list out all 10 of them, being very systematic in doing so to make sure we don't miss any or list any twice. Or we could try to count how many subsets of A *don't* have 3 elements in them. How many have no elements? Just 1 (the empty set). How many have 5? Again, just one. These are the cases in which we say "no" to all elements, or "yes" to all elements. Okay, what about the subsets which contain a single element? There are 5 of these - we need to say "yes" to exactly one element, and there are 5 to choose from. Now this is also the number of subsets containing 4 elements - those are the ones for which we must say "no" to exactly one element.

So far we have counted 12 of the 32 subsets. We have not yet counted the subsets with cardinality 2 and with cardinality 3. There are a total of 20 subsets left to split up between these two groups. But the numbers must be the same! If we say "yes" to exactly two elements, that can be accomplished in exactly the same number of ways as the number of ways we can say "no" to exactly two elements. So the number of 2-element subsets is equal to the number of 3-element subsets. Together there are 20 of these subsets, so 10 each.

Number of elements:	0	1	2	3	4	5
Number of subsets:	1	5	10	10	5	1

Bit Strings

"Bit" is short for "binary digit," so a bit string is a string of binary digits. The binary digits are simply the numbers 0 and 1. So all of the following are bit strings:

$$1001 \quad 0 \quad 1111 \quad 1010101010$$

The number of bits (0's or 1's) in the string is the *length* of the string. So the strings above have lengths 4, 1, 4, and 10 respectively. We also can ask how many of the bits are 1's. The number of 1's in a bit string is the *weight* of the string. So the weights of the above strings are 2, 0, 4, and 5 respectively.

Bit Strings

- A n-bit string is a bit string of length n. That is, it is a string containing n symbols, each of which is a bit - a 0 or a 1.
- The *weight* of a bit string is the number of 1's in it.
- $\mathbf{B}^n$ is the *set* of all n-bit strings.
- $\mathbf{B}_k^n$ is the set of all n-bit strings of weight k.

For example the elements of the set $\mathbf{B}_2^3$ are the bit strings 011, 101, and 110. Those are the only strings containing three bits exactly two of which are 1's.

The counting questions: How many bit strings have length 5? How many of those have weight 3? In other words, we are asking for the cardinalities $|\mathbf{B}^5|$ and $|\mathbf{B}_3^5|$.

To find the number of 5-bit strings is easy. We have 5 bits, and each can either be a 0 or a 1. So there are 2 choices for the first bit, 2 choices for the second, and so on. By the multiplicative principle, there are $2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 = 2^5 = 32$ such strings.

Finding the number of 5-bit strings of weight 3 is harder. Think about how such a string could start. The first bit must be either a 0 or a 1. In the first case (the string starts with a 0) we must then decide on four more bits. To have a total of three 1's, of those four remaining bits there must be three 1's. In other words, we must include all 4-bit strings of weight 3. In the second case (the string starts with a 1) we still have four bits to choose, but now only two of them can be 1's. So we must use all the 4-bit strings of weight 2. In other words:

$$|\mathbf{B}_3^5| = |\mathbf{B}_2^4| + |\mathbf{B}_3^4|$$

This is an example of a *recurrence relation* - we represented one instance of our counting problem in terms of two simpler instances of the problem. It holds because the strings in $\mathbf{B}_3^5$ all have the form $1\mathbf{B}_2^4$ (that is, a 1 followed by a string from $\mathbf{B}_2^4$) or $0\mathbf{B}_3^4$. If only we knew the cardinalities of $\mathbf{B}_2^4$ and $\mathbf{B}_3^4$. Well of course

$$|\mathbf{B}_2^4| = |\mathbf{B}_1^3| + |\mathbf{B}_2^3| \quad \text{and} \quad |\mathbf{B}_3^4| = |\mathbf{B}_2^3| + |\mathbf{B}_3^3|$$

We can keep going down, but this should be good enough - $\mathbf{B}_1^3$ and $\mathbf{B}_2^3$ both contain 3 bit strings - we need to pick one of the three bits to be a 1 (three ways to do that) or one of the three bits to be a 0 (three ways to do that). $\mathbf{B}_3^3$ contains just one string: 111. Thus $|\mathbf{B}_2^4| = 6$ and $|\mathbf{B}_3^4| = 4$, which puts $\mathbf{B}_3^5$ at a total of 10 strings.

But wait - 32 and 10 were the answers to the counting questions about subsets. Coincidence? Not at all. Each bit string can be thought of as a *code* for a subset. For the set $A = \{1, 2, 3, 4, 5\}$ we would use 5-bit strings - one bit for each element of A. Each bit in the string is a 0 if its corresponding element of A is not in the subset, and a 1 if the element of A is in the subset. Remember, deciding the subset amounted to a sequence of five yes/no votes for the elements of A. Instead of yes, we put a 1; instead of no we put a 0.

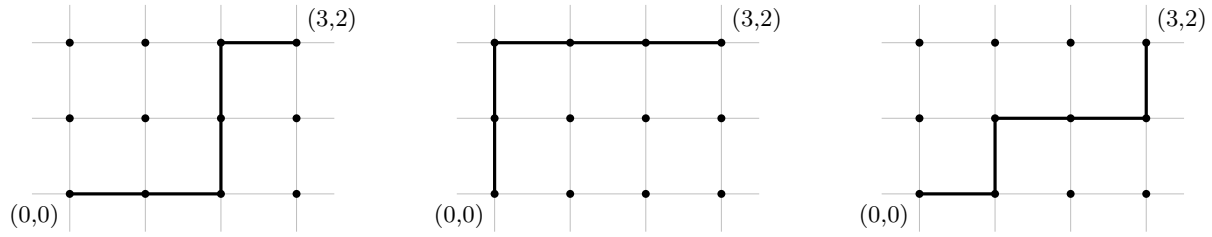
For example, the bit string 11001 represents the subset $\{1, 2, 5\}$ since the first, second and fifth bits are 1's. The subset $\{3, 5\}$ would be coded by the string 00101. What we really have here is a bijection from $\mathcal{P}(A)$ to $\mathbf{B}^5$.

Now for a subset to contain exactly three elements, the corresponding bit string must contain exactly three 1's. In other words, the weight must be 3. So counting the number of 3-element subsets of A is the same as counting the number 5-bit strings of weight 3.

Lattice Paths

The (integer) *lattice* is the set of all points in the Cartesian plain for which both the x and y coordinates are integers. If you like to draw graphs on graph paper, the lattice is all of the intersections of the grid lines.

A *lattice path* is one of the shortest possible paths connecting two points on the lattice, moving only horizontally and vertically. For example, here are three possible lattice paths from the points $(0, 0)$ to $(3, 2)$:

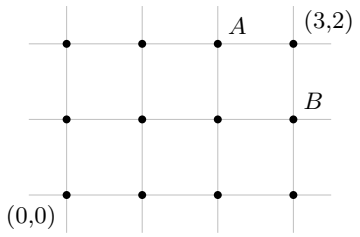


Notice that to ensure the path is the *shortest* possible, each move must be either to the right or up. Additionally, in this case, each path has *length* 5 - there are a total of 5 steps, each either right or up, which must be taken.

The counting question: how many lattice paths are there between $(0, 0)$ and $(3, 2)$? We could try to draw all of these, or instead of drawing them, maybe just list which direction we travel on each of the 5 steps. One path might be RRUUR, or maybe UURRR, or perhaps RURRU (those correspond to the three paths drawn above). So how many such strings of R's and U's are there?

Notice that each of these strings must contain 5 symbols. Exactly 3 of them must be R's (since our destination is 3 units to the right). This seems awfully familiar. In fact, what if we used 1's instead of R's and 0's instead of U's? Then we would just have 5-bit strings of weight 3. There are 10 of those. So there are 10 lattice paths from (0,0) to (3,2).

The correspondence between bit strings and lattice paths does not stop there. Here is another way to count lattice paths. Consider the lattice shown below:



Now any lattice path from (0,0) to (3,2) must pass through exactly one of A and B. The point A is 4 steps away from (0,0) and two of them are right. So the number of lattice paths to A is the same as the number of 4-bit strings of weight 2 - that is, 6. The point B is 4 steps away from (0,0) but now 3 of them are right. So the number of paths to point B is the same as the number of 4-bit strings of weight 3, namely 4. So the total number of paths to (3,2) is just $6 + 4$. This is the same way we calculated the number of 5-bit strings of weight 3. The point: the exact same recurrence relation exists for bit strings and for lattice paths.

Binomial Coefficients

Binomial coefficients are the coefficients in the expanded version of a binomial, such as $(x + y)^5$. What happens when we multiply such a binomial out? We will expand $(x + y)^n$ for various values of n. Each of these are done by multiplying everything out (i.e., FOIL-ing) and then collecting like terms.

$$(x + y)^1 = x + y$$
$$(x + y)^2 = x^2 + 2xy + y^2$$
$$(x + y)^3 = x^3 + 3x^2y + 3xy^2 + y^3$$
$$(x + y)^4 = x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4$$

In fact, there is a quicker way to expand the above binomials. For example, consider the next one, $(x + y)^5$. What we are really doing is multiplying out,

$$(x + y)(x + y)(x + y)(x + y)(x + y)$$

Now in the expansion, there will be only one x^5 term and one y^5 term. That is because to get an x^5, we need to use the x term in each of the copies of the binomial $(x + y)$, and similarly for y^5. What about x^4y? To get terms like this, we need to use four x's and one y. So we need exactly one of the five binomials to contribute a y. There are 5 choices for this, so there are 5 ways to get x^4y, so the coefficient of x^4y is 5. This is also the coefficient for xy^4 for the same (but opposite) reason - there are 5 ways to pick which of the 5 binomials contribute the single x. So far we have

$$(x + y)^5 = x^5 + 5x^4y + ?x^3y^2 + ?x^2y^3 + 5xy^4 + y^5$$

We still need the coefficients of x^3y^2 and x^2y^3. In both cases, we need to pick exactly 3 of the 5 binomials to contribute one variable, the other two to contribute the other. Wait. This sounds familiar. We have 5 things, each can be one of two things, and we need a total of 3 of one of them. That's just like taking 5 bits and making sure exactly 3 of them are 1's. So the coefficient of x^3y^2 (and also x^2y^3) will be exactly the same as the number of bit strings of length 5 and weight 3, which we found earlier to be 10. So we have:

$$(x + y)^5 = x^5 + 5x^4y + 10x^3y^2 + 10x^2y^3 + 5xy^4 + y^5$$

These numbers we keep seeing over and over again - the number of subsets of a particular size, the number of bit strings of a particular weight, the number of lattice paths, and the coefficients of these binomial products - these numbers will all be called *binomial coefficients*. We even have a special symbol for them: $\binom{n}{k}$.

Binomial Coefficients
For each integer $n \geq 0$ and integer k with $0 \leq k \leq n$ we have a number

$$\binom{n}{k}$$

read "n choose k." We have:

- $\binom{n}{k} = |\mathbf{B}_k^n|$, the number of n-bit strings of weight k.
- $\binom{n}{k}$ is the number of subsets of a set of size n each with cardinality k.
- $\binom{n}{k}$ is the number of lattice paths of length n containing k steps to the right.
- $\binom{n}{k}$ is the coefficient of $x^k y^{n-k}$ in the expansion of $(x+y)^n$.
- $\binom{n}{k}$ is the number of ways to select k objects from a total of n objects.

The last bullet point is usually taken as the definition of $\binom{n}{k}$ - we have n objects, and must choose k of them, so there are n choose k ways of doing this. Each of our counting problems above can be viewed in this way:

- How many subsets of $\{1,2,3,4,5\}$ contain exactly 3 elements? We must choose 3 of the 5 elements to be in our subset. There are $\binom{5}{3}$ ways to do this, so there are $\binom{5}{3}$ such subsets.
- How many bit strings have length 5 and weight 3? We must choose 3 of the 5 bits to be 1's. There are $\binom{5}{3}$ ways to do this, so there are $\binom{5}{3}$ such bit strings.
- How many lattice paths are there from (0,0) to (3,2)? We must choose 3 of the 5 steps to be towards the right. There are $\binom{5}{3}$ ways to do this, so there are $\binom{5}{3}$ such paths.
- What is the coefficient of x^3y^2 in the expansion of $(x+y)^5$? We must choose 3 of the 5 copies of the binomial to contribute an x. There are $\binom{5}{3}$ ways to do this, so the coefficient is $\binom{5}{3}$.

It should be clear that in each case above, we have the right answer. All we had to do is phrase the question correctly and it became obvious that $\binom{5}{3}$ is correct. However, this does not tell us that the answer is in fact 10 in each case. We will eventually find a formula for $\binom{n}{k}$ but for now, look back at how we arrived at the answer 10 in our counting problems above. It all came down to bit strings, and we have a recurrence relation for bit strings:

$$|\mathbf{B}_k^n| = |\mathbf{B}_{k-1}^{n-1}| + |\mathbf{B}_k^{n-1}|$$

Remember, this is because we can start the bit string with either a 1 or a 0. In both cases, we have now have $n-1$ more bits to pick. The strings starting with 1 must contain $k-1$ more 1's, while the strings starting with 0 still need k more 1's.

Now since $|\mathbf{B}_k^n| = \binom{n}{k}$, we can use the same recurrence relation for binomial coefficients:

Recurrence relation for $\binom{n}{k}$

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$$

1.2.1 Pascal's Triangle

Let's arrange the binomial coefficients $\binom{n}{k}$ into a triangle like follows:

$$\begin{array}{ccccccccc}
 & & & \binom{0}{0} & & & & & \\
 & & \binom{1}{0} & & \binom{1}{1} & & & & \\
 & \binom{2}{0} & & \binom{2}{1} & & \binom{2}{2} & & & \\
\binom{3}{0} & & \binom{3}{1} & & \binom{3}{2} & & \binom{3}{3} & & \\
\binom{4}{0} & & \binom{4}{1} & & \binom{4}{2} & & \binom{4}{3} & & \binom{4}{4}
\end{array}$$

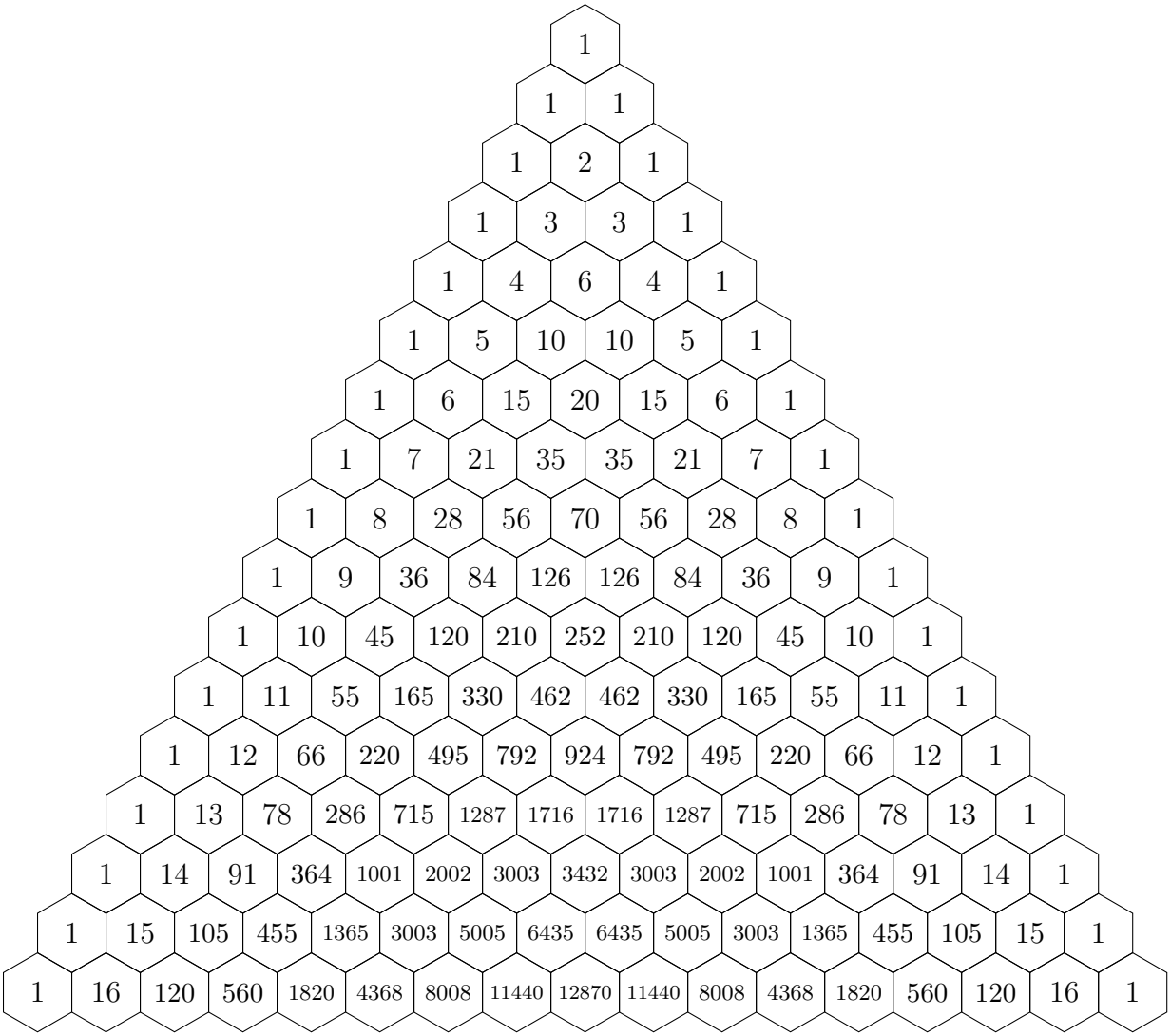
Of course we can continue this as far down as we like. The recurrence relation for $\binom{n}{k}$ tells us that each entry in the triangle is the sum of the two entries above it. The entries on the sides of the triangle are always 1. This is because $\binom{n}{0} = 1$ for all n - there is only one way to pick 0 of n objects - and $\binom{n}{n} = 1$ since there is one way to select all n out of n objects. Using the recurrence relation, and the fact that the sides of the triangle are 1's, we can easily replace all the entries above with the correct values of $\binom{n}{k}$. Doing so gives us Pascal's triangle.

We can use Pascal's triangle to calculate binomial coefficients. For example, using the triangle on the next page, we can find $\binom{12}{6} = 924$.

Exercises

1. How many 10-bit strings contain 6 or more 1's?
2. What is the coefficient of x^9 in the expansion of $(x+1)^{14} + x^3(x+2)^{15}$?
3. Let $S = \{1, 2, 3, 4, 5, 6\}$
   (a) How many subsets are there total?
   (b) How many subsets contain $\{2, 3, 5\}$ as a subset?
   (c) How many subsets of S contain no prime numbers?
   (d) How many subsets contain at least one odd number?
   (e) How many doubletons (i.e., subsets of two elements) contain exactly one even number?
4. How many shortest lattice paths start at (3,3) and
   (a) end at (10,10)?

Pascal's Triangle



(b) end at (10,10) and pass through (5,7)?

(c) end at (10,10) and avoid (5,7)?

1.3 Combinations and Permutations

A *permutation* is a (possible) rearrangement of objects. For example, there are 6 permutations of the letters *a, b, c*:

$$abc, \quad acb, \quad bac, \quad bca, \quad cab, \quad cba$$

We know that we have them all listed above - there are 3 choices for which letter we put first, then 2 choices for which letter comes next, which leaves only 1 choice for the last letter. The multiplicative principle says we multiply these numbers.

Example: How many permutations are there of the letters *a, b, c, d, e, f*?

Solution: We do NOT want to try to list all of these out. However, if we did, we would need to pick a letter to write down first. There are 6 choices for that letter. For each choice of first letter, there are 5 choices for the second letter (we cannot repeat the first letter), and for each of those, there are 4 choices for the third, 3 choices for the fourth, 2 choices for the fifth and finally only 1 choice for the last letter. So there are $6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 720$ permutations of the 6 letters. We could also write this as 6! (that is, 6 factorial).

Sometimes we do not want to permute all of the letters.

Example: How many 4 letter "words" can you make from the letters *a* - *f*, with no repeated letters?

Solution: This is just like the problem of permuting 4 letters, only now we have more choices for each letter. For the first letter, there are 6 choices. For each of those, there are 5 choices for the second letter. Then there are 4 choices for the third letter, and 3 choices for the last letter. The total number of words is $6 \cdot 5 \cdot 4 \cdot 3 = 360$. This is not 6!, but it is almost. The difference is that in 6! we multiply by 2 and 1, where as here we did not. So to cancel the 2 and 1, we could write $\frac{6!}{2!}$.

In general, we can ask how many permutations exist of k objects choosing those objects from a larger n-many objects. (In the example above, $k = 4$, and $n = 6$). We write this number $P(n, k)$. From the example above we see that to compute $P(n, k)$ we must apply the multiplicative principle to k numbers, starting with n and counting backwards. So for example

$$P(10, 4) = 10 \cdot 9 \cdot 8 \cdot 7$$

Notice again that $P(10, 4)$ starts out looking like 10!, but we stop after 7. So

$$P(10, 4) = \frac{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = \frac{10!}{6!}$$

Careful - the factorial in the denominator is not 4! but rather $(10 - 4)!$.

Permutations
$P(n,k)$ is the number of permutations of k out of n objects.

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

Note that when $n = k$, we have $P(n,n) = \frac{n!}{(n-n)!} = n!$ (since $0! = 1$). This makes sense - we already had that $n!$ gives the number of permutations of all n objects.

Here is another way to find the number of permutations of k out of n objects: first select which k objects will be in the permutation, then count how many ways there are to arrange them. Once you have selected the k objects, we know there are $k!$ ways to arrange (permute) them. But how do you select k objects from the n? You have n objects, and you need to *choose* k of them. You can do that in $\binom{n}{k}$ ways. This says:

$$P(n,k) = \binom{n}{k} \cdot k!$$

Now since we have a closed formula for $P(n,k)$ already, we can substitute that in:

$$\frac{n!}{(n-k)!} = \binom{n}{k} \cdot k!$$

If we divide both sides by $k!$ we get a closed formula for $\binom{n}{k}$.

Closed formula for $\binom{n}{k}$

$$\binom{n}{k} = \frac{n!}{(n-k)!k!}$$

We say $P(n,k)$ counts permutations, and $\binom{n}{k}$ counts *combinations*. The formulas for each are very similar - there is just an extra $k!$ in the denominator of $\binom{n}{k}$. That extra $k!$ accounts for the fact that $\binom{n}{k}$ does not distinguish between the different orders that the k objects can appear in - we are just selecting (or choosing) the k objects, not arranging them. Perhaps "combination" is a misleading label - we don't mean it like a combination lock (where the order would definitely matter). Perhaps better: a combination of flavors - you just need to decide which flavors to combine, not the order in which to combine them.

To further illustrate the connection between combinations and permutations, we close with an example.

> **Example**: You decide to have a dinner party. Even though you are incredibly popular and have 14 different friends, you only have enough chairs to invite 6 of them.
>
> 1. How many choices do you have for which 6 friends to invite?
> 2. What if you need to decide not only which friends to invite but also in which order to invite them in? How many choices do you have then?

Solution:

1. You must simply choose 6 friends from a group of 14. This can be done in $\binom{14}{4}$ ways. We can find this number either by using Pascal's triangle or the closed formula $\frac{14!}{8!\cdot 6!} = 3003$.
2. Here you must count all the ways you can permute 6 friends chosen from a group of 14. So the answer is $P(14,6)$, which can be calculated as $\frac{14!}{8!} = 2192190$

How are these numbers related? Notice that $P(14,6)$ is *much* larger than $\binom{14}{6}$. This makes sense - $\binom{14}{6}$ picks 6 friends, but $P(14,6)$ arranges the 6 friends as well as picks them. In fact, we can say exactly how much larger $P(14,6)$ is. In both counting problems we choose 6 out of 14 friends. For the first one, we stop there - at 3003 ways. But for the second counting problem, each of those 3003 choices of 6 friends can be arranged in exactly 6! ways. So now we have $3003 \cdot 6!$ choices - and that is exactly 2192190.

Alternatively, look at the first problem another way. We want to select 6 out of 14 friends, but we do not care about the order they are selected in. To select 6 out of 14 friends, we might try this:

$$14 \cdot 13 \cdot 12 \cdot 11 \cdot 10 \cdot 9$$

This is a reasonable guess, since we have 14 choices for the first guest, then 13 for the second, and so on. But the guess is wrong (in fact, that product is exactly $2192190 = P(14,6)$). It distinguishes between the different orders in which we could invite the guests. To correct for this, we could divide by the number of different arrangements of the 6 guests (so that all of these would count as just one outcome). There are precisely 6! ways to arrange 6 guests, so the correct answer to the first question is

$$\frac{14 \cdot 13 \cdot 12 \cdot 11 \cdot 10 \cdot 9}{6!}$$

Note that another way to write this is

$$\frac{14!}{8! \cdot 6!}$$

which is what we had originally.

Exercises

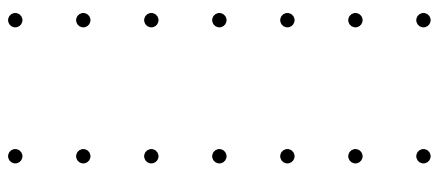
1. A pizza parlor offers 10 toppings.

   (a) How many 3-topping pizzas could they put on their menu? Assume double toppings are not allowed.

   (b) How many total pizzas are possible, with between zero and ten toppings (but not double toppings) allowed?

   (c) The pizza parlor will list the 10 toppings in two columns on their menu. How many ways can they arrange the toppings in the left column?

2. How many quadrilaterals can you draw using the dots below as vertices (corners)?



3. How many of the quadrilaterals possible in the previous problem are:

   (a) Squares? (b) Rectangles? (c) Parallelograms? (d) Trapezoids?

4. On a business retreat, your company of 20 businessmen go golfing.

   (a) You need to divide up into foursomes (groups of 4 people): a first foursome, a second foursome, and so on. How many ways can you do this?

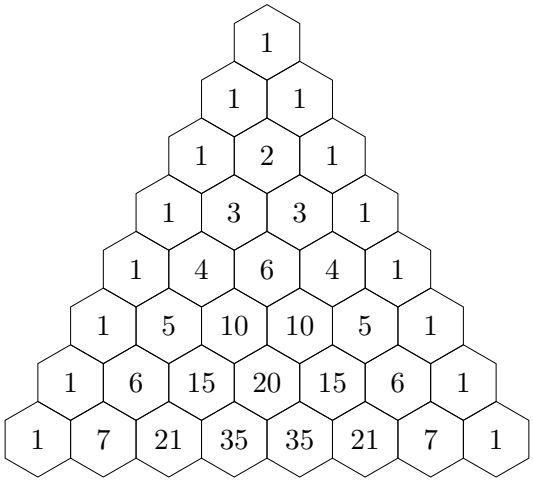
   (b) After all your hard work, you realize that in fact, you want each foursome to include one of the five CEO's. How many ways can you do this?

5. How many different seating arrangements are possible for King Arthur and his 9 knights around their round table?

1.4 Combinatorial Proofs

1.4.1 Patterns in Pascal's Triangle

Have a look again at Pascal's triangle. Forget for a moment where it comes from - just look at it as a mathematical object. What do you notice?



There are lots of patterns hidden away in the triangle - enough to fill a reasonably sized book. Here are just a few of the most obvious ones:

1. The entries on the border of the triangle are all 1.

2. Any entry not on the border is the sum of the two entries above it.

3. The triangle is symmetric - on any row, entries on the left side are mirrored on the right side.

4. The sum of all entries on a given row is a power of 2. (You should check this!)

We would like to state these observations in a more precise way, and then prove that they are correct. Now each entry in Pascal's triangle is in fact a binomial coefficient. The 1 on the very top of the triangle is $\binom{0}{0}$. The next row (which we will call row 1, even though it is not the top-most row) consists of $\binom{1}{0}$ and $\binom{1}{1}$. Row 4 (the row 1, 4, 6, 4, 1) consists of the binomial coefficients

$$\binom{4}{0} \quad \binom{4}{1} \quad \binom{4}{2} \quad \binom{4}{3} \quad \binom{4}{4}$$

Given this description of the elements in Pascal's triangle, we can rewrite the above observations as follows:

1. $\binom{n}{0} = 1$ and $\binom{n}{n} = 1$
2. $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$
3. $\binom{n}{k} = \binom{n}{n-k}$
4. $\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n} = 2^n$

Each of these are an example of a *binomial identity* - a identity (i.e., equation) involving binomial coefficients.

Our goal is to establish these identities - prove that they hold for all values of n and k. These proofs can be done in many ways. One option would be to give algebraic proofs, using the formula for $\binom{n}{k}$:

$$\binom{n}{k} = \frac{n!}{(n-k)!\,k!}$$

Here's how you might do that for the second identity above.

Example: Give an algebraic proof for the binomial identity

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$$

Proof. By the definition of $\binom{n}{k}$, we have

$$\binom{n-1}{k-1} = \frac{(n-1)!}{(n-1-(k-1))!(k-1)!} = \frac{(n-1)!}{(n-k)!(k-1)!} \text{ and } \binom{n-1}{k} = \frac{(n-1)!}{(n-1-k)!k!}$$

Thus, starting with the right hand side of the equation we are trying to establish:

$$\begin{aligned}
\binom{n-1}{k-1} + \binom{n-1}{k} &= \frac{(n-1)!}{(n-k)!(k-1)!} + \frac{(n-1)!}{(n-1-k)!\,k!} \\
&= \frac{(n-1)!k}{(n-k)!\,k!} + \frac{(n-1)!(n-k)}{(n-k)!\,k!} \\
&= \frac{(n-1)!(k+n-k)}{(n-k)!\,k!} \\
&= \frac{n!}{(n-k)!\,k!} \\
&= \binom{n}{k}
\end{aligned}$$

The second line above (where the common denominator is found) works because $k(k-1)! = k!$ and $(n-k)(n-k-1)! = (n-k)!$. QED

This is certainly a valid proof, but also is entirely useless. Even if you understand the proof perfectly, it does not tell you *why* the identity is true. In other words, the proof does not have any explanatory power. A better approach would be to explain what $\binom{n}{k}$ *means* and then say why that is also what $\binom{n-1}{k-1} + \binom{n-1}{k}$ means. Let's see how this works for the four identities we have, um, identified.

Example: Explain why $\binom{n}{0} = 1$ and $\binom{n}{n} = 1$.

Solution: What do these binomial coefficients tell us? Well, $\binom{n}{0}$ gives the number of ways to select 0 objects from a collection of n objects. There is only one way to do this - to not select any of the objects. Thus $\binom{n}{0} = 1$. Similarly, $\binom{n}{n}$ gives the number of ways to select n objects from a collection of n objects. There is only one way to do this - select all n objects. Thus $\binom{n}{n} = 1$.

Alternatively, we know that $\binom{n}{0}$ is the number of n-bit strings with weight 0 - that is, the number of strings made up of n bits, 0 of which are 1's. There is only one such string - the string of all 0's. So $\binom{n}{0} = 1$. Similarly $\binom{n}{n}$ is the number of n-bit strings with weight n. There is only one string - the string of all 1's - with this property.

Another way: $\binom{n}{0}$ gives the number of subsets of a set of size n containing 0 elements. There is only one such subset - the empty set. $\binom{n}{n}$ gives the number of subsets containing n elements. The only such subset is the original set (of all elements).

Example: Explain why $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$

Solution: The easiest way to see this is to consider bit strings. $\binom{n}{k}$ is the number of bit strings of length n containing k 1's. Of all of these strings, some start with a 1 and the rest start with a 0. First consider all the bit strings which start with a 1. After the 1, there must be $n-1$ more bits (to get the total length up to n) and exactly $k-1$ of them must be 1's (as we already have one, and we need k total). How many strings are there like that? There are exactly $\binom{n-1}{k-1}$ such bit strings, so of all the length n bit strings containing k 1's, $\binom{n-1}{k-1}$ of them start with a 1. Similarly, there are $\binom{n-1}{k}$ which start with a 0 - we still need $n-1$ bits and now k of them must be 1's. Since there are $\binom{n-1}{k}$ bit strings containing $n-1$ bits with k 1's, that is the number of length n bit strings with k 1's which start with a 0.

Another way: consider the question - how many ways can you select k pizza toppings from a menu containing n choices? One way to do this is just $\binom{n}{k}$. Another way to answer the same question is to first decide whether or not you want anchovies. If you do want anchovies, you still need to pick $k-1$ toppings, now from just $n-1$ choices. That can be done in $\binom{n-1}{k-1}$ ways. If you do not want anchovies, then you still need to select k toppings from $n-1$ choices (the anchovies are out). You can do that in $\binom{n-1}{k}$ ways. Since the choices with anchovies are disjoint from the choices without anchovies, the total choices are $\binom{n-1}{k-1} + \binom{n-1}{k}$. But wait. We answered the same question in two different ways - so the two answers must be the same. Thus $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$.

You can also explain (prove) this identity by counting subsets, or even lattice paths.

Example: Prove the binomial identity $\binom{n}{k} = \binom{n}{n-k}$.

Solution: Why is this true? $\binom{n}{k}$ counts the number of ways to select k things from n choices. On the other hand, $\binom{n}{n-k}$ counts the number of ways to select $n-k$ things from n choices. Are these really the same? Well, what if instead of selecting the $n-k$ things you choose to exclude them. How many ways are there to choose $n-k$ things to exclude from n choices. Clearly this is $\binom{n}{n-k}$ as well (it doesn't matter whether you include or exclude the things once you have chosen them). And if you exclude $n-k$ things, then you are including the other n things. So the set of outcomes should be the same.

Let's do the pizza counting example like we did above. How many ways are there to pick k toppings from a list of n choices? On the one hand, then answer is simply $\binom{n}{k}$. Alternatively, you could make a list of all the toppings you don't want. To end up with a pizza containing exactly k toppings, you need to pick $n-k$ toppings to not put on the pizza. You have $\binom{n}{n-k}$ choices for the toppings you don't want. Both of these ways give you a pizza with k toppings - in fact all the ways to get a pizza with k toppings. Thus these two answers must be the same: $\binom{n}{k} = \binom{n}{k-1}$.

You can also prove (explain) this identity using bit strings, subsets, or lattice paths. The bit string argument is nice: $\binom{n}{k}$ counts the number of bit strings of length n with k 1's. This is also the number of bit string of length n with k 0's though - just replace each 1 with a 0 and each 0 with a 1. But if a string of length n has k 0's, it must have $n-k$ 1's. And there are exactly $\binom{n}{n-k}$ strings of length n with $n-k$ 1's.

Example: Prove the binomial identity $\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n} = 2^n$

Solution: Let's do a "pizza proof" again. We need to find a question about pizza toppings which has 2^n as the answer. How about this: If a pizza joint offers n toppings, how many pizzas can you build using any number of toppings from no toppings to all toppings, using each topping at most once?

On one hand, the answer is 2^n - for each topping you can say "yes" or "no," so you have two choices for each topping.

On the other hand, divide the possible pizzas into disjoint groups - the pizzas with no toppings, the pizzas with one topping, the pizzas with two toppings, etc. If we want no toppings, there is only one pizza like that (the empty pizza, if you will) but it would be better to think of that number as $\binom{n}{0}$ - we choose 0 of the n toppings. How many pizzas have 1 topping? We need to choose 1 of the n toppings, so $\binom{n}{1}$. We have:

Pizzas with 0 toppings: $\binom{n}{0}$

Pizzas with 1 topping: $\binom{n}{1}$

Pizzas with 2 toppings: $\binom{n}{2}$

$\vdots$

Pizzas with n toppings: $\binom{n}{n}$.

The total number of possible pizzas will be the sum of these, which is exactly the left hand side of the identity we are trying to prove.

Again, we could have proved the identity using subsets, bit strings, or lattice paths (although the lattice path argument is a little tricky).

Hopefully this gives some idea of how explanatory proofs of binomial identities can go. It is worth pointing out that sometimes more traditional proofs are also nice.[3] For example, consider the following rather slick proof of the last identity.

Recall the binomial theorem:

$$(x+y)^n = \binom{n}{0}x^n + \binom{n}{1}x^{n-1}y + \binom{n}{2}x^{n-2}y^2 + \cdots + \binom{n}{n-1}x \cdot y^n + \binom{n}{n}y^n$$

Let $x = 1$ and $y = 1$. We get:

$$(1+1)^n = \binom{n}{0}1^n + \binom{n}{1}1^{n-1}1 + \binom{n}{2}1^{n-2}1^2 + \cdots + \binom{n}{n-1}1 \cdot 1^n + \binom{n}{n}1^n$$

Of course this simplifies to:

$$(2)^n = \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n-1} + \binom{n}{n}$$

Something fun to try: Let $x = 1$ and $y = 2$. Neat huh?

1.4.2 More Proofs

The explanatory proofs given in the above examples are typically called *combinatorial proofs*. In general, to give a combinatorial proof for a binomial identity, say $A = B$ you do the following:

1. Find a counting problem you will be able to answer in two ways.
2. Explain why one answer to the counting problem is A.
3. Explain why the other answer to the counting problem is B.

Since both A and B are the answers to the same question, we must have $A = B$.

The tricky thing is coming up with the question. This is not always obvious, but it gets easier the more counting problems you solve. You will start to recognize types of answers as the answers to types of questions. More often what will happen is you will be solving a counting problem and happen to think up two different ways of finding the answer. Now you have a binomial identity and the proof is right there - the proof *is* the problem you just solved together with your two solutions.

For example, consider this counting question:

> How many 10-letter words contain exactly four A's, three B's, two C's and one D?

Let's try to solve this problem. We have 10 spots for letters to go. Four of those need to be A's. We can pick the four A-spots in $\binom{10}{4}$ ways. Now where can we put the B's? Well there are only 6 spots left, we need to pick 3 of them. This can be done in $\binom{6}{3}$ ways. The two C's need to go in two of the 3 remaining spots, so we have $\binom{3}{2}$ ways of doing that. That leaves just one spot of the D, but we could write that 1 choice as $\binom{1}{1}$. Thus the answer is:

$$\binom{10}{4}\binom{6}{3}\binom{3}{2}\binom{1}{1}$$

[3] Most every binomial identity can be proved using mathematical induction, using the recursive definition for $\binom{n}{k}$. We will discuss induction later in the course.

But why stop there? We can find the answer another way too. First let's decide where to put the one D: we have 10 spots, we need to choose 1 of them, so this can be done in $\binom{10}{1}$ ways. Next, choose one of the $\binom{9}{2}$ ways to place the two C's. We now have 7 spots left, and three of them need to be filled with B's. There are $\binom{7}{3}$ ways to do this. Finally the A's can be placed in $\binom{4}{4}$ (that is, only one) ways. So another answer to the question is

$$\binom{10}{1}\binom{9}{2}\binom{7}{3}\binom{4}{4}$$

Interesting. This gives us the binomial identity:

$$\binom{10}{4}\binom{6}{3}\binom{3}{2}\binom{1}{1} = \binom{10}{1}\binom{9}{2}\binom{7}{3}\binom{4}{4}$$

Here are a couple of other binomial identities with combinatorial proofs.

Example: Prove the identity

$$1n + 2(n-1) + 3(n-2) + \cdots + (n-1)2 + n1 = \binom{n+2}{3}$$

Solution: To give a combinatorial proof we need to think up a question we can answer in two ways: one way needs to give the left-hand-side of the identity, the other way needs to be the right-hand-side of the identity. Our clue to what question to ask comes from the right hand side: $\binom{n+2}{3}$ counts the number of ways to select 3 things from a group of $n+2$ things. Let's name those things $1, 2, 3, \ldots, n+2$. In other words, we want to find 3-element subsets of those numbers (since order should not matter, subsets are exactly the right thing to think about). We will have to be a bit clever to explain why the left-hand-side also gives the number of these subsets. Here's the proof.

Proof. Consider the question "How many 3-element subsets are there of the set $\{1, 2, 3, \ldots, n+2\}$?" We answer this in two ways:

Answer 1: We must select 3 elements from the collection of $n+2$ elements. This can be done in $\binom{n+2}{3}$ ways.

Answer 2: Break this problem up into cases, by what the middle number in the subset is. That is, each subset is $\{a, b, c\}$, say written in increasing order. We count how many such subsets there are for each distinct value of b. The smallest possible value of b is 2, and the largest is $n+1$.

When $b = 2$, there are $1 \cdot n$ subsets - 1 choice a and n choices (3 through $n+2$) for c.

When $b = 3$, there are $2 \cdot (n-1)$ subsets - 2 choices for a and $n-1$ choices for c.

When $b = 4$, there are $3 \cdot (n-2)$ subsets - 3 choices for a and $n-2$ choices for c.

And so on. When $b = n+1$, there are n choices for a and only 1 choice for c, so $n \cdot 1$ subsets.

Therefore the total number of subsets is

$$1n + 2(n-1) + 3(n-2) + \cdots + (n-1)2 + n1.$$

Since Answer 1 and Answer 2 are answers to the same question, they must be equal. Therefore

$$1n + 2(n-1) + 3(n-2) + \cdots + (n-1)2 + n1 = \binom{n+2}{3}$$

QED

Example: Prove the binomial identity

$$\binom{n}{0}^2 + \binom{n}{1}^2 + \binom{n}{2}^2 + \cdots + \binom{n}{n}^2 = \binom{2n}{n}$$

Solution: We will give two different proofs of this fact. The first will be very similar to the previous example (counting subsets). The second proof is a little slicker - it will use lattice paths.

Proof. Consider the question: "How many pizzas can you make using n toppings when there are $2n$ toppings to choose from?"

Answer 1: There are $2n$ toppings, from which you must choose n. This can be done in $\binom{2n}{n}$ ways.

Answer 2: Divide the toppings into two groups of n toppings (perhaps n meats and n veggies). Any choice of n toppings must include some number from the first group and some number from the second group. Consider each possible number of meat toppings separately:

0 meats: $\binom{n}{0}\binom{n}{n}$, since you need to choose 0 of the n meats and n of the n veggies.

1 meat: $\binom{n}{1}\binom{n}{n-1}$, since you need 1 of n meats so $n-1$ of n veggies.

2 meats: $\binom{n}{2}\binom{n}{n-2}$. Choose 2 meats and the remaining $n-2$ toppings from the n veggies.

And so on. The last case is n meats, which can be done in $\binom{n}{n}\binom{n}{0}$ ways.

Thus the total number of pizzas possible is

$$\binom{n}{0}\binom{n}{n} + \binom{n}{1}\binom{n}{n-1} + \binom{n}{2}\binom{n}{n-2} + \cdots + \binom{n}{n}\binom{n}{0}$$

This is not quite the left hand side ... yet. Notice that $\binom{n}{n} = \binom{n}{0}$ and $\binom{n}{n-1} = \binom{n}{1}$ and so on, by the symmetry formula. Thus we do indeed get

$$\binom{n}{0}^2 + \binom{n}{1}^2 + \binom{n}{2}^2 + \cdots + \binom{n}{n}^2$$

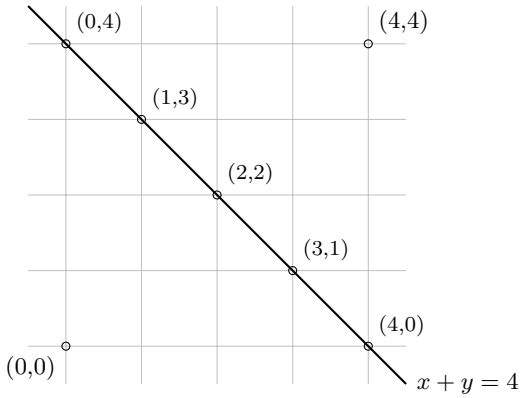
QED

Here's the second proof:

Proof. Consider the question: "How many lattice paths are there from $(0,0)$ to (n,n)?"

Answer 1: We must travel $2n$ units, and n of them must be in the up direction. Thus there are $\binom{2n}{n}$ paths.

Answer 2: Note that any path from $(0,0)$ to (n,n) must cross the line $x+y=n$. That is, must pass through exactly one of the points: $(0,n)$, $(1,n-1)$, $(2,n-2)$, ..., $(n,0)$. For example, this is what happens in the case $n=4$:



How many paths pass through $(0, n)$? To get to that point, you must travel n units, and 0 of them are to the right, so there are $\binom{n}{0}$ ways to get to $(0, n)$. From $(0, n)$ to (n, n) takes n steps, and 0 of them are up. So there are $\binom{n}{0}$ ways to get from $(0, n)$ to (n, n). Therefore there are $\binom{n}{0}\binom{n}{0}$ paths from $(0, 0)$ to (n, n) through the point $(0, n)$.

What about through $(1, n-1)$. There are $\binom{n}{1}$ paths to get there (n steps, 1 to the right) and $\binom{n}{1}$ paths to complete the journey to (n, n) (n steps, 1 up). So there are $\binom{n}{1}\binom{n}{1}$ paths from $(0, 0)$ to (n, n) through $(1, n-1)$.

In general, to get to (n, n) through the point $(k, n-k)$ we have $\binom{n}{k}$ paths to the mid point and then $\binom{n}{k}$ paths from the mid point to (n, n). So there are $\binom{n}{k}\binom{n}{k}$ paths from $(0, 0)$ to (n, n) through $(k, n-k)$.

All together then the total paths from $(0, 0)$ to (n, n) passing through exactly one of these mid points is

$$\binom{n}{0}^2 + \binom{n}{1}^2 + \binom{n}{2}^2 + \cdots + \binom{n}{n}^2$$

QED

Exercises

1. Give a combinatorial proof for the identity $1 + 2 + 3 + \cdots + n = \binom{n+1}{2}$.
2. A woman is getting married. She has 15 best friends but can only select 6 of them to be her bride's maids, one of which needs to be her maid of honor. How many ways can she do this?
   (a) What if she first selects the 6 bride's maids, and then selects one of them to be the maid of honor?
   (b) What if she first selects her maid of honor, and then 5 other bride's maids?
   (c) Explain why $6\binom{15}{6} = 15\binom{14}{5}$.
3. Consider the bit strings in $\mathbf{B}_2^6$ (bit strings of length 6 and weight 2).
   (a) How many of those bit strings start with 1?
   (b) How many of those bit strings start with 01?
   (c) How many of those bit strings start with 001?

(d) Are there any other strings we have not counted yet? Which ones, and how many are there?

(e) How many bit strings are there total in $\mathbf{B}_2^6$?

(f) What binomial identity have you just given a combinatorial proof for?

4. Let's count *ternary* digit strings - strings in which each digit can be 0, 1, or 2.

(a) How many ternary digit strings contain exactly n digits?

(b) How many ternary digit strings contain exactly n digits and n 2's.

(c) How many ternary digit strings contain exactly n digits and $n-1$ 2's. (Hint: where can you put the non-2 digit, and then what could it be?)

(d) How many ternary digit strings contain exactly n digits and $n-2$ 2's. (Hint: see previous hint)

(e) How many ternary digit strings contain exactly n digits and $n-k$ 2's.

(f) How many ternary digit strings contain exactly n digits and no 2's. (Hint: what kind of a string is this?)

(g) Use the above parts to give a combinatorial proof for the identity

$$\binom{n}{0} + 2\binom{n}{1} + 2^2\binom{n}{2} + 2^3\binom{n}{3} + \cdots + 2^n\binom{n}{n} = 3^n$$

5. Give a combinatorial proof for the identity $P(n,k) = \binom{n}{k}k!$

1.5 Stars and Bars

Consider the following counting problem:

> You have 7 cookies to give to 4 kids. How many ways can you do this?

Take a moment to think about how you might solve this problem. You may assume that it is acceptable to give a kid no cookies. Also, the cookies are all identical. And the order in which you give out the cookies does not matter.

Before solving the problem, here is a wrong answer. You might guess that the answer should be 4^7 because for each of the 7 cookies, there are 4 choices of kids to which you can give the cookie. This is reasonable, but wrong. To see why, consider a few possible outcomes: we could assign the first six cookies to kid A, and the seventh cookie to kid B. Another outcome would assign the first cookie to kid B and the six remaining cookies to kid A. Both outcomes are included in the 4^7 answer. But for our counting problem, both outcomes are really the same - kid A gets six cookies and kid B gets one cookie.

What do outcomes actually look like? How can we represent them? One approach would be to write a outcome as a string of four numbers like this:

3112

which represent the outcome in which the first kid gets 3 cookies, the second and third kid each get 1 cookie, and the fourth kid gets 2 cookies. Represented this way, the order in which the numbers occur matters - 1312 is a different outcome, because the first kid gets a one cookie instead of 3.

Each number in the string can be any integer between 0 and 7. But the answer is not 7^4 - we need the *sum* of the numbers to be 7.

Another way we might represent outcomes is to write a string of seven letters:

ABAADCD

which represents that the first cookie goes to kid A, the second cookie goes to kid B, the third and fourth cookies go to kid A, and so on. In fact, this outcome is identical to the previous one - A gets 3 cookies, B and C get 1 each and D gets 2. Each of the seven letters in the string can be any of the 4 possible letters (one for each kid), but the number of such strings is not 4^7, because here order does *not* matter. In fact, another way to write the same outcome is

AAABCDD

This will be the preferred representation of the outcome. Since we can write the letters in any order, we might as well write them in *non-decreasing* order for the purposes of counting. So we will write all the A's first, then all the B's, and so on.

Now think about how you could specify such an outcome. All we really need to do is say when to switch from one letter to the next. In terms of cookies, we need to say after how many cookies do we stop giving cookies to the first kid and start giving cookies to the second kid. And then after how many do we switch to the third kid? And after how many do we switch to the fourth? So yet another way to represent an outcome is like this:

$$***|*|*|**$$

Three cookies go to the first kid, then we switch and give one cookie to the second kid, then switch, one to the third kid, switch, two to the fourth kid. Notice that we need 7 stars and 3 bars - one star for each cookie, and one bar for each switch between kids, so one fewer bars than there are kids (we don't need to switch after the last kid - we are done).

Why have we done all of this? Simple: to count the number of ways to distribute 7 cookies to 4 kids, all we need to do is count how many *stars and bars* charts there are. But a stars and bars chart is just a string of symbols, some stars and some bars. If instead of stars and bars we would use 0's and 1's, it would just be a bit string. We know how to count those.

Before we get too excited, we should make sure that really *any* string of (in our case) 7 stars and 3 bars corresponds to a different way to distribute cookies to kids. In particular consider a string like this:

$$|***||****$$

Does that correspond to a cookie distribution? Yes. It represents the distribution in which kid A gets 0 cookies (because we switch to kid B before any stars), kid B gets three cookies (three stars before the next bar), kid C gets 0 cookies (no stars before the next bar) and kid D gets the remaining 4 cookies. No matter how the stars and bars are arranged, we can distribute cookies in that way. Also, given any way to distribute cookies, we can represent that with a stars and bars chart. For example, the distribution in which kid A gets 6 cookies and kid B gets 1 cookie has the following chart:

$$******|*||$$

Now after all that work we are finally ready to count. Each way to distribute cookies corresponds to a stars and bars chart with 7 stars and 3 bars. So there are 10 symbols, and we must choose 3 of them to be bars. Thus:

$$\text{There are } \binom{10}{3} \text{ ways to distribute 7 cookies to 4 kids.}$$

While we are at it, we can also answer a related question: how many ways are there to distribute 7 cookies to 4 kids so that each kid gets at least one cookie? What can you say about the corresponding stars and bars charts? The charts must start and end with at least one star (so that kids A and D) get cookies, and also no two bars can be adjacent (so that kids B and C are not skipped). One way to assure this is to only place bars in the spaces between the stars. With 7 stars, there are 6 spots between the stars, so we must choose 3 of those 6 spots to fill with bars. Thus there are $\binom{6}{3}$ ways to distribute 7 cookies to 4 kids giving at least one cookie to each kid.

Another (and more general) way to approach this modified problem is to first give each kid one cookie. Now the remaining 3 cookies can be distributed to the 4 kids without restrictions. So we have 3 stars and 3 bars for a total of 6 symbols, 3 of which must be bars. So again we see that there are $\binom{6}{3}$ ways to distribute the cookies.

Stars and bars can be used in counting problems other than kids and cookies. Here are a few examples.

Example: Your favorite mathematical pizza chain offers 10 toppings. How many pizzas can you make if you are allowed 6 toppings? The order of toppings does not matter, but now you are allowed repeats - so one possible pizza is: triple sausage, double pineapple, and onions.

Solution: We get 6 toppings (counting possible repeats). Represent each of these toppings as a star. There are 10 toppings to choose from, so we must switch from considering one topping to the next 9 times - these are the bars. Think of going down the menu one topping at a time: you see anchovies first, and skip to the next, sausage. You say yes to sausage 3 times then switch to the next topping on the list. You keep skipping until you get to pineapple, which you say yes to twice. Another switch and you are at onions - you say yes once. Then you keep switching until you get to the last topping, never saying yes again (since you already have said yes 6 times.

Now that we are confident that we have the right number of stars and bars, we answer the question simply: there are 6 stars and 9 bars, so 15 symbols. We need to pick 9 of them to be bars, so there number of pizzas possible is

$$\binom{15}{9}$$

Example: How many 7 digit phone numbers are there in which the digits are non-increasing? That is, every digit is less than or equal to the previous one.

Solution: We need to decide on 7 digits - so we will use 7 stars. The bars will represent a switch from each possible single digit number down the next smaller one. So the phone number 866-5221 is represented by the stars and bars chart

$$|*||**|*|||**|*|$$

There are 10 choices for each digit (0-9) so we must switch between choices 9 times. We have 7 stars and 9 bars, so the total number of phone numbers is

$$\binom{16}{9}$$

Example: How many integer solutions[4] are there to the equation

$$x_1 + x_2 + x_3 + x_4 + x_5 = 13$$

1. where $x_i \geq 0$ for each x_i?
2. where $x_i > 0$ for each x_i?
3. where $x_i \geq 2$ for each x_i?

Solution: This problem is just like giving 13 cookies to 5 kids - we need to say how many of the 13 units go to each of the 5 variables. In other words, we have 13 stars and 4 bars (the bars are like the "+" signs in the equation).

1. If x_i can be 0 or greater, we are in the standard case with no restrictions. So 13 stars and 4 bars can be arranged in $\binom{17}{4}$ ways.
2. Now each variable must be at least 1. So give one unit to each variable to satisfy that restriction. Now there are 8 stars left, and still 4 bars, so the number of solutions is $\binom{12}{4}$.
3. Now each variable must be 2 or greater. So before any counting, give each variable 2 free units. We now have 3 remaining stars and 4 bars, so there are $\binom{7}{4}$ solutions.

Exercises

1. After gym class you are tasked with putting the 14 identical dodge-balls away into 5 bins.
   (a) How many ways can you do this if there are no restrictions?
   (b) How many ways can you do this if each bin must contain at least one dodge-ball?
   (c) How many ways can you do this if no bin can hold more than 6 balls?
2. How many integer solutions are there to the equation $x + y + z = 8$ for which
   (a) x, y, and z are all positive?
   (b) x, y, and z are all non-negative?
   (c) x, y, and z are all greater than -3.
3. When playing Yahtzee, you roll five regular 6-sided dice. How many different outcomes are possible from a single roll? The order of the dice does not matter.
4. How many integer solutions to $x_1 + x_2 + x_3 + x_4 = 25$ are there fore which $x_1 \geq 1$, $x_2 \geq 2$, $x_3 \geq 3$ and $x_4 \geq 4$?

1.6 Functions

Have we now discovered all the rules needed in counting? Probably not. To see what we can do and what we still must figure out, let's shift focus and think about counting in another way. Consider again the problem if giving cookies to kids. One way to think about this problem is to say that we are *assigning* each cookie to a kid. Now in the case we have looked at so far, the cookies were

[4] An integer solution to an equation is a solution in which the unknown must have an integer value.

identical, so it did not matter which cookie went to which kid. But what if it did? Suppose we have 7 different cookies and 4 different kids - how many ways are there to give out the cookies?

Perhaps you already have a solution in mind. However, for the time being, we will consider this question more generally. What are we really doing here? We are *mapping* each cookie to a kid. This sounds like a function! It is. To help us completely characterize counting problems, we should look at them as problems about how to count functions.

1.6.1 Counting Functions

Suppose (yet again) that you have 7 cookies to give to 4 kids. Now the cookies are all different, so which cookie goes to which kid matters. How many ways are there to do this? What we really have is a function with domain the set of cookies and codomain the set of kids. So we might as well ask:

> **Example**: How many functions are there $f : \{1, 2, 3, 4, 5, 6, 7\} \to \{a, b, c, d\}$?
>
> *Solution*: We must assign $f(x)$ for each x in the domain. We have four choices for $f(1)$ - we could have $f(1) = a$, $f(1) = b$, $f(1) = c$ or $f(1) = d$. Similarly, we have four choices for $f(2)$, and four choices for $f(3)$ and so on. In fact, for each of the seven elements of the domain, we have four choices, so the number of functions here is 4^7.

We might also wonder how many of those functions are injective and how many are surjective. What would these mean in terms of cookies? The function would be injective if each kid got no more than one cookie. There is no way to do this, because we have more cookies than kids. The number of injective functions with this domain and codomain is 0. What if we switch the domain and codomain? So now we are mapping kids to cookies. Note that this is not the same, although it is a perfectly fine counting question (the answer is $P(7, 4)$, do you see why?).

On the other hand, a surjective function would be one in which each kid got at least one cookie. Can we count those? It turns out that counting surjective functions is hard. We will do it, but first back up and consider another simpler example.

> **Example**: How many functions are there $f : \{1, 2, 3, 4, 5\} \to \{a, b, c, d, e\}$? How many of those functions are injective?
>
> *Solution*: First we count all the functions with domain $\{1, 2, 3, 4, 5\}$ and codomain $\{a, b, c, d, e\}$. For f to be a function, it must assign each element of the domain to exactly one element of the codomain. In other words, we need to assign one of the elements of $\{a, b, c, d, e\}$ to $f(1)$. There are 5 choices for this. Similarly, there are five choices for $f(2)$. In fact, there are 5 choices for $f(x)$ for all five possible values of $x \in \{1, 2, 3, 4, 5\}$. Thus the total number of function is $5 \cdot 5 \cdot 5 \cdot 5 \cdot 5 = 5^5$.
>
> What about the injective (one-to-one) functions? Again, we have 5 choices for $f(1)$. However, once we assign a value to $f(1)$, we cannot assign that value to $f(2)$, or any other $f(x)$.[5] So for each of the 5 choices for $f(1)$, we only have 4 choices for $f(2)$, and then only 3 choices for $f(3)$ and only 2 choices for $f(4)$, leaving only 1 choice (the last element of the codomain) for $f(5)$. Therefore the number of one-to-one functions is 5!.

[5] Remember, a function is injective if and only if different elements from the domain go to *different* elements of the codomain.

Using the same counting techniques, you should be able to count the total number of functions from any finite size domain to any finite size codomain, and also count the one-to-one functions (as long as the size of the codomain is at least as large as the domain, otherwise there would be no one-to-one functions).

So what about surjective functions? Well in the previous example, there are exactly $5! = 120$ surjections. We know this because whenever the domain and codomain of a function are the same finite size, a function is injective if and only if it is surjective. In general, if the domain and codomain both contain n elements, then the number of surjective functions is $n!$, since this is the number of injective functions. Also, if the codomain is strictly larger than the domain, then the surjective functions are easy to count - there are 0 of them (since there will always be elements of the codomain left out of the range). But what if the size of the codomain is smaller than the size of the domain?

The idea is to count the functions which are *not* surjective, and them subtract that from the total number of functions. This works very well when the codomain has two elements in it.

Example: How many functions $f : \{1, 2, 3, 4, 5\} \to \{a, b\}$ are surjective?

Solution: There are 2^5 functions all together - there are two choices for where to send each of the 5 elements of the domain. Now of these, the functions which are *not* surjective must exclude one or more elements of the codomain from the range. So first, consider functions for which a is not in the range. This can only happen one way - everything gets sent to b. Alternatively, we could exclude b from the range. Then everything gets sent to a, so there is only one function like this. These are the only ways in which a function could not be onto (no function excludes both a and b from the range) so there are exactly $2^5 - 2$ surjective functions.

When there are three elements in the codomain, there are now three choices for a single element to exclude from the range. Additionally, we could pick pairs of two elements to exclude from the range, and we must make sure we don't over count these.

Example: How many functions $f : \{1, 2, 3, 4, 5\} \to \{a, b, c\}$ are surjective?

Solution: Again start with the total number of functions: 3^5 (as each of the five elements of the domain can go to any of three elements of the codomain). Now we count the functions which are *not* surjective.

Start by excluding a from the range. Then we have two choices (b or c) for where to send each of the five elements of the domain. Thus there are 2^5 functions which exclude a from the range. Similarly, there are 2^5 functions which exclude b, and another 2^5 which exclude c. Now have we counted all functions which are not surjective? Yes, but in fact, we have counted some multiple times. For example, the function which sends everything to c was one of the 2^5 functions we counted when we excluded a from the range, and also one of the 2^5 functions we counted when we excluded b from the range. We must subtract out all the functions which specifically exclude two elements from the range. There is 1 function when we exclude a and b, one function when we exclude a and c, and one function when we exclude b and c.

We are using PIE: to count the functions which are not surjective, we added up the functions which exclude a, b, and c separately, then subtracted the functions which exclude pairs of elements. We would then add back in the functions which exclude groups of three elements, except that there are no such functions. We find that the

number of functions which are *not* surjective is

$$2^5 + 2^5 + 2^5 - 1 - 1 - 1 + 0$$

Perhaps a more descriptive way to write this is

$$\binom{3}{1}2^5 - \binom{3}{2}1^5 + \binom{3}{3}0^5$$

since each of the 2^5's was the result of choosing 1 of the 3 elements of the codomain to exclude from the range, each of the three 1^5's was the result of choosing 2 of the 3 elements of the codomain to exclude. Writing 1^5 instead of 1 makes sense too: we have 1 choice of were to send each of the 5 elements of the domain.

Now we can finally count the number of surjective functions:

$$3^5 - \left[\binom{3}{1}2^5 - \binom{3}{2}1^5\right] = 150$$

You might worry that to count surjective functions when the codomain is larger than 3 elements would be too tedious - we need to use PIE but with more than 3 sets the formula for PIE is very long. However, we have lucked out. As we saw in the example above, the number of functions which exclude a single element from the range is the same no matter which single element is excluded. Similarly, the number of functions which exclude a pair of elements will be the same for every pair. With larger codomains, we will see the same behavior with groups of 3, 4, and more elements excluded. So instead of adding/subtracting each of these, we can simply add or subtract all of them at once, if you know how many there are. Here's what happens with 4 and 5 elements in the codomain.

Example:

1. How many functions $f : \{1, 2, 3, 4, 5\} \to \{a, b, c, d\}$ are surjective?
2. How many functions $f : \{1, 2, 3, 4, 5\} \to \{a, b, c, d, e\}$ are surjective?

Solution:

1. There are 4^5 functions all together - we will subtract the functions which are not surjective. We could exclude any one of the four elements of the codomain, and doing so will leave us with 3^5 functions. This counts too many so we subtract the functions which exclude two of the four elements of the codomain, each pair giving 2^5 functions. But this excludes too many, so we add back in the functions which exclude three of the four elements of the codomain, each triple giving 1^5 function. There are $\binom{4}{1}$ groups of functions excluding a single element, $\binom{4}{2}$ groups of functions excluding a pair of elements, and $\binom{4}{3}$ groups of functions excluding a triple of elements. This means that the number of functions which are *not* surjective is:
$$\binom{4}{1}3^5 - \binom{4}{2}2^5 + \binom{4}{3}1^5$$
We can now say that the number of functions which are surjective is:
$$4^5 - \left[\binom{4}{1}3^5 - \binom{4}{2}2^5 + \binom{4}{3}1^5\right]$$

2. The number of surjective functions is:

$$5^5 - \left[\binom{5}{1}4^5 - \binom{5}{2}3^5 + \binom{5}{3}2^5 - \binom{5}{4}1^5\right]$$

We took the total number of functions 5^5 and subtracted all that were not surjective. There were $\binom{5}{1}$ ways to select a single element from the codomain to exclude from the range, and for each there were 4^5 functions. But this double counts, so we use PIE and subtract functions excluding two elements from the range: there are $\binom{5}{2}$ choices for the two elements to exclude, and for each pair, 3^5 functions. This takes out too many functions, so we add back in functions which exclude 3 elements from the range: $\binom{5}{3}$ choices for which three to exclude, and then 2^5 functions for each choice of elements. Finally we take back out the 1 function which excludes 4 elements for each of the $\binom{5}{4}$ choices of 4 elements.

Exercises

1. Write out all functions $f : \{1, 2, 3\}$ to $\{a, b\}$. How many are there? How many are injective? How many are surjective? How many are both?
2. Write out all functions $f : \{1, 2\}$ to $\{a, b, c\}$. How many are there? How many are injective? How many are surjective? How many are both?
3. Consider functions $f : \{1, 2, 3, 4\} \to \{a, b, c, d, e, f\}$.
   (a) How many functions are there total?
   (b) How many functions are injective?
   (c) How many functions are surjective?
   (d) How many functions have the property that $f(1) \neq a$ or $f(2) \neq b$, or both?
4. Consider sets A and B with $|A| = 10$ and $|B| = 17$.
   (a) How many functions $f : A \to B$ are there?
   (b) How many functions $f : A \to B$ are injective?

1.7 Advanced Counting using PIE

Counting surjections used the Principle of Inclusion/Exclusion (PIE), which gives a method for finding the cardinality of the union of not necessarily disjoint sets. The idea is that to find how many things are in *one or more* of the sets A, B, and C we should just add up the number of things in each of these sets. However, if there is any overlap among the sets, those elements are counted multiple times. So we subtract the things in each intersection of a pair of sets. But doing this removes elements which are in all three sets once too often, so we need to add it back in.

For four or more sets, we do not write down a formula for PIE. Instead, we just think of the principle: add up all the elements in single sets, then subtract out things you counted twice (elements in the intersection of a *pair* of sets), then add back in elements you removed too often (elements in the intersection of groups of three sets), then take back out elements you added back in too often (elements in the intersection of groups of four sets), then add back in, take back out, add back in, etc.

Here are some additional examples of how it can be used.

1.7.1 Counting Derangements

A *derangement* of n elements $\{1, 2, 3, \ldots, n\}$ is a permutation in which no element is fixed. For example, there are 6 permutations of the three elements $\{1, 2, 3\}$:

$$123 \quad 132 \quad 213 \quad 231 \quad 312 \quad 321$$

but most of these have one or more elements fixed: 123 has all three elements fixed, 132 has the first element fixed (1 is in its original first position), and so on. In fact, the only derangements of three elements are

$$231 \text{ and } 312$$

If we go up to 4 elements, there are 24 permutations (because we have 4 choices for the first element, 3 choices for the second, 2 choices for the third leaving only 1 choice for the last - $4! = 24$). How many of these are derangements? If you list out all 24 permutations and eliminate those which are not derangements, you will be left with just 9 derangements. Let's see how we can get that number using PIE.

Example: How many derangements are there of 4 elements?

Solution: We count all permutations, and subtract those which are not derangements. There are $4! = 24$ permutations of 4 elements. Now for a permutation to *not* be a derangement, at least one of the 4 elements must be fixed. There are $\binom{4}{1}$ choices for which single element we fix. Once fixed, we need to find a permutation of the other three elements - there are 3! permutations on 3 elements. But now we have counted too many non-derangements, so we must subtract those permutations which fix two elements. There are $\binom{4}{2}$ choices for which two elements we fix, and then for each pair, 2! permutations of the remaining elements. But this subtracts too many, so add back in permutations which fix 3 elements, all $\binom{4}{3}1!$ of them. Finally subtract the $\binom{4}{4}0!$ permutations (recall $0! = 1$) which fix all four elements. All together we get that the number of derangements of 4 elements is:

$$4! - \left[\binom{4}{1}3! - \binom{4}{2}2! + \binom{4}{3}1! - \binom{4}{4}0!\right] = 24 - 15 = 9$$

Of course we can use a similar formula to count the derangements on any number of elements. However, the more elements we have, the longer the formula gets. It turns out that there is no closed formula for counting derangements or surjective functions for that matter. We need to use PIE.

Before moving on to another advanced counting technique, here is another example.

Example: Five gentlemen attend a party, leaving their hats at the door. At the end of the party, they hastily grab hats on their way out. How many different ways could this happen so that none of the gentlemen leave with their own hat?

Solution: We are counting derangements on 5 elements. There are 5! ways for the gentlemen to grab hats in any order - but many of these permutations will result in someone getting their own hat. So we subtract all the ways in which one or more of the men get their own hat. In other words, we subtract the non-derangements. Doing so requires PIE. Thus the answer is:

$$5! - \left[\binom{5}{1}4! - \binom{5}{2}3! + \binom{5}{3}2! - \binom{5}{4}1! + \binom{5}{5}0!\right]$$

1.7.2 Stars, Bars, and Pie

Another time PIE crops up is in stars and bars problems. Let's go back to our 7 cookies and 4 kids problem. Recall we want to distribute the 7 identical cookies to 4 (non-identical) kids. Now we place a restriction on it

Example: How many ways can you distribute 7 cookies to 4 kids so that no kid gets more than 2 cookies?

Solution: To answer this, we will subtract all the outcomes in which a kid gets 3 or more cookies. How many outcomes are there like that? We can force kid A to eat 3 or more cookies by giving him 3 cookies before we start. Doing so reduces the problem to one in which we have 4 cookies to give to 4 kids without any restrictions. In that case, we have 4 stars (the 4 remaining cookies) and 3 bars (one less than the number of kids) so we can distribute the cookies in $\binom{7}{3}$ ways. Of course we could choose any one of the 4 kids to give too many cookies, so it would appear that there are $\binom{4}{1}\binom{7}{3}$ ways to distribute the cookies giving too many to one kid. But in fact, we have over counted. We must get rid of the outcomes in which two kids have too many cookies. There are $\binom{4}{2}$ ways to select to 2 kids to give extra cookies. It takes 6 cookies to do this, leaving only 1 cookie. So we have 1 star and still 3 bars. The remaining 1 cookie can thus be distributed in $\binom{4}{3}$ ways (for each of the $\binom{4}{2}$ choices of which 2 kids to over-feed). We could continue in this fashion (using PIE) but there are no ways to give too many cookies to 3 kids - you don't have enough cookies.

All together we get that the number of ways to distribute 7 cookies to 4 kids without giving any kid more than 2 cookies is:

$$\binom{10}{3} - \left[\binom{4}{1}\binom{7}{3} - \binom{4}{2}\binom{4}{3}\right] = 120 - [140 - 24] = 4$$

This makes sense: the only way to distribute cookies with this restriction is for 3 kids to get 2 cookies and one kid to get 1. There are 4 choices for which kid gets the 1 cookie.

Example: Earlier we counted the number of solutions to the equation

$$x_1 + x_2 + x_3 + x_4 + x_5 = 13.$$

How many of those solutions have $0 \leq x_i \leq 3$ for each x_i?

Solution: We must subtract off the number of solutions in which one or more of the variables has a value greater than 3. We will need to use PIE because counting the number of solutions for which each of the five variables separately are greater than 3 counts solutions multiple times. Here is what we get:

- Total solutions: $\binom{17}{4}$

- Solutions where $x_1 > 3$: $\binom{13}{4}$ - give x_1 4 units first, then distribute the remaining 9 units to the 5 variables.
- Solutions where $x_1 > 3$ and $x_2 > 3$: $\binom{9}{4}$ - after you give 4 units to x_1 and another 4 to x_2, you only have 5 units left to distribute.
- Solutions where $x_1 > 3$, $x_2 > 3$ and $x_3 > 3$: $\binom{5}{4}$.
- Solutions where $x_1 > 3$, $x_2 > 3$, $x_3 > 3$, and $x_4 > 3$: 0.

We also need to account for the fact that we could choose any of the five variables in the place of x_1 above, any pair of variables in the place of x_1 and x_2 and so on. It is because of this that the double counting occurs, so we need to use PIE. All together we have that the number of solutions with $0 \leq x_i \leq 3$ is

$$\binom{17}{4} - \left[\binom{5}{1}\binom{13}{4} - \binom{5}{2}\binom{9}{4} + \binom{5}{3}\binom{5}{4}\right]$$

Exercises

1. Consider sets A and B with $|A| = 10$ and $|B| = 5$. How many functions $f : A \to B$ are surjective?

2. Let $A = \{1, 2, 3, 4, 5\}$. How many injective functions $f : A \to A$ have the property that for each $x \in A$, $f(x) \neq x$?

3. Ten ladies of a certain age drop off their red hats at the hat check of a museum. As they are leaving, the hat check attendant gives the hats back randomly. In how many ways can exactly six of the ladies receive their own hat (and the other four not)?

Chapter 2

Sequences

As the legend goes, there is a monastery in Hanoi with a great hall containing 3 tall pillars. Resting on the first pillar are 64 giant disks (or washers) - all different sizes, stacked from largest to smallest. The monks are charged with the following task: they must move the entire stack of disks to the third pillar. However, due to the size of the disks, the monks cannot move more than one at a time. Each disk must be placed on one of the pillars before the next disk is moved. And because the disks are so heavy and fragile, the monks may never place a larger disk on top of a smaller disk.

When the monks finally complete their task, the world shall come to an end. Your task: figure out how long before we need to start worrying about the end of the world.

This puzzle is called the *Tower of Hanoi*. You are tasked with finding the minimum number of moves to complete the puzzle. This certainly sounds like a counting problem. Perhaps you have an answer? If not, what else could we try? The answer depends on the number of disks you need to move. In fact, we could answer the puzzle first for 1 disk, then 2, then 3 and so on. If we list out all of the answers for each number of disks, we will get a *sequence* of numbers. The nth term in the sequence is the answer to the question, "what is the smallest number of moves required to complete the Tower of Hanoi puzzle with n disks?" You might wonder why we would create such a sequence, instead of just answering the question. The reason is, but looking at how the sequence of numbers grows, we gain insight into the problem. It is easy to count the number of moves required for small numbers of disks. We can then look for a pattern among the first few terms of the sequence. Hopefully this will suggest a method for finding the nth term - the answer to our question. Of course we will also need to verify that our suspected pattern is correct, and that this correct pattern really does give us the nth term we think it does, but it is impossible to prove that your formula is correct without having a formula to start with.

Of course sequences are also interesting mathematical objects to study in their own right. Let's see why.

2.1 Basics

A sequence is simply an ordered list of numbers. For example, here is a sequence: 0, 1, 2, 3, 4, 5, ... This is different from the set $\mathbb{N}$, because while the sequence is a complete list of every element in the set of natural numbers, in the sequence, we very much care what order the numbers come in. For this reason, when we use variables to represent terms in a sequence, they will look like this:

$$a_0, a_1, a_2, a_3, \ldots$$

We might replace the a with another letter, and sometimes we omit a_0, starting with a_1 instead. The numbers in the subscripts are called *indices* (the plural of *index*). We can think of the terms in the sequence as the outputs of a function with domain $\mathbb{N}$: the nth term in the sequence is $f(n) = a_n$.

Example: Can you find the next term in the following sequences?

1. $7, 7, 7, 7, 7, \ldots$
2. $3, -3, 3, -3, 3, \ldots$
3. $1, 5, 2, 10, 3, 15, \ldots$
4. $1, 2, 4, 8, 16, 32, \ldots$
5. $1, 4, 9, 16, 25, 36, \ldots$
6. $1, 2, 3, 5, 8, 13, 21, \ldots$
7. $1, 3, 6, 10, 15, 21, \ldots$
8. $2, 3, 5, 7, 11, 13, \ldots$
9. $3, 2, 1, 0, -1, \ldots$
10. $1, 1, 2, 6, \ldots$

Solution: No you cannot. You might guess that the next terms are:

1. 7
2. -3
3. 4
4. 64
5. 49
6. 34
7. 28
8. 17
9. -2
10. 24

In fact, those are the next terms of the sequences I had in mind when I made up the example. But there is no way to be sure they are correct.

That said, we will often do this. Given the first few terms of a sequence, we can ask what the pattern in the sequence suggests the next terms are.

Given that no number of initial terms in a sequence is enough to say for certain which sequence we are dealing with, we need to find another way to specify a sequence. There are two ways to do this.

Closed formula
A *closed formula* for a sequence $a_0, a_1, a_2, \ldots$ is a formula for a_n using a fixed finite number of operations on n. This is what you normally think of as a formula in n.

Recursive definition
A *recursive definition* (sometimes called an *inductive definition*) for a sequence $a_0, a_1, a_2, \ldots$ consists of a *recurrence relation*: an equation relating a term of the sequence to previous terms (terms with smaller index) and an *initial condition*: a list of a few terms of the sequence (one less than the number of terms in the recurrence relation).

It is easier to understand what is going on here with an example:

Example: Here are a few closed formulas for sequences:

- $a_n = n^2$
- $a_n = \frac{n(n+1)}{2}$
- $a_n = \frac{(\frac{1+\sqrt{5}}{2})^n - 1/(\frac{1+\sqrt{5}}{2})^n}{5}$.

Note in each case, if you are given n, you can calculate a_n directly - just plug in n.

Here are a few recursive definitions for sequences:

- $a_n = 2a_{n-1}$ with $a_0 = 1$.
- $a_n = 2a_{n-1}$ with $a_0 = 27$.

- $a_n = a_{n-1} + a_{n-2}$ with $a_0 = 0$ and $a_1 = 1$.

In these cases, if you are given n, you cannot calculate a_n directly, you first need to find a_{n-1} or a_{n-1} and a_{n-2}.

You might wonder why we would bother with recursive definitions for sequences - after all, it is harder to find a_n with a recursive definition than with a closed formula. This is true, but it is also harder to find a closed formula for a sequence than it is to find a recursive definition. So to find a useful closed formula, we might first find the recursive definition, then use that to find the closed formula.

This is not to say that recursive definitions aren't useful in finding a_n. You can always calculate a_n given a recursive definition - it might just take a while.

Example: Find a_6 in the sequence defined by $a_n = 2a_{n-1} - a_{n-2}$ with $a_0 = 3$ and $a_1 = 4$.

Solution: We know that $a_6 = 2a_5 - a_4$. So to find a_6 we need to find a_5 and a_4. Well

$$a_5 = 2a_4 - a_3 \qquad \text{and} \qquad a_4 = 2a_3 - a_2$$

so if we can only find a_3 and a_2 we would be set. Of course

$$a_3 = 2a_2 - a_1 \qquad \text{and} \qquad a_2 = 2a_1 - a_0$$

so we only need to find a_1 and a_0. But we are given these. Thus

$$\begin{aligned} a_0 &= 3 \\ a_1 &= 4 \\ a_2 &= 2 \cdot 4 - 3 = 5 \\ a_3 &= 2 \cdot 5 - 4 = 6 \\ a_4 &= 2 \cdot 6 - 5 = 7 \\ a_5 &= 2 \cdot 7 - 6 = 8 \\ a_6 &= 2 \cdot 8 - 7 = 9. \end{aligned}$$

Note that now we can guess a closed formula for the nth term of the sequence: $a_n = n + 3$. To be sure this will always work, we could plug in this formula into the recurrence relation:

$$2a_{n-1} - a_{n-2} = 2((n-1)+3) - ((n-2)+3) = 2n + 4 - n - 1 = n + 3 = a_n$$

Since $a_0 = 0 + 3 = 3$ and $a_1 = 1 + 3 = 4$ are the correct initial conditions, we have lucked upon the correct closed formula.

Finding closed formulas, or even recursive definitions, for sequences is not trivial - there is no one method for doing this. Just like in evaluating integrals or solving differential equations, it is useful to have a bag of tricks you can apply, but sometimes there is no easy answer.

One useful trick to keep in your bag is relating a given sequence to another sequence for which we already know the closed formula.

Example: Use the formulas $T_n = \frac{n(n+1)}{2}$ and $a_n = 2^n$ to find closed formulas for the following sequences. Assume the first term has index 1 (not 0).

1. 2, 4, 7, 11, 16, 22, ...
2. 2, 3, 5, 9, 17, 33,...
3. 2, 6, 12, 20, 30, 42,...
4. 6, 10, 15, 21, 28, ...
5. 1, 3, 7, 15, 31, ...
6. 3, 6, 12, 24, 48, ...
7. 6, 10, 18, 34, 66, ...
8. 15, 33, 57, 87, 123, ...

Solution: Before you say this is impossible, what we are asking for is simply to find a closed formula which agrees with all of the initial terms of the sequences. Of course there is no way to read into the mind of the person who wrote the numbers down, but we can at least do this.

Now the first few terms of T_n, starting with T_1 are $1, 3, 6, 10, 15, 21, \ldots$ (these are the triangular numbers). The first few terms of a_n (starting this time with a_0) are $1, 2, 4, 8, 16, \ldots$. Now let's try to find formulas for the given sequences.

1. 2, 4, 7, 11, 16, 22, ...- Note that if subtract 1 from each term, we get the sequence T_n. So this sequence is $T_n + 1$. Therefore a closed formula is $\frac{n(n+1)}{2} + 1$. A quick check of the first few n confirms we have it right.
2. 2, 3, 5, 9, 17, 33,...- This sequence is the result of adding 1 to each term in a_n. So we might guess the closed formula $2^n + 1$. If we try this though, we get the first terms $2^1 + 1 = 3$ and $2^2 + 1 = 5$. We are off, because a_n started with $n = 0$, and now we are starting with $n = 1$. So we shift: the closed formula is $2^{n-1} + 1$.
3. 2, 6, 12, 20, 30, 42,...- Notice that all these terms are even. What happens if we factor out a 2? We get T_n! So this sequence has closed formula $n(n + 1)$.
4. 6, 10, 15, 21, 28, ...- These are all triangular numbers. However, we are starting with 6 as our first term instead of as our third term. So if we could plug in 3 instead of 1 into the formula for T_n, we would be set. Therefore the closed formula is $\frac{(n+2)(n+3)}{2}$ (where $n + 3$ came from $(n + 2) + 1$)
5. 1, 3, 7, 15, 31, ...- Try adding one to each term and we get powers of 2. You might guess this because each term is approximately twice the previous term. Closed formula: $2^n - 1$.
6. 3, 6, 12, 24, 48, ...- These numbers are all multiples of 3. Let's try dividing each by 3. Doing so gives 1, 2, 4, 8, ... Aha. We get the closed formula $3 \cdot 2^{n-1}$.
7. 6, 10, 18, 34, 66, ...- To get from one term to the next, we almost double each term. So maybe we can relate this back to 2^n. Yes, each term is 2 more than a power of 2. So we get $2^{n+1} + 2$ (the $n+1$ is because the first term is 2 more than 2^2, not 2^1). Alternatively, we could have related this sequence to the second sequence in this example: starting with 3, 5, 9, 17, ... we see that this sequence is twice the terms from that sequence (omitting the 2). That sequence had closed formula $2^{n-1} + 1$. To make the 3 first, we would write $2^n + 1$. Our sequence would be twice this, so $2(2^n + 1)$, which of course is the same as we got before.

8. 15, 33, 57, 87, 123, ...- Try dividing each term by 3. That gives the sequence $5, 11, 19, 29, 41, \ldots$. Now add one: $12, 20, 30, 42, \ldots$ - which is sequence 3 in this example, except starting with 6 instead of 2. So let's start with the formula for sequence 3: $n(n+1)$. To start with the 6, we shift: $(n+1)(n+2)$. But this is one too many, so subtract 1: $(n+1)(n+2)-1$. That gives us our sequence, but divided by 3. So we want $3((n+1)(n+2)-1)$.

Exercises

1. Find the closed formula for each of the following sequences by relating them to a well know sequence. Assume the first term given is a_1.
   (a) $2, 5, 10, 17, 26, \ldots$
   (b) $0, 2, 5, 9, 14, 20, \ldots$
   (c) $8, 12, 17, 23, 30, \ldots$
   (d) $1, 5, 23, 119, 719, \ldots$
2. The Fibonacci sequence is $0, 1, 1, 2, 3, 5, 8, 13, \ldots$ (where $F_0 = 0$).
   (a) Give the recursive definition for the sequence.
   (b) Write out the first few terms of the sequence of partial sums.
   (c) Give a closed formula for the sequence of partial sums in terms of F_n (for example, you might say $F_0 + F_1 + \cdots + F_n = 3F_{n-1}^2 + n$, although that is definitely not correct).
3. Write out the first few terms of the sequence given by $a_1 = 3$; $a_n = 2a_{n-1} + 4$. Then find a recursive definition for the sequence $10, 24, 52, 108, \ldots$.
4. Write out the first few terms of the sequence given by $a_n = n^2 - 3n + 1$. Then find a closed formula for the sequence (starting with a_1) $0, 2, 6, 12, 20, \ldots$.

2.2 Arithmetic and Geometric Sequences

We now turn to the question of finding closed formulas for particular types of sequences.

Arithmetic Sequences
If the terms of a sequence differ by a constant, we say the sequence is *arithmetic*.
If the first term (a_1) of the sequence is a and the common difference is d, then we have,
Recursive definition: $a_n = a_{n-1} + d$ with $a_1 = a$.
Closed formula: $a_n = a + d(n-1)$.

How do we know this? For the recursive definition, we need to specify a_1. Then we need to express a_n in terms of a_{n-1}. If we call the first term a, then $a_1 = a$. For the recurrence relation, by the definition of an arithmetic sequence, the difference between successive terms is some constant, say d. So $a_n - a_{n-1} = d$, or in other words,

$$a_1 = a \qquad a_n = a_{n-1} + d$$

To find a closed formula, first try writing out the first write out the sequence using the recursive definition without simplifying:

$$a, a+d, a+d+d, a+d+d+d, \ldots$$

We see that to find the nth term, we need to start with a and then add d a bunch of times. In fact, add it $n-1$ times. Thus $a_n = a + d(n-1)$.

Example: Find recursive definitions and closed formulas for the sequences:

1. $2, 5, 8, 11, 14, \ldots$.
2. $50, 43, 36, 29, \ldots$.

Solution: First we should check that these sequences really are arithmetic. Doing so will reveal the common difference d.

1. $5-2=3$, $8-5=3$, etc. To get from each term to the next, we add three, so $d=3$. The recursive definition is therefore $a_n = a_{n-1} + 3$ with $a_1 = 2$. The closed formula is $a_n = 2 + 3(n-1)$.
2. Here the common difference is -7, since we add -7 to 50 to get 43, and so on. Thus we have a recursive definition of $a_n = a_{n-1} - 7$ with $a_1 = 50$. The closed formula is $a_n = 50 - 7(n-1)$.

What about sequences like $2, 6, 18, 54, \ldots$? This is not arithmetic, because the difference between terms is not constant. However, the *ratio* between successive terms is constant. We call such sequences *geometric*.

The recursive definition for the geometric progression with first term a and common ratio r is

$$a_1 = a \qquad a_n = a_{n-1} \cdot r$$

This makes sense. To get the next term we multiply the previous term by r. We can find the closed formula like we did for the arithmetic progression. Write $a_1 = a$, $a_2 = a \cdot r$, $a_3 = a_2 \cdot r = a \cdot r \cdot r$ and so on. We must multiply the first term a by r a number of times - $n-1$ to be precise. We get $a_n = a \cdot r^{n-1}$.

Geometric Sequences
A sequence is called *geometric* if the ratio between successive terms is constant. Suppose the first term a_1 is a and the common ratio is r. Then we have,
Recursive definition: $a_n = ra_{n-1}$ with $a_1 = a$.
Closed formula: $a_n = a \cdot r^{n-1}$.

Example: Find the recursive and closed formula for the sequences:

1. $3, 6, 12, 24, 48, \ldots$
2. $27, 9, 3, 1, 1/3, \ldots$

Solution: Again, we should first check that these sequences really are geometric - but doing so will tell us what r is.

1. $6/3 = 2$, $12/6 = 2$, $24/12 = 2$, etc. Yes, to get from any term to the next, we multiply by $r = 2$. So the recursive definition is $a_n = 2a_{n-1}$ with $a_1 = 3$. The closed formula is $a_n = 3 \cdot 2^{n-1}$.
2. $9/27 = 1/3$ and $3/9 = 1/3$ and so on. The common ratio is $r = 1/3$. So the sequence has recursive definition $a_n = \frac{1}{3}a_{n-1}$ with $a_1 = 27$ and closed formula $a_n = 27 \cdot \frac{1}{3}^{n-1}$.

2.2.1 Sums of Arithmetic and Geometric Sequences

Look at the sequence $1, 3, 6, 10, 15, \ldots$. We have a formula for this already (they are the triangular numbers) but let's see if we can derive the formula anew. First, is this sequence arithmetic? No, since $3 - 1 = 2$ and $6 - 3 = 3 \neq 2$, so there is no common difference. Is the sequence geometric? No. $3/1 = 3$ but $6/3 = 2$, so there is no common ratio. What to do?

Notice thought that the differences between terms are a arithmetic sequence: $2, 3, 4, 5, 6, \ldots$. This says that the nth term of the sequence $1, 3, 5, 10, 15, \ldots$ is the *sum* of the first n terms in the sequence $1, 2, 3, 4, 5, \ldots$. We say that the first sequence is the *sequence of partial sums* of the second sequence (partial sums because we are not taking the sum of all infinitely many terms - this might be familiar from second semester calculus). If we know how to add up the terms of an arithmetic sequence, we could use this to find a closed formula for a sequence whose differences are the terms of that arithmetic sequence. In fact, this is how we found the formula $\frac{n(n+1)}{2}$ for the triangular numbers.

We could use a similar technique to find a closed formula for the sequence $2, 3, 5, 9, 17, \ldots$ - the differences are $1, 2, 4, 8, \ldots$, a geometric sequence. If we had a method for summing geometric sequences, we could get a formula for sequences with geometric differences.

Summing Arithmetic Sequences: Reverse and Add

Here is a trick that allows us to quickly find the sum of an arithmetic sequence.

Example: Find the sum: $2 + 5 + 8 + 11 + 14 + \cdots + 470$.

Solution: The idea is to mimic the trick we used to find the formula for triangular numbers. If we add the first and last, we get 472. The second term and second-to-last term also add up to 472. To make this clearer, we might express this as follows. Call the sum S. Then,

$$\begin{array}{rrcccccccc} & S = & 2 & + & 5 & + & 8 & + \cdots + & 467 & + & 470 \\ + & S = & 470 & + & 467 & + & 464 & + \cdots + & 5 & + & 2 \\ \hline & 2S = & 472 & + & 472 & + & 472 & + \cdots + & 472 & + & 472 \end{array}$$

To find $2S$ then we add 472 to itself a number of times. What number? We need to decide how many terms (summands) are in the sum. The nth term in the sum can be expressed as $2 + 3(n-1)$ - this is an arithmetic sequence after all. We want to find n when $2 + 3(n-1) = 470$. Solving for n gives $n = 157$. So n ranges from 1 to 157, giving 157 terms in the sum. This is the number of 472's in the sum for $2S$. Thus

$$2S = 157 \cdot 472 = 74104$$

It is now easy to find S:

$$S = 74104/2 = 37052$$

This will work in general, for any sum of *arithmetic* sequences. Call the sum S. Reverse and add. This produces a single number added to itself many times. Find the number of times. Multiply. Divided by 2. Done.

Example: Find a closed formula for $6 + 10 + 14 + \cdots + (4n - 2)$.

Solution: Again, we have a sum of an arithmetic sequence. We need to know how many terms are in the sequence. Clearly each term in the sequence has the form $4k - 2$ (as evidenced by the last term). For which values of k though? To get 6, $k = 2$. To get $4n - 2$ take $k = n$. So to find the number of terms, we need to know how many integers are in the range $2, 3, \ldots, n$. The answer is $n - 1$. (There are n numbers from 1 to n, so one less if we start with 2.)

Now use the reverse and add trick:

$$\begin{array}{rrcccccccccc} & S = & 6 & + & 10 & + & 14 & +\cdots+ & 4n-6 & + & 4n-2 \\ + & S = & 4n-2 & + & 4n-6 & + & 4n-10 & +\cdots+ & 10 & + & 6 \\ \hline & 2S = & 4n+4 & + & 4n+4 & + & 4n+4 & +\cdots+ & 4n+4 & + & 4n+4 \end{array}$$

Since there are $n - 2$ terms, we get

$$2S = (n-2)(4n+4) \qquad \text{so} \qquad S = \frac{(n-2)(4n+4)}{2}$$

Summing Geometric Sequences: Multiply, Shift and Subtract

To find the sum of a geometric sequence, we cannot use the previous trick. Do you see why? The reason we got the same term added to itself many times is because there was a constant difference. So as we added that difference in one direction, we subtracted the difference going the other way, leaving a constant total. For geometric sums, we have a different trick.

Example: What is $3 + 6 + 12 + 24 + \cdots + 12288$?

Solution: Multiply each term by the common ratio (2). You get $2S = 6 + 12 + 24 + \cdots + 24576$. Now subtract: $2S - S = -3 + 24576 = 24573$. Since $2S - S = S$, we have our answer.

To see what happened in the above example, try writing it this way:

$$\begin{array}{rrl} & S & = 3 + 6 + 12 + 24 + \cdots + 12288 \\ - & 2S & = \quad 6 + 12 + 24 + \cdots + 12288 + 24576 \\ \hline & -S & = 3 + 0 + 0 + 0 + \cdots + 0 - 24576 \end{array}$$

Then divide both sides by -1 and we have the same result for S. The idea is, by multiplying the sum by the common ratio, each term becomes the next term. We shift over the sum to get the subtraction to mostly cancel out, leaving just the first term and new last term.

Example: Find a closed formula for $S(n) = 2 + 10 + 50 + \cdots + 2 \cdot 5^n$.

Solution: The common ratio is 5. So we have

$$\begin{array}{rrl} & S & = 2 + 10 + 50 + \cdots + 2 \cdot 5^n \\ - & 5S & = \quad 10 + 50 + \cdots + 2 \cdot 5^n + 2 \cdot 5^{n+1} \\ \hline & -4S & = 2 - 2 \cdot 5^{n+1} \end{array}$$

Thus $S = \dfrac{2 - 2 \cdot 5^{n+1}}{-4}$

Even though this might seem like a new technique, you have probably used it before.

Example: Express $0.464646\ldots$ as a fraction.

Solution: Let $N = 0.46464646\ldots$. Consider $100N$. We get:

$$\begin{array}{rl} N & = 0.4646464\ldots \\ -0.01N & = 0.00464646\ldots \\ \hline 0.99N & = 0.46 \end{array}$$

So $N = \frac{46}{99}$. What have we done? We viewed the repeating decimal $0.464646\ldots$ as a sum of the geometric sequence $0.46, 0.0046, 0.000046, \ldots$ The common ratio is .01. The only real difference is that we are now computing an *infinite* geometric sum, we we do not have the extra term on the end to consider.

$\sum$ and $\prod$ notation

To simplify writing out sums, we will use notation like $\displaystyle\sum_{k=1}^{n} a_k$. This means add up the a_ks where k changes from 1 to n.

Example: Use $\sum$ notation to rewrite the sums:

1. $1 + 2 + 3 + 4 + \cdots + 100$
2. $1 + 2 + 4 + 8 + \cdots + 2^{50}$
3. $6 + 10 + 14 + \cdots + (4n - 2)$.

Solution:

1. $\displaystyle\sum_{k=1}^{100} k$
2. $\displaystyle\sum_{k=0}^{50} 2^k$
3. $\displaystyle\sum_{k=2}^{n} (4k - 2)$

If we want to multiply the a_k instead, we would write $\displaystyle\prod_{k=1}^{n} a_k$. For example, $\displaystyle\prod_{k=1}^{n} k = n!$.

Exercises

1. Consider the sequence $8, 14, 20, 26, \ldots,$.
   (a) What is the next term in the sequence?
   (b) Find a formula for the nth term of this sequence, assuming $a_1 = 8$.
   (c) Find the sum of the first 100 terms of the sequence: $\sum_{k=1}^{100} a_k$.
2. Consider the sequence $1, 7, 13, 19, \ldots, 6n + 7$.
   (a) How many terms are there in the sequence?
   (b) What is the second-to-last term?

(c) Find the sum of all the terms in the sequence.

3. Find $5 + 7 + 9 + 11 + \cdots + 521$.
4. Find $5 + 15 + 45 + \cdots + 5 \cdot 3^{20}$
5. Find $1 - \frac{2}{3} + \frac{4}{9} - \cdots + \frac{2^{30}}{3^{30}}$
6. Find x and y such that $27, x, y, 1$ is part of an arithmetic sequence. Then find x and y so that the sequence is part of a geometric sequence. (Warning: x and y might not be integers.)
7. Use summation ($\sum$) or product ($\prod$) notation to rewrite the following.
   (a) $2 + 4 + 6 + 8 + \cdots + 2n$
   (b) $1 + 5 + 9 + 13 + \cdots + 425$
   (c) $1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots + \frac{1}{50}$
   (d) $2 \cdot 4 \cdot 6 \cdot \cdots \cdot 2n$
   (e) $(\frac{1}{2})(\frac{2}{3})(\frac{3}{4}) \cdots (\frac{100}{101})$
8. Expand the following sums and products. That is, write them out the long way.
   (a) $\displaystyle\sum_{k=1}^{100}(3 + 4k)$
   (b) $\displaystyle\sum_{k=0}^{n} 2^k$
   (c) $\displaystyle\sum_{k=2}^{50} \frac{1}{(k^2 - 1)}$
   (d) $\displaystyle\prod_{k=2}^{100} \frac{k^2}{(k^2 - 1}$
   (e) $\displaystyle\prod_{k=0}^{n}(2 + 3k)$

2.3 Polynomial Fitting

So far we have seen methods for finding the closed formulas for certain types of sequences - arithmetic and geometric. Since we know how to compute the sum of the first n terms of arithmetic and geometric sequences, we can compute the closed formulas for sequences which have an arithmetic (or geometric) sequence of differences between terms. But what if we consider a sequence which is the sum of the first n terms of a sequence which is the sum of an arithmetic sequence!

Before we get too carried away, let's consider an example: How many squares (of all sizes) are there on a chessboard? A chessboard consists of 64 squares, but we also want to consider squares of longer side length. Even though we are only considering an 8×8 board, there is already a lot to count. So instead, let us build a sequence: the first term will be the number of squares on a 1×1 board, the second term will be the number of squares on a 2×2 board, and so on. After a little thought, we arrive at the sequence

$$1, 5, 14, 30, 55, \ldots$$

This sequence is not arithmetic (or geometric for that matter), but perhaps it's sequence of differences is. For differences we get

$$4, 9, 16, 25, \ldots$$

Not a huge surprise: one way to count the number of squares in a 4×4 chessboard is to notice that there are 16 squares with side length 1, 9 with side length 2, 4 with side length 3 and 1 with side length 4. So the original sequence is just the sum of squares. Now this sequence of differences is not arithmetic since it's sequence of differences (the differences of the differences of the original sequence) is not constant. In fact, this sequence of *second differences* is

$$5, 7, 9, \ldots$$

which *is* an arithmetic sequence - *its* sequence of differences is constant (2, 2, 2,...). Notice that our original sequence had *third differences* (that is, differences of differences of differences of the original) constant. We will call such a sequence Δ^3. The sequence $1, 4, 9, 16, \ldots$ has second differences constant, so it will be a Δ^2 sequence. In general, we will say a sequence is a Δ^k sequence if the kth differences are constant.

Now Δ^0 sequences are constant, so a closed formula for them is easy to compute (it's just the constant). Δ^1 sequences are arithmetic - we have a method for finding closed formulas for them as well. Δ^2 sequences are sums of arithmetic sequences - we can find formulas for these as well. But notice that the format of the closed formula for a Δ^2 sequence is always quadratic - the squares are Δ^2 with closed formula $a_n = n^2$, the triangular numbers (also Δ^2) have closed formula $a_n = \frac{n(n+1)}{2}$, which when multiplied out gives you an n^2 term as well. It appears that every time we increase the complexity of the sequence - that is, increase the number of differences before we get constants - we also increase the degree of the polynomial used for the closed formula. We go from constant to linear to quadratic. This makes sense. The sequence of differences between terms tells us something about the rate of growth of the sequence. If a sequence is growing at a constant rate, then the formula for the sequence will be linear. If the sequence is growing at a rate which itself is growing at a constant rate, then the formula is quadratic. You have seen this elsewhere - if a function has a constant second derivative (rate of change) then the function must be quadratic.

This works in general:

Finite Differences
The closed formula for a sequence will be a degree k polynomial if and only if the sequence is Δ^k - that is, the kth sequence of differences is constant.

This tells us that the sequence $1, 5, 14, 30, 55, \ldots$ will have a cubic (degree 3 polynomial) for its closed formula.

Now once we know what format the closed formula for a sequence will take, it is much easier to actually find the closed formula. In the case that the closed formula is a degree k polynomial, we just need $k + 1$ data points to "fit" the polynomial to the date.

Example: Find a formula for the sequence $3, 7, 14, 24, \ldots$. Assume $a_1 = 3$.

Solution: First, check to see if the formula has constant differences at some level. The sequence of first differences is $4, 7, 10, \ldots$ which is arithmetic, so the sequence of second differences is constant. The sequence is Δ^2, so the formula for a_n will be a degree 2 polynomial. That is, we know that for some constants a, b, and c,

$$a_n = an^2 + bn + c$$

Now to find a, b, and c. First, it would be nice to know what a_0 is, since plugging in $n = 0$ simplifies the above formula greatly. In this case, $a_0 = 2$ (work backwards from the sequence of constant differences). Thus

$$a_0 = 2 = a \cdot 0^2 + b \cdot 0 + c$$

so $c - 2$. Now plug in $n = 1$ and $n = 2$. We get

$$a_1 = 3 = a + b + 2$$

$$a_2 = 7 = a4 + b2 + 2.$$

At this point we have two (linear) equations and two unknowns, so we can solve the system for a and b. We find $a = \frac{3}{2}$ and $b = \frac{-1}{2}$, so $a_n = \frac{3}{2}n^2 - \frac{1}{2}n + 2$.

Example: Find a closed formula for the number of squares on an $n \times n$ chessboard.

Solution: We have seen that the sequence $1, 5, 14, 30, 55, \ldots$ is Δ^3, so we are looking for a degree 3 polynomial. That is,

$$a_n = an^3 + bn^2 + cn + d$$

We can find d if we know what a_0 is. Working backwards from the third differences, we find $a_0 = 0$ (which makes sense - there are no squares on a 0×0 chessboard). Thus $d = 0$. Now plug in $n = 1$, $n = 2$, and $n = 3$:

$$\begin{aligned} 1 &= a + b + c \\ 5 &= 8a + 4b + 2c \\ 14 &= 27a + 9b + 3c \end{aligned}$$

If we solve this system of equations (using elimination, or an augmented matrix, or a computer) we get $a = \frac{1}{3}$, $b = \frac{1}{2}$ and $c = \frac{1}{6}$. Therefore the number of squares on an $n \times n$ chessboard is $a_n = \frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n$. Note: since the squares on a chessboard problem really is asking for the sum of squares, we now have a nice formula for $\sum_{k=1}^{n} k^2$.

Not all sequences will have polynomials as their closed formula. That is, this process doesn't always work.

Example: Determine whether the following sequences can be described by a polynomial, and if so, of what degree.

1. $1, 2, 4, 8, 16, \ldots$
2. $0, 7, 50, 183, 484, 1055, \ldots$
3. $1, 1, 2, 3, 5, 8, 13, \ldots$

Solution:

1. The sequence of first differences is $1, 2, 4, 8, 16 \ldots$. This is identical to the original sequence, so taking any additional differences will not give anything different either. So there is no number of differences you could take to get constants, so the sequence is not Δ^k for any k. Therefore the closed formula for the sequence is not a polynomial. (In fact, we know the closed formula is $a_n = 2^{n-1}$, not a polynomial.)
2. First differences: $7, 43, 133, 301, 571, \ldots$. Second differences: $36, 90, 168, 270, \ldots$. Third difference: $54, 78, 102, \ldots$. Fourth differences: $24, 24, \ldots$ - constant. Thus the sequence is Δ^4, so the closed formula is a degree 4 polynomial.
3. This is the Fibonacci sequence. The first differences are $0, 1, 1, 2, 3, 5, 8, \ldots$, the second differences are $1, 0, 1, 1, 2, 3, 5 \ldots$ - we notice that after the first few terms, we get the original sequence back. So there will never be constant differences, so the closed formula for the Fibonacci sequence is not a polynomial.

Exercises

1. Use polynomial fitting to find the formula for the nth term of the following sequences:
   (a) 2, 5, 11, 21, 36,...
   (b) 0, 2, 6, 12, 20, ...
2. Can you use polynomial fitting to find the formula for the nth term of the sequence 4, 7, 11, 18, 29, 47, ...? Explain why or why not.

2.4 Solving Recurrence Relations

We have seen that it is often easier to find recursive definitions than closed formulas. Lucky for us, there are a few techniques for converting recursive definitions to closed formulas. Doing so is called solving a *recurrence relation*. Recall that the recurrence relation is the part of a recursive definition besides the initial conditions. For example, the recurrence relation for the Fibonacci sequence is $F_n = F_{n-1} + F_{n-2}$. (This, together with the initial conditions $F_0 = 0$ and $F_1 = 1$ give the entire recursive definition for the sequence.)

Example: Find a recurrence relation and initial conditions for the sequence $1, 5, 17, 53, 161, 485 \ldots$.

Solution: Finding the recurrence relation would be easier if we had some context for the problem (like the Tower of Hanoi, for example). Alas, we have only the sequence. Remember, the recurrence relation tells you how to get from previous terms to future terms. What is going on here? We could look at the differences between terms: $4, 12, 36, 108, \ldots$. Notice that these are growing by a factor of 3. Is the original sequence as well? $1 \cdot 3 = 3$, $5 \cdot 3 = 15$, $17 \cdot 3 = 51$ and so on. It appears that we always end up with 2 less than the next term. Aha!

So $a_n = 3a_{n-1} + 2$ is our recurrence relation. The initial condition is $a_1 = 1$.

We are going to try to *solve* the recurrence relations. By this we mean something very similar to solving differential equations: we want to find a function of n (a closed formula) which satisfies the recurrence relation, as well as the initial condition. (Note: recurrence relations are sometimes called difference equations since they can describe the difference between terms - this highlights

the relation to differential equations further.) Just like for differential equations, finding a solution might be tricky, but checking that the solution is correct is easy - just plug it in.

Example: Check that $a_n = 2^n + 1$ is a solution to the recurrence relation $a_n = 2a_{n-1} - 1$ with $a_1 = 3$.

Solution: First, it is easy to check the initial condition: a_1 should be $2^1 + 1$ according to our closed formula. But $2^1 + 1 = 3$, which is what we want. To check that our proposed solution satisfies the the recurrence relation, try plugging it in.

$$2a_{n-1} - 1 = 2(2^{n-1} + 1) - 1 = 2^n + 2 - 1 = 2^n + 1 = a_n$$

That's what our recurrence relation says! We have a solution.

Sometimes we can be clever and solve a recurrence relation by inspection. We generate the sequence using the recurrence relation and keep track of what we are doing so that we can see how to jump to finding just the a_n term. Here are two examples of how you might do that.

Telescoping refers to the phenomenon when many terms in a large sum cancel out - so the sum "telescopes." For example:

$$(2-1) + (3-2) + (3-4) + \cdots + (100-99) + (101-100) = -1 + 101$$

because every third term looks like: $2 + -2 = 0$, and then $3 + -3 = 0$ and so on.

We can use this behavior to solve recurrence relations. Here is an example.

Example: Solve the recurrence relation $a_n = a_{n-1} + n$ with initial term $a_0 = 4$.

Solution: To get a feel for the recurrence relation, write out the first few terms of the sequence: $4, 5, 7, 10, 14, 19, \ldots$. Look at the difference between terms. $a_1 - a_0 = 1$ and $a_2 - a_1 = 2$ and so on. The key thing here is that the difference between terms is n. We can write this explicitly: $a_n - a_{n-1} = n$. Of course, we could have arrived at this conclusion directly from the recurrence relation - just subtract a_{n-1} from both sides.

Now use this equation over and over again, changing n each time:

$$\begin{aligned} a_1 - a_0 &= 1 \\ a_2 - a_1 &= 2 \\ a_3 - a_2 &= 3 \\ \vdots \quad &\quad \vdots \\ a_n - a_{n-1} &= n \end{aligned}$$

Now add all these equations together. On the right hand side, we get the sum $1 + 2 + 3 + \cdots + n$. We already know that that can be simplified to $\frac{n(n+1)}{2}$. What happens on the left hand side? We get

$$(a_1 - a_0) + (a_2 - a_1) + (a_3 - a_2) + \cdots (a_{n-1} - a_{n-2}) + (a_n - a_{n-1})$$

This sum telescopes. We are left with only the $-a_0$ from the first equation and the a_n from the last equation. Putting this all together we have $-a_0 + a_n = \frac{n(n+1)}{2}$ or

$a_n = \frac{n(n+1)}{2} + a_0$. But we know that $a_0 = 4$. So the solution to the recurrence relation, subject to the initial condition is

$$a_n = \frac{n(n+1)}{2} + 4$$

(Now that we know that, we should notice that yes, the sequence is indeed the result of adding 4 to each of the triangular numbers.)

The above example shows a way to solve recurrence relations of the form $a_n = a_{n-1} + f(n)$ where $\sum_{k=1}^{n} f(k)$ has a known closed formula. If you rewrite the recurrence relation as $a_n - a_{n-1} = f(n)$, and then add up all the different equations with n ranging between 1 and n, the left hand side will always give you $a_n - a_0$. The right hand side will be $\sum_{k=1}^{n} f(k)$, which is why we need to know the closed formula for that sum.

However, telescoping will not help us with a recursion such as $a_n = 3a_{n-1} + 2$ - the left hand side will not telescope, since you will have $-3a_{n-1}$'s but only one a_{n-1}. However, we can still be clever if we use *iteration*.

We have already seen an example of iteration when we found the closed formula for arithmetic and geometric sequences. The idea is, we *iterate* the process of finding the next term, starting with the known initial condition, up until we have a_n. Then we simplify. In the arithmetic sequence example, we simplified by multiplying d by the number of times we add it to a when we get to a_n, to get from $a_n = a + d + d + d + \cdots + d$ to $a_n = a + d(n-1)$.

To see how this works, let's go through the same example we used for telescoping, but this time use iteration.

Example: Use iteration to solve the recurrence relation $a_n = a_{n-1} + n$ with $a_0 = 4$.

Solution: Again, start by writing down the recurrence relation when $n = 1$. This time, don't subtract the a_{n-1} terms to the other side.

$$a_1 = a_0 + 1$$

Now $a_2 = a_1 + 2$, but we know what a_1 is. We get

$$a_2 = (a_0 + 1) + 2$$

Now go to $a_3 = a_2 + 3$, using our known value of a_2:

$$a_3 = ((a_0 + 1) + 2) + 3$$

We notice a pattern. Each time, we take the previous term and add the current index. So

$$a_n = ((((a_0 + 1) + 2) + 3) + \cdots + n - 1) + n$$

Regrouping terms, we notice that a_n is just a_0 plus the sum of the integers from 1 to n. So, since $a_0 = 4$,

$$a_n = 4 + \frac{n(n+1)}{2}$$

Of course in this case we still needed to know formula for the sum of $1, \ldots, n$. Let's try iteration with a sequence for which telescoping doesn't work.

Example: Solve the recurrence relation $a_n = 3a_{n-1} + 2$ subject to $a_0 = 1$.

Solution: Again, we iterate the recurrence relation, building up to the index n.

$$\begin{aligned}
a_1 &= 3a_0 + 2 \\
a_2 &= 3(a_1) + 2 = 3(3a_0 + 2) + 2 \\
a_3 &= 3[a_2] + 2 = 3[3(3a_0 + 2) + 2] + 2 \\
&\vdots \quad \vdots \qquad \vdots \\
a_n &= 3(a_{n-1}) + 2 = 3(3(3(3 \cdots (3a_0 + 2) + 2) + 2) \cdots + 2) + 2
\end{aligned}$$

It is difficult to see what is happening here because we have to distribute all those 3's. Let' try again, this time simplifying a bit as we go.

$$\begin{aligned}
a_1 &= 3a_0 + 2 \\
a_2 &= 3(a_1) + 2 = 3(3a_0 + 2) + 2 = 3^2 a_0 + 2 \cdot 3 + 2 \\
a_3 &= 3[a_2] + 2 = 3[3^2 a_0 + 2 \cdot 3 + 2] + 2 = 3^3 a_0 + 2 \cdot 3^2 + 2 \cdot 3 + 2 \\
&\vdots \quad \vdots \qquad \vdots \\
a_n &= 3(a_{n-1}) + 2 = 3(3^{n-1} a_0 + 2 \cdot 3^{n-2} + \cdots + 2) + 2 \\
&= 3^n a_0 + 2 \cdot 3^{n-1} + 2 \cdot 3^{n-2} + \cdots + 2 \cdot 3 + 2
\end{aligned}$$

Now we simplify. $a_0 = 1$, so we have $3^n + (\cdots)$. Note that every other term has a 2 in it - in fact, we have a geometric sum with first term 2 and common ratio 3. We have seen how to simplify $2 + 2 \cdot 3 + 2 \cdot 3^2 + \cdots + 2 \cdot 3^{n-1}$. We get $\frac{2-2\cdot 3^n}{-2}$ which simplifies to $3^n - 1$. Putting this together with the first 3^n term gives our closed formula:

$$a_n = 2 \cdot 3^n - 1$$

Iteration can be messy, but when the recurrence relation only refers to one previous term (and maybe some function of n) it can work well. However, trying to iterate a recurrence relation such as $a_n = 2a_{n-1} + 3a_{n-2}$ will be way to complicated - we would need to keep track of two sets of previous terms, each of which were expressed by two previous terms, and so on. The length of the formula would grow exponentially (double each time, in fact). Luckily there happens to be a method for solving recurrence relations which works very well on relations like this.

2.4.1 The Characteristic Root Technique

Suppose we want to solve a recurrence relation expressed as a combination of the two previous terms, such as $a_n = a_{n-1} + 6a_{n-2}$. In other words, we want to find a function of n which satisfies $a_n - a_{n-1} - 6a_{n-2} = 0$. Now iteration is too complicated, but think just for a second what would happen if we *did* iterate. In each step, we would, among other things, multiply a previous iteration by 6. So our closed formula would have include 6 multiplied some number of times. Thus it is reasonable to guess the solution will contain parts that look geometric. So perhaps the solution will take the form r^n. for some constant r.

The nice thing is, we know how to check whether a formula is actually a solution to a recurrence relation - plug it in. What happens if we plug in r^n into the recursion above? We get

$$r^n - r^{n-1} - 6r^{n-2} = 0$$

Now solve for r:

$$r^{n-2}(r^2 - r - 6) = 0$$

so by factoring, $r = -2$ or $r = 3$. This tells us that $a_n = (-2)^n$ is a solution to the recurrence relation, as is $a_n = 3^n$. Which one is correct? They both are, unless we specify initial conditions. Notice we could also have $a_n = (-2)^n + 3^n$. Or $a_n = 7(-2)^n + 4 \cdot 3^n$. In fact, for any a and b, $a_n = a(-2)^n + b3^n$ is a solution - try plugging that into the recurrence relation. To find the values of a and b, use the initial conditions.

The points us in the direction of a more general technique for solving recurrence relations. Notice we will always be able to factor out the r^{n-2} as we did above. So we really only care about the other part. We call this other part the *characteristic equation* for the recurrence relation. We are interested in finding the roots of the characteristic equation, which are called (surprise) the characteristic roots.

Characteristic Roots
Given a recurrence relation $a_n + \alpha a_{n-1} + \beta a_{n-2} = 0$, the characteristic polynomial is

$$x^2 + \alpha x + \beta$$

giving the *characteristic equation*:

$$x^2 + \alpha x + \beta = 0$$

If r_1 and r_2 are two distinct roots of the characteristic polynomial (i.e, solutions to the characteristic equation), then the solution to the recurrence relation is

$$a_n = ar_1^n + br_2^n$$

where a and b are constants determined by the initial conditions.

Example: Solve the recurrence relation $a_n = 7a_{n-1} - 10a_{n-2}$ with $a_0 = 2$ and $a_1 = 3$.

Solution: Rewrite the recurrence relation $a_n - 7a_{n-1} + 10a_{n-2} = 0$. Now form the characteristic equation:

$$x^2 - 7x + 10 = 0$$

and solve for x:

$$(x-2)(x-5) = 0$$

so $x = 2$ and $x = 5$ are the characteristic roots. We therefore know that the solution to the recurrence relation will have the form

$$a_n = a2^n + b5^n$$

To find a and b, plug in $n = 0$ and $n = 1$ to get a system of two equations with two unknowns:

$$2 = a2^0 + b5^0 = a + b$$
$$3 = a2^1 + b5^1 = 2a + 5b$$

Solving this system gives $a = \frac{7}{3}$ and $b = -\frac{1}{3}$ so the solution to the recurrence relation is

$$a_n = \frac{7}{3}2^n - \frac{1}{3}3^n$$

Perhaps the most famous recurrence relation is $F_n = F_{n-1} + F_{n-2}$, which together with the initial conditions $F_0 = 0$ and $F_1 = 1$ defines the Fibonacci sequence. But notice that this is precisely the type of recurrence relation on which we can use the characteristic root technique. When you do, the only thing that changes is that the characteristic equation does not factor - you need to use the quadratic formula to find the characteristic roots. In fact, doing so gives the third most famous irrational number - φ, the golden ratio.

Before leaving the characteristic root technique, we should think about what all might happen when you solve the characteristic equation. Above we have an example in which the characteristic polynomial has two distinct roots. These roots can be integers, or perhaps irrational numbers (requiring the quadratic formula to find them). In these cases, we know what the solution to the recurrence relation looks like.

However, it is possible for the characteristic polynomial to only have one root - say the characteristic polynomial factors as $(x-r)^2$. It is still the case that r^n would be a solution to the recurrence relation, but we won't be able to find solutions for all initial conditions using the general form $a_n = ar_1^n + br_2^n$, since we can't distinguish between r_1^n and r_2^n - they are the same (repeated) root. We are in luck though:

Characteristic Root Technique for Repeated Roots
Suppose the recurrence relation $a_n = \alpha a_{n-1} + \beta a_{n-2}$ has a characteristic polynomial with only one root r. Then the solution to the recurrence relation is

$$a_n = ar^n + bnr^n$$

where a and b are constants determined by the initial conditions.

Notice the extra n in bnr^n. This allows us to solve for the constants a and b from the initial conditions.

Example: Solve the recurrence relation $a_n = 6a_{n-1} - 9a_{n-2}$ with initial conditions $a_0 = 1$ and $a_1 = 4$.

Solution: The characteristic polynomial is $x^2 - 6x + 9$. We solve the characteristic equation

$$x^2 - 6x + 9 = 0$$

by factoring:

$$(x-3)^2 = 0$$

so $x = 3$ is the only characteristic root. Therefore we know that the solution to the recurrence relation has the form

$$a_n = a3^n + bn3^n$$

for some constants a and b. Now use the initial conditions:

$$a_0 = 1 = a3^0 + b \cdot 0 \cdot 3^0 = a$$
$$a_1 = 4 = a \cdot 3 + b \cdot 1 \cdot 3 = 3a + 3b$$

Since $a = 1$, we find that $b = \frac{1}{3}$. Therefore the solution to the recurrence relation is

$$a_n = 3^n + \frac{1}{3}n3^n$$

Although we will not consider examples more complicated than these, this characteristic root technique can be applied to much more complicated recurrence relations. For example, $a_n = 2a_{n-1} + a_{n-2} - 3a_{n-3}$ has characteristic polynomial $x^3 - 2x^2 - x + 3$. Assuming you see how to factor such a degree 3 (or more) polynomial you can easily find the characteristic roots and as such solve the recurrence relation (the solution would look like $a_n = ar_1^n + br_2^n + cr_3^n$ if there were 3 distinct roots). It is also possible to solve recurrence relations of the form $a_n = \alpha a_{n-1} + \beta a_{n-2} + C$ for some constant C. Additionally, if the characteristic roots could be complex numbers - this is acceptable as well.

Exercises

1. Find the next 2 terms in the sequence $3, 5, 11, 21, 43, 85 \ldots$. Then give a recursive definition for the sequence. Finally, use the characteristic root technique to find a closed formula for the sequence.
2. Solve the recurrence relation $a_n = a_{n-1} + 2^n$ with $a_0 = 5$.
3. Show that 4^n is a solution to the recurrence relation $a_n = 3a_{n-1} + 4a_{n-2}$.
4. Find the solution to the recurrence relation $a_n = 3a_{n-1} + 4a_{n-2}$ with initial terms $a_0 = 2$ and $a_1 = 3$.

5. Find the solution to the recurrence relation $a_n = 3a_{n-1} + 4a_{n-2}$ with initial terms $a_0 = 5$ and $a_1 = 8$.

6. Solve the recurrence relation $a_n = 2a_{n-1} - a_{n-2}$.

   (a) What is the solution if the initial terms are $a_0 = 1$ and $a_1 = 2$?

   (b) What do the initial terms need to be in order for $a_9 = 30$?

   (c) For which x are there initial terms which make $a_9 = x$?

7. Solve the recurrence relation $a_n = 3a_{n-1} + 10a_{n-2}$ with initial terms $a_0 = 4$ and $a_1 = 1$.

2.5 Induction

Mathematical induction is a proof technique, not unlike direct proof or proof by contradiction or combinatorial proof.[1] In other words, induction is a style of argument we use to convince ourselves and others that a mathematical statement is always true. Many mathematical statements can be proven by simply explaining what they mean. Others are very difficult to prove – in fact, there are relatively simple mathematical statements which nobody yet knows how to prove. To facilitate the discovery of proofs, it is important to be familiar with some standard styles of arguments. Induction is one such style. Let's start with an example.

2.5.1 Stamps

Here is a question about making different amounts of postage:

> You need to mail a package, but don't yet know how much postage you will need. You have a large supply of 8-cent stamps and 5-cent stamps. Which amounts of postage can you make exactly using these stamps? Which amounts are impossible to make?

Try this problem on your own first. Perhaps in investigating the problem you picked some amounts of postage, and then figured out whether you could make that amount using just 8-cent and 5-cent stamps. Perhaps you did this in order - can you make 1 cents of postage? Can you make 2 cents? 3 cents? And so on. If this is what you did, you were actually answering a *sequence* of questions. We have methods for dealing with sequences. Let's see if that helps.

Actually, we will not make a sequence of questions, but rather a sequence of statements. Let $P(n)$ be the statement "you can make n cents of postage using just 8-cent and 5-cent stamps." Since each $P(n)$ is a statement, it is either true or false. So if we form the sequence of statements

$$P(1), P(2), P(3), P(4), \ldots$$

the sequence will consist of T's (for true) and F's (for false). In our particular case the sequence starts

$$F, F, F, F, T, F, F, T, F, F, T, F, F, T, \ldots$$

because $P(1), P(2), P(3), P(4)$ are all false (you cannot make 1, 2, 3, or 4 cents of postage) but $P(5)$ is true (use one 5-cent stamp), and so on.

Let's think a bit about how we could find the value of $P(n)$ for some specific value of n (the "value" will be either T or F). How did we find the value of the nth term of a sequence of numbers?

[1] You might or might not be familiar with these yet, but you will be soon.

How did we find a_n? There were two ways we could do this: either there was a closed formula for a_n, so we could plug in n into the formula and get our output value, or we had a recursive definition for the sequence, so we could use the previous terms of the sequence to compute the nth term. When dealing with sequences of statements, we could use either of these techniques as well. Maybe there is a way to use n itself to determine whether we can make n cents of postage. That would be something like a closed formula. Or instead we could use the previous terms in the sequence (of statements) to determine whether we can make n cents of postage. That is, if we know the value of $P(n-1)$, can we get from that to the value of $P(n)$? That would be something like a recursive definition for the sequence. Remember, finding recursive definitions for sequences was often easier than finding closed formulas. The same is true here.

Suppose I told you that $P(43)$ was true (it is). Can you determine from this fact the value of $P(44)$ (whether it true or false)? Yes you can. Even if we don't know how exactly we made 43 cents out of the 5-cent and 8-cent stamps, we do know that there was some way to do it. What if that way used at least three 5-cent stamps (making 15 cents)? We could replace those three 5-cent stamps with two 8-cent stamps (making 16 cents). The total postage has gone up by 1, so we have a way to make 44 cents, so $P(44)$ is true. Of course, we assumed that we had at least three 5-cent stamps. What if we didn't? Then we must have at least three 8-cent stamps (making 24 cents). If we replace those three 8-cent stamps with five 5-cent stamps (making 25 cents) then again we have bumped up our total by 1 cent so we can make 44 cents, so $P(44)$ is true.

Notice that we have not said how to make 44 cents, just that we can, on the basis that we can make 43 cents. How do we know we can make 43 cents? Perhaps because we know we can make 42 cents, which we know we can do because we know we can make 41 cents, and so on. It's a recursion! As with a recursive definition of a numerical sequence, we must specify our initial value. In this case, the initial value is "$P(1)$ is false." That's not good, since our recurrence relation just says that $P(k+1)$ is true *if* $P(k)$ is also true. We need to start the process with a true $P(k)$. So instead, we might want to use "$P(31)$ is true" as the initial condition.

Putting this all together we arrive at the following fact: it is possible to (exactly) make any amount of postage greater than 27 cents using just 5-cent and 8-cent stamps.[2] In other words, $P(k)$ is true for any $k \geq 28$. To prove this, we could do the following:

1. Demonstrate that $P(28)$ is true.

2. Prove that if $P(k)$ is true, then $P(k+1)$ is true (for any $k \geq 28$).

Suppose we have done this. Then we know that the 28th term of the sequence above is a T (using (1) - the initial condition or *base case*), and that every term after the 28th is T also (using (2) - the recursive part or *inductive case*). Here is what the proof would actually look like.

Proof. Let $P(n)$ be the statement "it is possible to make exactly n cents of postage using 5-cent and 8-cent stamps." We will show $P(n)$ is true for all $n \geq 28$.

First, we show that $P(28)$ is true: $28 = 4 \cdot 5 + 1 \cdot 8$, so we can make 28 cents using four 5-cent stamps and one 8-cent stamp.

Now suppose $P(k)$ is true for some arbitrary $k \geq 28$. Then it is possible to make k cents using 5-cent and 8-cent stamps. Note that since $k \geq 28$, it cannot be that we use less than three 5-cent stamps *and* less than three 8-cent stamps: using two of each would give only 26 cents. Now if we have made k cents using at least three 5-cent stamps, replace three 5-cent stamps by two 8-cent stamps. This replaces 15 cents of postage with 16 cents, moving from a total of k cents to $k+1$

[2] This is not claiming that there are no amounts less than 27 cents which can also be made - there are.

cents - so $P(k+1)$ is true. On the other hand, if we have made k cents using at least three 8-cent stamps, then we can replace three 8-cent stamps with five 5-cent stamps, moving from 24 cents to 25 cents, giving a total of $k+1$ cents of postage. So in this case as well $P(k+1)$ is true.

Therefore, by the principle of mathematical induction, $P(n)$ is true for all $n \geq 28$. QED

2.5.2 Formalizing proofs

What we did in the stamp example above works for many types of problems. Proof by induction is useful when trying to prove statements about all natural numbers, or all natural numbers greater than some fixed first case (like 28 in the example above), and in some other situations besides. In particular, induction should be used when there is some way to go from one case to the next - when you can see how to always "do one more."

This is a big idea. Thinking about a problem *inductively* can give new insight into the problem. For example, to really understand the stamp problem, you should think about how any amount of postage (greater than 28 cents) can be made (this in non-inductive reasoning) and also how the ways in which postage can be made *changes* as the amount increases (inductive reasoning). When you are asked to provide a proof by induction, you are being asked to think about the problem *dynamically*; how does increasing n change the problem?

But there is another side to proofs by induction as well. In mathematics, it is not enough to understand a problem, you must also be able to communicate the problem to others. Like any discipline, mathematics has standard language and style, allowing mathematicians to share their ideas efficiently. Proofs by induction have a certain formal style, and being able to write in this style is important. It allows us to keep our ideas organized and might even help us with formulating a proof.

Here is the general structure of a proof by mathematical induction:

Induction Proof Structure
Start by saying what the statement is which you want to prove: "Let $P(n)$ be the statement..."
To prove that $P(n)$ is true for all $n \geq 0$, you must prove two facts:

1. Base case: Prove that $P(0)$ is true. You do this directly. This is often easy.
2. Inductive case: Prove that $P(k) \rightarrow P(k+1)$ for all $k \geq 0$. That is, prove that for any $k \geq 0$ if $P(k)$ is true, then $P(k+1)$ is true as well. This is the proof of an if...then... statement, so you can assume $P(k)$ is true ($P(k)$ is called the *inductive hypothesis*). You must then explain why $P(k+1)$ is also true, given that assumption.

Assuming you are successful on both parts above, you can conclude, "Therefore by the principle of mathematical induction, $P(n)$ is true for all $n \geq 0$."

Sometimes the statement $P(n)$ will only be true for values of $n \geq 4$, for example, or some other value. In such cases, replace all the 0's above with 4's (or the other value).

The other advantage of formalizing inductive proofs is it allows us to verify that the logic behind this style of argument is valid. Why does induction work? Think of a row of dominoes set up standing on their edges. We want to argue that in a minute, all the dominoes will have fallen down. For this to happen, you will need to push the first domino - that is the base case. It will also have to be that the dominoes are close enough together that when any particular domino falls, it will cause the next domino to fall - that is the inductive case. If both of these conditions are met -

you push the first domino over and each domino will cause the next to fall, then all the dominoes will fall.

Induction is powerful! Think how much easier it is to knock over dominoes when you don't have to push over each domino yourself. You just start the chain reaction, and the rely on the relative nearness of the dominoes to take care of the rest.

Think about our study of sequences. It is easier to find recursive definitions for sequences than closed formulas - going from one case to the next is easier than going directly to a particular case. That is what is so great about induction. Instead of going directly to the (arbitrary) case for n, we just need to say how to get from one case to the next.

When you are asked to prove a statement by mathematical induction, you should first think about *why* the statement is true, using inductive reasoning. Explain why induction is the right thing to do, and roughly why the inductive case will work. Then, sit down and write out a careful, formal proof using the structure above.

2.5.3 Examples

Here are some examples of proof my mathematical induction.

Example: Prove for each natural number $n \geq 1$ that $1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}$.

Solution: First, let's think inductively about this equation. In fact, we know this is true for other reasons (reverse and add comes to mind). But why might induction be applicable? The left hand side is adding up the numbers from 1 to n. If we know how to do that, adding just one more term $(n + 1)$ would not be that hard. For example, if $n = 100$, suppose we know that the sum of the first 100 numbers is 5050 (so $1 + 2 + 3 + \cdots + 100 = 5050$, which is true). Now to find the sum of the first 101 numbers, it makes more sense to just add 101 to 5050, instead of computing the entire sum again. We would have $1 + 2 + 3 + \cdots + 100 + 101 = 5050 + 101 = 5151$. In fact, it would always be easy to add just one more term. This is why we should use induction.

Now the formal proof.

Proof. Let $P(n)$ be the statement $1 + 2 + 3 + \cdots + n = \frac{n(n+2)}{2}$. We will show that $P(n)$ is true for all natural numbers $n \geq 1$.

Base case: $P(1)$ is the statement $1 = \frac{1(1+1)}{2}$ which is clearly true.

Inductive case: Let $k \geq 1$ be a natural number. Assume (for induction) that $P(k)$ is true. That means $1 + 2 + 3 + \cdots + k = \frac{k(k+1)}{2}$. We will prove that $P(k + 1)$ is true as well. That is, we must prove that $1 + 2 + 3 + \cdots + k + (k + 1) = \frac{(k+1)(k+2)}{2}$. To prove this equation, start with the left hand side:

$$1 + 2 + 3 + \cdots + k + (k + 1) = \frac{k(k + 1)}{2} + (k + 1)$$

by the induction hypothesis. Now simplifying:

$$\frac{k(k + 1)}{2} + k + 1 = \frac{k(k + 1)}{2} + \frac{2(k + 1)}{2} = \frac{k(k + 1) + 2(k + 1)}{2} = \frac{(k + 2)(k + 1)}{2}$$

Thus $P(k + 1)$ is true, so by the principle of mathematical induction $P(n)$ is true for all natural numbers $n \geq 1$. QED

Note that in the part of the proof in which we proved $P(k+1)$ from $P(k)$, we used the equation $P(k)$ - this was the inductive hypothesis. This is very easy to do with proving a fact about a sum like this. Sometimes it is not as easy to see where to use the inductive hypothesis.

Example: Prove that for all $n \in \mathbb{N}$, $6^n - 1$ is a multiple of 5.

Solution: Again, start by understanding the dynamics of the problem. What does increasing n do? Let's try with a few examples. If $n = 1$, then yes, $6^1 - 1 = 5$ is a multiple of 5. What does incrementing n to 2 look like? We get $6^2 - 1 = 35$, which again is a multiple of 5. Next, $n = 3$: but instead of just finding $6^3 - 1$, what did the increase in n do? We will still subtract 1, but now we are multiplying by another 6 first. Viewed another way, we are multiplying a number which is one more than a multiple of 5 by 6 (because $6^2 - 1$ is a multiple of 5, so 6^2 is one more than a multiple of 5). What do numbers which are one more than a multiple of 5 look like? The must have last digit 1 or 6. What happens when you multiply such a number by 6? Depends on the number, but in any case, the last digit of the new number must be a 6. And then if you subtract 1, you get last digit 5, so a multiple of 5.

The point is, every time we multiply by just one more six, we still get a number with last digit 6, so subtracting 1 gives us a multiple of 5. Now the formal proof:

Proof. Let $P(n)$ be the statement, "$6^n - 1$ is a multiple of 5." We will prove that $P(n)$ is true for all $n \in \mathbb{N}$.

Base case: $P(0)$ is true: $6^0 - 1 = 0$ which is a multiple of 5.

Inductive case: Let k be an arbitrary natural number. Assume, for induction, that $P(k)$ is true. That is, $6^k - 1$ is a multiple of 5. Then $6^k - 1 = 5j$ for some integer j. This means that $6^k = 5j + 1$. Multiply both sides by 6:

$$6^{k+1} = 6(5j + 1) = 30j + 6$$

But we want to know about $6^{k+1} - 1$, so subtract 1 from both sides:

$$6^{k+1} - 1 = 30j + 5$$

Of course $30j + 5 = 5(6j + 1)$, so is a multiple of 5.

Therefore $6^{k+1} - 1$ is a multiple of 5, or in other words, $P(k + 1)$ is true. Thus, by the principle of mathematical induction $P(n)$ is true for all $n \in \mathbb{N}$. QED

We had to be a little bit clever (i.e., use some algebra) to locate the $6^k - 1$ inside of $6^{k+1} - 1$ before we could apply the inductive hypothesis. This is what can make inductive proofs challenging.

In the two examples above, we started with $n = 1$ or $n = 0$. We can start later if we need to.

Example: Prove that $n^2 < 2^n$ for all integers $n \geq 5$.

Solution: First, the idea of the argument. What happens when we increase n by 1? On the left hand side, we increase the base of the square - we go to the next square number. On the right hand side, we increase the power of 2. This means we double the number. So the question is, how does doubling a number relate to increasing to the next square? Think about what the difference of two consecutive squares looks like. We have $(n + 1)^2 - n^2$. This factors:

$$(n + 1)^2 - n^2 = (n + 1 - n)(n + 1 + n) = 2n + 1$$

(this should not be a surprise - we already know that the sum of consecutive odd numbers gives the perfect squares). But doubling the right hand side increases it by 2^n, since $2^{n+1} = 2^n + 2^n$. Of course when n is large enough, $2^n > 2n + 1$.

What we are saying here is that each time n increases, the left hand side grows by less than the right hand side. So if the left hand side starts smaller (as it does when $n = 5$), it will never catch up. Now the formal proof:

Proof. Let $P(n)$ be the statement $n^2 < 2^n$.[3] We will prove $P(n)$ is true for all integers $n \geq 5$.

Base case: $P(5)$ is the statement $5^2 < 2^5$. Since $5^2 = 25$ and $2^5 = 32$, we see that $P(5)$ is indeed true.

Inductive case: Let $k \geq 5$ be an arbitrary integer. Assume, for induction, that $P(k)$ is true. That is, assume $k^2 < 2^k$. We will prove that $P(k + 1)$ is true, i.e., $(k + 1)^2 < 2^{k+1}$. To prove such an inequality, start with the left hand side and work towards the right hand side:

$$\begin{aligned}
(k+1)^2 &= k^2 + 2k + 1 \\
&< 2^k + 2k + 1 && \ldots \text{by the inductive hypothesis} \\
&< 2^k + 2^k && \ldots \text{since } 2k + 1 < 2^k \text{ for } k \geq 5 \\
&= 2^{k+1}
\end{aligned}$$

Following the equalities and inequalities through, we get $(k + 1)^2 < 2^{k+1}$, in other words, $P(k + 1)$. Therefore by the principle of mathematical induction, $P(n)$ is true for all $n \geq 5$. QED

The previous example might remind you of the *racetrack principle* from calculus, which says that if $f(a) < g(a)$, and $f'(x) < g'(x)$ for $x > a$, then $f(x) < g(x)$ for $x > a$. Same idea: the larger function is increasing at a faster rate than the smaller function, so the larger function will stay larger. In discrete math, we don't have derivatives, so we look at differences - when n increase by 1. Thus induction is the way to go.

Warning:

With great power, comes great responsibility. Induction isn't magic. It seems very powerful to be able to assume $P(k)$ is true - after all, we are trying to prove $P(n)$ is true (and the only difference is in the variable: k vs. n). Are we assuming that what we want to prove is true? Not really. We assume $P(k)$ is true only for the sake of proving that $P(k + 1)$ is true.

Still you might start to believe that you can prove anything with induction. Consider this incorrect "proof" that every Canadian has the same eye color: Let $P(n)$ be the statement that any n Canadians have the same eye color. $P(1)$ is true, since everyone has the same eye color as themselves. Now assume $P(k)$ is true. That is, assume that in any group of k Canadians, everyone has the same eye color. Now consider an arbitrary group of $k + 1$ Canadians. The first k of these must all have the same eye color, since $P(k)$ is true. Also, the last k of these must have the same eye color, since $P(k)$ is true. So in fact, everyone the group must have the same eye color. Thus $P(k + 1)$ is true. So by the principle of mathematical induction, $P(n)$ is true for all n.

[3] $P(n)$ is not the statement "$n^2 < 2^n$ for all integers $n \geq 5$" - you do not include the quantifier on n in the statement of $P(n)$.

Clearly something went wrong. The problem is that the proof that $P(k)$ implies $P(k+1)$ assumes that $k \geq 2$. We have only shown $P(1)$ is true. In fact, $P(2)$ is false.

2.5.4 Strong Induction

Consider the following classic puzzle:

> You have 9 coins which all look identical. However, you know that one of the coins is counterfeit, and weighs slightly less than the other coins (which all have the same weight). The weight difference is not great enough to tell by holding the coins, but luckily you have an old-fashioned balance scale (like the one held by Lady Justice). What is the smallest number of weighings needed to find the counterfeit coin?

If you have never tried to solve this puzzle, you should try now. Don't read on until you have given it some thought.

Need a hint? The answer is 2. Why?

Here's what you do. Put 3 coins on each side of the balance scale, leaving 3 off. Now if the scales are balanced, then the light coin must be one of the 3 you didn't weigh. If the scale tips, then the light coin is one of the three on the higher side. Now you have eliminated all but 3 coins using just one weighing. For the second weighing, put one of the three suspect coins on each side of the scales (leaving one off). Same idea: if the scales balance, then the unweighed coin is the counterfeit. Otherwise, the side of the scale that tipped up contains the light coin. Neat huh?

Let's generalize? If you had more than 9 coins to start with, say you had n coins, how many weighings would you need? A reasonable guess would be that each time you weigh the coins, you can eliminate 2/3 of them. So using 3 weighings, you could find the counterfeit coin among up to 27 coins. A forth weighing would be required for 28 to 81 coins.

Conjecture. *The minimum number of weighings required to find the counterfeit coin among n coins is $\log_3(n)$, rounded up. In other words, if x is the smallest integer such that $n \leq 3^x$, then the minimum number of weighings is x.*

How do we know this is true? If you have n coins, divide them into 3 equal piles (if this is not possible, make two piles the same size, and the third pile one more or less as needed). Put one pile on each side of the scale, leaving one pile off (if the piles were not equal, weigh the piles with an equal number of coins against each other). If the scales balance, the counterfeit coin is in the unweighed pile. If the scales tip, the counterfeit coin is in the light pile. In any event, we have now reduced the number of coins to roughly $n/3$.

What now? Well obviously we do this again. And again, until we find our bad coin. Here is where induction comes in: instead of going through the process over and over again, we can just say that we have reduced the problem to an earlier case, for which we assume our conjecture holds. For example if $n = 81$, we do the weighing as described, leaving us with $n/3 = 27$ coins. By induction, we already know how to complete the problem for 27 coins, so we are done.

This "reducing to a simpler case" is a type of induction. But not exactly the same sort of induction we have encountered so far. In ordinary induction, we assume our result holds for k, and prove that it holds for the next higher case: $k+1$. There is not much difference in showing that the result holds for the case k assuming that it holds for the previous case $(k-1)$. But what we are doing here, is showing that the result holds for the case k, assuming it works for *all* smaller cases. Is this okay? In fact, it is. Think of the dominoes: to prove that they all fall down, it is enough to prove that the first one falls, and that if all the dominoes up to the kth one have fallen, then the kth one will fall too.

The advantage is that we now have a stronger inductive hypothesis. We can assume that $P(1)$, $P(2)$, $P(3)$,... $P(k)$ is true, just to show that $P(k+1)$ is true. Previously, we just assumed $P(k)$ for this purpose.

It is slightly easier to change our variables for strong induction. Here is what the formal proof would look like:

Strong Induction Proof Structure
Again, start by saying what you want to prove: "Let $P(n)$ be the statement..." Then establish two facts:

1. Base case: Prove that $P(0)$ is true.
2. Inductive case: Assume $P(k)$ is true for all $k < n$. Prove that $P(n)$ is true.

Conclude, "therefore, by strong induction, $P(n)$ is true for all $n > 0$."

Of course, it is acceptable to replace 0 with a larger base case if needed.[4]

Now let's prove our conjecture about the coin weighing puzzle.

Proof. Let $P(n)$ be the statement, "if x is the smallest natural number such that $n \leq 3^x$, then it is possible to find the light coin in no more than x weighings."

Base case: Consider $P(2)$. Here x is 1, since $2 \leq 3^1$. It is possible to find the heavy coin using just 1 weighing - put the coins on the scale.

Inductive case: Suppose we have n coins, and assume $P(k)$ is true for all $k < n$. Now divide the n coins into three piles - equal if possible, otherwise two equal with the third either one more or one less than the other two. Put two equal piles on each side of the scales. If the scales balance, then the counterfeit coin is in the unweighed pile. If the scales tip, the coin is in the pile on the lighter side of the scale. In any event, we now know the counterfeit coin is one of $k = n/3$ or $k = n/3 \pm 1$ coins. What's more, if x is the smallest number such that $n \leq 3^x$, then $k \leq 3^{x-1}$. By the inductive hypothesis, it is now possible to find the counterfeit coin in $x - 1$ more weighings. Thus $P(n)$ is true.

Therefore, by strong induction, $P(n)$ is true for all $n \geq 2$. QED

Here is a more mathematically relevant example.

Example: Prove that any natural number greater than 1 is either prime or can be written as the product of primes.

Solution: First, the idea: if we take some number n, maybe it is prime. If so, we are done. If not, then it is composite, so it is the product of two smaller numbers. Each of these factors is smaller than n (but at least 2), so we can repeat the argument with these numbers. We have reduced to a smaller case.

Now the formal proof:

Proof. Let $P(n)$ be the statement, "n is either prime or can be written as the product of primes." We will prove $P(n)$ is true for all $n \geq 2$.

Base case: $P(2)$ is true because 2 is indeed prime.

Inductive case: assume $P(k)$ is true for all $k < n$. We want to show that $P(n)$ is true. That is, we want to show that n is either prime or is the product of primes. If n is prime, we are done. If not, then n has more than 2 divisors, so we can write $n = m_1 \cdot m_2$, with m_1 and m_2 less than n (and greater than 1). By the inductive

[4] Technically, strong induction does not require you to prove a separate base case. This is because when proving the inductive case, you must show that $P(0)$ is true, assuming $P(k)$ is true for all $k < 0$. But this is not any help - so you end up proving $P(0)$ anyway. To be on the safe side, we will always include the base case separately.

hypothesis, m_1 and m_2 are each either prime or can be written as the product of primes. In either case, we have that n is written as the product of primes.

Thus by the strong induction, $P(n)$ is true for all $n \geq 2$. QED

Exercises

1. Use induction to prove for all $n \in \mathbb{N}$ that $\sum_{k=0}^{n} 2^k = 2^{n+1} - 1$.
2. Prove that $7^n - 1$ is a multiple of 6 for all $n \in \mathbb{N}$.
3. Prove that $1 + 3 + 5 + \cdots + (2n-1) = n^2$ for all $n \geq 1$.
4. Prove that $F_0 + F_2 + F_4 + \cdots + F_{2n} = F_{2n+1} - 1$ where F_n is the nth Fibonacci number.
5. Prove that $2^n < n!$ for all $n \geq 4$. (Recall, $n! = 1 \cdot 2 \cdot 3 \cdot \cdots \cdot n$.)
6. What is wrong with the following "proof" of the "fact" that $n + 3 = n + 7$ for all values of n (besides of course that the thing it is claiming to prove is false)?

   Proof. Let $P(n)$ be the statement that $n + 3 = n + 7$. We will prove that $P(n)$ is true for all $n \in \mathbb{N}$. Assume, for induction that $P(k)$ is true. That is, $k + 3 = k + 7$. We must show that $P(k+1)$ is true. Now since $k+3 = k+7$, add 1 to both sides. This gives $k+3+1 = k+7+1$. Regrouping $(k + 1) + 3 = (k + 1) + 7$. But this is simply $P(k + 1)$. Thus by the principle of mathematical induction $P(n)$ is true for all $n \in \mathbb{N}$. QED

7. The proof in the previous problem does not work. But if we modify the "fact," we can get a working proof. Prove that $n + 3 < n + 7$ for all values of $n \in \mathbb{N}$. You can do this proof with algebra (without induction), but the goal of this exercise is to write out a valid induction proof.
8. Find the flaw in the following "proof" of the "fact" that $n < 100$ for every $n \in \mathbb{N}$.

   Proof. Let $P(n)$ be the statement $n < 100$. We will prove $P(n)$ is true for all $n \in \mathbb{N}$. First we establish the base case: when $n = 0$, $P(n)$ is true, because $0 < 100$. Now for the inductive step, assume $P(k)$ is true. That is, $k < 100$. Now if $k < 100$, then k is some number, like 80. Of course $80 + 1 = 81$ which is still less than 100. So $k + 1 < 100$ as well. But this is what $P(k + 1)$ claims, so we have shown that $P(k) \to P(k + 1)$. Thus by the principle of mathematical induction, $P(n)$ is true for all $n \in \mathbb{N}$. QED

9. While the above proof does not work (it better not - the statement it is trying to prove is false!) we can prove something similar. Prove that there is a strictly increasing sequence $a_1, a_2, a_3, \ldots$ of numbers (not necessarily integers) such that $a_n < 100$ for all $n \in \mathbb{N}$. (By *strictly increasing* we mean $a_n < a_{n+1}$ for all n - so each term must be larger than the last.)
10. What is wrong with the following "proof" of the "fact" that for all $n \in \mathbb{N}$, the number $n^2 + n$ is odd?

Proof. Let $P(n)$ be the statement "$n^2 + n$ is odd." We will prove that $P(n)$ is true for all $n \in \mathbb{N}$. Suppose for induction that $P(k)$ is true, that is, that $k^2 + k$ is odd. Now consider the statement $P(k+1)$. Now $(k+1)^2 + (k+1) = k^2 + 2k + 1 + k + 1 = k^2 + k + 2k + 2$. By the inductive hypothesis, $k^2 + k$ is odd, and of course $2k + 2$ is even. An odd plus an even is always odd, so therefore $(k+1)^2 + (k+1)$ is odd. Therefore by the principle of mathematical induction, $P(n)$ is true for all $n \in \mathbb{N}$. QED

11. Now give a valid proof (by induction - even though you might be able to do so without using induction) of the statement, "for all $n \in \mathbb{N}$, the number $n^2 + n$ is even."

12. Prove that there is a sequence of positive real numbers $a_1, a_2, a_3, \ldots$ such that the partial sum $a_1 + a_2 + a_3 + \cdots + a_n$ is strictly less than 2 for all $n \in \mathbb{N}$. Hint: think about how you could define what a_{k+1} is to make the induction argument work.

13. Prove that every natural number is either a power of 2, or can be written as the sum of distinct powers of 2.

14. Use induction to prove that if n people all shake hands with each other, that the total number of handshakes is $\frac{n(n-1)}{2}$.

15. Suppose that a particular real number x has the property that $x + \frac{1}{x}$ is an integer. Prove that $x^n + \frac{1}{x^n}$ is an integer for all natural numbers n.

16. Use induction to prove that $\displaystyle\sum_{k=0}^{n} \binom{n}{k} = 2^n$. That is, the sum of the nth row of Pascal's Triangle is 2^n.

17. Use induction to prove $\binom{4}{0} + \binom{5}{1} + \binom{6}{2} + \cdots + \binom{4+n}{n} = \binom{5+n}{n}$. (This is an example of the hockey stick theorem.)

18. Use the product rule for logarithms ($\log(ab) = \log(a) + \log(b)$) to prove, by induction on n, that $\log(a^n) = n \log(a)$, for all natural numbers $n \geq 2$.

19. Let $f_1, f_2, \ldots, f_n$ be differentiable functions. Prove, using induction, that

$$(f_1 + f_2 + \cdots + f_n)' = f_1' + f_2' + \cdots + f_n'$$

You may assume $(f + g)' = f' + g'$ for any differentiable functions f and g.

20. Suppose $f_1, f_2, \ldots, f_n$ are differentiable functions. Use mathematical induction to prove the generalized product rule:

$$(f_1 f_2 f_3 \cdots f_n)' = f_1' f_2 f_3 \cdots f_n + f_1 f_2' f_3 \cdots f_n + f_1 f_2 f_3' \cdots f_n + \cdots + f_1 f_2 f_3 \cdots f_n'$$

You may assume the product rule for two functions is true.

Chapter 3

Logic and Proofs

Logic is the study of consequence. Given a few mathematical statements or facts, we would like to be able to draw some conclusions. For example, if I told you that a particular real valued function was continuous on the interval $[0, 1]$, and $f(0) = -1$ and $f(1) = 5$, can we conclude that there is some point between $[0, 1]$ where the graph of the function crosses the x-axis? Yes, we can, thanks to the Intermediate Value Theorem from Calculus. Can we conclude that there is exactly one point? No. Whenever we find an "answer" in math, we really have a (perhaps hidden) argument - given the situation we are in, we can conclude the answer is the case. Of course real mathematics is about proving general statements (like the Intermediate Value Theorem), and this too is done via an argument, usually called a proof. We start with some given conditions - these are the premises of our argument. From these we find a consequence of interest - our conclusion.

The problem is, as you no doubt know from arguing with friends, not all arguments are *good* arguments. A "bad" argument is one in which the conclusion does not follow from the premises - the conclusion is not a consequence of the premises. Logic is the study of what makes an argument good or bad. In other words, logic aims to determine in which cases a conclusion is, or is not, a consequence of a set of premises.

By the way, "argument" is actually a technical term in math (and philosophy – another discipline which studies logic):

Definition 2. An *argument* is a set of statements, one of which is called the *conclusion* and the rest of which are called *premises*. An argument is said to be *valid* if the conclusion must be true whenever the premises are all true. An argument is *invalid* if it is not valid; it is possible for all the premises to be true and the conclusion to be false.

For example, consider the following two arguments:

If Edith eats her vegetables, then she can have a cookie.
Edith eats her vegetables.

$\therefore$ Edith gets a cookie.

Florence must eat her vegetables in order to get a cookie.
Florence eats her vegetables.

$\therefore$ Florence gets a cookie.

(The symbol "$\therefore$" means "therefore")

Are these arguments valid? Hopefully you agree that the first one is but the second one is not. Logic tells us why. How? By analyzing the structure of the statements in the argument. Notice that the two arguments above look almost identical. Edith and Florence both eat their vegetables. In both cases there is a connection between eating of vegetables and cookies. But we claim that it is valid to conclude that Edith gets a cookie, but not that Florence does. The difference must be in the connection between eating vegetables and getting cookies. We need to be good at reading and comprehending these sentences. Do the two sentences mean the same thing? Unfortunately, when talking in everyday language, we are often sloppy, and you might be tempted to say they are equivalent. But notice that just because Florence *must* eat her vegetables, we have not said

that doing so would be enough - she might also need to clean her room, for example. In everyday (non-mathematical) practice, you might say that the "other direction" was implied. We don't ever get to say that.

Our goal in studying logic is to gain intuition for which arguments are valid and which are invalid. This will require us to become better at reading and writing mathematics – a worthy goal in its own right. So let's get started.

3.1 Propositional Logic

A proposition is simply a statement. Propositional logic studies the ways statements can interact with each other. It is important to remember that propositional logic does not really care about the content of the statements. For example, from a propositional logic statement point, the claims, "if the moon is made of cheese then basketballs are round," and, "if spiders have eight legs then Sam walks with a limp" are exactly the same. They are both statements of the form, "if ⟨something⟩, then ⟨something else⟩."

Here's a question: is it true that when playing Monopoly, if you get more doubles than any other player you will lose, or that if you lose you must have bought the most properties? We will answer this question, and won't need to know anything about Monopoly. Instead we will look at the logical form of the statement. First though, let's back up and make sure we are very clear on some basics.

Definition 3. A *statement* is any declarative sentence which is either true or false.

Example: These are statements:

- Telephone numbers in the USA have 10 digits.
- The moon is made of cheese.
- 42 is a perfect square.
- Every even number greater than 2 can be expressed as the sum of two primes.

And these are not:

- Would you like some cake?
- The sum of two squares.
- $1 + 3 + 5 + 7 + \cdots + 2n + 1$.
- Go to your room!
- This sentence is false.[1]
- That's what she said.

The reason the last sentence is not a statement is because it contains variables ("that" and "she"). Unless those are specified, the sentence cannot be true or false, and as such not a statement. Other examples of this: $x + 3 = 7$. Depending on x, this is either true or false, but as it stands it it neither.

You can build more complicated statements out of simpler ones using *logical connectives*. For example, this is a statement:

Telephone numbers in the USA have 10 digits and 42 is a perfect square.

[1] This is a tricky one. Remember, a sentence is only a statement if it is either true or false. Here, the sentence is not false, for if it were, it would be true. It is not true, for that would make it false.

Note that we can break this down into two smaller statements. The two shorter statements are *connected* by an "and." We will consider 5 connectives: "and" (Sam is a man and Chris is a woman), "or" (Sam is a man or Chris is a woman), "if... then..." (if Sam is a man, then Chris is a woman), "if and only if" (Sam is a man if and only if Chris is a woman), and "not" (Sam is not a man).

Since we rarely care about the content of the individual statements, we can replace them with variables. We use capital letters in the middle of the alphabet for these *propositional* (or *sentential*) variables: $P, Q, R, S \ldots$. We also have symbols for the logical connectives: $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$, $\neg$.

Logical Connectives

- $P \wedge Q$ means P and Q, called a *conjunction*.
- $P \vee Q$ means P or Q, called a *disjunction*.
- $P \rightarrow Q$ means if P then Q, called an *implication* or *conditional*.
- $P \leftrightarrow Q$ means P if and only if Q, called a *biconditional*.
- $\neg P$ means not P, called a *negation*.

The logical connectives allow us to construct longer statements out of simpler statements. But the result is still a statement - it is either true or false. The *truth value* can be determined by the truth or falsity of the parts, depending on the connectives.

Truth Conditions for Connectives

- $P \wedge Q$ is true when both P and Q are true
- $P \vee Q$ is true when P or Q or both are true.
- $P \rightarrow Q$ is true when P is false or Q is true or both.
- $P \leftrightarrow Q$ is true when P and Q are both true, or both false.
- $\neg P$ is true when P is false.

I think all of these are obvious, except for $P \rightarrow Q$. Consider the statement,

> If Bob gets a 90 on the final, then Bob will pass the class.

This is definitely an implication: P is the statement "Bob gets a 90 on the final" and Q is the statement "Bob will pass the class." Suppose I made that statement to Bob – in what circumstances would it be fair to call me a liar? What if Bob really did get a 90 on the final, and he did pass the class? Then I have not lied; my statement is true. But if Bob did get a 90 on the final and did not pass the class, then I lied, making the statement false. The tricky case is this: what if Bob did not get a 90 on the final? Maybe he passes the class, maybe he doesn't - did I lie in either

case? I think not. In these last two cases, P was false, and so was the statement $P \to Q$. In the first case, Q was true, and so was $P \to Q$. So $P \to Q$ is true when either P is false or Q is true. Perhaps an easier way to look at it is this: $P \to Q$ is *false* in only one case: if P is true and Q is false. Otherwise, $P \to Q$ is true. Admittedly, there are times in English when this is not how "if..., then..." works. However, in mathematics, we *define* the implication to work this way.

3.1.1 Truth Tables

Our question about Monopoly is to determine whether the following statement is true:

> If you get more doubles than any other player then you will lose, or if you lose then you must have bought the most properties.

In other words, we need to decide when the statement $(P \to Q) \vee (Q \to R)$ is true. Using the rules above, we see that either $P \to Q$ is true or $Q \to R$ is true (or both). Those are true if either P is false or Q is true (in the first case) and Q is false or R is true (in the second case). So... yeah, it gets kind of messy. Luckily, we can make a chart to keep track of all the possibilities. Enter truth tables. The idea is this: on each row, we list a possible combination of T's and F's (standing of course, for true and false) for each of the sentential variables, and then mark down whether the statement in question is true or false in that case. We do this for every possible combination of T's and F's. Then we can clearly see in which cases the statement is true or false. For complicated statements, we will first fill in values for each part of the statement, as a way of breaking up our task into smaller, more manageable pieces.

All you really need to know is the truth tables for each of the logical connectives. Here they are:

P	Q	$P \wedge Q$
T	T	T
T	F	F
F	T	F
F	F	F

P	Q	$P \vee Q$
T	T	T
T	F	T
F	T	T
F	F	F

P	Q	$P \to Q$
T	T	T
T	F	F
F	T	T
F	F	T

P	Q	$P \leftrightarrow Q$
T	T	T
T	F	F
F	T	F
F	F	T

The truth table for negation looks like this:

P	$\neg P$
T	F
F	T

None of these truth tables should come as a surprise - they are all just restating the definitions of the connectives. Let's try another one:

Example: Make a truth table for the statement $\neg P \vee Q$.

Solution: Note that this statement is not $\neg(P \vee Q)$, the negation belongs to P alone. Here is the truth table:

P	Q	$\neg P$	$\neg P \vee Q$
T	T	F	T
T	F	F	F
F	T	T	T
F	F	T	T

We added a column for $\neg P$ to make filling out the last column easier. The entries in the $\neg P$ column were determined by the entries in the P column. Then to fill in

the final column, look only at the column for Q and the column for $\neg P$ and use the rule for $\vee$.

You might notice that the final column is identical to the final column in the truth table for $P \to Q$. Since we listed the possible values for P and Q in the same (in fact, standard) order, this says that $\neg P \vee Q$ and $P \to Q$ are *logically equivalent*

Now let's answer our question about monopoly:

Example: Analyze the statement, "if you get more doubles than any other player you will lose, or that if you lose you must have bought the most properties," using truth tables.

Solution: Represent the statement in symbols as $(P \to Q) \vee (Q \to R)$, where P is the statement "you get more doubles than any other player," Q is the statement "you will lose," and R is the statement "you must have bought the most properties." Now make a truth table.

The truth table needs to contain 8 rows in order to account for every possible combination of truth and falsity among the three statements. Here is the full truth table:

P	Q	R	$P \to Q$	$Q \to R$	$(P \to Q) \vee (Q \to R)$
T	T	T	T	T	T
T	T	F	T	F	T
T	F	T	F	T	T
T	F	F	F	T	T
F	T	T	T	T	T
F	T	F	T	F	T
F	F	T	T	T	T
F	F	F	T	T	T

The first three columns are simply a systematic listing of all possible combinations of T and F for the three statements (do you see how you would list the 16 possible combinations for four statements?). The next two columns are determined by the values of P, Q, and R and the definition of implication. Then, the last column is determined by the values in the previous two columns and the definition of $\vee$. It is this final column we care about.

Notice that in each of the eight possible cases, the statement in question is true. So our statement about monopoly is true (regardless of how many properties you own, how many doubles you roll, or whether you win or lose).

The statement about monopoly is an example of a *tautology* - a statement which is true on the basis of its logical form alone. Tautologies are always true but they don't tell us much about the world. No knowledge about monopoly was required to determine that the statement was true. In fact, it is equally true that "If the moon is made of cheese, then Elvis is still alive, or if Elvis is still alive, then unicorns have 5 legs."

3.1.2 Deductions

Earlier we claimed that the following was a valid argument:

> If Edith eats her vegetables, then she can have a cookie. Edith ate her vegetables. Therefore Edith gets a cookie.

How do we know this is valid? Let's look at the form of the statements. Let P denote "Edith eats her vegetables" and Q denote "Edith can have a cookie." The logical form of the argument is then:

$$\begin{array}{ll} & P \to Q \\ & P \\ \hline \therefore & Q \end{array}$$

This is an example of a *deduction rule* - a logical form of an argument which is always valid. This one is a particularly famous rule called *modus ponens*. Are you convinced that it is a valid deduction rule? If not, consider the following truth table:

P	Q	$P \to Q$
T	T	T
T	F	F
F	T	T
F	F	T

This is just the truth table for $P \to Q$, but what matters here is that all the lines in the deduction rule have their own column in the truth table. Now remember that an argument is valid provided the conclusion must be true given that the premises are true. The premises in this case are $P \to Q$ and P. Which *rows* of the truth table correspond to both of these being true? P is true in the first two rows, and of those, only the first row has $P \to Q$ true as well. And low-and-behold, in this one case, Q is true as well. So if $P \to Q$ and P are both true, we see that Q must be true as well.

Here are a few more examples.

Example: Show that

$$\begin{array}{ll} & P \to Q \\ & \neg P \to Q \\ \hline \therefore & Q \end{array}$$

is a valid deduction rule.

Solution: We make a truth table which contains all the lines of the argument form:

P	Q	$P \to Q$	$\neg P$	$\neg P \to Q$
T	T	T	F	T
T	F	F	F	T
F	T	T	T	T
F	F	T	T	F

(we include a column for $\neg P$ just as a step to help getting the column for $\neg P \to Q$).

Now look at all the rows for which both $P \to Q$ and $\neg P \to Q$ are true. This happens only in rows 1 and 3. Hey! In those rows Q is true as well, so the argument form is valid (it is a valid deduction rule).

Example: Decide whether

$$\begin{array}{ll} & (P \to R) \vee (Q \to R) \\ \hline \therefore & (P \vee Q) \to R \end{array}$$

is a valid deduction rule.

Solution: Let's make a truth table containing both statements.

P	Q	R	$P \vee Q$	$P \to R$	$Q \to R$	$(P \vee Q) \to R$	$(P \to R) \vee (Q \to R)$
T	T	T	T	T	T	T	T
T	T	F	T	F	F	F	F
T	F	T	T	T	T	T	T
T	F	F	T	F	T	F	T
F	T	T	T	T	T	T	T
F	T	F	T	T	F	F	T
F	F	T	F	T	T	T	T
F	F	F	F	T	T	T	T

Look at the fourth row. In this case, $(P \to R) \vee (Q \to R)$ is true, but $(P \vee Q) \to R$ is false. Therefore the argument form is not valid (it is not a valid deduction rule). The truth table tells us more: our premise is true when P is true and Q and R are both false, but the conclusion is false in this case. The same problem occurs when Q is true and P and R are false (row 6).

Notice that if we switch the premise and conclusion, then we do have a valid argument - whenever $(P \vee Q) \to R$ is true, so is $(P \to R) \vee (Q \to R)$. Another way to say this is to state that the statement

$$[(P \vee Q) \to R] \to [(P \to R) \vee (Q \to R)]$$

is a tautology.

Exercises

1. Consider the statement about a party, "If it's your birthday or there will be cake, then there will be cake."

   (a) Translate the above statement into symbols. Clearly state which statement is P and which is Q.

   (b) Make a truth table for the statement.

   (c) Assuming the statement is true, what (if anything) can you conclude if there will be cake?

   (d) Assuming the statement is true, what (if anything) can you conclude if there will not be cake?

   (e) Suppose you found out that the statement was a lie. What can you conclude?

2. Suppose P and Q are the statements: P: Jack passed math. Q: Jill passed math.

   (a) Translate "Jack and Jill both passed math" into symbols.

   (b) Translate "If Jack passed math, then Jill did not" into symbols.

   (c) Translate "$P \vee Q$" into English.

   (d) Translate "$\neg(P \wedge Q) \to Q$" into English.

   (e) Suppose you know that if Jack passed math, then so did Jill. What can you conclude if you know that:

      (a) Jill passed math?

      (b) Jill did not pass math?

3. Geoff Poshingten is out at a fancy pizza joint, and decides to order a calzone. When the waiter asks what he would like in it, he replies, "I want either pepperoni or sausage, and if I have sausage, I must also include quail. Oh, and if I have pepperoni or quail then I must also have ricotta cheese."

   (a) Translate Geoff's order into logical symbols.

   (b) The waiter knows that Geoff is either a liar or a truth-teller (so either everything he says is false, or everything is true). Which is it?

   (c) What, if anything, can the waiter conclude about the ingredients in Geoff's desired calzone?

4. Make a truth table for the statement $(P \vee Q) \to (P \wedge Q)$.

5. Make a truth table for the statement $\neg P \wedge (Q \to P)$. What can you conclude about P and Q if you know the statement is true?

6. Make a truth table for the statement $\neg P \to (Q \wedge R)$.

7. Determine if the following argument form is valid: $\begin{array}{ll} & P \vee Q \\ & \neg P \\ \hline \therefore & Q \end{array}$.

8. Determine if the following argument form is valid: $\begin{array}{ll} & P \to (Q \vee R) \\ & \neg(P \to Q) \\ \hline \therefore & R \end{array}$

9. Determine if the following argument form is valid: $\begin{array}{ll} & (P \wedge Q) \to R \\ & \neg P \vee \neg Q \\ \hline \therefore & \neg R \end{array}$

3.2 Rephrasing - Logical Equivalence

When reading or writing a proof, or even just trying to understand a mathematical statement, it can be very helpful to rephrase the statement. But how do you know you are doing so correctly? How do you know the two statements are equivalent?

One way is to make a truth table for each and ensure that the final columns of both are identical. We saw earlier that $P \to Q$ is logically equivalent to $\neg P \vee Q$ because their truth tables agreed. Now we can just remember this fact. If we see the statement, "if Sam is a man then Chris is a woman," we can instead think of it as "Sam is a woman or Chris is a woman." You might also be tempted to rephrase further: "Sam or Chris is a woman." This is okay of course, but that this second rephrasing is allowed is due to the meaning of "is," not any of our logical connectives.

Here are some common logical equivalences which can help rephrase mathematical statements:

Double Negation

$\neg\neg P$ is logically equivalent to P

Example: "It is not the case that c is not odd" means "c is odd."

No surprise there. Now let's see how negation plays with conjunctions and disjunctions.

De Morgan's Laws

$$\neg(P \wedge Q) \text{ is logically equivalent to } \neg P \vee \neg Q$$

$$\neg(P \vee Q) \text{ is logically equivalent to } \neg P \wedge \neg Q$$

Example: "c is not even or c is not prime" means "c is not both odd and prime"

Do you believe De Morgan's laws? If not, make a truth table for each of them. I think most of us get these right most of the time without thinking about them too hard. If I told you that I had popcorn and goobers at the movies, but then you found out it was opposite day (so my statement was false) then you would agree, I hope, that I either did not have popcorn *or* did not have goobers (or didn't have either). You would not insist that I could not have had either.

I should warn you that often in English, we are sloppy about our and's, or's and not's. When you write about mathematics, you should be careful and write what you mean. If you are not sure what to write, rephrasing carefully using De Morgan's laws can help you make sure that statement matches your intended meaning.

Now some rules for implications:

Negation of Implication

$$\neg(P \to Q) \text{ is logically equivalent to } P \wedge \neg Q$$

In words: the only way for an implication to be false is for the "if" part to be true and the "then" part to be false.

This is very important, and not obvious - implications are tricky. But look at the truth table for $P \to Q$:

P	Q	$P \to Q$
T	T	T
T	F	F
F	T	T
F	F	T

There is only one way for the implication to be false - P is true and Q is false. Another way to see that this is true is by using De Morgan's Laws. We saw earlier that $P \to Q$ can be rephrased as $\neg P \vee Q$ so we have

$$\neg(P \to Q) \text{ is logically equivalent to } \neg(\neg P \vee Q)$$

But by De Morgan's laws, $\neg(\neg P \vee Q)$ is equivalent to $\neg\neg P \wedge \neg Q$. By double negation $\neg\neg P$ is the same as P.

While we are thinking about implications, we should talk about the converse and contrapositive:

Converse and Contrapositive

- The *converse* of an implication $P \to Q$ is the implication $Q \to P$. The converse is **NOT** logically equivalent to the original implication.
- The *contrapositive* of an implication $P \to Q$ is the statement $\neg Q \to \neg P$. An implication and its contrapositive are logically equivalent.

A related, but lesser used, term is the *inverse* of an implication. The inverse of $P \to Q$ is $\neg P \to \neg Q$. Notice that the inverse of an implication is the contrapositive of the converse. Read that one more time. Good. So since implications and their contrapositives are logically equivalent, the inverse and converse of an implication are logically equivalent to each other, but not to the original implication.

Example: Suppose I tell Sue that if she gets a 93% on her final, she will get an A in the class. Assuming that what I said is true, what can you conclude in the following cases:

(a) Sue gets a 93% on her final.

(b) Sue gets an A in the class.

(c) Sue does not get a 93% on her final.

(d) Sue does not get an A in the class.

Solution: Note first that whenever $P \to Q$ and P are both true statements, Q must be true as well. For this problem, take P to mean "Sue gets a 93% on her final" and Q to mean "Sue will get an A in the class."

(a) We have $P \to Q$ and P so Q follows - Sue gets an A.

(b) You cannot conclude anything – Sue could have gotten the A because she did extra credit for example. Notice that we do not know that if Sue gets an A, then she gets a 93% on her final. That is the converse of the original implication, so it might or might not be true.

(c) The inverse of $P \to Q$ is $\neg P \to \neg Q$, which states that if Sue does not get a 93% on the final then she will not get an A in the class. But this does not follow from the original implication. Again, we can conclude nothing – Sue could have done extra credit.

(d) What would happen if Sue does not get an A but *did* get a 93% on the final. Then P would be true and Q would be false. But this makes the implication $P \to Q$ false! So it must be that Sue did not get a 93% on the final. Notice now we have the implication $\neg Q \to \neg P$ which is the contrapositive of $P \to Q$. Since $P \to Q$ is assumed to be true, we know $\neg Q \to \neg P$ is true as well – they are equivalent.

As we said above, an implication is not logically equivalent to its converse. Given particular statements P and Q, the statements $P \to Q$ and $Q \to P$ could both be true, both be false, or one

could be true and the other false (in either order). Now if both $P \to Q$ and $Q \to P$ are true, then we say that P and Q are equivalent. In fact, we have:

If and only if

$$P \leftrightarrow Q \text{ is logically equivalent to } (P \to Q) \wedge (Q \to P).$$

Example: given an integer n, it is true that n is even if and only if n^2 is even. That is, if n is even, then n^2 is even, as well as the converse: if n^2 is even then n is even.

You can think of "if and only if" statements as having two parts: an implication and its converse. We might say one is the "if" part, and the other is the "only if" part. We also sometimes say that "if and only if" statements have two directions: a forward direction ($P \to Q$) and a backwards direction ($P \leftarrow Q$, which is really just sloppy notation for $Q \to P$).

Let's think a little about which part is which. Is $P \to Q$ the "if" part or the "only if" part? Perhaps we should look at an example.

> **Example**: Suppose it is true that I sing if and only if I'm in the shower. We know this means that both if I sing, then I'm in the shower, and also the converse - that if I'm in the shower, then I sing. Let P be the statement, "I sing," and Q be, "I'm in the shower." So $P \to Q$ is the statement "if I sing, then I'm in the shower." Which part of the if and only if statement is this?
>
> What we are really asking is what is the meaning of "I sing if I'm in the shower" and "I sing only if I'm in the shower." When is the first one (the "if" part) *false*? When I am in the shower but not singing. That is the same condition on being false as the statement "if I'm in the shower, then I sing." So the "if" part is $Q \to P$. On the other hand, to say, "I sing only if I'm in the shower" is equivalent to saying "if I sing, then I'm in the shower," so the only if part is $P \to Q$.

It is not terribly important to know which part is the if or only if part, but this does get at something very very important: THERE ARE MANY WAYS TO STATE AN IMPLICATION! The problem is, since these are all different ways of saying the same implication, we cannot use truth tables to analyze the situation. Instead, we just need good English skills.

> **Example**: Rephrase the implication, "if I dream, then I am asleep" in as many different ways as possible. Then do the same for the converse.
>
> *Solution*: The following are all equivalent to the original implication:
>
> 1. I am asleep if I dream.
> 2. I dream only if I am asleep.
> 3. In order to dream, I must be asleep.
> 4. To dream, it is necessary that I am asleep.
> 5. To be asleep, it is sufficient to dream.
> 6. I am not dreaming unless I am asleep.
>
> The following are equivalent to the converse - if I am asleep, then I dream:
>
> 1. I dream if I am asleep.

2. I am asleep only if I dream.
3. It is necessary that I dream in order to be asleep.
4. It is sufficient that I be asleep in order to dream.
5. If I don't dream, then I'm not asleep.

Hopefully you agree with the above example. We include the "necessary and sufficient" versions because those are common when discussing mathematics. In fact, let's agree once and for all what they mean:

Necessary and Sufficient

- "P is necessary for Q" means $Q \to P$.
- "P is sufficient for Q" means $P \to Q$.
- If P is necessary and sufficient for Q, then $P \leftrightarrow Q$.

To be honest, I have trouble with these if I'm not very careful. I find it helps to have an example in mind:

Example: Recall from calculus, if a function is differentiable at a point c, then it is continuous at c, but that the converse of this statement is not true (for example, $f(x) = |x|$ at the point 0). Restate this fact using necessary and sufficient language.

Solution: It is true that in order for a function to be differentiable at a point c, it is necessary for the function to be continuous at c. However, it is not necessary that a function be differentiable at c for it to be continuous at c.

It is true that to be continuous at a point c, is is sufficient that the function be differentiable at c. However, it is not the case that being continuous at c is sufficient for a function to be differentiable at c.

Exercises

1. Determine whether the following two statements are logically equivalent: $\neg(P \to Q)$ and $P \wedge \neg Q$. Explain how you know you are correct.
2. Are the statements $P \to (Q \vee R)$ and $(P \to Q) \vee (P \to R)$ logically equivalent?
3. Consider the statement "If Oscar eats Chinese food, then he drinks milk."
   (a) Write the converse of the statement.
   (b) Write the contrapositive of the statement.
   (c) Is it possible for the contrapositive to be false? If it was, what would that tell you?
   (d) Suppose the original statement is true, and that Oscar drinks milk. Can you conclude anything (about his eating Chinese food)? Explain.

(e) Suppose the original statement is true, and that Oscar does not drink milk. Can you conclude anything (about his eating Chinese food)? Explain.

4. Simplify the following statements (so that negation only appears right before variables).

   (a) $\neg(P \to \neg Q)$

   (b) $(\neg P \vee \neg Q) \to \neg(\neg Q \wedge R)$

   (c) $\neg((P \to \neg Q) \vee \neg(R \wedge \neg R))$

   (d) It is false that if Sam is not a man then Chris is a woman, and that Chris is not a woman.

5. Which of the following statements are equivalent to the implication, "if you win the lottery, then you will be rich," and which are equivalent to the converse of the implication?

   (a) Either you win the lottery or else you are not rich.

   (b) Either you don't win the lottery or else you are rich.

   (c) You will win the lottery and be rich.

   (d) You will be rich if you win the lottery.

   (e) You will win the lottery if you are rich.

   (f) It is necessary for you to win the lottery to be rich.

   (g) It is sufficient to with the lottery to be rich.

   (h) You will be rich only if you win the lottery.

   (i) Unless you win the lottery, you won't be rich.

   (j) If you are rich, you must have one the lottery.

   (k) If you are not rich, then you did not win the lottery.

   (l) You will win the lottery if and only if you are rich.

6. Consider the implication, "if you clean your room, then you can watch TV." Rephrase the implication in as many ways as possible. Then do the same for the converse.

3.3 Quantifiers and Predicate Logic

So far we have seen how statements can be combined with logical symbols. This is helpful when trying to understand a complicated mathematical statement - you can determine under which conditions the complicated statement is true. Additionally, we have been able to analyze the logical form of arguments to decide which arguments are valid and which are not. However, the types of statements we have been able to make so far has be sorely limited. For example, consider a classic argument:

All men are mortal.
Socrates is a man.
Therefore, Socrates is mortal.

This is clearly a valid argument - it is an example of a *syllogism*. Historically, the study of logic began with Aristotle who worked out all possible forms of syllogisms and decided which were valid and which were not. We will not do that here. However, this is an important example because it highlights a limitation of the propositional logic we have studied so far. Can we use propositional logic to analyze the argument?

The trouble is that we don't have a way to translate "all men are mortal." It looks like an implication - being a man implies you are mortal. So maybe it is $P \to Q$. But what is P? We could rephrase: "if Socrates is a man, then Socrates is mortal." Now we have a valid argument form we have seen before. But it is not quite the same. (What if the argument was: All men are mortal, all mortals have hair, therefore all men have hair - also valid and Socrates has nothing to do with it.) Or perhaps we could go with, "for every thing there is, if the thing is a man, then the thing is mortal." Looks promising, but we still can't let P be "the thing is a man" - that is not a statement because "thing" is a variable.

One way to sort this mess out is to introduce a new sort of logic called *predicate* logic. This is the logic of properties. Doing so will allow us to discuss how properties of various things are related. Above, if a thing has the property of being a man, then it has the property of being mortal. We can then *quantify* over what things we talk about. The the example above, *all* things.

Another way to accomplish much of the same goal is to use set theory. The idea here is that we want talk about collections of things - for example, the collection of all men, and the collection of all mortal things. We can then express "all men are mortal" by saying that the set of men is a subset of the set of mortals. Then we claim that Socrates is a member of the set of men, so therefore is also a member of the set of mortals.

One last example to highlight these two different approaches before delving into details. We all agree that all squares are rectangles. The set theory approach would be to consider the set of squares and the set of rectangles, and point out that one is a subset of the other (the squares are a subset of the rectangles). The predicate logic approach would be to consider the properties of "being a square" and of "being a rectangle" and assign these to predicates - say S and R. We would then say $\forall x(S(x) \to R(x))$ - for all things, if the thing is a square, then it is a rectangle. So having the property of being a square implies having the property of being a rectangle.

Now some details.

Consider the statement "for all integers a and b, if ab is even, then a is even or b is even." If we use propositional logic to analyze this statement, what should P be?

You might want to say P is "ab is even" so Q can be "a is even" and R can be "b is even" and say that the whole statement is therefore of the form $P \to (Q \vee R)$. This does not work! One reason is that "ab is even" is not a statement - it contains free variables, so it is not true or false (until the variables have values). Another reason is that we have just lost the "for all integers a and b" from our statement. We can remedy both these problems using *Predicate Logic*.

3.3.1 Predicates

We can think of predicates as properties of objects. For example, consider the predicate E which we will use to mean "is even." Being even is a property of some numbers, so E needs to be applied to something. We will adopt the notation $E(x)$ to mean x is even. (Some books would write Ex instead.) Notice that if we put a number in for x, then this becomes a statement - and as such can be true or false. So $E(2)$ is true, and $E(3)$ is false. A predicate is like a function with codomain equal to the set of truth values $\{T, F\}$. On the other hand $E(x)$ is not true or false, since we don't know what x is. If we have a variable floating around like that, we say the expression is merely a formula, and not a statement.

Since $E(2)$ is a statement (a proposition), we can apply propositional logic to it. Consider

$$E(2) \wedge \neg E(3)$$

which is a true statement, because it is both the case that 2 is even and that 3 is not even. What we have done here is capture the logical form (using connectives) of the statement "2 is even and 3 is not" as well as the mathematical content (using predicates).

Notice that we can only assert even-ness of a single number at a time. That is to say, E is a *one-place* predicate. There are also predicates which assert a property of two or more numbers (or other objects) at the same time. Consider the *two-place* predicate "is less than." Perhaps we will use the variable L. Now we can say $L(2, 3)$, which is true because 2 is less than 3. Of course we are already have a symbol for this: $2 < 3$. However, what about "divides evenly into" as a predicate? We can say $D(2, 10)$ is true because 2 divides evenly into 10, while $D(3, 10)$ is false since there is a remainder when you divide 10 by 3. Incidentally, there is a standard mathematical symbol for this: $2|10$ is read "2 divides 10."

Predicates can be as complicated and have as many places as we want or need. For example, we could $R(x, y, z, u, v, w)$ be the predicate asserting that x, y, z are distinct natural numbers whose only common factor is u, the difference between x and y is v and the difference between y and z is w. This is a silly and most likely useless example, but it is an example of a predicate. It is true of some ordered lists of six numbers (6-tuples), and false of others. Additionally, predicates need not have anything to do with numbers: we could let $F(a, b, c, d)$ be the predicate that asserts that a and b are the only two children of mother c and father d.

3.3.2 Quantifiers

Perhaps the most important reason to use predicate logic is that doing so allows for quantification. We can now express statements like "ever natural number is either even or odd," and "there is a natural number such that no number is less than it." Think back to Calculus and the Mean Value Theorem. It states that for every function f and every interval (a, b), if f is continuous on the interval $[a, b]$ and differentiable on the interval (a, b), then there exists a number c such that $a \leq c \leq b$ and $f'(c)(b - a) = f(b) - f(a)$. Using the correct predicates and quantifiers, we could express this statement entirely in symbols.

There are two quantifiers we will be interested in: existential and universal.

Quantifiers

- The existential quantifier is $\exists$ and is read "there exists" or "there is." For example,

  $$\exists x(x < 0)$$

  asserts that there is a number less than 0.

- The universal quantifier is $\forall$ and is read "for all" or "every." For example,

  $$\forall x(x \geq 0)$$

  asserts that every number is greater than or equal to 0.

Are these statements true? Well, first notice that they cannot both be true. In fact, they assert exactly the opposite of each other. (Note that $x < y \leftrightarrow \neg(x \geq y)$ – although you might wonder what x and y are here, so it might be better to say $\forall x \forall y\,(x < y \leftrightarrow \neg(x \geq y))$.) Which one is it though? The answer depends entirely on our domain of discourse - the universe over which we quantify. Usually, this universe is clear from the context. If we are only discussing the natural numbers, then $\forall x \ldots$ means "for every natural number $x \ldots$." On the other hand, in calculus we care about the real numbers, so it would mean "for every real number $x \ldots$." If the context is not clear, we might right $\forall x \in \mathbb{N} \ldots$ to mean "for every natural number $x \ldots$." Of course, for the two statements above, the second is true of the natural numbers, the first is true for any larger universe.[2]

Some more examples: to say "every natural number is either even or odd," we would write, using E and O as the predicates for even and odd respectively:

$$\forall x(E(x) \vee O(x))$$

To say "there is a number such that no number is less than it" we would write:

$$\exists x \forall y(y \geq x)$$

Actually, I did a little translation before I wrote that down. The above statement would be literally read "there is a number such that every number is greater than or equal to it." This of course amounts to the same thing. However, if I wanted to be exact, I could have also written:

$$\exists x \neg \exists y(y < x).$$

Notice also that say that there is a number for which no number is smaller is equivalent to saying that it is not the case that for every number there is a number smaller than it:

$$\neg \forall x \exists y(y < x).$$

That these three statements are equivalent is no coincidence. To understand what is going on, we will need to better understand how quantification interacts with the logical connectives, specifically negation.

3.3.3 Quantifiers and Connectives

What does it mean to say that it is false that there is something that has a certain property? Well, it means that everything does not have that property. What does it mean for it to be false that everything has a certain property? It means that there is something that doesn't have the property. So in symbols, we have the following

Quantifiers and Negation

$$\neg \forall x P(x) \text{ is equivalent to } \exists x \neg P(x)$$

and

$$\neg \exists x P(x) \text{ is equivalent to } \forall x \neg P(x).$$

[2] Remember, we take the natural numbers to be 0, 1, 2, 3, ...

In other words, to move a negation symbol past a quantifier, you must switch the quantifier. This can be done multiple times:

$$\neg\exists x\forall y\exists z P(x,y,z) \text{ is equivalent to } \forall x\exists y\forall z\neg P(x,y,z).$$

Now we also know how to move negation symbols through other connectives (using De Morgan's Laws) so it is always possible to rewrite a statement so that the only negation symbols that appear are right in front of a predicate. This hints at the possibility of having a standard form for all predicate statements. However, to get this we must also understand how to move quantifiers through connectives.

Before we get too excited, note that we only need to worry about two connectives: $\wedge$ and $\vee$. This is because we can rewrite $p \to q$ as $\neg p \vee q$ (they are logically equivalent) and $p \leftrightarrow q$ as $(p \wedge q) \vee (\neg p \wedge \neg q)$ (also logically equivalent).

Let us consider an example to see what can happen.

Example: Let E be the predicate for being even, and O for being odd. Consider:

$$\exists x E(x) \wedge \exists x O(x),$$

which says that there is a number which is even and a number which is odd. This is of course true. However there is no number which is both even and odd, so

$$\exists x(E(x) \wedge O(x))$$

is false. Note also that

$$\exists x(E(x) \vee O(x))$$

while true, is not really the same thing – if O is instead the predicate for "is less than 0" then the original statement is false, but this new one is true (of the natural numbers). Changing the quantifier also doesn't help:

$$\forall x(E(x) \wedge O(x))$$

is false. So what can we do?

The problem is that in the original sentence, the variable x is doing double duty. We want to express the fact that there is an even number and an odd number. But that even number is in no way related to that odd number. So we might as well have said

$$\exists x E(x) \wedge \exists y O(y).$$

Now we can move the quantifiers out:

$$\exists x\exists y(E(x) \wedge O(y)).$$

The same thing works with $\vee$ and for $\forall$ with either connective. As long as there is no repeat in quantified variables we can move the quantifiers outside of conjunctions and disjunctions.

A warning though: you cannot do this for $\to$, at least not directly.[3] Let's see what happens.

[3] We must be similarly careful with $\leftrightarrow$

Example: Consider

$$\forall x P(x) \to \exists y Q(y)$$

for some predicates P and Q. This sentence is **not** the same as

$$\forall x \exists y (P(x) \to Q(y)).$$

Remember that $P \to Q$ is the same as $\neg P \vee Q$. So the original sentence is really

$$\neg \forall x P(x) \vee \exists y Q(y).$$

Before we move the quantifiers out, we must move the $\forall x$ past the negation sign, which switches it to a $\exists x$:

$$\exists x \neg P(x) \vee \exists y Q(y).$$

Then we can finish by writing,

$$\exists x \exists y (\neg P(x) \vee Q(y))$$

or equivalently

$$\exists x \exists y (P(x) \to Q(y)).$$

Exercises

1. Translate into symbols. Use $E(x)$ for "x is even" and $O(x)$ for "x is odd."
   (a) No number is both even and odd.
   (b) One more than any even number is an odd number.
   (c) There is prime number that is even.
   (d) Between any two numbers there is a third number.
   (e) There is no number between a number and one more than that number.
2. Translate into English:
   (a) $\forall x (E(x) \to E(x+2))$
   (b) $\forall x \exists y (\sin(x) = y)$
   (c) $\forall y \exists x (\sin(x) = y)$
   (d) $\forall x \forall y (x^3 = y^3 \to x = y)$
3. Simplify the statements (so negation appears only directly next to predicates).
   (a) $\neg \exists x \forall y (\neg O(x) \vee E(y))$
   (b) $\neg \forall x \neg \forall y \neg (x < y \wedge \exists z (x < z \vee y < z))$
   (c) There is a number n for which no other number is either less n than or equal to n.
   (d) It is false that for every number n there are two other numbers which n is between.

3.4 Proofs

Anyone who doesn't believe there is creativity in mathematics clearly has not tried to write proofs. Finding a way to convince the world that a particular statement is necessarily true is a mighty undertaking and can often be quite challenging. There is not guaranteed path to success in the search for proofs. For example, in the summer of 1742, a German mathematician by the name of Christian Goldbach wondered whether every even integer greater than 2 could be written as the sum of two primes. Centuries later, we still don't have a proof of this apparent fact (computers have checked that "Goldbach's Conjecture" holds for all numbers less than 4×10^{18}, which leaves only infinitely many more numbers to check).

Writing proofs is a bit of an art. Like any art, to be truly great at it, you need some sort of inspiration. But that's not to say we cannot learn some basic techniques as a foundation. Just as musicians can learn proper fingering, and painters can learn the proper way to hold a brush, we can look at the proper way to construct arguments. A good place to start might be to study a classic.

Theorem. *There are infinitely many primes.*

Proof. Suppose this were not the case - that there are only finitely many primes. Then there must be a last, largest prime, call it p. Consider the number

$$N = p! + 1 = (p \cdot (p-1) \cdot \cdots 3 \cdot 2 \cdot 1) + 1.$$

Now N is certainly larger than p. Also, N is not divisible by any number less than or equal to p, since every number less than or equal to p divides $p!$. Thus the prime factorization of N contains prime numbers (possibly just N itself) all greater than p. Therefor p is not the largest prime, a contradiction. Therefore there are infinitely many primes. QED

This proof is an example of a *proof by contradiction* - one of the standard styles of mathematical proof. First and foremost, the proof is an argument - a sequence of statements, the last of which is the *conclusion* which follows from the previous statements. The argument is valid - the conclusion must be true if the premises are true. Let's go through the proof line by line.

1. Suppose there are only finitely many primes.

   this is a premise - note the use of "suppose"

2. There must be a largest prime, call it p.

   follows from line 1, by the definition of "finitely many"

3. Let $N = p! + 1$.

   basically just notation, although this is the inspired part of the proof; looking at $p! + 1$ is the key insight

4. N is larger than p

   by the definition of $p!$

5. N is not divisible by any number less than or equal to p

   by the definition, $p!$ is divisible by each number less than or equal to p, so $p! + 1$ is not

6. The prime factorization of N contains prime numbers greater than p

   since N is divisible by each prime number in the prime factorization of N, and by line 5.

7. Therefore p is not the largest prime.

   by line 6 - N is divisible by a prime larger than p

8. This is a contradiction.

   from line 2 and line 7: the largest prime is p and there is a prime larger than p

9. Therefore there are infinitely many primes

   from line 1 and line 8: our only premise lead to a contradiction, so the premise is false

We should say a bit more about the last line. Up through line 8, we have a valid argument with the premise "there are only finitely many primes" and the conclusion "there is a prime larger than the largest prime." This is a valid argument - each line follows from previous lines - so if the premises are true, then the conclusion *must* be true. However, the conclusion is **NOT** true - it is a contradiction, so necessarily false. The only way out: the premise must be false.

The sort of line-by-line analysis we did above is a great way to really understand what is going on. Whenever you come across a proof in a textbook, you really should make sure you understand what each line is saying and why it is true. However, it is equally important to understand the overall structure of the proof. This is where using tools from logic is helpful. Luckily there are a relatively small number of standard proof styles that keep showing up again and again. Being familiar with these can help understand proof, as well as give ideas of how to write your own.

Direct Proof

The simplest (from a logic perspective) style of proof is a *direct proof*. Often all that is required to prove something is a systematic explanation of what everything means. Direct proofs are especially useful when proving implications. The general format to prove $P \to Q$ is this:

Assume P. Explain, explain,..., explain. Therefore Q.

Often we want to prove universal statements, perhaps of the form $\forall x(P(x) \to Q(x))$. Again, we will want to assume $P(x)$ is true and deduce from that $Q(x)$. But what about the x? We want this to work for *all* x. We accomplish this by fixing x to be an arbitrary element (of the sort we are interested in).

Here are a few examples. First, we will set up the proof structure for a direct proof, then fill in the details.

Example: Prove: For all integers n, if n is even, then n^2 is even.

Solution: The format of the proof with be this: Let n be an arbitrary integer. Assume that n is even. Explain explain explain. Therefore n^2 is even.

To fill in the details, we will basically just explain what it means for n to be even, and then see what that means for n^2. Here is a complete proof.

Proof. Let n be an arbitrary integer. Suppose n is even. Then $n = 2k$ for some integer k. Now $n^2 = (2k)^2 = 4k^2 = 2(2k^2)$. Since $2k^2$ is an integer, n^2 is even. QED

Example: Prove: For all integers a, b, and c, if $a|b$ and $b|c$ then $a|c$. Here $x|y$, read "x divides y" means that y is a multiple of x (so x will divide into y without remainder).

Solution: Even before we know what the divides symbol means, we can set up a direct proof for this statement. It will go something like this: Let a, b, and c be arbitrary integers. Assume that $a|b$ and $b|c$. Dot dot dot. Therefore $a|c$.

How do we connect the dots? We need to say what our hypothesis ($a|b$ and $b|c$) really means and why this gives us what the conclusion ($a|c$) really means. Another way to say that $a|b$ is to say that $b = ka$ for some integer k (that is, that b is a multiple of a). What are we going for? That $c = la$, for some integer l (because we want c to be a multiple of a). Here is the complete proof.

Proof. Let a, b, and c be integers. Assume that $a|b$ and $b|c$. In other words, b is a multiple of a and c is a multiple of b. So there are integers k and j such that $b = ka$ and $c = jb$. Combining these (through substitution) we get that $c = jka$. But jk is an integer, so this says that c is a multiple of a. Therefore $a|c$. QED

Proof by Contrapositive

Recall that an implication $P \to Q$ is logically equivalent to its contrapositive $\neg Q \to \neg P$. There are plenty of examples of statements which are hard to prove directly, but whose contrapositive can easily be proved direct. This is all that proof by contrapositive does. It gives a direct proof of the contrapositive of the implication, which since is equivalent to the original implication, is also a proof of that.

In terms of the skeleton of the proof, to prove $P \to Q$ by contrapositive, you will have,

Assume $\neg Q$. Explain, explain, ..., explain. Therefore $\neg P$.

As before, if there are variables and quantifiers, we set them to be arbitrary elements of our domain. Here are a couple examples.

Example: Is the statement "for all integers n, if n^2 is even, then n is even" true?

Solution: Note this is the converse of the statement we proved directly above. That is just a coincidence – it does not help us prove it at all. From trying a few examples, it definitely appears this is true. So let's proof it.

A direct proof of this statement would require fixing an arbitrary n and assuming that n^2 is even. But it is not at all clear how this would allow us to conclude anything about n - just because $n^2 = 2k$ does not in itself suggest how we could write n as a multiple of 2.

Let's try something else: write the contrapositive of the statement. We get, for all integers n, if n is odd then n^2 is odd. This looks much more promising. Our proof will look something like this:

Let n be an arbitrary integer. Suppose that n is not even. This means that.... In other words.... But this is the same as saying .... Therefore n^2 is not even.

Now we fill in the details.

Proof. We will prove the contrapositive. Let n be an arbitrary integer. Suppose that n is not even, and thus odd. Then $n = 2k + 1$ for some integer k. Now

$n^2 = (2k+1)^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1$. Since $2k^2 + 2k$ is an integer, we see that n^2 is odd and therefore not even. QED

Example: Prove: for all integers a and b, if $a + b$ is odd, then a is odd or b is odd.

Solution: The problem with trying a direct proof is that it will be hard to separate a and b from knowing something about $a + b$. It is much easier to combine them. Our proof will have the following format:

Let a and b be integers. Assume that a and b are both even. la la la. Therefore $a + b$ is even.

Now for a complete proof.

Proof. Let a and b be integers. Assume that a and b are even. Then $a = 2k$ and $b = 2l$ for some integers k and l. Now $a + b = 2k + 2l = 2(k + 1)$. Since $k + l$ is an integer, we see that $a + b$ is even, completing the proof. QED

Note that our assumption that a and b are even is really the negation of a or b is odd. We used De Morgan's law here.

We have seen how to prove some statements in the form of implications: either directly or by contrapositive. Some statements aren't written as implications to begin with.

Example: Consider the statement, for every prime number p, either $p = 2$ or p is odd. We can rephrase this: for every prime number p, if $p \neq 2$, then p is odd. Now try to prove it.

Proof. Let p be an arbitrary prime number. Assume p is not odd. So p is divisible by 2. Since p is prime, it must have exactly two divisors, and it has 2 as a divisor, so p must be divisible by only 1 and 2. Therefore $p = 2$. This completes the proof (by contrapositive). QED

Proof by Contradiction

There might be statements which really cannot be rephrased as implications. For example, "$\sqrt{2}$ is irrational." In this case, it is hard to know where to start. What can we assume? Well, say we want to prove the statement P. Now what if we could prove that $\neg P \to Q$ where Q was false? If this implication is true, and Q is false, what can we say about $\neg P$? It must be false as well - which makes P true!

This is why proof by contradiction works. If we can prove that $\neg P$ leads to a contradiction, then the only conclusion is that $\neg P$ is false, so P is true. That's what we wanted to prove. In other words, if it is impossible for P to be false, P must be true.

Here are a couple examples of proofs by contradiction.

Example: Prove that $\sqrt{2}$ is irrational.

Proof. Suppose not. Then $\sqrt{2}$ is equal to a fraction $\frac{a}{b}$. Without loss of generality, assume $\frac{a}{b}$ is in lowest terms (otherwise reduce the fraction). So,

$$2 = \frac{a^2}{b^2}$$

$$2b^2 = a^2$$

Thus a^2 is even, and as such a is even. So $a = 2k$ for some integer k, and $a^2 = 4k^2$. We then have,

$$2b^2 = 4k^2$$

$$b^2 = 2k^2$$

Thus b^2 is even, and as such b is even. Since a is also even, we see that $\frac{a}{b}$ is not in lowest terms, a contradiction. Thus $\sqrt{2}$ is irrational. QED

Example: Prove: there are no integers x and y such that $x^2 = 4y + 2$.

Proof. We proceed by contradiction. So suppose there *are* integers x and y such that $x^2 = 4y + 2 = 2(2y + 1)$. So x^2 is even. We have seen that this implies that x is even. So $x = 2k$ for some integer k. Then $x^2 = 4k^2$. This in turn gives $2k^2 = (2y + 1)$. But $2k^2$ is even, and $2y+1$ is odd, so these cannot be equal. Thus we have a contradiction, so there must not be any integers x and y such that $x^2 = 4y + 2$. QED

Example: The Pigeon Hole Principle: if more than n pigeons fly into n pigeon holes, then at least one pigeon hole will contain at least two pigeons. Prove this!

Proof. Suppose, contrary to stipulation, that each of the pigeon holes contain at most one pigeon. Then at most, there will be n pigeons. But we assumed that there are more than n pigeons, so this is impossible. Thus there must be a pigeon hole with more than one pigeon. QED

Note that while we phrased this proof as a proof by contradiction, we could have also used a proof by contrapositive as our contradiction was simply the negation of the hypothesis. Sometimes this will happen, in which case you can use either style of proof. There are examples however where the contradiction occurs "far away" from the original statement.

Proof by (counter)-Example

It is almost NEVER okay to prove a statement with just an example. Certainly none of the statements proved above can be proved through an example. This is because in each of those cases we are trying to prove that something holds of <u>all</u> integers. We claim that n^2 being even implies that n is even, *no matter what integer* n we pick. Showing that this works for $n = 4$ is not even close to enough.

This cannot be stressed enough. If you are trying to prove a statement of the form $\forall x P(x)$, you can absolutely NOT prove this with an example.[4]

However, existential statements can be proven this way. If we want to prove that there is an integer n such that $n^2 - n + 41$ is not prime, all we need to do is find one. This might seem like a silly thing to want to prove until you try a few values for n.

n	1	2	3	4	5	6	7
$n^2 - n + 41$	41	43	47	53	61	71	83

So far we have gotten only primes. You might be tempted to conjecture, "For all positive integers n, the number $n^2 - n + 41$ is prime." If you wanted to prove this, you would need to use a direct proof, a proof by contrapositive, or another style of proof, but certainly it is not enough to give even 7 examples. In fact, we can prove this conjecture is *false* by proving its negation: "There is a positive integer n such that $n^2 - n + 41$ is not prime." Since this is an existential statement, it suffices to show that there does indeed exist such a number.

In fact, we can quickly see that $n = 41$ will give 41^2 which is certainly not prime. You might say that this is a counter-example to the conjecture that $n^2 - n + 41$ is always even. Since so many statements in mathematics are universal and so their negations are existential, we can often prove that a statement is false by providing a counter-example.

Example: Above we proved, "for all integers a and b, if $a + b$ is odd, then a is odd or b is odd." Is the converse true?

Solution: The converse is the statement, "for all integers a and b, if a is odd or b is odd, then $a + b$ is odd." This is false! How do we prove it is false? We need to prove the negation of the converse. Let's look at the symbols. The converse is

$$\forall a \forall b((O(a) \vee O(b)) \rightarrow O(a + b))$$

We want to prove the negation:

$$\neg\forall a \forall b((O(a) \vee O(b)) \rightarrow O(a + b))$$

Simplify using the rules from the previous sections:

$$\exists a \exists b((O(a) \vee O(b)) \wedge \neg O(a + b))$$

As the negation passed by the quantifiers, they changed from $\forall$ to $\exists$. We then needed to take the negation of an implication, which is equivalent to asserting the if part and not the then part.

Now we know what to do. To prove that the converse is false we need to find two integers a and b so that a is odd or b is odd, but $a + b$ is not odd (so even). That's easy: 1 and 3. (remember, or means one or the other or both). Both of these are odd, but $1 + 3 = 4$ is not odd.

[4] This is not to say that looking at examples is a waste of time. Doing so will often give you an idea of how to write a proof. But the examples do not belong in the proof.

Proof by Cases

We could go on and on and on about different proof styles (we haven't even mentioned induction or combinatorial proofs here), but instead we will end with one final useful technique: proof by cases. The idea is this: to prove that P is true, we prove that $Q \to P$ and $\neg Q \to P$ for some statement Q. So not matter what, whether or not Q is true, we know that P is true. In fact, we could generalize this. Suppose we want to prove P. We know that at least one of the statements $Q_1, Q_2, \ldots, Q_n$ are true. If we can show that $Q_1 \to P$ and $Q_2 \to P$ and so on all the way to $Q_n \to P$, then we can conclude P. The key thing is that we want to be sure that one of our cases (the Q_i's) must be true no matter what.

If that last paragraph was confusing, hopefully an example will make things better.

Example: Prove: for any integer n, the number $(n^3 - n)$ is even.

Solution: It is hard to know where to start this, because we don't know much of anything about n. We might be able to prove that $n^3 - n$ is even if we knew that n was even. In fact, we could probably prove that $n^3 - n$ was even if n was odd. But since n must either be even or odd, this will be enough. Here's the proof.

Proof. We consider two cases: if n is even or if n is odd.

Case 1: n is even. Then $n = 2k$ for some integer k. This give

$$\begin{aligned} n^3 - n &= 8k^3 - 2k \\ &= 2(4k^2 - k) \end{aligned}$$

and since $4k^2 - k$ is an integer, this says that $n^3 - n$ is even.

Case 2: n is odd. Then $n = 2k + 1$ for some integer k. This gives

$$\begin{aligned} n^3 - n &= (2k+1)^3 - (2k+1) \\ &= 8k^3 + 6k^2 + 6k + 1 - 2k - 1 \\ &= 2(4k^3 + 3k^2 + 2k) \end{aligned}$$

and since $4k^3 + 3k^2 + 2k$ is an integer, we see that $n^3 - n$ is even again.

Since $n^3 - n$ is even in both exhaustive cases, we see that $n^3 - n$ is indeed always even. QED

Exercises

1. Consider the statement "for all integers a and b, if $a + b$ is even, then a and b are even"
   (a) Write the contrapositive of the statement
   (b) Write the converse of the statement
   (c) Write the negation of the statement.
   (d) Is the original statement true or false? Prove your answer.
   (e) Is the contrapositive of the original statement true or false? Prove your answer.
   (f) Is the converse of the original statement true or false? Prove your answer.

(g) Is the negation of the original statement true or false? Prove your answer.

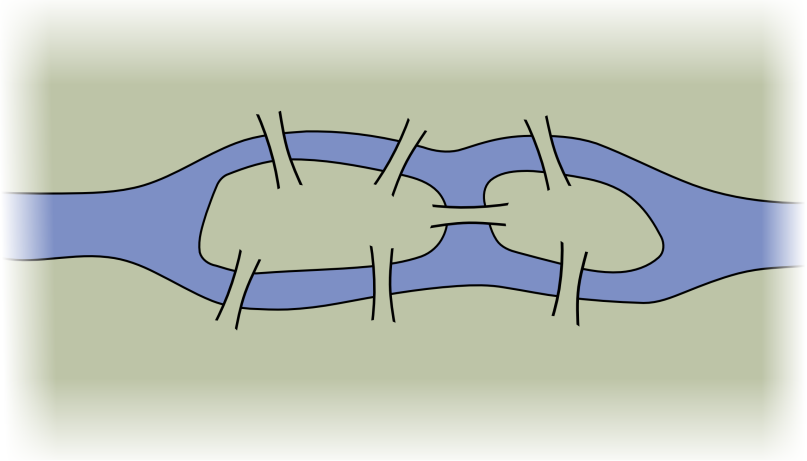
2. Prove that $\sqrt{3}$ is irrational.

Chapter 4

Graph Theory

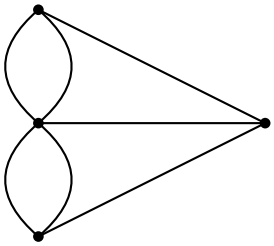
Graph Theory is a relatively new area of mathematics, first studied by the super famous mathematician Leonhard Euler in 1735. Since then it has blossomed in to a powerful tool used in nearly every branch of science and is currently an active area of mathematics research. We will begin our study with the problem that started it all: The Seven Bridges of Königsberg.

> In the time of Euler, in the town of Königsberg in Prussia, there was a river containing two islands. The islands were connected to the banks of the river by seven bridges (as seen below). The bridges were very beautiful, and on their days off, townspeople would spend time walking over the bridges. As time passed, a question arose: was it possible to plan a walk so that you cross each bridge once and only once? Euler was able to answer this question. Are you?



Try finding a path which uses each bridge exactly once.

Here is another problem: below is a drawing of four dots connected by some lines. Is it possible to trace over each line once and only once (without lifting up your pencil)? You must start and end on one of the dots.



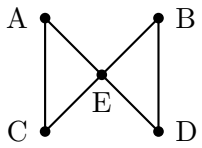
There is an obvious connection between these two problems - any path through the dot and line drawing corresponds exactly to a path over the bridges of Königsberg.

Pictures like the dot and line drawing are called *graphs*. Graphs are made up of a collection of dots (called *vertices*) and lines connecting those dots (called *edges*). The nice thing about looking at graphs instead of pictures of rivers, islands and bridges is that we now have a mathematical object to study. A *set* of vertices and a set of edges (in fact, we can take the set of edges to be a set of two element subsets from the set of vertices). We have distilled the "important" parts of the bridge picture for the purposes of the problem - it does not matter how big the islands are, what the bridges are made out of, if the river contains alligators, etc. All that matters is which land masses are connected to which other land masses, and how many times. This was the great insight that Euler had.

We will return to the question of finding paths through graphs later. But first, here are a few other situations you can represent with graphs.

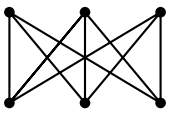
Example: Al, Bob, Cam, Dan, and Euclid are all members of the social networking website *Facebook*. The site allows members to be "friends" with each other. It turns out that Al and Cam are friends, as are Bob and Dan. Euclid is friends with everyone. Represent this situation with a graph.

Solution: Each person will be represented by a vertex and each friendship will be represented by an edge - that is, there will be an edge between two vertices if and only if the people represented by those vertices are friends. We get the following graph:



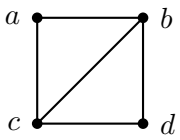
Example: Each of three houses must be connected to each of three utilities. Is it possible to do this without any of the utility lines crossing?

Solution: We will answer this question later. For now, notice how we would ask this question in the context of graph theory. We are really asking whether it is possible to redraw the graph below without any edges crossing (except at vertices).

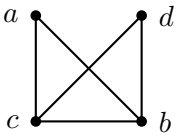


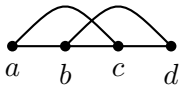
4.1 Basics

While we almost always think of graphs as pictures, these are really just visual representations of mathematical objects. In fact, all a graph is is a set of vertices some pairs of which are "related" by an edge. For example, we can describe a particular graph like this: the vertices are the letters $\{a, b, c, d\}$ and the edges are the pairs $\{(a, b), (a, c), (b, c), (b, d), (c, d)\}$. Technically the edges are 2-elements subsets of the set of vertices, so we should write (a, b) as $\{a, b\}$ but we use parentheses to ease reading. We could have described the graph in words as follows: we have four vertices, a, b, c, and d, and a is connected (by an edge) to b and c, b is connected to both c and d, and c is also connected to d. One way to draw this graph is this:



However we could also have drawn the graph differently. For example either of these:





Viewed as pictures, the graphs above are the same whether or not the vertices are labeled as they are, or at all. In fact, two graphs are equal precisely when there is a way to label the vertices so that all the pairs of vertices which have edges in one graph also have edges in the other graph, and *vice versa*. Two vertices which are connected by an edge are called *adjacent*.

Notice that the graph above has the property that no pair of vertices is connected more than once, and no vertex is connected to itself. Graphs like these are sometimes called *simple*, although we will just call them *graphs*. This is because our definition for a graph says that the edges form a set of 2-element subsets of the vertices. Remember that it doesn't make sense to say a set contains an element more than once. So no pair of vertices can be connected by an edge more than once. And since each edge must be a set containing two vertices, we cannot have a single vertex connected to itself by an edge.

There are times we want to consider double (or more) edges and single edge loops. For example, the "graph" we drew for the Bridges of Königsberg problem had double edges because there really are two bridges connecting a particular island to the near shore. We will call these objects *multigraphs*. This is a good name: a multiset is a set in which we are allowed to include a single element multiple times.

The graph above is also *connected* - you can get from any vertex to any other vertex by following some path of edges (a and d are connected by a path two edges long). A graph that is not connected can be thought of as two separate graphs drawn close together. Unless otherwise stated, we will assume all graphs are connected.

The graph above does not have an edge between a and d. Thus it is possible to add an edge to the graph. If we add all possible edges, then the resulting graph is said to be *complete*. That is, a graph is complete if every pair of vertices is connected by an edge. Since a graph is determined completely by which vertices are adjacent to which other vertices, there is only one complete graph with a given number of vertices. We give these a special name: K_n is the complete graph on n vertices.

Each vertex in K_n is adjacent to $n-1$ other vertices. We call the number of edges emanating from a given vertex the *degree* of that vertex. So every vertex in K_n has degree $n-1$. How many edges does K_n have? One might think the answer should be $n(n-1)$, since we count $n-1$ edges n times (once for each vertex). However, each edge is adjacent to 2 vertices, so we count every edge exactly twice in this way. Thus there are $n(n-1)/2$ edges in K_n. Alternatively, we can say there are $\binom{n}{2}$ edges, since to draw an edge we must choose 2 of the n vertices.

In general, if we know the degrees of all the vertices in a graph we can find the number of edges. The sum of the degrees of all vertices will always be twice the number of edges, since each edge adds to the degree of two vertices. Notice this means that the sum of the degrees of all vertices in any graph must be even!

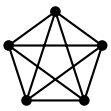
Example: At a recent math seminar, 9 mathematicians greeted each other by shaking hands. Is it possible that each mathematician shook hands with exactly 7 people at the seminar?

Solution: It seems like this should be possible - each mathematician chooses one person to not shake hands with. But it is impossible. We are asking whether a graph with 9 vertices can have each vertex have degree 7. If such a graph existed, the sum of the degrees of the vertices would be $9 \cdot 7 = 63$. This would be twice the number of edges (handshakes) so this says that the graph would have 31.5 edges. That is impossible - edges only come in whole numbers. Thus at least one (in fact an odd number) of the mathematicians must have shaken hands with an *even* number of people at the seminar.

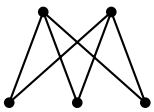
One final definition: we say a graph is *bipartite* if the vertices can be divided into two sets, A and B, with no two vertices in A adjacent and no two vertices in B adjacent. Of course the vertices in A can be adjacent to some or all of the vertices in B. If all vertices in A and adjacent to all the vertices in B, then the graph is a *complete bipartite graph*, and gets a special name: $K_{m,n}$, where $|A| = m$ and $|B| = n$. The graph in the houses and utilities puzzle is $K_{3,3}$.

Named Graphs
Some graphs are used more than others, and get special names.

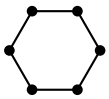
- K_n: the complete graph on n vertices.
- $K_{m,n}$: the complete bipartite graph with sets of m and n vertices.
- C_n: the cycle graph on n vertices - just one big loop.
- P_n: the path graph on n vertices - just one long path.



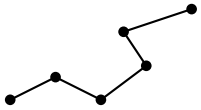
K_5



$K_{2,3}$



C_6



P_6

Graph Theory Definitions

Graph: A collection of *vertices*, some of which are connected by *edges*. More precisely, a pair of sets V and E where V is a set of vertices and E is a set of 2-element subsets of V.

Adjacent: Two vertices are *adjacent* if they are connected by an edge. Two edges are *adjacent* if they share a vertex.

Bipartite graph: A graph for which it is possible to divide the vertices into two disjoint sets such that there are no edges between any two vertices in the same set.

Complete bipartite graph: A bipartite graph for which every vertex in the first set is adjacent to every vertex in the second set.

Complete graph: A graph with edges connecting every pair of vertices.

Connected: A graph is *connected* if there is a path from any vertex to any other vertex.

Cycle: A path (see below) that starts and stops at the same vertex, but contains no other repeated vertices.

Degree of a vertex: The number of edges connected to a vertex is called the *degree* of the vertex.

Euler path: A path which uses each edge exactly once.

Euler circuit: An Euler path which starts and stops at the same vertex.

Multigraph: A *multigraph* is just like a graph but can contain multiple edges between two vertices as well as single edge loops (that is an edge from a vertex to itself).

Path: A sequence of vertices such that consecutive vertices (in the sequence) are adjacent (in the graph). A path in which no vertex is repeated is called *simple*.

Planar: A graph is planar if it is possible to draw it (in the plane) without any edges crossing.

Subgraph: We say that H is a subgraph of G if every vertex and edge of H is also a vertex or edge of G. We say H is an *induced* subgraph of G if every vertex of H is a vertex of G and for pair of vertices in H are adjacent in H if and only if they are adjacent in G.

Tree: A (connected) graph with no cycles. (A non-connected graph with no cycles is called a *forest*.) The vertices in a tree with degree 1 are called *leaves*.

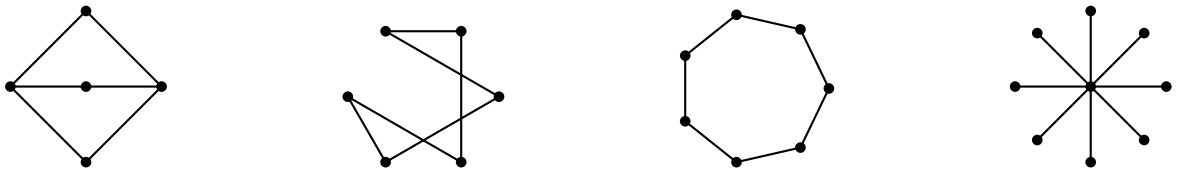
Exercises

1. If 10 people each shake hands with each other, how many handshakes took place? What does this question have to do with graph theory?

2. Among a group of 5 people, is it possible for everyone to be friends with exactly 2 of the people in the group? What about 3 of the people in the group?

3. Is it possible for two *different* graphs to have the same number of vertices and the same number of edges? What if the degrees of the vertices in the two graphs are the same (so both graphs have vertices with degrees 1, 2, 2, 3, and 4, for example)? Draw two such graphs or explain why not.

4. Which of the graphs below are bipartite?

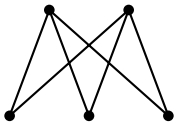


5. For which n is the graph C_n bipartite?

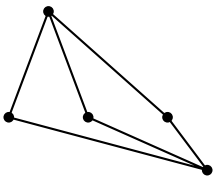
4.2 Planar Graphs

When is it possible to draw a graph so that none of the edges cross, except for at vertices? If this is possible, we say the graph is *planar* (since you can draw it on the *plane*).

Notice that the definition of planar includes the phrase "it is possible to." This means that even if a graph does not look like it is planar, it still might be. Perhaps you can redraw it in a way in which no edges cross. For example, this is a planar graph:



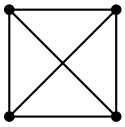
That is because we can redraw it like this:

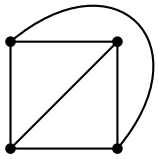


The graphs are the same graph, so if one is planar, the other must be too. The original drawing of the graph was not a *planar representation* of the graph.

Now when a planar graph is drawn without edges crossing, the graph divides the plane into regions. We will call each region a *face*. The graph above has 3 faces (yes, we **do** include the "outside" region as a face). The number of faces does not change no matter how you draw the graph (as long as you do so without the edges crossing), so it makes sense to ascribe the number of faces as a property of the planar graph.

A warning: you can only count faces when the graph is drawn in a planar way. For example, consider these two representations of the same graph:





If you try to count faces using the graph on the left, you might say there are 5 faces (including the outside). But drawing the graph with a planar representation shows that in fact there are only 4 faces.

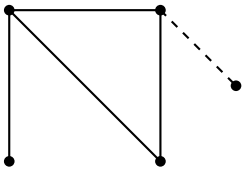
There is a connection between the number of vertices (V), the number of edges (E) and the number of faces (F) in any connected planar graph. This relationship is called Euler's Formula.

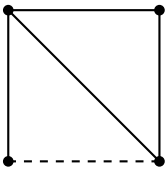
Euler's Formula for Planar Graphs
For any (connected) planar graph with V vertices, E edges and F faces, we have

$$V - E + F = 2$$

Why is Euler's formula true? One way to convince yourself of its validity is to draw a planar graph step by step. Start with the graph P_2:

Any connected graph (besides just a single isolated vertex) must contain this subgraph. Now build up to your graph by adding edges and vertices. Each step will consist of either adding a new vertex connected by a new edge to part of your graph (so creating a new "spike") or by connecting two vertices already in the graph with a new edge (completing a circuit).



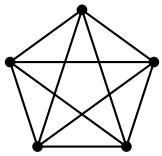


What do these "moves" do? When adding the spike, the number of edges increases by 1, the number of vertices increases by one, and the number of faces remains the same. But this means that $V - E + F$ does not change. Completing a circuit adds one edge, adds one face, and keeps the number of vertices the same. So again, $V - E + F$ does not change.

Since we can build any graph using a combination of these two moves, and doing so never changes the quantity $V - E + F$, that quantity will be the same for all graphs. But notice that our starting graph P_2 has $V = 2$, $E = 1$ and $F = 1$, so $V - E + F = 2$.

Non-planar graphs

Not all graphs are planar. If there are too many edges and too few vertices, then some of the edges will need to intersect. The first time this happens is in K_5.



If you try to redraw this without edges crossing, you quickly get into trouble. There seems to be one edge too many. In fact, we can prove that no matter how you draw it, K_5 will always have edges crossing.

Theorem. *K_5 is not planar.*

Proof. The proof is by contradiction. So assume that K_5 were planar. Then the graph would satisfy Euler's formula for planar graphs. K_5 has 5 vertices and 10 edges, so we get

$$5 - 10 + F = 2$$

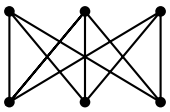
which says that if the graph were drawn without any edges crossing, there would be $F = 7$ faces.

Now consider how many edges surround each face. Each face must be surrounded by at least 3 edges (since K_5 is simple - it contains no double edges or loops). Let B be the total number of *boundaries* around all the faces in the graph. Thus we have that $B \geq 3F$. But also $B = 2E$, since each edge is used as a boundary exactly twice. Putting this together we get

$$3F \leq 2E$$

But this is impossible, since we have already determined that $F = 7$ and $E = 10$, and $21 \not\leq 20$. This is a contradiction so in fact K_5 is not planar. QED

The other simplest graph which is not planar is $K_{3,3}$



Proving that $K_{3,3}$ is not planar answers the houses and utilities puzzle - it is not possible to connect each of three houses to each of three utilities without the lines crossing.

Theorem. *$K_{3,3}$ is not planar.*

Proof. Again, we proceed by contradiction. Suppose $K_{3,3}$ were planar. Then by Euler's formula there will be 5 faces, since $V = 6$, $E = 9$, and $6 - 9 + F = 2$.

How many boundaries surround these 5 faces? Let B be this number. Since each edge is used as a boundary twice, we have $B = 2E$. Also, $B \geq 4F$ since each face is surrounded by 4 or more boundaries. We know this is true because the graph is simple (so there are no faces surrounded by 1 or 2 boundaries) and because $K_{3,3}$ is bipartite, so does not contain any 3-edge cycles. Thus

$$4F \leq 2E$$

But this would say that $20 \leq 18$, which is clearly false. Thus $K_{3,3}$ is not planar. QED

Note the similarities and differences in these proofs. Both are proofs by contradiction, and both start with using Euler's formula to derive the (supposed) number of faces in the graph. Then we find a relationship between the number of faces and the number of edges based on how many edges surround each face. This is the part that changes - in the proof for K_5, we got $3F \leq 2E$ and for $K_{3,3}$ we go $4F \leq 2E$. The coefficient of F is the key. It is the smallest number of edges which could surround any face. If some number of edges surround a face, then these edges form a circuit. So that number is the size of the smallest circuit in the graph.

In general, if we let g be the size of the smallest cycle in a graph (g stands for *girth*, which is the technical term for this) then for any planar graph we have $gF \leq 2E$. When this disagrees with Euler's formula, we know for sure that the graph cannot be planar.

Exercises

1. Is it possible for a planar graph to have 6 vertices, 10 edges and 5 faces? Explain.
2. The graph G has 6 vertices with degrees $2, 2, 3, 4, 4, 5$. How many edges does G have? Could G be planar? If so, how many faces would it have.
3. If a graph has 10 vertices and 10 edges and contains an Euler circuit, must it be planar? How many faces would it have?

4.3 Coloring

We move now to perhaps the most famous graph theory problem - how to color maps.

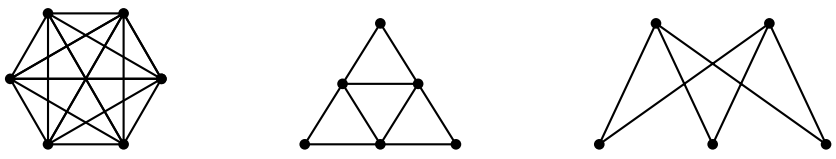
> **Question:** Given any map of countries, states, counties, etc. how many colors are needed to color each region on the map so that neighboring regions are colored differently?

Actual map makers usually use around seven colors - for one thing, they require watery regions to be a specific color, and with a lot of colors it is easier to find a permissible coloring. But we want to know whether there is a smaller palette that will work for any map.

How is this related to graph theory? Well, if we place a vertex in the center of each region (say in the capital of each state) and then connect two vertices if their states share a border, we get a graph. The coloring regions on the map corresponds to coloring the vertices of the graph. Since neighboring regions cannot be colored the same, our graph cannot have vertices colored the same when those vertices are adjacent (connected by an edge).

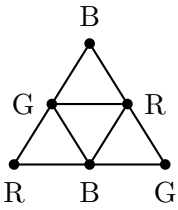
In general, given any graph, a coloring of the vertices is called (not surprisingly) a *vertex coloring*. If the vertex coloring has the property that adjacent vertices are colored differently, then the coloring is called *proper*. Every graph has a proper vertex coloring - for example, you could color every vertex with a different color. But often you can do better. The smallest number of colors needed to get a proper vertex coloring is called the *chromatic number* of the graph.

Example: Find the chromatic number of the graphs below.



Solution: The graph on the left is K_6. The only way to properly color the graph is to give every vertex a different color (since every vertex is adjacent to every other vertex). Thus the chromatic number is 6.

The middle graph can be properly colored with just 3 colors (Red, Blue, and Green). For example:



There is no way to color it with just two colors, since there are three vertices mutually adjacent (i.e., a triangle). Thus the chromatic number is 3.

The graph on the right is just $K_{2,3}$. As with all bipartite graphs, this graph has chromatic number 2 - color the vertices in the top row red and the vertices on the bottom row blue.

It appears that graphs can have any chromatic number. It should not come as a surprise that K_n has chromatic number n. So how could there possibly be an answer to the original map coloring question? If the chromatic number of graph can be arbitrarily large, then it seems like there would be no upper bound to the number of colors needed for any map. But there is.

The key observation is that while it is true that for any number n, there is a graph with chromatic number n, only some graphs arrive as representations of maps. If you convert a map to a graph, the edges between vertices correspond to borders between the countries. So you should be able to connect vertices in such a way where the edges do not cross. In other words, the graphs representing maps are all *planar*!

So the question is, what is the larges chromatic number of any planar graph? The answer is one of the best know theorems of mathematics:

Theorem. *The Four Color Theorem*
If G is a planar graph, then the chromatic number of G is less than or equal to 4. Thus any map can be colored with 4 or fewer colors.

We will not prove this theorem. Really. Even though the theorem is easy to state and understand, the proof is not. In fact, there is currently no "easy" known proof of the theorem. The current best proof still requires powerful computers to check an *unavoidable set* of 633 *reducible configurations*. The idea is that every graph must contain one of these reducible configurations (this fact also needs to be checked by a computer) and that reducible configurations can in fact be colored in 4 or fewer colors.

Non-planar graphs

PUT IN STUFF ABOUT COLORING IN GENERAL, INCLUDING APPLICATIONS.

Other colorings

MAYBE ADD VIZINGS THEOREM??

The chromatic number of a graph tells us about coloring vertices - but we could also ask about coloring edges. What if we colored every edge of a graph either red or blue. Can we do so without,

say, creating a triangle of same like colored edges (i.e., an all red or all blue triangle - we say the triangle is *mono-chromatic*)? Certainly for some graphs the answer is yes - try doing so for K_4. What about K_5? K_6? How far can we go?

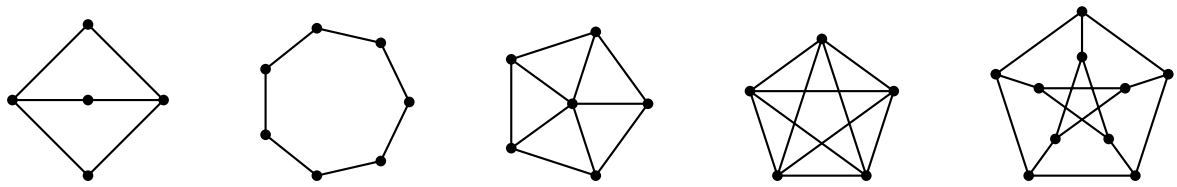
The answer the above problem is known - I encourage you to try to solve it. We could extend the question though - what if we had three colors? What if we were trying to avoid other graphs. The surprising fact is that very little is known about these questions. For example, we know that you need to go up to K_{17} in order to force a mono-chromatic triangle using three colors, but nobody knows how big you need to go with more colors. Similarly, we know that using two colors K_{18} is the smallest graph that forces a mono-chromatic copy of K_4, but the best we have to force a mono-chromatic K_5 is a range - somewhere from K_{43} to K_{49}.

Exercises

1. What is the smallest number of colors you need to properly color the vertices of $K_{4,5}$. That is, find the chromatic number of the graph.

2. Draw a graph with chromatic number 6 (i.e., which requires 6 colors to properly color the vertices). Could your graph be planar? Explain.

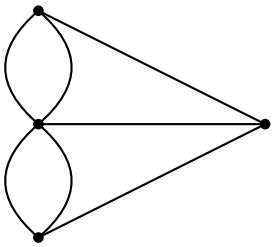
3. Find the chromatic number of each of the following graphs.



4.4 Euler Paths and Circuits

If we start at a vertex and trace along edges to get to other vertices, we create a *path* on the graph. If the path travels along every edge exactly once, then the path is called an *Euler path* (or *Eulerian path*). If in addition, the starting and ending vertices are the same (so you trace along every edge exactly once and end up where you started) then the path is called an *Euler circuit*. Of course if a graph is not connected, there is no hope of finding such a path or circuit. For the rest of this section, assume all graphs are connected.

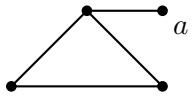
The bridges of Königsberg problem is really a question about the existence of Euler paths. There will be a route that crosses every bridge exactly once if and only if the graph below has an Euler path:



This graph is small enough that we could actually check every possible path and in doing so convince ourselves that there is no Euler path (let alone an Euler circuit). On small graphs which

do have an Euler path, it is usually not difficult to find one. Our goal is to find a quick way to check whether a graph has an Euler path or circuit, even if the graph is quite large.

If you have not already, be sure to try the worksheet on Euler paths. It is quite instructive to build graphs which do and do not have Euler paths. One way to guarantee that a graph does *not* have an Euler circuit is to include a "spike" - a vertex of degree 1.



The vertex a has degree 1, and if you try to make an Euler circuit, you see that you will get stuck at the vertex. It is a dead end. That is, unless you start there. But then there is no way to return, so there is no hope of finding an Euler circuit. There is however an Euler path - you can start at the vertex a, then loop around the triangle. You will end at the vertex of degree 3.

You run into a similar problem whenever you have a vertex of odd degree. If you start at such a vertex, you will not be able to end there (after visiting every edge exactly once). After using one edge to leave the starting vertex, you will be left with an even number of edges emanating from the vertex. Half of these could be used for returning to the vertex, the other half for leaving. So you return, then leave. Return, then leave. The only way to use up all the edges is to use the last one by leaving the vertex. On the other hand, if you have a vertex with odd degree that you do not start a path at, then you will eventually get stuck at that vertex. The path will use pairs of edges connected the the vertex to arrive and leave again. Eventually all but one of these edges will be used up, leaving only an edge to arrive by, and none to leave again.

What all this says is that if a graph has an Euler path and two vertices with odd degree, then the Euler path must start at one of the odd degree vertices and end at the other. In such a situation, every other vertex *must* have an even degree - since we need an equal number of edges to get to those vertices as to leave them. How could we have an Euler circuit? The graph could not have an odd degree vertex - an Euler path would have to start there or end there, but not both. Thus for a graph to have an Euler circuit, all vertices must have even degree.

The converse is also true: if all the vertices of a graph have even degree, then the graph has an Euler circuit, and if there are exactly two vertices with odd degree, the graph has an Euler path. To prove this is a little tricky, but the basic idea is that you will never get stuck because there is an "outbound" edge for every "inbound" edge at every vertex. If you try to make an Euler path and miss some edges, you will always be able to "splice in" a circuit using the edges you previously missed.

Euler Paths and Circuits

- A graph has an Euler circuit if and only if the degree of every vertex is even.
- A graph has an Euler path if and only if there are at most two vertices with odd degree.

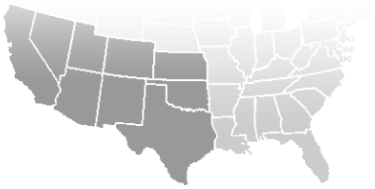
Since the bridges of Königsberg graph has all four vertices with odd degree, there is no Euler path through the graph.

Hamilton paths

Suppose you wanted to tour Königsberg in such a way where you visit each land mass (the two islands and both banks) exactly once. This can be done. In graph theory terms we are asking whether there is a path which visits every vertex exactly once. Such a path is called a *Hamilton path* (or Hamiltonian path). It appears that finding Hamilton paths would be easier - graphs often have more edges than vertices, so there are fewer requirements to be met. However, nobody knows whether this is true. There is no known simple test for whether a graph has a Hamilton path. For small graphs this is not a problem, but as the size of the graph grows, it gets harder and harder to check wither there is a Hamilton path. In fact, this is an example of a question which as far as we know is too difficult for computers to solve - it is an example of a problem which is NP-complete.

Exercises

1. You and your friends want to tour the southwest by car. You will visit the nine states below, with the following rather odd rule: you must cross each border between neighboring states exactly once (so, for example, you must cross the Colorado-Utah border exactly once). Can you do it? If so, does it matter where you start your road trip? What fact about graph theory solves this problem?



2. Which of the following graphs contain an Euler path? Which contain an Euler circuit?
   (a) K_4
   (b) K_5.
   (c) $K_{5,7}$
   (d) $K_{2,7}$
   (e) C_7
   (f) P_7

3. For which n does the graph K_n contain an Euler circuit? Explain.

4. For which m and n does the graph $K_{m,n}$ contain an Euler path? An Euler circuit? Explain.

4.5 Matching in Bipartite Graphs

We conclude with one more example of a graph theory problem to illustrate the variety and vastness of the subject.

Suppose you have a (not-necessarily-complete) bipartite graph G. This will consist of two sets of vertices A and B with some edges connecting some vertices of A to some vertices in B (but of course, no edges between two vertices both in A or both in B). A *matching of* A is a subset of the edges for which each vertex of A belongs to exactly one edge of the subset, an no vertex in B belongs to more than one edge in the subset. In practice we will assume that $|A| = |B|$ (the

two sets have the same number of vertices) so this says that every vertex in the graph belongs to exactly one edge in the matching.

Some context might make this easier to understand. Think of the vertices in A as representing students in a class, and the vertices in B as representing presentation topics. We put an edge from a vertex $a \in A$ to a vertex $b \in B$ if student a would like to present on topic b. Of course, some students would want to present on more than one topic, so their vertex would have degree greater than 1. As the teacher you want to assign each student their own unique topic. Thus you want to find a matching of A: you pick some subset of the edges so that each student gets matched up with exactly one topic, and no topic gets matched to two students.[1]

The question is: when does a bipartite graph contain a matching of A? To begin to answer this question, consider what could prevent the graph from containing a matching. This will not necessarily tell us a condition when the graph *does* have a matching, but at least it is a start.

One way G could not have a matching is if there is a vertex in A not adjacent to any vertex in B (so having degree 0). What else? What if two students both like the same one topic, and no others? Then after assigning that one topic to the first student, there is nothing left for the second student to like, so it is very much as if the second student has degree 0. Or what if three students like only two topics between them. Again, after assigning one student a topic, we reduce this down to the previous case of two students liking only one topic. We can continue this way with more and more students.

It should be clear at this point that if there is every a group of n students who as a group like $n-1$ or fewer topics, then no matching is possible. This is true for any value of n, and any group of n students.

To make this more graph-theoretic, say you have a set $S \subseteq A$ of vertices. Define $N(S)$ to be the set of all the *neighbors* of vertices in S. That is, $N(S)$ contains all the vertices (in B) which are adjacent to at least one of the vertices in S. (In the student/topic graph, $N(S)$ is the set of topics liked by the students of S.) Our discussion above can be summarized as follows:

Matching Condition
If a bipartite graph $G = \{A, B\}$ has a matching of A, then

$$|N(S)| \geq |S|$$

for all $S \subseteq A$.

Is the converse true? That is, suppose G satisfies the matching condition $|N(S)| \geq |S|$ for all $S \subseteq A$ (that is, every set of vertices has at least as many neighbors than vertices in the set). Does that mean that there is a matching? Surprisingly yes. The obvious necessary condition is also sufficient.[2] This is a theorem first proved by Philip Hall in 1935.[3]

Theorem (Hall's Marriage Theorem)**.** *Let G be a bipartite graph with sets A and B. Then G has a matching of A if and only if*

$$|N(S)| \geq |S|$$

[1] The standard example for matchings used to be the *marriage problem* in which A consisted of the men in the town, B the women, and an edge represented a marriage that was agreeable to both parties. A matching then represented a way for the town elders to marry off everyone in the town, no polygamy allowed. We have chosen a more progressive context for the sake of political correctness.

[2] This happens often in graph theory. If you can avoid the obvious counter-examples, you often get what you want.

[3] There is also an infinite version of the theorem which was proved by Marshal Hall, Jr. The name is a coincidence though as the two Halls are not related.

for all $S \subseteq A$.

There are quite a few different proofs of this theorem – a quick internet search will get you started.

In addition to its application to marriage and student presentation topics, matchings have applications all over the place. We finish with one such example.

Example: Suppose you deal 52 regular playing cards into 13 piles of 4 cards each. Prove that you can always select one card from each pile to get one of each of the 13 card values Ace, 2, 3, ..., 10, Jack, Queen, and King.

Solution: Doing this directly would be difficult, but we can use the matching condition to help. Construct a graph G with 13 vertices in the set A, each representing one of the 13 card values, and 13 vertices in the set B, each representing one of the 13 piles. Draw an edge between a vertex $a \in A$ to a vertex $b \in B$ if a card with value a is in the pile b. Notice that we are just looking for a matching of A; each value needs to be found in the piles exactly once.

We will have a matching if the matching condition holds. Given any set of card values (a set $S \subseteq A$) we must show that $|N(S)| \geq |S|$. That is, the number of piles that contain those values is at least the number of different values. But what if it wasn't? Say $|S| = k$. If $|N(S)| < k$, then we would have fewer than $4k$ different cards in those piles (since each pile contains 4 cards). But there are $4k$ cards with the k different values, so at least one of these cards must be in another pile, a contradiction. Thus the matching condition holds, so there is a matching, as required.

Appendix A

Additional Topics

A.1 Generating Functions

There is an extremely powerful tool in discrete mathematics – a method of manipulating sequences called the generating function. The idea is this: instead of an infinite sequence (for example: $2, 3, 5, 8, 12, \ldots$) we look at a single function which encodes the sequence. But not a function which gives the nth term as output. Instead, a function whose power series (like from calculus 2) "displays" the terms of the sequence. So for example, we would look at the power series $2 + 3x + 5x^2 + 8x^3 + 12x^4 + \cdots$ which displays (as coefficients) the sequence $2, 3, 5, 8, 12, \ldots$.

An infinite power series is simply an infinite sum of terms of the form $c_n x^n$ were c_n is some constant. So we might write a power series like this:

$$\sum_{k=0}^{\infty} c_k x^k$$

or expanded like this

$$c_0 + c_1 x + c_2 x^2 + c_3 x^3 + c_4 x^4 + c_5 x^5 + \cdots$$

When viewed in the context of generating functions, we call such a power series a *generating series*. The generating series generates the sequence

$$c_0, c_1, c_2, c_3, c_4, c_5, \ldots$$

that is, the sequence generated by a generating series is simply the sequence of *coefficients* of the infinite polynomial.

> **Example**: What sequence is represented by the generating series $3 + 8x^2 + x^3 + \frac{x^5}{7} + 100x^6 + \cdots$?
>
> *Solution*: We just read off the coefficients of each x^n term. So $a_0 = 3$ since the coefficient of x^0 is 3 ($x^0 = 1$ so this is the constant term). What is a_1? It is NOT 8, since 8 is the coefficient of x^2, so 8 is the term a_2 of the sequence. To find a_1 we need to look for the coefficient of x^1 which in this case is 0. So $a_1 = 0$. We have $a_3 = 1$, $a_4 = 0$, and $a_5 = \frac{1}{7}$. So we have the sequence
>
> $$3, 0, 8, 1, \frac{1}{7}, 100, \ldots$$
>
> Note, when discussing generating functions, we always start our sequence with a_0.

Now you might very naturally ask why we would do such a thing. One reason is that encoding a sequence with a power series helps us keep track of which term is which in the sequence. For example, if we write the sequence $1, 3, 4, 6, 9, \ldots, 24, 41, \ldots$ it is impossible to which term 24 is (even if we agreed that the first term was supposed to be a_0). However, if wrote the generating series instead we would have $1 + 3x + 4x^2 + 6x^3 + 9x^4 + \cdots + 24x^{17} + 41x^{18} + \cdots$. Now it is clear that

24 is the 17th term of the sequence (that is, $a_{17} = 24$). Of course to get this benefit we could have displayed our sequence in any number of ways - perhaps $\boxed{1}_0\boxed{3}_1\boxed{4}_2\boxed{6}_3\boxed{9}_4\cdots\boxed{24}_{17}\boxed{41}_{18}\cdots$ - but we do not do this. The reason is that the generating series looks like an ordinary power series (although we are interpreting it differently) so we can do things with it that we ordinarily do with power series - such as write down what it converges to.

For example, from calculus we know that the power series $1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \cdots + \frac{x^n}{n!} + \cdots$ converges to the function e^x. So we can use e^x as a way of talking about the sequence of coefficients of the power series for e^x. When we write down a nice compact function which has an infinite power series that we view as a generating series, then we call that function a *generating function*. In this example we would say,

$$1, 1, \frac{1}{2}, \frac{1}{6}, \frac{1}{24}, \ldots, \frac{1}{n!}, \ldots \text{ has generating function } e^x$$

A.1.1 Building Generating Functions

The e^x example is very specific - we have that one rather odd sequence, and the only reason we know its generating function is because we happen to know the Taylor series for e^x. Our goal now is to gather some tricks to build the generating function when given a particular sequence.

Let's see what the generating functions are for some very simple sequences. The simplest of all: 1, 1, 1, 1, 1, ... What does the *generating series* look like? It is simply $1 + x + x^2 + x^3 + x^4 + \cdots$. Now, can we find a closed formula for this power series? Yes! This particular series is really just a geometric series with common ratio x. So if we use our "multiply, shift and subtract" technique, we have

$$\begin{aligned} S &= 1 + x + x^2 + x^3 + \cdots \\ -xS &= \quad x + x^2 + x^3 + x^4 + \cdots \\ (1-x)S &= 1 \end{aligned}$$

Therefore we see that

$$1 + x + x^2 + x^3 \cdots = \frac{1}{1-x}$$

You might remember from calculus that this is only true on the interval of convergence for the power series - when $|x| < 1$. That is true for us, but we don't care - we are never going to plug anything in for x, so as long as there is some value of x for which the generating function and generating series agree, we are happy. And in this case we are happy:

The generating function for $1, 1, 1, 1, 1, 1, \ldots$ is $\dfrac{1}{1-x}$

Now let's use this basic generating function to find generating functions for more sequences. What if we replace x by $-x$. We get

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + \cdots \text{ which generates } 1, -1, 1, -1, \ldots$$

If we replace x by $3x$ we get

$$\frac{1}{1-3x} = 1 + 3x + 9x^2 + 27x^3 + \cdots \text{ which generates } 1, 3, 9, 27, \ldots$$

So by replacing the x in $\frac{1}{1-x}$ we can get generating functions for a variety of sequences. But not all. For example, you cannot plug in anything for x to get the generating function for $2, 2, 2, 2, \ldots$. However, we are not lost yet. Notice that each term of $2, 2, 2, 2, \ldots$ is the result of multiplying the terms of $1, 1, 1, 1, \ldots$ by a constant (2). So let's multiply the generating function by 2 as well.

$$\frac{2}{1-x} = 2 + 2x + 2x^2 + 2x^3 + \cdots \text{ which generates } 2, 2, 2, 2, \ldots$$

Similarly, to find the generating function for the sequence $3, 9, 27, 81, \ldots$, we note that this sequence is the result of multiplying each term of $1, 3, 9, 27, \ldots$ by 3. Since we have the generating function for $1, 3, 9, 27, \ldots$ we can say

$$\frac{3}{1-3x} = 3 \cdot 1 + 3 \cdot 3x + 3 \cdot 9x^2 + 3 \cdot 27x^3 + \cdots \text{ which generates } 3, 9, 27, 81, \ldots$$

What about the sequence $2, 4, 10, 28, 82, \ldots$? Here the terms are always 1 more than powers of 3. We have added the sequences $1, 1, 1, 1, \ldots$ and $1, 3, 9, 27, \ldots$ term by term. Therefore we can get a generating function by adding the respective generating functions:

$$2 + 4x + 10x^2 + 28x^3 + \cdots = (1+1) + (1+3)x + (1+9)x^2 + (1+27)x^3 + \cdots = \frac{1}{1-x} + \frac{1}{1-3x}$$

The fun does not stop there: if we replace x in our original generating function by x^2 we get

$$\frac{1}{1-x^2} = 1 + x^2 + x^4 + x^6 \cdots \text{ which generates } 1, 0, 1, 0, 1, 0, \ldots$$

How could we get $0, 1, 0, 1, 0, 1, \ldots$? Start with the previous sequence and *shift* it over by 1. But how do you do this? To see how shifting works, let's first try to get the generating function for the sequence $0, 1, 3, 9, 27, \ldots$. We know that $\frac{1}{1-3x} = 1 + 3x + 9x^2 + 27x^3 + \cdots$. To get the zero out front, we need the generating series to look like $x + 3x^2 + 9x^3 + 27x^4 + \cdots$ (so there is no constant term). Multiplying by x has this effect. So the generating function for $0, 1, 3, 9, 27, \ldots$ is $\frac{x}{1-3x}$. This will also work to get the generating function for $0, 1, 0, 1, 0, 1, \ldots$:

$$\frac{x}{1-x^2} = x + x^3 + x^5 + \cdots \text{ which generates } 0, 1, 0, 1, 0, 1 \ldots$$

What would happen if we add (term by term) the sequences $1, 0, 1, 0, 1, 0, \ldots$ and $0, 1, 0, 1, 0, 1, \ldots$. We should get $1, 1, 1, 1, 1, 1 \ldots$. What happens when we add the generating functions? It works (try it)!

$$\frac{1}{1-x^2} + \frac{x}{1-x^2} = \frac{1}{1-x}$$

Here's a tricky one: what happens if you take the derivative of $\frac{1}{1-x}$? We simply get $\frac{1}{(1-x)^2}$. But if we differentiate term by term in the power series, we get $(1+x+x^2+x^3+\cdots)' = 1+2x+3x^2+4x^3+\cdots$ which is the generating series for $1, 2, 3, 4, \ldots$. This says

The generating function for $1, 2, 3, 4, 5, \ldots$ is $\dfrac{1}{(1-x)^2}$

What happens if we take a second derivative: $\frac{2}{(1-x)^3} = 2 + 6x + 12x^2 + 20x^3 + \cdots$. So $\frac{1}{(1-x)^3} = 1 + 3x + 6x^2 + 10x^3 + \cdots$ is a generating function for the triangular numbers.

A.1.2 Differencing

We have seen how to find generating functions from $\frac{1}{1-x}$ using multiplication (by a constant or by x), substitution, addition, and differentiation. To use each of these, you must notice a way to transform the sequence $1, 1, 1, 1, 1 \ldots$ into your desired sequence. This is not always easy. It is also not really the way we have been looking at analyzing sequences. One thing we have considered often is the sequence of differences between terms of a sequence. This will turn out to be helpful in finding generating functions as well. The idea is that the sequence of differences is often simpler than the original sequence. So if we know a generating function for the differences, we would like to use this to find a generating function for the original sequence. First, we must figure out how to relate a generating series to the generating series for the sequence of differences.

For example, consider the sequence $2, 4, 10, 28, 82, \ldots$. How could we move to the sequence of first differences: $2, 6, 18, 54, \ldots$? We want to subtract 2 from the 4, 4 from the 10, 10 from the 28, and so on. So if we subtract (term by term) the sequence $0, 2, 4, 10, 28, \ldots$ from $2, 4, 10, 28 \ldots$, we will be set. Of course it is easy to find the generating function for $0, 2, 4, 10, 28 \ldots$ (multiply the generating function for $2, 4, 10, 28, \ldots$ by x) - then just subtract. Use A to represent the generating function for $2, 4, 10, 28, 82, \ldots$ Then:

$$\begin{aligned} A &= 2 + 4x + 10x^2 + 28x^3 + 82x^4 + \cdots \\ \underline{-xA &= 0 + 2x + 4x^2 + 10x^3 + 28x^4 + 82x^5 + \cdots} \\ (1-x)A &= 2 + 2x + 6x^2 + 18x^3 + 54x^4 + \cdots \end{aligned}$$

Now we don't get exactly the sequence of differences - but we get something close. In this particular case, we already know the generating function A (we found it in the previous section) but most of the time we will use this differencing technique to *find* A: if we have the generating function for the sequence of differences, we can then solve for A. Here is an example.

Example: Find a generating function for $1, 3, 5, 7, 9, \ldots$.

Solution: We notice that the sequence of differences is constant, and we know how to find the generating function for any constant sequence. So call the generating function for $1, 3, 5, 7, 9, \ldots$ simply A. We have

$$\begin{aligned} A &= 1 + 3x + 5x^2 + 7x^3 + 9x^4 + \cdots \\ \underline{-xA &= 0 + x + 3x^2 + 5x^3 + 7x^4 + 9x^5 + \cdots} \\ (1-x)A &= 1 + 2x + 2x^2 + 2x^3 + 2x^4 + \cdots \end{aligned}$$

Now we know that $2x + 2x^2 + 2x^3 + 2x^4 + \cdots = \dfrac{2x}{1-x}$. Thus

$$(1-x)A = 1 + \frac{2x}{1-x}$$

Now solve for A:

$$A = \frac{1}{1-x} + \frac{2x}{(1-x)^2} = \frac{1+x}{(1-x)^2}$$

Does this makes sense? Before we simplified the two fractions into one, we were adding the generating function for the sequence $1, 1, 1, 1, \ldots$ to the generating function for the sequence $0, 2, 4, 6, 8, 10, \ldots$ (remember $\frac{1}{(1-x)^2}$ generates $1, 2, 3, 4, 5, \ldots$ - multiplying by

$2x$ shifts it over, putting the zero out front, and doubles each term). If we add these term by term, we get the correct sequence $1, 3, 5, 7, 9, \ldots$.

Now that we have a generating function for the odd numbers, we can use that to find the generating function for the squares.

Example: Find the generating function for $1, 4, 9, 16, \ldots$.

Solution: Again we call the generating function for the sequence A. Use differencing:

$$\begin{aligned} A &= 1 + 4x + 9x^2 + 16x^3 + \cdots \\ -xA &= 0 + x + 4x^2 + 9x^3 + 16x^4 + \cdots \\ (1-x)A &= 1 + 3x + 5x^2 + 7x^3 + \cdots \end{aligned}$$

Since $1 + 3x + 5x^2 + 7x^3 + \cdots = \dfrac{1+x}{(1-x)^2}$ we have $A = \dfrac{1+x}{(1-x)^3}$

In each of the examples above, found the difference between consecutive terms which gave use a sequence of differences we knew a generating function for. We can generalize this to more complicated relationships between terms of the sequence. For example, what if we know that the sequence satisfies the recurrence relation $a_n = 3a_{n-1} - 2a_{n-2}$? In other words, if we take a term of the sequence and subtracted 3 times the previous term and then added 2 times the term before that, we would get 0 (since $a_n - 3a_{n-1} + 2a_{n-2} = 0$). That will will for all but the first two terms of the sequence. So after the first two terms, the sequence of results of these calculations would be a sequence of 0's, which we definitely know a generating function for.

Example: The sequence $1, 3, 7, 15, 31, 63, \ldots$ satisfies the recurrence relation $a_n = 3a_{n-1} - 2a_{n-2}$. Find the generating function for the sequence.

Solution: Call the generating function for the sequence A. We have

$$\begin{aligned} A &= 1 + 3x + 7x^2 + 15x^3 + 31x^4 + \cdots + a_n x^n + \cdots \\ -3xA &= 0 - 3x - 9x^2 - 21x^3 - 45x^4 - \cdots - 3a_{n-1}x^n - \cdots \\ +\quad 2x^2A &= 0 + 0x + 2x^2 + 6x^3 + 14x^4 + \cdots + 2a_{n-2}x^n + \cdots \\ (1 - 3x + 2x^2)A &= 1 \end{aligned}$$

Let us see what happened there - we multiplied A by $-3x$ which shifts every term over one spot and multiplies them by -3. On the third line, we multiplied x by $2x^2$, which shifted every term over two spots and multiplied them by 2. When we add the corresponding terms up, we are taking each term, subtracting 3 times the previous term, and adding 2 times the term before that. You can see that for the initial terms this does indeed give $0x^n$. This will happen for each term because $a_n - 3a_{n-1} + 2a_{n-2} = 0$. In general, we might have two terms from the beginning of the generating series, although in this case the second term happens to be 0 as well.

Now we just need to solve for A:

$$A = \frac{1}{1 - 3x + 2x^2}$$

A.1.3 Multiplication - Partial Sums

What happens to the sequences when you multiply two generating functions? Let's see: $A = a_0 + a_1x + a_2x^2 + \cdots$ and $B = b_0 + b_1x + b_2x^2 + \cdots$. To multiply A and B, we need to do a lot of distributing (infinite FOIL?) but keep in mind we will regroup and only need to write down the first few terms to see the pattern. What is the constant term? a_0b_0. What is the coefficient of x? $a_0b_1 + a_1b_0$. And so on. We get:

$$AB = a_0b_0 + (a_0b_1 + a_1b_0)x + (a_0b_2 + a_1b_1 + a_2b_0)x^2 + (a_0b_3 + a_1b_2 + a_2b_1 + a_3b_0)x^3 + \cdots$$

Example: "Multiply" the sequence $1, 2, 3, 4, \ldots$ by the sequence $1, 2, 4, 8, 16, \ldots$.

Solution: The new constant term is just $1 \cdot 1$. The next term will be $1 \cdot 2 + 2 \cdot 1 = 4$. The next term: $1 \cdot 4 + 2 \cdot 2 + 3 \cdot 1 = 11$. One more: $1 \cdot 8 + 2 \cdot 4 + 3 \cdot 2 + 4 \cdot 1 = 28$. The resulting sequence is

$$1, 4, 11, 28, 57, \ldots$$

Since the generating function for $1, 2, 3, 4, \ldots$ is $\frac{1}{(1-x)^2}$ and the generating function for $1, 2, 4, 8, 16, \ldots$ is $\frac{1}{1-2x}$, we have that the generating function for $1, 4, 11, 28, 57, \ldots$ is $\frac{1}{(1-x)^2(1-2x)}$

Now what happens when you multiply a sequence by $1, 1, 1, \ldots$? Try it with $1, 2, 3, 4, 5 \ldots$. The first term is $1 \cdot 1 = 1$. Then $1 \cdot 2 + 1 \cdot 1 = 3$. Then $1 \cdot 3 + 1 \cdot 2 + 1 \cdot 1 = 6$. The next term will be 10. We are getting the triangular numbers. More precisely we get the sequence of partial sums of $1, 2, 3, 4, 5, \ldots$. In terms of generating functions, we take $\frac{1}{1-x}$ (generating $1, 1, 1, 1, 1 \ldots$) and multiply it by $\frac{1}{(1-x)^2}$ (generating $1, 2, 3, 4, 5, \ldots$) and this give $\frac{1}{(1-x)^3}$. This should not be a surprise - we found the same generating function for the triangular numbers earlier.

The point is, if you need to find a generating function for the sum of the first n terms of a particular sequence, and you know the generating function for *that* sequence, you can multiply it by $\frac{1}{1-x}$. This makes sense - to go back from the sequence of partial sums to the original sequence, you look at the sequence of differences. When you get the sequence of differences you end up multiplying by $1 - x$ - that is, dividing by $\frac{1}{1-x}$. Multiplying by $\frac{1}{1-x}$ gives partial sums, dividing by $11 - x$ gives differences.

A.1.4 Solving Recurrence Relations with Generating Functions

We end with an example of one of the many reasons studying generating functions is helpful - we can use generating functions to solve recurrence relations.

Example: Solve the recurrence relation $a_n = 3a_{n-1} - 2a_{n-2}$ with initial conditions $a_0 = 1$ and $a_1 = 3$.

Solution: We saw in an example above that this recurrence relation gives the sequence $1, 3, 7, 15, 31, 63, \ldots$ which has generating function $\dfrac{1}{1 - 3x + 2x^2}$. We did this by calling the generating function A and then computing $A - 3xA + 2x^2A$ which was just 1, since every other term canceled out.

But how does knowing the generating function help us? Well, we must first break up the generating function into two simpler ones. For this, we need partial fraction decomposition. Start by factoring the denominator:

$$\frac{1}{1-3x+2x^2} = \frac{1}{(1-x)(1-2x)}$$

Now partial fraction decomposition tells us that we can write this faction as the sum of two fractions (we decompose the given fraction):

$$\frac{1}{(1-x)(1-2x)} = \frac{a}{1-x} + \frac{b}{1-2x} \quad \text{for some constants } a \text{ and } b$$

To find a and b we add the two decomposed fractions using a common denominator. This gives

$$\frac{1}{(1-x)(1-2x)} = \frac{a(1-2x)+b(1-x)}{(1-x)(1-2x)}$$

so

$$1 = a(1-2x)+b(1-x)$$

This must be true for all values of x. If $x = 1$, then the equation becomes $1 = -a$ so $a = -1$. When $x = \frac{1}{2}$ we get $1 = b/2$ so $b = 2$. This tells us that we can decompose the fraction like this:

$$\frac{1}{(1-x)(1-2x)} = \frac{-1}{1-x} + \frac{2}{1-2x}$$

This completes the partial fraction decomposition. But now notice that these two fractions look like generating functions we know. In fact, we should be able to expand each of them.

$$\frac{-1}{1-x} = -1-x-x^2-x^3-x^4-\cdots \text{ which generates } -1,-1,-1,-1,-1,\ldots$$

$$\frac{2}{1-2x} = 2+4x+8x^2+16x^3+32x^4+\cdots \text{ which generates } 2,4,8,16,32,\ldots$$

We can in fact give a closed formula for the nth term of each of these sequences. The first is just $a_n = -1$. The second is $a_n = 2^{n+1}$. The sequence we are interested in is just the sum of these, so the solution to the recurrence relation is

$$a_n = 2^{n+1} - 1$$

So now we can add generating functions to our list of methods for solving recurrence relations - although we do need to know how to do partial fraction decomposition.

Exercises

1. Find the generating function for each of the following sequences by relating them back to a sequence with known generating function.

   (a) $4,4,4,4,4,\ldots$

   (b) $2,4,6,8,10,\ldots$

   (c) $0,0,0,2,4,6,8,10,\ldots$

(d) $1, 5, 25, 125, \ldots$

(e) $1, -3, 9, -27, 81, \ldots$

(f) $1, 0, 5, 0, 25, 0, 125, 0, \ldots$

(g) $0, 1, 0, 0, 2, 0, 0, 3, 0, 0, 4, 0, 0, 5, \ldots$

2. Find the sequence generated by the following generating functions:

(a) $\dfrac{4x}{1-x}$

(b) $\dfrac{1}{1-4x}$

(c) $\dfrac{x}{1+x}$

(d) $\dfrac{3x}{(1+x)^2}$

(e) $\dfrac{1+x+x^2}{(1-x)^2}$ (Hint: multiplication)

3. Show how you can get the generating function for the triangular numbers in three different ways:

(a) Take two derivatives of the generating function for $1, 1, 1, 1, 1, \ldots$

(b) Use differencing.

(c) Multiply two known generating functions.

4. Use differencing to find the generating function for $4, 5, 7, 10, 14, 19, 25, \ldots$

5. Find a generating function for the sequence with recurrence relation $a_n = 3a_{n-1} - a_{n-2}$ with initial terms $a_0 = 1$ and $a_1 = 5$.

6. Use the recurrence relation for the Fibonacci numbers to find the generating function for the Fibonacci sequence.

7. Use multiplication to find the generating function for the sequence of partial sums of Fibonacci numbers, $S_0, S_1, S_2, \ldots$ where $S_0 = F_0$, $S_1 = F_0 + F_1$, $S_2 = F_0 + F_1 + F_2$, $S_3 = F_0 + F_1 + F_2 + F_3$ and so on.

8. Find the generating function for the sequence with closed formula $a_n = 2(5^n) + 7(-3)^n$.

9. Find a closed formula for the nth term of the sequence with generating function $\dfrac{3x}{1-4x} + \dfrac{1}{1-x}$.

10. Find a_7 for the sequence with generating function $\dfrac{2}{(1-x)^2} \cdot \dfrac{x}{1-x-x^2}$

A.2 Introduction to Number Theory

So far in this course, we have used the natural numbers to solve problems. This was the right set of numbers to work with in discrete mathematics because we always dealt with a whole number of things. The natural numbers have been a tool. Now let's take a moment to inspect that tool. What mathematical discoveries can we make *about* the natural numbers themselves?

This is the main question of number theory, a huge, ancient, complex, and above all, beautiful branch of mathematics. Historically, number theory was known as the Queen of Mathematics - it was very much a branch of *pure* mathematics, studied for its own sake instead of as a means to understanding real world applications. However, this has changed in recent years, as applications of number theory have been unearthed. Probably the most well known example of this is RSA cryptography, one of the methods used in encrypt data on the internet. It is number theory that makes this possible.

So what sorts of questions belong to the realm of number theory? Here is a motivating example. Recall in our study of induction, we asked:

Which amounts of postage can be made exactly using just 5-cent and 8-cent stamps?

We were able to prove that *any* amount greater than 27 cents could be made. You might wonder what would happen if we changed the denomination of the stamps. What if we instead had 4- and 9-cent stamps? Would there be some amount after which all amounts would be possible? Well, again, we could replace two 4-cent stamps with a 9-cent stamp, or three 9-cent stamps with seven 4-cent stamps. In each case we can create one more cent of postage. Using this as the inductive case would allow us to prove that any amount of postage greater than 23 cents can be made.

What if we had 2-cent and 4-cent stamps. Here it looks less promising. If we take some number of 2-cent stamps and some number of 4-cent stamps, what can we say about the total? Could it ever be odd? Doesn't look like it.

So *why* does 5 and 8 work, 4 and 9 work, but 2 and 4 not work? What is it about these numbers? If I gave you a pair of numbers, could you tell me right away if they would work or not? We will answer these questions, and more, after first investigating some simpler properties of numbers themselves.

A.2.1 Divisibility

It is easy to add and multiply natural numbers. If we extend our focus to all integers, then subtraction is also easy (we need the negative numbers so we can subtract any number from any other numbers, even larger from smaller). Division is the first operation that presents a challenge. If we wanted to extend our set of numbers so any division would be possible (maybe excluding division by 0) we would need to look at the rational numbers - the set of all numbers which can be written as fractions. This would be going too far, so we will refuse this option.

In fact, it is a good thing that not every number can be divided by other numbers. This helps us understand the structure of the natural numbers and opens the door to many interesting questions and applications.

If given numbers a and b, it is possible that $a \div b$ gives a whole number. In this case, we say that b *divides* a, in symbols, we write $b \mid a$. If this holds, then b is a divisor or factor of a, and a is a multiple of b. This also means that if $b \mid a$, then $a = bk$ for some integer k (this is saying a is some multiple of b).

The Divisibility Relation
Given integers m and n, we say,

$$m \mid n \qquad \text{“}m \textit{ divides } n\text{”}$$

provided $n \div m$ is an integer. Thus the following assertions mean the same thing:

1. $m \mid n$
2. $n = mk$ for some integer k
3. m is a factor (or divisor) of n
4. n is a multiple of m.

Notice that $m \mid n$ is a statement - it is either true or false. On the other hand, $n \div m = n/m$ is some number. If we want to claim that n/m is not an integer, so m does not divide n, then we can write $m \nmid n$.

Example: Decide whether each of the statements below are true or false.

1. $4 \mid 20$
2. $20 \mid 4$
3. $0 \mid 5$
4. $5 \mid 0$
5. $7 \mid 7$
6. $1 \mid 37$
7. $-3 \mid 12$
8. $8 \mid 12$
9. $1642 \mid 136299$

Solution:

1. True. 4 “goes into” 20 five times without remainder. In other words, $20 \div 4 = 5$, an integer. We could also justify this by saying that 20 is a multiple of 4: $20 = 4 \cdot 5$.
2. False. While 20 is a multiple of 4, it is false that 4 is a multiple of 20.
3. False. $5 \div 0$ is not even defined, let alone an integer.
4. True. In fact, $x \mid 0$ is true for all x. This is because 0 is a multiple of every number: $0 = x \cdot 0$.
5. True. In fact, $x \mid x$ is true for all x.
6. True. 1 divides every number (other than 0).
7. True. Negative numbers work just fine for the divisibility relation. Here $12 = -3 \cdot 4$. It is also true that $3 \mid -12$ and that $-3 \mid -12$.
8. False. Both 8 and 12 are divisible by 4, but this does not mean that 12 is divisible by 8.
9. False.

This last example raises a question: how might one decide whether $m \mid n$? Of course, if you had a trusted calculator, you could ask it for the value of $n \div m$ - if it spits out anything other than an integer, you know $m \nmid n$. This seems a little like cheating though: we don't have division, so should we really use division to check divisibility?

While we don't really know how to divide, we do know how to multiply. So we might try multiplying m by larger and larger numbers until we get close to n. How close? Well, we want to be sure that if we multiply m by the next larger integer, we go over n.

For example, let's try this to decide whether $1642 \mid 136299$. Start finding multiples of 1642:

$$1642 \cdot 2 = 3284 \qquad 1642 \cdot 3 = 4926 \qquad 1642 \cdot 4 = 6568 \qquad \cdots$$

All of these are well less than 136299. I suppose we can jump ahead a bit:

$$1642 \cdot 50 = 82100 \qquad 1642 \cdot 80 = 131360 \qquad 1642 \cdot 85 = 139570$$

Ah, so we need to look somewhere between 80 and 85. Try 83:

$$1642 \cdot 83 = 136286$$

Is this the best we can do? How far are we from our desired 136299? If we subtract, we get $136299 - 136286 = 13$. So we know we cannot go up to 84, that will be too much. In other words, we have found that

$$136299 = 83 \cdot 1642 + 13$$

Since $13 < 1642$, we can now safely say that $1642 \nmid 136299$.

It turns out that the process we went through above can be repeated for any pair of numbers. We can always write the number a as some multiple of the number b plus some remainder. We know this because we know about *division with remainder* from elementary school. This is just a way of saying it using multiplication. Due to the procedural nature that can be used to find the remainder, this fact is usually called the *division algorithm*:

The Division Algorithm
Given any two integers a and b, we can always find an integer q such that

$$a = qb + r$$

where r is an integer satisfying $0 \leq r < |b|$

The idea is we can always take a large enough multiple of b so that the remainder r is as small as possible. Note that we do allow the possibility of $r = 0$. In this case, we get $b \div a$.

A.2.2 Remainder Classes

The division algorithm tells us that there are only b possible remainders when dividing by b. If we fix this divisor, we can group integers by the remainder. Each group is called a *remainder class modulo b* (or sometimes *residue class*).

Example: Describe the remainder classes modulo 5.

Solution: We want to classify numbers by what their remainder would be when divided by 5. From the division algorithm, we know there will be exactly 5 remainder classes, because there are only 5 choices for what r could be ($0 \leq r < 5$).

First consider $r = 0$. Here we are looking for all the numbers divisible by 5. In other words, the multiples of 5. We get the infinite set:

$$\{\ldots, -15, -10, -5, 0, 5, 10, 15, 20, \ldots\}$$

Notice we also include negative integers.

Next consider $r = 1$. Which integers, when divided by 5, have remainder 1? Well, certainly 1, does, as does 6, and 11. Negatives? Here we must be careful: -6 does NOT have remainder 1. We can write $-6 = -2 \cdot 5 + 4$ or $-6 = -1 \cdot 5 - 1$, but only one of these is a "correct" instance of the division algorithm: $r = 4$ since we need r to be non-negative. So in fact, to get $r = 1$, we would could have -4, or -9, etc. Thus we get the remainder class:

$$\{\ldots, -14, -9, -4, 1, 6, 11, 16, 21, \ldots$$

There are three more to go. The remainder classes for 2, 3, and 4 are, respectively:

$$\{\ldots, -13, -8, -3, 2, 7, 12, 17, 22, \ldots\}$$

$$\{\ldots, -12, -7, -2, 3, 8, 13, 18, 23, \ldots\}$$

$$\{\ldots, -11, -6, -1, 4, 9, 14, 19, 24, \ldots\}$$

Note that in the example above, *every* integer is in exactly one remainder class. The technical way to say this is that the remainder classes modulo b form a *partition* of the integers.[1] The most important fact about partitions, is that it is possible to define an *equivalence relation* from a partition: this is a relationship between pairs of numbers which acts in all the important ways like the "equals" relationship.[2]

All fun technical language aside, the idea is really simple. If two numbers belong to the same remainder class, then in some way, they are the same. That is, they are the same *up to division by* b. In the case where $b = 5$ above, the numbers 8 and 23, while not the same number, are the same when it comes to dividing by 5, because both have remainder 3.

It matters what the divisor is: 8 and 23 are the same up to division by 5, but not up to division by 7, since 8 has remainder of 1 when divided by 7 while 23 has a remainder of 2.

With all this in mind, let's introduce some notation. We want to say that 8 and 23 are basically the same, even though they are not equal. So it would be wrong to say $8 = 23$. Instead, we write $8 \equiv 23$. But this is not always true - it works if we are thinking division by 5, so that needs to be there as well. So what we will actually write is this:

$$8 \equiv 23 \pmod 5$$

which is read, "8 is congruent to 23 modulo 5." Of course then we could observe that

$$8 \not\equiv 23 \pmod 7$$

Congruence Modulo n
We say a *is congruent to* b *modulo* n, and write,

$$a \equiv b \pmod n$$

provided a and b have the same remainder when divided by n. In other words, provided a and b belong to the same remainder class modulo n.

[1] It is possible to develop a mathematical theory of partitions, prove statements about all partitions in general and then apply those observations to our case here.

[2] Again, there is a mathematical theory of equivalence relations which applies in many more instances than the one we look at here.

Many books define congruence modulo n slightly differently. They say that $a \equiv b \pmod{n}$ if and only if $n \mid a - b$. In other words, two numbers are congruent modulo n, if their difference is a multiple of n. So which definition is correct? Turns out, it doesn't matter - they are equivalent.

To see why, consider two numbers a and b which are congruent modulo n. Then a and b have the same remainder when divided by n. So we have:

$$a = q_1 n + r \qquad\qquad b = q_2 n + r$$

Here the two r's really are the same. Consider what we get when we take the difference of a and b:

$$a - b = q_1 n + r - (q_2 n + r) = q_1 n - q_2 n = (q_1 - q_2)n$$

So $a - b$ is a multiple of n, or equivalently, $n \mid a - b$.

On the other hand, if we assume first that $n \mid a - b$, so $a - b = kn$, then consider what happens if we divide each term by n. Dividing a by n will leave some remainder, as will dividing b by n. However, dividing kn by n will leave 0 remainder. So the remainders on the left hand side must cancel out. That is, the remainders must be the same.

Thus we have:

Congruence and Divisibility
For any integers a, b, and n, we have

$$a \equiv b \pmod{n} \qquad \text{if and only if} \qquad n \mid a - b$$

It will also be useful to switch back and forth between congruences and regular equations. The above fact helps with this. We know that $a \equiv b \pmod{n}$ if and only if $n \mid a - b$, if and only if $a - b = kn$ for some integer k. Rearranging that equation, we get $a = b + kn$. In other words, if a and b are congruent modulo n, then a is b more than some multiple of n. This conforms with our earlier observation that all the numbers in a particular remainder class are the same amount larger than the multiples of n.

Congruence and Equality
For any integers a, b, and n, we have

$$a \equiv b \pmod{n} \qquad \text{if and only if} \qquad a = b + kn \text{ for some integer } k$$

A.2.3 Properties of Congruence

We said earlier that congruence modulo n behaves in many important ways the same way equality does. Specifically, we could prove that congruence modulo n is an *equivalence relation*, which would require checking the following three facts:

Congruence Modulo n is an Equivalence Relation
Given any integers a, b, c, and n, the following hold:

1. $a \equiv a \pmod{n}$
2. If $a \equiv b \pmod{n}$ then $b \equiv a \pmod{n}$
3. If $a \equiv b \pmod{n}$ and $b \equiv c \pmod{n}$, then $a \equiv c \pmod{n}$

In other words, congruence modulo n is reflexive, symmetric, and transitive, so is an equivalence relation.

You should take a minute to convince yourself that each of the properties above actually hold of congruence. Try explaining each using both the remainder and divisibility definitions.

Next let's consider how congruence behaves when doing basic arithmetic. We already know that if you subtract two congruent numbers, the result will be congruent to 0 (be a multiple of n). What if we add something congruent to 1 to something congruent to 2? Will we get something congruent to 3?

Congruence and Arithmetic
Suppose $a \equiv b \pmod{n}$ and $c \equiv d \pmod{n}$. Then the following hold:

1. $a + c \equiv b + d \pmod{n}$
2. $a - c \equiv b - d \pmod{n}$
3. $ac \equiv bd \pmod{n}$

The above facts might be written a little strangely, but the idea is simple. If we have a true congruence, and we add the same thing to both sides, the result is still a true congruence. This sounds like we are saying:

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

Of course that is true as well - it is the special case where $c = d$. But what we have works in more generality. Think of congruence as being "basically equal." If we have two numbers which are basically equal, and we add basically the same thing to both sides, the result will basically be equal.

This seems reasonable. Is it really true? Let's prove the first fact.

Proof. Suppose $a \equiv b \pmod{n}$ and $c \equiv d \pmod{n}$. That means $a = b + kn$ and $c = d + jn$ for integers k and j. Add these equations:

$$a + c = b + d + kn + jn.$$

But $kn + jn = (k + j)n$, which is just a multiple of n. So $a + c = b + d + (j + k)n$, or in other words, $a + c \equiv b + d \pmod{n}$ QED

The other two facts can be proved in a similar way.

One of the important consequences of these facts about congruences, is that we can basically replace any number in a congruence with any other number it is congruent to. Here are some examples to see how (and why) that works:

Example: Find the remainder of 3491 divided by 9.

Solution: We could do long division, but there is another way. We want to find x such that $x \equiv 3491 \pmod 9$. Now $3491 = 3000 + 400 + 90 + 1$. Of course $90 \equiv 0 \pmod 9$, so we can replace the 90 in the sum with 0. Why is this okay? We are actually subtracting the "same" thing from both sides:

$$\begin{array}{rl} & x \equiv 3000 + 400 + 90 + 1 \pmod 9 \\ - & 0 \equiv 90 \pmod 9 \\ \hline & x \equiv 3000 + 400 + 0 + 1 \pmod 9 \end{array}$$

Next, note that $400 = 4 \cdot 100$, and $100 \equiv 1 \pmod 9$ (since $9 \mid 99$). So we can in fact replace the 400 with simply a 4. Again, we are appealing to our claim that we can replace congruent elements, but we are really appealing to property 3 about the arithmetic of congruence: we know $100 \equiv 1 \pmod 9$, so if we multiply both sides by 4, we get $400 \equiv 4 \pmod 9$.

Similarly, we can replace 3000 with 3, since $1000 = 1 + 999 \equiv 1 \pmod 9$. So our original congruence becomes

$$x \equiv 3 + 4 + 0 + 1 \pmod 9$$

$$x \equiv 8 \pmod 9$$

Therefore 3491 divided by 9 has remainder 8.

The above example should convince you that the well known divisibility test for 9 is true: the sum of the digits of a number is divisible by 9 if and only if the original number is divisible by 9. In fact, we now know something more: any number is congruent to the sum of its digits, modulo 9.[3]

Let's try another.

Example: Find the remainder when 3^{123} is divided by 7.

Solution: Of course, we are working with congruence because we want to find the smallest positive x such that $x \equiv 3^{123} \pmod 7$. Now first write $3^{123} = (3^3)^{41}$. We have:

$$3^{123} = 27^{41} \equiv 6^{41} \pmod 7$$

since $27 \equiv 6 \pmod 7$. Notice further that $6^2 = 36$ is congruent to 1 modulo 7. Thus we can simplify further:

$$6^{41} = 6 \cdot (6^2)^{20} \equiv 6 \cdot 1^{20} \pmod 7$$

But $1^{20} = 1$, so we are done:

$$3^{123} \equiv 6 \pmod 7$$

The above example works. We are using the fact that if $a \equiv b \pmod n$, then $a^p \equiv b^p \pmod n$. This is just applying property 3 a bunch of times.

[3] This works for 3 as well, but definitely not for any modulus in general.

So far we have seen how to add, subtract and multiply with congruences. What about division? There is a reason we have waited to discuss it. It turns out that we cannot simply divide. In other words, even if $ad \equiv bd \pmod{n}$, we do not know that $a \equiv b \pmod{n}$. Consider, for example:

$$18 \equiv 42 \pmod{8}$$

This is true. Now 18 and 42 are both divisible by 6. However,

$$3 \not\equiv 7 \pmod{8}$$

While this doesn't work, note that $3 \equiv 7 \pmod{4}$. We cannot divide 8 by 6, but we can divide 8 by the greatest common factor of 8 and 6. Will this always work?

Suppose $ad \equiv bd \pmod{n}$ In other words, we have $ad = bd + kn$ for some integer k. Of course ad is divisible by d, as is bd. So kn must also be divisible by d. Now if n and d have no common factors (other than 1), then we must have $d \mid k$. But in general, if we try to divide kn by d, we don't know that we will get an integer multiple of n. Some of the n might get divided as well. To be safe, let's divide as much of n as we can. Take the largest factor of both d and n, and cancel that out from n. The rest of the factors of d will come from k, no problem.

We will call the largest factor of both d and n, the $\gcd(d, n)$, for greatest common divisor. In our example above, $\gcd(6, 8) = 2$ since the greatest divisor common to 6 and 8 is 2.

Congruence and Division
Suppose $ad \equiv bd \pmod{n}$. Then $a \equiv b \pmod{\frac{n}{\gcd(d,n)}}$

If d and n have no common factors[4] then $\gcd(d, n) = 1$, so $a \equiv b \pmod{n}$.

Example: Simplify the following congruences using division: (a) $24 \equiv 39 \pmod{5}$ and (b) $24 \equiv 39 \pmod{15}$.

Solution: (a) Both 24 and 39 are divisible by 3, and 3 and 5 have no common factors, so we get

$$8 \equiv 13 \pmod{5}$$

(b) Again, we can divide by 3. However, if doing so blindly gives us $8 \equiv 13 \pmod{15}$ which is no longer true. Instead, we must also divide the modulus 15 by the greatest common factor of 3 and 15, which is 3. Again we get

$$8 \equiv 13 \pmod{5}$$

A.2.4 Solving Congruences

Now that we have some algebraic rules to govern congruence relations, we can attempt to solve for an unknown in a congruence. For example, is there a value of x that satisfies,

$$3x + 2 \equiv 4 \pmod{5},$$

and if so, what is it?

In this example, since the modulus is small, we could simply try every possible value for x. There are really one 5 to consider, since any integer that satisfied the congruence could be replaced

with any other integer it was congruent to modulo 5. Here, $x = 4$ gives 14 which is indeed congruent to 4 modulo 5. This means that $x = 9$ and $x = 14$ and $x = 19$ and so on will each also be a solution because as we saw above, replacing any number in a congruence with a congruent number does not change the truth of the congruence.

So in this example, simply compute $3x + 2$ for values of $x \in \{0, 1, 2, 3, 4\}$. This gives 2, 5, 8, 11, and 14 respectively, for which only 14 is congruent to 4.

Let's also see how you could solve this using our rules for algebra of congruences. Such an approach would be much simpler that the trail and error tactic if the modulus was larger. First, we know we can subtract 2 from both sides:

$$3x \equiv 2 \pmod 5.$$

Then to divide both sides by 3, we first add 0 to both sides. Of course, on the right hand side, we want that 0 to be a 10 (yes, 10 really is 0 - they are congruent modulo 5). This gives,

$$3x \equiv 12 \pmod 5.$$

Now divide both sides by 3. Since $\gcd 3, 5 = 1$, we do not need to change the modulus:

$$x \equiv 4 \pmod 5.$$

Notice that this in fact gives the *general solution*: not only can $x = 4$, but x can be any number which is congruent to 4. We can leave it like this, or write "$x = 4 + 5k$ for any integer k."

Here are a few of more examples of this process.

Example: Solve the following congruences for x.

1. $7x \equiv 12 \pmod{13}$
2. $84x - 38 \equiv 79 \pmod{15}$
3. $20x \equiv 23 \pmod{14}$

Solution:

1. All we need to do here is divide both sides by 7. We add 13 to the right hand side repeatedly until we get a multiple of 7 (adding 13 is the same as adding 0, so this is legal). We get 25, 38, 51, 64, 77 - got it. So we have:

$$\begin{aligned} 7x &\equiv 12 \pmod{13} \\ 7x &\equiv 77 \pmod{13} \\ x &\equiv 11 \pmod{13} \end{aligned}$$

2. Here, since we have numbers larger than the modulus, we can reduce them prior to applying any algebra. We have $84 \equiv 9$, $38 \equiv 8$ and $79 \equiv 4$. Thus,

$$\begin{aligned} 84x - 38 &\equiv 79 \pmod{15} \\ 9x - 8 &\equiv 4 \pmod{15} \\ 9x &\equiv 12 \pmod{15} \\ 9x &\equiv 72 \pmod{15} \end{aligned}$$

We got the 72 by adding $0 \equiv 60 \pmod{15}$ to both sides of the congruence. Now divide both sides by 9. However, since $\gcd 9, 15 = 3$, we must divide the modulus by 3 as well:

$$x \equiv 8 \pmod 5$$

So the solutions are those values which are congruent to 8, or equivalently 3, modulo 5. This means that in some sense there are 3 solutions modulo 15: 3, 8, and 13. We can write the solution:

$$x \equiv 3 \pmod{15}; \quad x \equiv 8 \pmod{15}; \quad x \equiv 13 \pmod{15}$$

3. First reduce modulo 14:

$$6x \equiv 9 \pmod{14}$$

We could now divide both sides by 3, or try to increase 9 by a multiple of 14 to get a multiple of 6. If we divide by 3, we get,

$$2x \equiv 3 \pmod{14}$$

Now try adding multiples of 14 to 3, in hopes of getting a number we can divide by 2. This will not work! Every time we add 14 to the right side, the result will still be odd. We will never get an even number, so we will never be able to divide by 2. Thus there are no solutions to the congruence.

The last congruence above illustrates the way in which congruences might not have solutions. We could have seen this immediately in fact. Look at the original congruence:

$$20x \equiv 23 \pmod{14}$$

If we write this as an equation, we get

$$20x = 23 + 14k$$

or equivalently $20x - 14k = 23$. This is a linear Diophantine equation, and one for which we can easily see there will be no solution. The left hand side will always be even, but the right hand side is odd. A similar problem would occur if the right hand side was divisible by *any* number the left hand side was not.

So in general, given the congruence

$$ax \equiv b \pmod{n},$$

if a and n are divisible by a number which b is not divisible by, then there will be no solutions. In fact, we really only need to check one divisor of a and n: the greatest common divisor. Thus, a more compact way to say this is:

Congruences with no solutions
If $\gcd(a, n) \nmid b$, then $ax \equiv b \pmod{n}$ has no solutions.

A.2.5 Solving Linear Diophantine Equations

We now have the tools need to solve linear Diophantine equations such as

$$51x + 87y = 123.$$

The general strategy will be to convert the equation to a congruence, then solve that congruence.[5] Let's work this particular example to see how this might go.

First, check if perhaps there are no solutions because a divisor of 51 and 87 is not a divisor of 123. Really, we just need to check whether $\gcd(51, 87) \mid 123$. This greatest common divisor is 3, and yes $3 \mid 123$. At this point, we might as well factor out this greatest common divisor. So instead, we will solve:

$$17x + 29y = 41$$

Now observe that if there are going to be solutions, then for those values of x and y, the two sides of the equation must have the same remainder as each other, no matter what we divide by. In particular, if we divide both sides by 17, we must get the same remainder. Thus we can safely write,

$$17x + 29y \equiv 41 \pmod{17}$$

We choose 17 because $17x$ will have remainder 0. This will allow us to reduce the congruence to just one variable. We could have also moved to a congruence modulo 29, although there is usually a good reason to select the smaller choice, as this will allow us to reduce the other coefficient. In our case, we reduce the congruence as follows:

$$\begin{aligned} 17x + 29y &\equiv 41 \pmod{17} \\ 0x + 12y &\equiv 7 \pmod{17} \\ 12y &\equiv 24 \pmod{17} \\ y &\equiv 2 \pmod{17} \end{aligned}$$

Now at this point we know $y = 2 + 17k$ will work for any integer k. If we haven't made a mistake, we should be able to plug this back into our original Diophantine equation to find x:

$$\begin{aligned} 17x + 29(2 + 17k) &= 41 \\ 17x &= -17 - 29 \cdot 17k \\ x &= -1 - 29k \end{aligned}$$

We have now found all solutions to the Diophantine equation. For each k, $x = -1 - 29k$ and $y = 2 + 17k$ will satisfy the equation. We could check this for a few cases. If $k = 0$, the solution is $(-1, 2)$, and yes, $-17 + 2 \cdot 29 = 41$. If $k = 3$, the solution is $(-88, 53)$. If $k = -2$, we get $(57, -32)$.

To summarize this process, to solve $ax + by = c$, we,

1. Divide both sides of the equation by $\gcd a, b$ (if this does not leave the right hand side as an integer, there are no solutions). Let's assume that $ax + by = c$ has already been reduced in this way.

2. Pick the smaller of a and b (here, assume it is b), and convert to a congruence modulo b:

   $$ax + by \equiv c \pmod{b}$$

   This will reduce to a congruence with one variable, x:

   $$ax \equiv c \pmod{b}$$

[5] This is certainly not the only way to proceed. A more common technique would be to apply the *Euclidean algorithm*. Our way can be a little faster, and is presented here primarily for variety.

3. Solve the congruence as we did in the previous section. Write your solution as an equation, such as,
$$x = n + kb$$
4. Plug this into the original Diophantine equation, and solve for y.
5. If we want to know solutions in a particular range (for example, $0 \leq x, y \leq 20$), pick different values of k until you have all required solutions.

Here is another example.

Example: How can you make \$6.37 using just 5-cent and 8-cent stamps? What is the smallest and largest number of stamps you could use?

Solution: First we need a Diophantine equation. We will work in numbers of cents. Let x be the number of 5-cent stamps, and y be the number of 8-cent stamps. We have:
$$5x + 8y = 637.$$

Convert to a congruence and solve:
$$\begin{aligned} 8y &\equiv 367 \pmod 5 \\ 8y &\equiv 2 \pmod 5 \\ 8y &\equiv 32 \pmod 5 \\ y &\equiv 4 \pmod 5 \end{aligned}$$

Thus $y = 4 + 5k$. Then $5x + 8(4 + 5k) = 637$, so $x = 121 - 8k$.

This says that one way to make \$6.37 is to take 121 of the 5-cent stamps and 4 of the 8-cent stamps. To find the smallest and largest number of stamps, try different values of k.

k	(x, y)	Stamps
-1	(129, -1)	not possible
0	(121, 4)	125
1	(113, 9)	122
2	(105, 13)	119
⋮	⋮	⋮

Of course this is no surprise - having the most stamps means we have as many 5-cent stamps as possible, and to get the smallest number of stamps would require have the least number of 5-cent stamps. To minimize the number of 5-cent stamps, we want to pick k so that $121 - 8k$ is as small as possible (but still positive). When $k = 15$, we have $x = 1$ and $y = 79$.

Therefore, to make \$6.37, you can us as few as 80 stamps (1 5-cent stamp and 79 8-cent stamps) or as many as 125 stamps (121 5-cent stamps and 4 8-cent stamps).

Using this method, as long as you can solve linear congruences in one variable, you can solve linear Diophantine equations of two variables. There are times though that solving the linear congruence is a lot of work. For example, suppose you need to solve,
$$13x \equiv 6 \pmod{51}$$

You *could* keep adding 51 to the right side until you get a multiple of 13: You would get 57, 108, 159, 210, 261, 312, and 312 is the first of these that is divisible by 13. This works, but is really too much work. Instead we could convert *back* to a Diophantine equation:

$$13x = 6 + 51k$$

Now solve *this* like we have in this section. Write it as a congruence modulo 13:

$$\begin{aligned} 0 &\equiv 6 + 51k \pmod{13} \\ -12k &\equiv 6 \pmod{13} \\ 2k &\equiv -1 \pmod{13} \\ 2k &\equiv 12 \pmod{13} \\ k &\equiv 6 \pmod{13} \end{aligned}$$

so $k = 6 + 13j$. Now go back and figure out x:

$$\begin{aligned} 13x &= 6 + 51(6 + 13j) \\ x &= 24 + 51j \end{aligned}$$

Of course you could do this switching back and forth between congruences and Diophantine equations as many times as you like. If you *only* used this technique, you would essentially replicate the Euclidean algorithm (a more standard way to solve Diophantine equations).

Exercises

1. Suppose a, b, and c are integers. Prove that if $a \mid b$, then $a \mid bc$.
2. Suppose a, b, and c are integers. Prove that if $a \mid b$ and $a \mid c$ then $a \mid b + c$ and $a \mid b - c$.
3. Write out the remainder classes for $n = 4$.
4. Let a, b, c, and n be integers. Prove that if $a \equiv b \pmod{n}$ and $c \equiv d \pmod{n}$, then $a - c \equiv b - d \pmod{n}$.
5. Find the remainder of 3^{456} when divided by
   (a) 2
   (b) 5
   (c) 7
   (d) 9
6. Determine which of the following congruences have solutions, and find any solutions (between 0 and the modulus) by trial and error.
   (a) $4x \equiv 5 \pmod{6}$
   (b) $4x \equiv 5 \pmod{7}$
   (c) $6x \equiv 3 \pmod{9}$
   (d) $6x \equiv 4 \pmod{9}$
   (e) $x^2 \equiv 2 \pmod{4}$

(f) $x^2 \equiv 2 \pmod 7$

7. Solve the following congruences (describe the general solution).

   (a) $5x + 8 \equiv 11 \pmod{22}$

   (b) $6x \equiv 4 \pmod{10}$

   (c) $4x \equiv 24 \pmod{30}$

   (d) $341x \equiv 2941 \pmod 9$

8. I'm thinking of a number. If you multiply my number by 7, add 5, and divide the result by 11, you will be left with a remainder of 2. What remainder would you get if you divided my original number by 11?

9. Solve the following linear Diophantine equations, using modular arithmetic (describe the general solutions).

   (a) $6x + 10y = 32$

   (b) $17x + 8y = 31$

   (c) $35x + 47y = 1$

10. You have a 13 oz. bottle and a 20 oz. bottle, with which you wish to measure exactly 2 oz. However, you have a limited supply of water. If any water enters either bottle and then gets dumped out, it is gone forever. What is the least amount of water you can start with and still complete the task?

Appendix B

Solutions to Exercises

Solutions for Section 0.2

0.2.1. (a) $A \cap B = \{3, 4, 5\}$.

(b) $A \cup B = \{1, 2, 3, 4, 5, 6, 7\}$.

(c) $A \setminus B = \{1, 2\}$.

(d) $A \times B = \{(1, 2), (1, 3), (1, 5), (2, 2), (2, 3), (2, 5), (3, 2), (3, 3), (3, 5), (4, 2), (4, 3), (4, 5), (5, 2), (5, 3), (5, 5)\}$.

(e) Yes.

(f) No.

0.2.2. (a) $A \cap B = \{4, 6, 8, 10, 12\}$

(b) $A \cup B = \{x \in \mathbb{N} : (3 \leq x \leq 13) \vee x \text{ is even}\}$. (the set of all natural numbers which are either even or between 3 and 13 inclusive).

(c) $B \cap C = \emptyset$.

(d) $B \cup C = \mathbb{N}$.

0.2.3. For example, $A = \{2, 3, 5, 7, 8\}$ and $B = \{3, 5\}$.

0.2.4. Let $A = \{1, 2, 3\}$ and $B = \{1, 2, 3, 4, 5, \{1, 2, 3\}\}$

0.2.5. (a) No.

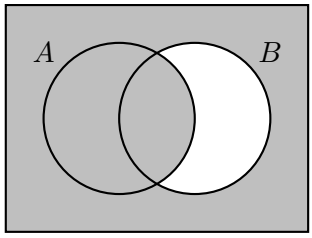
(b) No.

(c) $2\mathbb{Z} \cap 3\mathbb{Z}$ is the set of all integers which are multiples of both 2 and 3 (so multiples of 6). Therefore $2\mathbb{Z} \cap 3\mathbb{Z} = \{x \in \mathbb{Z} : \exists y \in \mathbb{Z}(x = 6y)\}$.

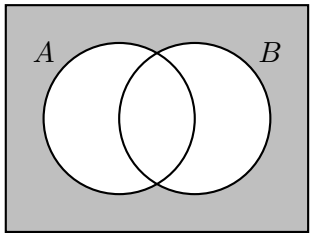
(d) $2\mathbb{Z} \cup 3\mathbb{Z}$.

0.2.6. The set of primes.

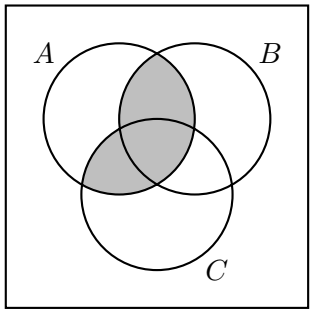
0.2.7. (a) $A \cup \overline{B}$:



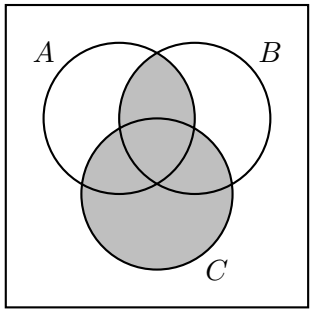
(b) $\overline{(A \cup B)}$:



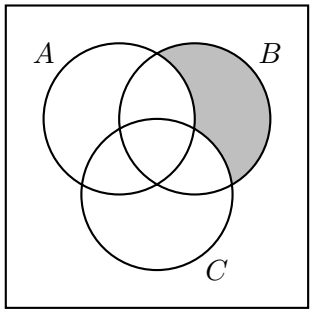
(c) $A \cap (B \cup C)$:



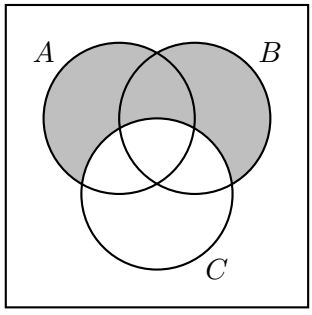
(d) $(A \cap B) \cup C$:



(e) $\overline{A} \cap B \cap \overline{C}$:



(f) $(A \cup B) \setminus C$:



0.2.8. For example, $A \cup B \cap \overline{(A \cap B)}$. Note that $\overline{A \cap B}$ would almost work, but also contain the area outside of both circles.

0.2.9. (a) 34.

(b) 103.

(c) 8.

0.2.10. $\mathcal{P}(A) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\}$.

0.2.11. There are 10 singletons. There are 45 doubletons (because $45 = 9 + 8 + 7 + \cdots + 2 + 1$).

0.2.12. $\{2, 3, 5\}, \{1, 2, 3, 5\}, \{2, 3, 4, 5\}, \{2, 3, 5, 6\}, \{1, 2, 3, 4, 5\}, \{1, 2, 3, 5, 6\}, \{2, 3, 4, 5, 6\}$, and $\{1, 2, 3, 4, 5, 6\}$.

0.2.13. For example $A = \{1, 2, 3, 4\}$ and $B = \{5, 6, 7, 8, 9\}$.

0.2.14. For example, $A = \{1, 2, 3\}$ and $B = \{2, 3, 4, 5\}$.

0.2.15. No. There must be 5 elements in common to both sets. Since there are 10 distinct elements all together in A and B, this means that between A and B, there must be 5 elements which they do not have in common (some in A but not in B, some in B but not in A). But 5 is odd, so to have $|A| = |B|$, we would need 7.5 elements in each set, which is impossible.

0.2.16. If R is the set of red cards and F is the set of face cards, we want to find $|R \cup F|$. This is not simply $|R| + |F|$ because there are 6 cards which are both red and a face card; $|R \cap F| = 6$. We find $|R \cup F| = 32$.

Solutions for Section 0.3

0.3.1. There are 8 different functions. For example, $f(1) = a$, $f(2) = a$, $f(3) = a$; or $f(1) = a$, $f(2) = b$, $f(3) = a$, and so on. None of the functions are injective. Exactly 6 of the functions are surjective. No functions are both (since no functions here are injective).

0.3.2. There are nine functions - you have a choice of three outputs for $f(1)$, and for each, you have three choices for the output $f(2)$. Of these functions, 6 are injective, 0 are surjective, and 0 are both.

0.3.3. (a) f is not injective, since $f(2) = f(5)$ - two different inputs have the same output.

(b) f is surjective, since every element of the codomain is an element of the range.

0.3.4. (a) f is not injective, since $f(1) = 3$ and $f(4) = 3$.

(b) f is not surjective, since there is no input which gives 2 as an output.

0.3.5. (a) f is injective, but not surjective.

(b) f is injective and surjective.

(c) f is injective, but not surjective.

(d) f is not injective, but is surjective.

0.3.6. (a) f is not injective. To prove this, we must simply find two different elements of the domain which map to the same element of the codomain. Since $f(\{1\}) = 1$ and $f(\{2\}) = 1$, we see that f is not injective.

(b) f is not surjective. The largest subset of A is A itself, and $|A| = 10$. So no natural number greater than 10 will ever be an output.

(c) $f^{-1}(1) = \{\{1\}, \{2\}, \{3\}, \ldots \{10\}\}$ (the set of all the singleton subsets of A).

(d) $f^{-1}(0) = \{\emptyset\}$. Note, it would be wrong to write $f^{-1}(0) = \emptyset$ - that would claim that there is no input which has 0 as an output.

(e) $f^{-1}(12) = \emptyset$, since there are no subsets of A with cardinality 12.

0.3.7. (a) $f^{-1}(3) = \{003, 030, 300, 012, 021, 102, 201, 120, 210, 111\}$

(b) $f^{-1}(28) = \emptyset$ (since the largest sum of three digits is $9 + 9 + 9 = 27$)

(c) Part (a) proves that f is not injective - the output 3 is assigned to 10 different inputs.

(d) Part (b) proves that f is not surjective - there is an element of the codomain (28) which is assigned to no inputs.

0.3.8. (a) $|f^{-1}(3)| \leq 1$. In other words, either $f^{-1}(3)$ is the emptyset or is a set containing exactly one element. Injective functions cannot have two elements from the domain both map to 3.

(b) $|f^{-1}(3)| \geq 1$. In other words, $f^{-1}(3)$ is a set containing at least one elements, possibly more. Surjective functions cannot have nothing mapping to 3.

(c) $|f^{-1}(3)| = 1$. There is exactly one element from X which gets mapped to 3, so $f^{-1}(3)$ is the set containing that one element.

0.3.9. X can really be any set, as long as $f(x) = 0$ or $f(x) = 1$ for every $x \in X$. For example, $X = \mathbb{N}$ and $f(n) = 0$ works.

0.3.10. (a) $|X| \leq |Y|$ - otherwise two or more of the elements of X would need to map to the same element of Y.

(b) $|X| \geq |Y|$ - otherwise there would be one or more elements of Y which were never an output.

(c) $|X| = |Y|$. This is the only way for both of the above to occur.

0.3.11. (a) Yes. (Can you give an example?)

(b) Yes.

(c) Yes.

(d) Yes.

(e) No.

(f) No.

0.3.12. (a) f is injective.

Proof. Let x and y be elements of the domain $\mathbb{Z}$. Assume $f(x) = f(y)$. If x and y are both even, then $f(x) = x+1$ and $f(y) = y+1$. Since $f(x) = f(y)$, we have $x+1 = y+1$ which implies that $x = y$. Similarly, if x and y are both odd, then $x - 3 = y - 3$ so again $x = y$. The only other possibility is that x is even an y is odd (or visa-versa). But then $x + 1$ would be odd and $y - 3$ would be even, so it cannot be that $f(x) = f(y)$. Therefore if $f(x) = f(y)$ we then have $x = y$, which proves that f is injective. QED

(b) f is surjective.

Proof. Let y be an element of the codomain $\mathbb{Z}$. We will show there is an element n of the domain ($\mathbb{Z}$) such that $f(n) = y$. There are two cases. First, if y is even, then let $n = y + 3$. Since y is even, n is odd, so $f(n) = n - 3 = y + 3 - 3 = y$ as desired. Second, if y is odd, then let $n = y - 1$. Since y is odd, n is even, so $f(n) = n + 1 = y - 1 + 1 = y$ as needed. Therefore f is surjective. QED

0.3.13. Yes, this is a function, if you choose the domain and codomain correctly. The domain will be the set of students, and the codomain will be the set of possible grades. The function is almost certainly not injective, because it is likely that two students will get the same grade. The function might be surjective - it will be if there is at least one student who gets each grade.

0.3.14. Yes, as long as the set of cards is the domain and the set of players is the codomain. The function is not injective because multiple cards go to each player. It is surjective since all players get cards.

0.3.15. This cannot be a function. If the domain were the set of cards, then it is not a function because not every card gets dealt to a player. If the domain were the set of players, it would not be a function because a single player would get mapped to multiple cards. Since this is not a function, it doesn't make sense to say whether it is injective/surjective/bijective.

Solutions for Section 1.1

1.1.1. 255.

1.1.2. 8.

1.1.3. 15.

1.1.4. (a) $2^8 = 256$. You have two choices for each tie - wear it or don't.

(b) You have 7 choices for regular ties (the 8 choices less the "no regular tie" option) and 31 choices for bow ties (32 total minus the "no bow tie" option). Thus total you have $7 \cdot 31 = 217$.

(c) $\binom{3}{2}\binom{5}{3} = 30$

(d) $5! = 120$

1.1.5. (a) 16 is the number of choices you have if you want to watch one movie, either a comedy or horror flick.

(b) 63 is the number of choices you have if you will watch two movies, first a comedy and then a horror.

1.1.6. $0 \leq |A \cap B| \leq 10$ and $15 \leq |A \cup B| \leq 25$.

1.1.7. $|A \cup B| + |A \cap B| = 13$

1.1.8. 39.

1.1.9. $|(A \cup C) \cap \overline{B}| = 44$. Use a Venn diagram.

1.1.10. One possibility: $(A \cup B) \cap C$.

1.1.11. (a) 8^5, since you select from 8 letters 5 times.

(b) $8 \cdot 7 \cdot 6 \cdot 5 \cdot 4$. After selecting a letter, you have fewer letters to select for the next one.

(c) 64 - you need to select the 4th and 5th letters.

(d) $64 + 64 - 0 = 128$. There are 64 words which start with "aha" and another 64 words that end with "bah." Perhaps we over counted the words that both start with "aha" and end with "bah" but since the words are only 5 letters long, there are no such words.

(e) $(8 \cdot 7 \cdot 6 \cdot 5 \cdot 4) - 3 \cdot (5 \cdot 4) = 6660$ - all the words minus the bad ones. The taboo word can be in any of three positions (starting with letter 1, 2, or 3) and for each position we must choose the other two letters (from the remaining 5 letters)

Solutions for Section 1.2

1.2.1. $\binom{10}{6} + \binom{10}{7} + \binom{10}{8} + \binom{10}{9} + \binom{10}{10} = 386$

1.2.2. Use the binomial theorem. $\binom{14}{9} + \binom{15}{6}2^9$.

1.2.3. (a) $2^6 = 64$

(b) $2^3 = 8$. We need to select yes/no for each of the remaining three elements.

(c) $2^3 = 8$. We need to decide yes/no for the three non-prime elements.

(d) $2^6 - 2^3 = 56$. There are 8 subsets which do not contain any odd numbers.

(e) 9. We need to select one odd (3 choices) and one even (3 choices).

1.2.4. (a) $\binom{14}{7}$

(b) $\binom{6}{2}\binom{8}{5}$

(c) $\binom{14}{7} - \binom{6}{2}\binom{8}{5}$

Solutions for Section 1.3

1.3.1. (a) $\binom{10}{3}$

(b) 2^{10}

(c) $P(10, 5)$

1.3.2. $\binom{7}{2}\binom{7}{2}$

1.3.3. (a) 5 (you need to skip one dot the top and the bottom).

(b) $\binom{7}{2}$ - once you select the two dots on the top, the bottom two are determined.

(c) This is tricky - you need to worry about running out of space. One way to count: break into cases by the location of the top left corner. You get $\binom{7}{2} + (\binom{7}{2} - 1) + (\binom{7}{2} - 3) + (\binom{7}{2} - 6) + (\binom{7}{2} - 10) + (\binom{7}{2} - 15)$

(d) All of them

1.3.4. (a) $\binom{20}{4}\binom{16}{4}\binom{12}{4}\binom{8}{4}\binom{4}{4}$

(b) $5!\binom{15}{3}\binom{12}{3}\binom{9}{3}\binom{6}{3}\binom{3}{3}$

1.3.5. 9! (there are 10 people seated around the table, but it does not matter where King Arthur sits, only who sits to his left, two seats to his left, and so on).

Solutions for Section 1.4

1.4.1. *Proof.* Question: How many subsets of $A = 1, 2, 3, \ldots, n + 1$ contain exactly two elements?

Answer 1: We must choose 2 elements from $n + 1$ choices, so there are $\binom{n+1}{2}$ subsets.

Answer 2: We break this question down into cases, based on what the larger of the two elements in the subset is. The larger element can't be 1, since we need at least one element smaller than it.

Larger element is 2: there is 1 choice for the smaller element.

Larger element is 3: there are 2 choices for the smaller element.

Larger element is 4: there are 3 choices for the smaller element.

And so on. When the larger element is $n+1$, there are n choices for the smaller element. Since each two element subset must be in exactly one of these cases, the total number of two element subsets is $1+2+3+\cdots+n$.

Answer 1 and answer 2 are both correct, so they must be equal. Therefore

$$1+2+3+\cdots+n=\binom{n+1}{2}$$

QED

1.4.2. (a) She has $\binom{15}{6}$ ways to select the 6 bride's maids, and then for each way, has 6 choices for the maid of honor. Thus she has $\binom{15}{6}6$ choices.

(b) She has 15 choices for who will be her maid of honor. Then she needs to select 5 of the remaining 14 friends to be bride's maids, which she can do in $\binom{14}{5}$ ways. Thus she has $15\binom{14}{5}$ choices.

(c) We have answered the question (how many wedding parties can the bride choose from) in two ways. The first way gives the left hand side of the identity and the second way gives the right hand side of the identity. Therefore the identity holds.

1.4.3. (a) After the 1, we need to find a 5-bit string with one 1. There are $\binom{5}{1}$ ways to do this.

(b) $\binom{4}{1}$ (we need to pick 1 of the remaining 4 slots to be the second 1).

(c) $\binom{3}{1}$

(d) Yes. We still need strings starting with 0001 (there are $\binom{2}{1}$ of these) and strings starting 00001 (there is only $\binom{1}{1}=1$ of these).

(e) $\binom{6}{2}$

(f) An example of the Hockey Stick Theorem:

$$\binom{1}{1}+\binom{2}{1}+\binom{3}{1}+\binom{4}{1}+\binom{5}{1}=\binom{6}{2}$$

1.4.4. (a) 3^n, since there are 3 choices for each of the n digits.

(b) 1, since all the digits need to be 2's. However, we might write this as $\binom{n}{0}$.

(c) There are $\binom{n}{1}$ places to put the non-2 digit. That digit can be either a 0 or a 1, so there are $2\binom{n}{1}$ such strings.

(d) We must choose two slots to fill with 0's or 1's. There are $\binom{n}{2}$ ways to do that. Once the slots are picked, we have two choices for the first slot (0 or 1) and two choices for the second slot (0 or 1). So there are a total of $2^2\binom{n}{2}$ such strings.

(e) There are $\binom{n}{k}$ ways to pick which slots don't have the 2's. Then those slots can be filled in 2^k ways (0 or 1 for each slot). So there are $2^k\binom{n}{k}$ such strings.

(f) These strings contain just 0's and 1's - so they are bit strings. There are 2^n bit strings. But keeping with the pattern above, we might write this as $2^n\binom{n}{n}$.

(g) We answer the question of how many length n ternary digit strings there are in two ways. First, each digit can be one of three choices, so the total number of strings is 3^n. On the other hand, we could break the question down into cases by how many of the digits are 2's. If they are all 2's, then there are $\binom{n}{0}$ strings. If all but one is a 2, then there are $2\binom{n}{1}$ strings. If all but 2 of the digits are 2's, then there are $2^2\binom{n}{2}$ strings - we choose 2 of the n digits to be non-2, and then there are 2 choices for each of those digits. And so on for every possible number of 2's in the string.

1.4.5. *Proof.* Question: How many k-letter words can you make using n different letters without repeating any letter?

Answer 1: There are n choices for the first letter, $n-1$ choices for the second letter, $n-2$ choices for the third letter, and so on until $n-(k-1)$ choices for the kth letter (since $k-1$ letters have already been assigned at that point). The product of these numbers can be written $\frac{n!}{(n-k)!}$ which is $P(n,k)$.

Answer 2: First pick k letters to be in the word from the n choices. This can be done in $\binom{n}{k}$ ways. Now arrange those letters into a word - there are k choices for the first letter, $k-1$ choices for the second, and so on, for a total of $k!$ arrangements of the k letters. Thus the total number of words is $\binom{n}{k}k!$. QED

Solutions for Section 1.5

1.5.1. (a) $\binom{18}{4}$. Each outcome can be represented by a sequence of 14 stars and 4 bars.

(b) $\binom{13}{4}$. First put one ball in each bin. This leaves 9 stars and 4 bars.

(c) $\binom{18}{4}-\left[\binom{5}{1}\binom{11}{4}-\binom{5}{2}\binom{4}{4}\right]$. Subtract all the distributions for which one or more bins contain 7 or more balls.

1.5.2. (a) $\binom{7}{2}$. After each variable gets 1 star for free, we are left with 5 stars and 2 bars.

(b) $\binom{10}{2}$. We have 8 stars and 2 bars.

(c) $\binom{19}{2}$. This problem is equivalent to finding the number of solutions to $x'+y'+z'=17$ where x', y' and z' are non-negative. (In fact, we really just do a substitution. Let $x=x'-3$, $y=y'-3$ and $z=z'-3$).

1.5.3. $\binom{10}{5}$. We have 5 stars (the five dice) and 5 bars (the five switches between the number 1-6).

1.5.4. $\binom{18}{3}$. Distribute 10 units to the variables before finding all solutions to $x_1'+x_2'+x_3'+x_4'=15$ in non-negative integers.

Solutions for Section 1.6

1.6.1. There are 8 different functions. For example, $f(1)=a$, $f(2)=a$, $f(3)=a$; or $f(1)=a$, $f(2)=b$, $f(3)=a$, and so on. None of the functions are injective. Exactly 6 of the functions are surjective. No functions are both (since no functions here are injective).

1.6.2. There are nine functions - you have a choice of three outputs for $f(1)$, and for each, you have three choices for the output $f(2)$. Of these functions, 6 are injective, 0 are surjective, and 0 are both.

1.6.3. (a) $6^4=1296$, since there are six choices of where to send each of the 4 elements of the domain.

(b) $P(6,4) = 6 \cdot 5 \cdot 4 \cdot 3 = 360$, since outputs cannot be repeated.

(c) None.

(d) There are $5 \cdot 6^3$ functions for which $f(1) \neq a$ and another $5 \cdot 6^3$ functions for which $f(2) \neq b$. There are $5^2 \cdot 6^2$ functions for which both $f(1) \neq a$ and $f(2) \neq b$. So the total number of functions for which $f(1) \neq a$ or $f(2) \neq b$ or both is

$$5 \cdot 6^3 + 5 \cdot 6^3 - 5^2 \cdot 6^2 = 1260$$

1.6.4. (a) 17^{10}

(b) $P(17, 10)$

Solutions for Section 1.7

1.7.1. $5^{10} - \left[\binom{5}{1}4^{10} - \binom{5}{2}3^{10} + \binom{5}{3}2^{10} - \binom{5}{4}1^{10}\right]$

1.7.2. $5! - \left[\binom{5}{1}4! - \binom{5}{2}3! + \binom{5}{3}2! - \binom{5}{4}1! + \binom{5}{5}0!\right]$. This is a sneaky way to as for the number of derangements on 5 elements.

1.7.3. $\binom{10}{6}\left(4! - \left[\binom{4}{1}3! - \binom{4}{2}2! + \binom{4}{3}1! - \binom{4}{4}0!\right]\right)$. We choose 6 of the 10 ladies to get their own hat, and the multiply by the number of ways the remaining hats can be deranged.

Solutions for Section 2.1

2.1.1. (a) $a_n = n^2 + 1$

(b) $a_n = \frac{n(n+1)}{2} - 1$

(c) $a_n = \frac{(n+2)(n+3)}{2} + 2$

(d) $a_n = (n+1)! - 1$ (where $n! = 1 \cdot 2 \cdot 3 \cdots n$)

2.1.2. (a) $F_n = F_{n-1} + F_{n-2}$ with $F_0 = 0$ and $F_1 = 1$.

(b) $0, 1, 2, 4, 7, 12, 20, \ldots$

(c) $F_0 + F_1 + \cdots + F_n = F_{n+2} - 1$

2.1.3. $3, 10, 24, 52, 108, \ldots$. The recursive definition for $10, 24, 52, \ldots$ is $a_n = 2a_{n-1} + 4$ with $a_1 = 10$.

2.1.4. $-1, -1, 1, 5, 11, 19, \ldots$ Thus the sequence $0, 2, 6, 12, 20, \ldots$ has closed formula $a_n = (n+1)^2 - 3(n+1) + 2$.

Solutions for Section 2.2

2.2.1. (a) 32.

(b) $a_n = 8 + 6(n-1)$

(c) 30500.

2.2.2. (a) $n + 2$ terms.

(b) $6n + 1$.

(c) $\frac{(6n+8)(n+2)}{2}$

2.2.3. 68117

2.2.4. $\frac{5-5\cdot 3^{21}}{-2}$

2.2.5. $\frac{1+\frac{2^{31}}{3^{31}}}{5/3}$

2.2.6. For arithmetic: $x = 55/3$, $y = 29/3$. For geometric: $x = 9$ and $y = 3$.

2.2.7. (a) $\displaystyle\sum_{k=1}^{n} 2k$

(b) $\displaystyle\sum_{k=1}^{107}(1+4(k-1))$

(c) $\displaystyle\sum_{k=1}^{50}\frac{1}{k}$

(d) $\displaystyle\prod_{k=1}^{n} 2k$

(e) $\displaystyle\prod_{k=1}^{100}\frac{k}{k+1}$

2.2.8. (a) $\displaystyle\sum_{k=1}^{100}(3+4k) = 7+11+15+\cdots+403$

(b) $\displaystyle\sum_{k=0}^{n} 2^k = 1+2+4+8+\cdots+2^n$

(c) $\displaystyle\sum_{k=2}^{50}\frac{1}{(k^2-1)} = 1+\frac{1}{3}+\frac{1}{8}+\frac{1}{15}+\cdots+\frac{1}{2499}$

(d) $\displaystyle\prod_{k=2}^{100}\frac{k^2}{(k^2-1} = \frac{4}{3}\cdot\frac{9}{8}\cdot\frac{16}{15}\cdots\frac{10000}{9999}$

(e) $\displaystyle\prod_{k=0}^{n}(2+3k) = (2)(5)(8)(11)(14)\cdots(2+3n)$

Solutions for Section 2.3

2.3.1. (a) Hint: third differences are constant, so $a_n = an^3+bn^2+cn+d$. Use the terms of the sequence to solve for a, b, c, and d.

(b) $a_n = n^2 - n$

2.3.2. No. The sequence of differences is the same as the original sequence so no differences will be constant.

Solutions for Section 2.4

2.4.1. 171 and 341. $a_n = a_{n-1}+2a_{n-2}$ with $a_0 = 3$ and $a_1 = 5$. Closed formula: $a_n = \frac{8}{3}2^n+\frac{1}{3}(-1)^n$

2.4.2. By telescoping or iteration. $a_n = 3+2^{n+1}$

2.4.3. We claim $a_n = 4^n$ works. Plug it in: $4^n = 3(4^{n-1}) + 4(4^{n-2})$. This works - just simplify the right hand side.

2.4.4. By the Characteristic Root Technique. $a_n = 4^n + (-1)^n$.

2.4.5. $a_n = \frac{13}{5}4^n + \frac{12}{5}(-1)^n$

2.4.6. The general solution is $a_n = a + bn$ where a and b depend on the initial conditions.

(a) $a_n = 1 + n$

(b) For example, we could have $a_0 = 21$ and $a_1 = 22$.

(c) For every x - take $a_0 = x - 9$ and $a_1 = x - 8$.

2.4.7. $a_n = \frac{19}{7}(-2)^n + \frac{9}{7}5^n$

Solutions for Section 2.5

2.5.1. *Proof.* We must prove that $1 + 2 + 2^2 + 2^3 + \cdots + 2^n = 2^{n+1} - 1$ for all $n \in \mathbb{N}$. Thus let $P(n)$ be the statement $1 + 2 + 2^2 + \cdots + 2^n = 2^{n+1} - 1$. We will prove that $P(n)$ is true for all $n \in \mathbb{N}$.

First we establish the base case, $P(0)$, which claims that $1 = 2^{0+1} - 1$. Since $2^1 - 1 = 2 - 1 = 1$, we see that $P(0)$ is true.

Now for the inductive case. Assume that $P(k)$ is true for an arbitrary $k \in \mathbb{N}$. That is, $1 + 2 + 2^2 + \cdots + 2^k = 2^{k+1} - 1$. We must show that $P(k+1)$ is true (i.e., that $1 + 2 + 2^2 + \cdots + 2^{k+1} = 2^{k+2} - 1$). To do this, we start with the left hand side of $P(k+1)$ and work to the right hand side:

$$\begin{aligned} 1 + 2 + 2^2 + \cdots + 2^k + 2^{k+1} &= 2^{k+1} - 1 + 2^{k+1} && \text{by the inductive hypothesis} \\ &= 2 \cdot 2^{k+1} - 1 \\ &= 2^{k+2} - 1 \end{aligned}$$

Thus $P(k+1)$ is true so by the principle of mathematical induction, $P(n)$ is true for all $n \in \mathbb{N}$. QED

2.5.2. *Proof.* Let $P(n)$ be the statement "$7^n - 1$ is a multiple of 6." We will show $P(n)$ is true for all $n \in \mathbb{N}$.

First we establish the base case, $P(0)$. Since $7^0 - 1 = 0$, and 0 is a multiple of 6, $P(0)$ is true.

Now for the inductive case. Assume $P(k)$ holds for an arbitrary $k \in \mathbb{N}$. That is, $7^k - 1$ is a multiple of 6, or in other words, $7^k - 1 = 6j$ for some integer j. Now consider $7^{k+1} - 1$:

$$\begin{aligned} 7^{k+1} - 1 &= 7^{k+1} - 7 + 6 && \text{by cleverness: } -1 = -7 + 6 \\ &= 7(7^k - 1) + 6 && \text{factor out a 7 from the first two terms} \\ &= 7(6j) + 6 && \text{by the inductive hypothesis} \\ &= 6(7j + 1) && \text{factor out a 6} \end{aligned}$$

Therefore $7^{k+1} - 1$ is a multiple of 6, or in other words, $P(k+1)$ is true. Therefore by the principle of mathematical induction, $P(n)$ is true for all $n \in \mathbb{N}$. QED

2.5.3. *Proof.* Let $P(n)$ be the statement $1+3+5+\cdots+(2n-1)=n^2$. We will prove that $P(n)$ is true for all $n \geq 1$.

First the base case, $P(1)$. We have $1=1^2$ which is true, so $P(1)$ is established.

Now the inductive case. Assume that $P(k)$ is true for some fixed arbitrary $k \geq 1$. That is, $1+3+5+\cdots+(2k-1)=k^2$. We will now prove that $P(k+1)$ is also true (i.e., that $1+3+5+\cdots+(2k+1)=(k+1)^2$). We start with the left hand side of $P(k+1)$ and work to the right hand side:

$$\begin{aligned} 1+3+5+\cdots+(2k-1)+(2k+1) &= k^2+(2k+1) && \text{by the induction hypothesis} \\ &= (k+1)^2 && \text{by factoring} \end{aligned}$$

Thus $P(k+1)$ holds, so by the principle of mathematical induction, $P(n)$ is true for all $n \geq 1$. QED

2.5.4. *Proof.* Let $P(n)$ be the statement $F_0+F_2+F_4+\cdots+F_{2n}=F_{2n+1}-1$. We will show that $P(n)$ is true for all $n \geq 0$. First the base case is easy because $F_0=0$ and $F_1=1$ so $F_0=F_1-1$. Now consider the inductive case. Assume $P(k)$ is true, that is, assume $F_0+F_2+F_4+\cdots+F_{2k}=F_{2k+1}-1$. To establish $P(k+1)$ we work from left to right:

$$\begin{aligned} F_0+F_2+F_4+\cdots+F_{2k}+F_{2k+2} &= F_{2k+1}-1+F_{2k+2} && \text{by the inductive hypothesis} \\ &= F_{2k+1}+F_{2k+2}-1 \\ &= F_{2k+3}-1 && \text{by the recursive definition of the Fibonacci numbers} \end{aligned}$$

Therefore $F_0+F_2+F_4+\cdots+F_{2k+2}=F_{2k+3}-1$, which is to say $P(k+1)$ holds. Therefore by the principle of mathematical induction, $P(n)$ is true for all $n \geq 0$. QED

2.5.5. *Proof.* Let $P(n)$ be the statement $2^n < n!$. We will show $P(n)$ is true for all $n \geq 4$. First, we check the base case and see that yes, $2^4 < 4!$ (as $16 < 24$) so $P(4)$ is true. Now for the inductive case. Assume $P(k)$ is true for an arbitrary $k \geq 4$. That is, $2^k < k!$. Now consider $P(k+1)$: $2^{k+1} < (k+1)!$. To prove this, we start with the left side and work to the right side.

$$\begin{aligned} 2^{k+1} &= 2 \cdot 2^k \\ &< 2 \cdot k! && \text{by the inductive hypothesis} \\ &< (k+1) \cdot k! && \text{since } k+1 > 2 \\ &= (k+1)! \end{aligned}$$

Therefore $2^{k+1} < (k+1)!$ so we have established $P(k+1)$. Thus by the principle of mathematical induction $P(n)$ is true for all $n \geq 4$. QED

2.5.6. The only problem is that we never established the base case. Of course, when $n=0$, $0+3 \neq 0+7$.

2.5.7. *Proof.* Let $P(n)$ be the statement that $n+3 < n+7$. We will prove that $P(n)$ is true for all $n \in \mathbb{N}$. First, note that the base case holds: $0+3 < 0+7$. Now assume for induction that $P(k)$ is true. That is, $k+3 < k+7$. We must show that $P(k+1)$ is true. Now since $k+3 < k+7$, add 1 to both sides. This gives $k+3+1 < k+7+1$. Regrouping $(k+1)+3 < (k+1)+7$. But this is simply $P(k+1)$. Thus by the principle of mathematical induction $P(n)$ is true for all $n \in \mathbb{N}$. QED

2.5.8. The problem here is that while $P(0)$ is true, and while $P(k) \to P(k+1)$ for *some* values of k, there is at least one value of k (namely $k = 99$) when that implication fails. For a valid proof by induction, $P(k) \to P(k+1)$ must be true for all values of k greater than or equal to the base case.

2.5.9. *Proof.* Let $P(n)$ be the statement "there is a strictly increasing sequence $a_1, a_2, a_3, \ldots, a_n$ with $a_n < 100$." We will prove $P(n)$ is true for all $n \geq 1$. First we establish the base case: $P(1)$ says there is a single number a_1 with $a_1 < 100$. This is true - take $a_1 = 0$. Now for the inductive step, assume $P(k)$ is true. That is there exists a strictly increasing sequence $a_1, a_2, a_3, \ldots, a_k$ with $a_k < 100$. Now consider this sequence, plus one more term, a_{k+1} which is greater than a_k but less than 100. Such a number exists - for example, the average between a_k and 100. So then $P(k+1)$ is true, so we have shown that $P(k) \to P(k+1)$. Thus by the principle of mathematical induction, $P(n)$ is true for all $n \in \mathbb{N}$. QED

2.5.10. We once again failed to establish the base case: when $n = 0$, $n^2 + n = 0$ which is even, not odd.

2.5.11. *Proof.* Let $P(n)$ be the statement "$n^2 + n$ is even." We will prove that $P(n)$ is true for all $n \in \mathbb{N}$. First the base case: when $n = 0$, we have $n^2 + n = 0$ which is even, so $P(0)$ is true. Now suppose for induction that $P(k)$ is true, that is, that $k^2 + k$ is even. Now consider the statement $P(k+1)$. Now $(k+1)^2 + (k+1) = k^2 + 2k + 1 + k + 1 = k^2 + k + 2k + 2$. By the inductive hypothesis, $k^2 + k$ is even, and of course $2k + 2$ is even. An even plus an even is always even, so therefore $(k+1)^2 + (k+1)$ is even. Therefore by the principle of mathematical induction, $P(n)$ is true for all $n \in \mathbb{N}$. QED

2.5.12. Further hint: the idea is to define the sequence so that a_n is less than the distance between the previous partial sum and 2. That way when you add it into the next partial sum, the partial sum is still less than 2. You could do this ahead of time, or use a clever $P(n)$ in the induction proof. Let $P(n)$ be the statement, "there is a sequence of positive real numbers $a_1, a_2, a_3, \ldots, a_n$ such that $a_1 + a_2 + a_3 + \cdots + a_n < 2$." The base case should be easy (just pick $a_1 < 2$). For the inductive case, you know that $a_1 + a_2 + \cdots + a_k < 2$ so you just need to argue that you can find some a_{k+1} small enough to have $a_1 + a_2 + \cdots + a_k + a_{k+1} < 2$.

2.5.13. The base case should be easy - 0 is a power of 2. For the inductive case, you actually want to use strong induction. Suppose k is either a power of 2 or can be written as the sum of distinct powers of 2, for any $k < n$. Now if n is a power of 2, we are done. If not, subtract the largest power of 2 from n possible. You get $n - 2^x$, which is a smaller number, in fact smaller than both n and 2^x. Thus $n - 2^x$ is either a power of 2 or can be written as the sum of distinct powers of 2, but none of them are going to be 2^x, so the together with 2^x we have written n as the sum of distinct powers of 2.

2.5.14. If $n = 2$, this should work out (so their's your base case). If we assume it works for k people (that the number of handshakes is $\frac{k(k-1)}{2}$, what happens if a $k+1$st person shows up. How many *new* handshakes take place? Now make this into a formal induction argument.

Note, we have already proven this without using induction, but this is fun too.

2.5.15. When $n = 0$, we get $x^0 + \frac{1}{x^0} = 2$ and when $n = 1$, $x + \frac{1}{x}$ is an integer, so the base case holds. Now assume the result holds for all natural numbers $n < k$. In particular, we know

that $x^{k-1} + \frac{1}{x^{k-1}}$ and $x + \frac{1}{x}$ are both integers. Thus their product is also an integer. But,

$$\begin{aligned}\left(x^{k-1} + \frac{1}{x^{k-1}}\right)\left(x + \frac{1}{x}\right) &= x^k + \frac{x^{k-1}}{x} + \frac{x}{x^{k-1}} + \frac{1}{x^k} \\ &= x^k + \frac{1}{x^k} + x^{k-2} + \frac{1}{x^{k-2}}\end{aligned}$$

Note also that $x^{k-2} + \frac{1}{x^{k-2}}$ is an integer by the induction hypothesis, so we can conclude that $x^k + \frac{1}{x^k}$ is an integer.

2.5.16. Here's the idea: since every entry in Pascal's Triangle is the sum of the two entries above it, we can get the $k+1$st row by adding up all the pairs of entry from the kth row. But doing this uses each entry on the kth row twice. Thus each time we drop to the next row, we double the total. Of course, row 0 has sum $1 = 2^0$ (the base case). Now try to make this precise with a formal induction proof. You will use the fact that $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$ for the inductive case.

2.5.17. To see why this works, try it on a copy of Pascal's triangle. We are adding up the entries along a diagonal, starting with the 1 on the left hand side of the 4th row. Suppose we add up the first 5 entries on this diagonal. The claim is that the sum is the entry below and to the left of the last of these 5 entries. Note that if this is true, and we instead add up the first 6 entries, we will need to add the entry one spot to the right of the previous sum. But these two together give the entry below them, which is below and left of the last of the 6 entries on the diagonal.

If you follow that, you can see what is going on. But it is not a great proof. A formal induction proof is needed:

Proof. Let $P(n)$ be the statement $\binom{4}{0} + \binom{5}{1} + \binom{6}{2} + \cdots + \binom{4+n}{n} = \binom{5+n}{n}$. For the base case, consider $n = 0$. This says $\binom{4}{0} = \binom{5}{0}$. Since these are both 1, the base case is true. Now for the inductive case, suppose $P(k)$ is true. That is, $\binom{4}{0} + \binom{5}{1} + \binom{6}{2} + \cdots + \binom{4+k}{k} = \binom{5+k}{k}$. If we add $\binom{4+k+1}{k+1}$ to both sides, we get

$$\binom{4}{0} + \binom{5}{1} + \binom{6}{2} + \cdots + \binom{4+k}{k} + \binom{5+k}{k+1} = \binom{5+k}{k} + \binom{5+k}{k+1}$$

But $\binom{5+k}{k} + \binom{5+k}{k+1} = \binom{5+k+1}{k+1}$. In other words, we have

$$\binom{4}{0} + \binom{5}{1} + \binom{6}{2} + \cdots + \binom{4+k}{k} + \binom{5+k}{k+1} = \binom{5+k+1}{k+1}$$

which is to say that $P(k+1)$ is true.

Therefore, by the principle of mathematical induction, $P(n)$ is true for all $n \geq 0$. QED

2.5.18. The idea here is that if we take the logarithm of a^n, we can increase n by 1 if we multiply by another a (inside the logarithm). This results in adding 1 more $\log(a)$ to the total.

Proof. Let $P(n)$ be the statement $\log(a^n) = n\log(a)$. The base case, $P(2)$ is true, because $\log(a^2) = \log(a \cdot a) = \log(a) + \log(a) = 2\log(a)$, by the product rule for logarithms.

Now assume, for induction, that $P(k)$ is true. That is, $\log(a^k) = k\log(a)$. Consider $\log(a^{k+1})$. We have

$$\log(a^{k+1}) = \log(a^k \cdot a) = \log(a^k) + \log(a) = k\log(a) + \log(a)$$

with the last equality due to the inductive hypothesis. But this simplifies to $(k+1)\log(a)$, establishing $P(k+1)$.

Therefore by the principle of mathematical induction, $P(n)$ is true for all $n \geq 2$. QED

2.5.19. Hint: You are allowed to assume the base case. For the inductive case, group all but the last function together as one sum of functions, then apply the usual sum of derivatives rule, and then the inductive hypothesis.

2.5.20. Hint: for the inductive step, we know by the product rule for two functions that

$$(f_1 f_2 f_3 \cdots f_k f_{k+1})' = (f_1 f_2 f_3 \cdots f_k)' f_{k+1} + (f_1 f_2 f_3 \cdots f_k) f'_{k+1}$$

Then use the inductive hypothesis on the first summand, and distribute.

Solutions for Section 3.1

3.1.1. (a) P: it's your birthday; Q: there will be cake. $(P \vee Q) \to Q$

(b) Hint: you should get three T's and one F.

(c) Only that there will be cake.

(d) It's NOT your birthday!

(e) It's your birthday, but the cake is a lie.

3.1.2. (a) $P \wedge Q$

(b) $P \to \neg Q$

(c) Jack passed math or Jill passed math (or both).

(d) If Jack and Jill did not both pass math, then Jill did.

(e) (a) Nothing else.

(b) Jack did not pass math either.

3.1.3. (a) Three statements: $P \vee S$, $S \to Q$, $(P \vee Q) \to R$. You could also connect the first two with a $\wedge$.

(b) He cannot be lying about all three sentences, so he is telling the truth.

(c) No matter what, Geoff wants ricotta. If he doesn't have quail, then he must have pepperoni but not sausage.

3.1.4.

P	Q	$(P \vee Q) \to (P \wedge Q)$
T	T	T
T	F	F
F	T	F
F	F	T

3.1.5.

P	Q	$\neg P \wedge (Q \to P)$
T	T	F
T	F	F
F	T	F
F	F	T

If the statement is true, then both P and Q are false.

3.1.6. Hint: Like above, only now you will need 8 rows instead of just 4.

3.1.7. The argument is valid. To see this, make a truth table which contains $P \vee Q$ and $\neg P$ (and P and Q of course). Look at the truth value of Q in each of the rows that have $P \vee Q$ and $\neg P$ true.

3.1.8. The argument form is valid. Again, make a truth table containing the premises and conclusion - look at the rows for which the premises are true.

3.1.9. The argument is NOT valid. If you make a truth table containing the premises and conclusion, there will be a row with both premises true but the conclusion false. For example, if P and Q are false and R is true, then $P \wedge Q$ is false, so $(P \wedge Q) \rightarrow R$ is true. Also $\neg P$ is true, so $\neg P \vee \neg Q$ is true. However, $\neg R$ is false.

Solutions for Section 3.2

3.2.1. Make a truth table for each and compare. The statements are logically equivalent.

3.2.2. Again, make two truth tables. The statements are logically equivalent.

3.2.3. (a) If Oscar drinks milk, then he eats Chinese food.

(b) If Oscar does not drink milk, then he does not eat Chinese food.

(c) Yes. The original statement would be false too.

(d) Nothing. The converse need not be true.

(e) He does not eat Chinese food. The contrapositive would be true.

3.2.4. (a) $P \wedge Q$

(b) $(P \vee Q) \vee (Q \wedge \neg R)$

(c) F. Or $(P \wedge Q) \wedge (R \wedge \neg R)$

(d) Either Sam is a woman and Chris is a man, or Chris is a woman.

3.2.5. The statements are equivalent to the...

(a) converse.

(b) implication.

(c) neither.

(d) implication.

(e) converse.

(f) converse.

(g) implication.

(h) converse.

(i) converse.

(j) converse (in fact, this IS the converse).

(k) implication (the statement is the contrapositive of the implication).

(l) neither.

3.2.6. Hint: of course there are many answers. It helps to assume that the statement is true and the converse is NOT true. Think about what that means in the real world and then start saying it in different ways. Some ideas: use necessary and sufficient language, use "only if," consider negations, use "or else" language.

Solutions for Section 3.3

3.3.1. (a) $\neg\exists x(E(x) \wedge O(x))$

(b) $\forall x(E(x) \rightarrow O(x+1))$

(c) $\exists x(P(x) \wedge E(x))$ (where $P(x)$ means "x is prime")

(d) $\forall x \forall y \exists z(x < z < y \vee y < z < x)$

(e) $\forall x \neg \exists y(x < y < x+1)$

3.3.2. (a) Any even number plus 2 is an even number.

(b) For any x there is a y such that $\sin(x) = y$. In other words, every number x is in the domain of sine.

(c) For every y there is an x such that $\sin(x) = y$. In other words, every number y is in the range of sine (which is false).

(d) For any numbers, if the cubes of two numbers are equal, then the numbers are equal.

3.3.3. (a) $\forall x \exists y(O(x) \wedge \neg E(y))$

(b) $\exists x \forall y(x \geq y \vee \forall z(x \geq z \wedge y \geq z))$

(c) There is a number n for which every other number is strictly greater than n.

(d) There is a number n which is not between any other two numbers.

Solutions for Section 3.4

3.4.1. (a) For all integers a and b, if a or b are not even, then $a+b$ is not even.

(b) For all integers a and b, if a and b are even, then $a+b$ is even.

(c) There are numbers a and b such that $a+b$ is even but a and b are not both even.

(d) False. For example, $a = 3$ and $b = 5$. $a + b = 8$, but neither a nor b are even.

(e) False, since it is equivalent to the original statement.

(f) True. Let a and b be integers. Assume both are even. Then $a = 2k$ and $b = 2j$ for some integers k and j. But then $a + b = 2k + 2j = 2(k + j)$ which is even.

(g) True, since the statement is false.

3.4.2. *Proof.* Suppose $\sqrt{3}$ were rational. Then $\sqrt{3} = \frac{a}{b}$ for some integers a and $b \neq 0$. Without loss of generality, assume $\frac{a}{b}$ is reduced. Now

$$3 = \frac{a^2}{b^2}$$

$$b^2 3 = a^2$$

So a^2 is a multiple of 3. This can only happen if a is a multiple of 3, so $a = 3k$ for some integer k. Then we have

$$b^2 3 = 9k^2$$

$$b^2 = 3k^2$$

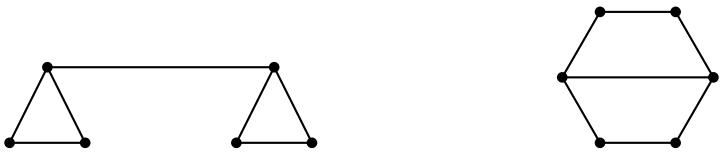
So b^2 is a multiple of 3, making b a multiple of 3 as well. But this contradicts our assumption that $\frac{a}{b}$ is in lowest terms. QED

Solutions for Section 4.1

4.1.1. This is asking for the number of edges in K_{10}. Each vertex (person) has degree (shook hands with) 9 (people). So the sum of the degrees is 90. However, the degrees count each edge (handshake) twice, so there are 45 edges in the graph. That is how many handshakes took place.

4.1.2. It is possible for everyone to be friends with exactly 2 people - you could arrange the 5 people in a circle and say that everyone is friends with the two people on either side of them (so you get the graph C_5). However, it is not possible for everyone to be friends with 3 people - that would lead to a graph with an odd number of odd degree vertices which is impossible - the sum of the degrees must be even.

4.1.3. Yes. For example, both graphs below contain 6 vertices, 7 edges, and have degrees (2,2,2,2,3,3).



4.1.4. Three of the graphs are bipartite. The one which is not is C_7 (second from the right).

4.1.5. C_n is bipartite if and only if $n = 1$ or is even.

Solutions for Section 4.2

4.2.1. No. A (connected) planar graph must satisfy Euler's formula: $V-E+F = 2$. Here $V-E+F = 6 - 10 + 5 = 1$.

4.2.2. G has 10 edges. It could be planar, and then it would have 6 faces.

4.2.3. Yes. According to Euler's formula it would have 2 faces. It does. The only such graph is C_{10}.

Solutions for Section 4.3

4.3.1. 2, since the graph is bipartite. One color for the top set of vertices, another color for the bottom set of vertices.

4.3.2. For example, K_6. If the chromatic number is 6, then the graph is not planar - the 4-color theorem states that all planar graphs can be colored with 4 or fewer colors.

4.3.3. The chromatic numbers are 2, 3, 4, 5, and 3 respectively from left to right.

Solutions for Section 4.4

4.4.1. This is a question about finding Euler paths. Draw a graph with a vertex in each state, and connect vertices if their states share a border. Exactly two vertices will have odd degree - the vertices for Nevada and Utah. Thus you must start your road trip at in one of those states and end it in the other.

4.4.2. (a) K_4 does not have an Euler path or circuit.

(b) K_5 has an Euler circuit (so also an Euler path).

(c) $K_{5,7}$ does not have an Euler path or circuit.

(d) $K_{2,7}$ has an Euler path but not an Euler circuit.

(e) C_7 has an Euler circuit (it is a circuit graph!)

(f) P_7 has an Euler path but no Euler circuit.

4.4.3. When n is odd, K_n contains an Euler circuit. This is because every vertex has degree $n-1$, so an odd n results in all degrees being even.

4.4.4. If both m and n are even, then $K_{m,n}$ has an Euler circuit. When both are odd, there is no Euler path or circuit. If one is 2 and the other is odd, then there is an Euler path but not an Euler circuit.

Solutions for Section A.1

A.1.1. (a) $\dfrac{4}{1-x}$

(b) $\dfrac{2}{(1-x)^2}$

(c) $\dfrac{2x^3}{(1-x}^2$

(d) $\dfrac{1}{1-5x}$

(e) $\dfrac{1}{1+3x}$

(f) $\dfrac{1}{1-5x^2}$

(g) $\dfrac{x}{(1-x^3)^2}$

A.1.2. (a) $0, 4, 4, 4, 4, 4, \ldots$

(b) $1, 4, 16, 64, 256, \ldots$

(c) $0, 1, -1, 1, -1, 1, -1, \ldots$

(d) $0, 3, -6, 9, -12, 15, -18, \ldots$

(e) $1, 3, 6, 9, 12, 15, \ldots$

A.1.3. (a) The second derivative of $\dfrac{1}{1-x}$ is $\dfrac{2}{(1-x)^3}$ which expands to $2+6x+12x^2+20x^3+30x^4+\cdots$. Dividing by 2 gives the generating function for the triangular numbers.

(b) Compute $A-xA$ and you get $1+2x+3x^2+4x^3+\cdots$ which can be written as $\dfrac{1}{(1-x)^2}$. Solving for A gives the correct generating function.

(c) The triangular numbers are the sum of the first n numbers $1, 2, 3, 4, \ldots$. To get the sequence of partial sums, we multiply by $\frac{1}{1-x}$. so this gives the correct generating function again.

A.1.4. Call the generating function A. Compute $A-xA = 4+x+2x^2+3x^3+4x^4+\cdots$. Thus $A-xA = 4+\dfrac{x}{(1-x)^2}$. Solving for A gives $\dfrac{4}{1-x}+\dfrac{x}{(1-x)^3}$.

A.1.5. $\dfrac{1+2x}{1-3x+x^2}$

A.1.6. Compute $A - xA - x^2A$ and the solve for A. The generating function will be $\dfrac{x}{1-x-x^2}$.

A.1.7. $\dfrac{x}{(1-x)(1-x-x^2)}$

A.1.8. $\dfrac{2}{1-5x} + \dfrac{7}{1+3x}$.

A.1.9. $a_n = 3 \cdot 4^{n-1} + 1$

A.1.10. Hint: you should "multiply" the two sequences. Answer: 158.

Solutions for Section A.2

A.2.1. *Proof.* Suppose $a \mid b$. Then b is a multiple of a, or in other words, $b = ak$ for some k. But then $bc = akc$, and since kc is an integer, this says bc is a multiple of a. In other words, $a \mid bc$. QED

A.2.2. *Proof.* Assume $a \mid b$ and $a \mid c$. This means that b and c are both multiples of a, so $b = am$ and $c = an$ for integers m and n. Then $b + c = am + an = a(m + n)$, so $b + c$ is a multiple of a, or equivalently, $a \mid b + c$. Similarly, $b - c = am - an = a(m - n)$, so $b - c$ is a multiple of a, which is to say $a \mid b - c$. QED

A.2.3. $\{\ldots, -8, -4, 0, 4, 8, 12, \ldots\}$, $\{\ldots, -7, -3, 1, 5, 9, 13, \ldots\}$, $\{\ldots, -6, -2, 2, 6, 10, 14, \ldots\}$, and $\{\ldots, -5, -1, 3, 7, 11, 15, \ldots\}$.

A.2.4. *Proof.* Assume $a \equiv b \pmod{n}$ and $c \equiv d \pmod{n}$. This means $a = b + kn$ and $c = d + jn$ for some integers k and j. Consider $a - c$. We have:

$$a - c = b + kn - (d + jn) = b - d + (k - j)n$$

In other words, $a - c$ is $b - d$ more than some multiple of n, so $a - c \equiv b - d \pmod{n}$. QED

A.2.5. (a) $3^{456} \equiv 1^{456} = 1 \pmod{2}$.

(b) $3^{456} = 9^{228} \equiv (-1)^{228} = 1 \pmod{5}$

(c) $3^{456} = 9^{228} \equiv 2^{228} = 8^{76} \equiv 1^{76} = 1 \pmod{7}$

(d) $3^{456} = 9^{228} \equiv 0^{228} = 0 \pmod{9}$

A.2.6. For all of these, just plug in all integers between 0 and the modulus to see which, if any, work.

(a) No solutions.

(b) $x = 3$.

(c) $x = 2$, $x = 5$, $x = 8$.

(d) No solutions.

(e) No solutions.

(f) $x = 3$.

A.2.7. (a) $x = 5 + 22k$ for $k \in \mathbb{Z}$.

(b) $x = 4 + 5k$ for $k \in \mathbb{Z}$.

(c) $x = 6 + 15k$ for $k \in \mathbb{Z}$.

(d) Hint: first reduce each number modulo 9, which can be done by adding up the digits of the numbers. Answer: $x = 2 + 9k$ for $k \in \mathbb{Z}$.

A.2.8. We must solve $7x + 5 \equiv 2 \pmod{11}$. This gives $x \equiv 9 \pmod{11}$. In general, $x = 9 + 11k$, but when you divide any such x by 11, the remainder will be 9.

A.2.9. (a) Divide through by 2: $3x + 5y = 16$. Convert to a congruence, modulo 3: $5y \equiv 16 \pmod{3}$, which reduces to $2y \equiv 1 \pmod{3}$. So $y \equiv 2 \pmod{3}$ or $y = 2 + 3k$. Plug this back into $3x + 5y = 16$ and solve for x, to get $x = 2 - 5k$. So the general solution is $x = 2 - 5k$ and $y = 2 + 3k$ for $k \in \mathbb{Z}$.

(b) $x = 7 + 8k$ and $y = -11 - 17k$ for $k \in \mathbb{Z}$.

(c) $x = -4 - 47k$ and $y = 3 + 35k$ for $k \in \mathbb{Z}$.

A.2.10. First, solve the Diophantine equation $13x + 20y = 2$. The general solution is $x = -6 - 20k$ and $y = 4 + 13k$. Now if $k = 0$, this correspond to filling the 20 oz. bottle 4 times, and emptying the 13 oz. bottle 6 times, which would require 80 oz. of water. Increasing k would require considerably more water. Perhaps $k = -1$ would be better? Then we would have $x = -6 + 20 = 14$ and $y = 4 - 13 = -11$, which describes the solution where we fill the 13 oz. bottle 14 times, and empty the 20 oz. bottle 11 times. This would require 182 oz. of water. Thus the most efficient procedure is to repeatedly fill the 20 oz bottle, emptying it into the 13 oz bottle, and discarding full 13 oz. bottles, which requires 80 oz. of water.