Introduction to Mathematical Thinking

RENZO CAVALIERI

Notes for Students of MATH 235

Fort Collins, Spring 2020

DEPARTMENT OF MATHEMATICS, COLORADO STATE UNIVERSITY, FORT COLLINS, CO, 80523-1874, USA. *Email:* renzo@math.colostate.edu

Contents

Introduction

Where do we start?

In no human activity we ever start at the beginning. For example, if you want to play a videogame, you don't program it, if you want to program a computer you don't build it, if you want to build a computer you don't build the parts, if you want to build the parts you don't process the metals and plastic from raw materials, and if you want to process the raw materials you don't extract them. If you wanted to play a videogame that you built from scratch you may end up investing your whole life into that goal, and maybe still not achieve it.

Math seems different. To many of us, one of the appeals of math is that it seems like you don't have to take anything off the shelf. You make some definitions, maybe choose some axioms ("rules of the game"), and then construct your theory via logical steps. This has (and is still) been a central goal of modern mathematics, but alas, even in mathematics, we don't start all the way at the beginning. For a very simple reason, which is that starting at the very beginning is very difficult, and one can only get there after they have substantially trained their skills.

In this class, we will ASSUME as given the following:

(1) the notion of numbers (natural, integer, rational and real), and the operations of addition and multiplication;
(2) the notion of set (roughly speaking a collection of "things" called elements).

We all "know" (or think we do) what these things are, but if you try and think of precise definitions you will find it very hard. What happened is that we got "used" to using these concepts instead of really understanding what they are... just like driving a car instead of building it. For a first class in mathematical thinking, I think it is best to start driving, and then whomever may be interested in the future may look under the hood.

What does it mean to write a "complete" proof?

A proof establishes the truth of a mathematical statement. A mathematical statement consists of a bunch of hypotheses, which are the things that you assume to be true, and of a statement called thesis that you want to deduce from the hypotheses. Note that sometimes the hypothesis are hidden.

For example, in the mathematical statement:

> *For every natural number n, the sum* $1+3+5+...+(2n-1) = n^2$

we have written the thesis. The hypotheses are that we know what natural numbers are, and how addition works. Often, hypotheses that are part of the framework of mathematics are omitted from the statement.

Proving that statement means writing a string of sentences that establish the validity of the above formula. There are many different ways to write a complete and correct proof, and no algorithm that will efficiently work in every case. This makes the task more difficult, but in a sense also more fun, because we are often led to looking for not just “a” proof, but a “nice” proof, where nice could mean many things: elegant, slick, efficient, or illuminating, for example...

The string of sentences must take us from the hypothesis to the thesis by a sequence of logical implications. There is a big difference between writing a proof for a computer or for a human being... in this class we are not learning to write “proofs that a computer can understand”, but “proofs that a human can understand”... and here there is some unavoidable ambiguity in what it means to write a complete proof... the way that I would write a proof for you is certainly different from the way I would write the same proof for another mathematical researcher, where I can assume that they have much more background and I am allowed to not have to say everything. Since you are beginners, try to err on the safe side. If in doubt, give extra motivation at each step. At the same time, be careful that you are not just repeating yourself without adding any content.

Let me give a silly and not very mathematical example. Suppose you want to provide an argument for the sentence:

If it rains, then it is a good idea to have an umbrella.

This is not an argument:

If it rains, it rains. Which means it is raining. Then it is a good idea to have an umbrella. Because you know, it rains. So it is a good idea to have an umbrella. Because it is a good idea.

A much more complete argument is the shorter:

If it rains, then it is a good idea to have an umbrella, since that prevents you from getting wet.

UNIT 1

Examples of proofs

1. Key Ideas

Proving a mathematical statement means to give a sequence of implications that show how the **thesis** (what is stated to be true) follows from the **hypothesis** (what is assumed to be true). There are many different ways to write a proof, and trying to constrain proof writing to an algorithm will sooner or later be limiting. Also, it is often useful to "translate" a mathematical statement into an equivalent statement where we have better tools to show its validity. Rather than continuing with a general but abstract discussion, let us illustrate this with one example. Consider the mathematical statement:

(S): The sum of the first n odd natural numbers is equal to n^2.

We can rewrite this statement symbolically as follows:

$$1 + 3 + 5 + \ldots + (2n - 1) = n^2, \tag{1}$$

or if we want to be even more slick,

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

Before we proceed any further, let us pause for a second to become aware of some hidden things to which we will pay a lot of attention during this semester:

hypotheses: In the statement above there seem to be no hypothesis, but in fact we have just hidden what we consider common knowledge: that we agree on what a natural number is, on what an odd number is, that we agree on how the operations of addition and multiplication work and that the symbol n^2 means $n \cdot n$. These are the hidden hypotheses of this mathematical statement.

quantifier: In order for statement (S) to be true, formula (2) must hold **for every possible choice of a natural number** n. So, if one wanted to be absolutely complete, one should rewrite it as

> For every natural number n, the sum of the first n odd natural numbers is equal to n^2.

clear pattern: In equation (1) we take advantage of the fact that we are communicating among humans and not to a computer. Writing $1 + 3 + 5$ establishes a pattern which is easily recognizable by just about anyone. That is why it is acceptable to put "..." to indicate that one should continue with the same pattern until the number $(2n - 1)$.

Now let us get to work and prove statement (S). OK, hold on... before we actually start the proof, there is one more important thing to do... being skeptical! We are going to **test** the statement for a few values of n. This is useful for two reasons:

(1) It provides a check that there was not a typo, or that you were given as a task to try and prove a false statement (never trust your teachers).
(2) Testing a statement often allows you to understand the statement better, and often suggests how to prove it.

So let us test our statement for n up to 5:

$n = 1$:
$$1 = 1^2,$$
$n = 2$:
$$1 + 3 = 4 = 2^2,$$
$n = 3$:
$$1 + 3 + 5 = 9 = 3^2,$$
$n = 4$:
$$1 + 3 + 5 + 7 = 16 = 4^2,$$
$n = 5$:
$$1 + 3 + 5 + 7 + 9 = 25 = 5^2.$$

Allright, it looks like I was not trying to fool you. But here is an interesting observation that one can make by looking at the above list of checks. At each line, you have a summation which is equal to the previous line plus one more term. Which in practice means that if you have already checked that $1 + 3 + 5 + 7 = 16$ in the fourth line, in the fifth line you could save yourself some energy by computing directly $16 + 9$ as opposed to going back to doing all over $1 + 3 + 5 + 7 + 9$. This idea is at the basis of a proof technique called **induction**. Let us first show the proof of statement (S) using this technique, and then make some general comments about induction.

1.1. Proof by induction.

We first establish the **base case**. Statement (S) is true for $n = 1$: the sum of the first odd natural number is 1, which is equal to 1^2.

We now assume statement (S) to be true for n equal to a particular but unspecified value $n = N$. This becomes a hypothesis (called the **inductive hypothesis**), which means that we can use it as a true fact, in order to establish the validity of statement (S) for $n = N + 1$. (this is called the **inductive step**).

Let us be very explicit: to perform the inductive step we assume that
$$(H) : 1 + 3 + 5 + \ldots + (2N - 1) = N^2$$
is true, and we show that it follows that
$$(T) : 1 + 3 + 5 + \ldots + (2N - 1) + (2N + 1) = (N + 1)^2$$
must also be true.

It is often convenient to start from what you have to show to be true. We are going to start from the left hand side of (T), and use a sequence of algebraic maniplulations together with the use of (H) to show that it equals the right hand side of (T).

Start from the left hand side of (T). By associativity of addition we have:

$$1 + 3 + 5 + \ldots + (2N - 1) + (2N + 1) = [1 + 3 + 5 + \ldots + (2N - 1)] + (2N + 1)$$

In the square parentheses (and in blue if you have colors) we have the left hand side of (H). Since we are assuming (H) to be a true statement, we can replace the square parenthesis with the right hand side of (H):

$$[1 + 3 + 5 + \ldots + (2N - 1)] + (2N + 1) = N^2 + (2N + 1).$$

Now by foiling we have

$$N^2 + (2N + 1) = (N + 1)^2,$$

which is the right hand side of (T). This concludes the proof by induction of statement (S).

Induction. What just happened? How did we just prove statement (S)? There is a perfect analogy between a proof by induction and the game of domino: you should imagine that statement (S) is in fact an infinite number of statements, one for each value of n (in the test above we saw the first 5). Imagine each of these statements is the tile of a domino game, put one after the other. A statement being true corresponds to a tile falling. When we show the **inductive step**, we are saying that if any one tile falls, then it knocks down the next one. Establishing the base case amounts to showing that the first tile falls. These two together show that every tile is gonna fall, because the first knocks down the second, the second the third, and so on forever.

Here is a formal description of the technique of proof by induction. Given a collection of mathematical statements $S(n)$, indexed by natural numbers n. A proof by induction consists in establishing the following two steps:

base case: show that $S(1)$ is true.

inductive step: assume that $S(N)$ is true for some unspecified number N, and show that it must be the case that $S(N+1)$ is also true.

By applying the inductive step for $N = 1$, (you have shown $S(1)$ to be true in the base case), you obtain that $S(2)$ is true. But now you can apply the inductive step for $N = 2$ (which you just learned was true), to obtain that $S(3)$ is true, and so on and so forth...

1.2. Proof by translation to a geometric problem.

We now prove statement (S) by translating it to a geometric statement which we then show to be true.

Consider a collection of beads arranged on the plane on all points (x, y) such that x, y are integers, and $1 \leq x \leq n$, $1 \leq y \leq n$, as illustrated in Figure 1.1. Since each row contains n beads, and there are n rows, there are exactly

n^2 beads in this arrangement. Now count the beads as shown by the red lines in the drawing: for every integer i, you group together the beads with coordinates $i = x \le y$ and those with coordinates $i = y \le x$. There are $2i - 1$ beads in this group. Since the total number of beads in the grid is equal to the sum of the beads in the various red groups, we have shown that

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

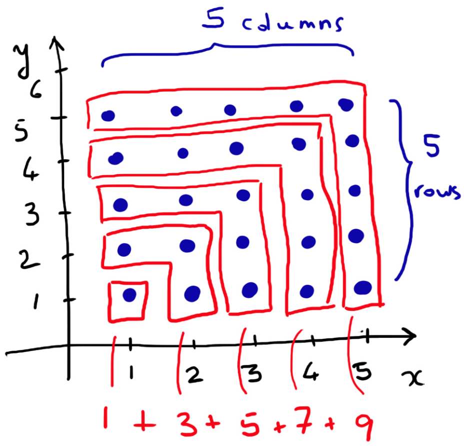


FIGURE 1.1. A figure that illustrates the argument of the geometric proof of statement (S) for $n = 5$. Remember: a figure is not a proof, but a figure helps understanding a proof by clarifying and establishing a geometric pattern which we are using in the proof of our argument.

1.3. Summing integers and counting handshakes. We define two integers that depend on n.

$H(n)$: the number of handshakes that happen if, in a group of n persons, everyone shakes hands with everyone else exactly once.

$G(m)$: the sum of the first m natural numbers. In symbols,

$$G(m) = \sum_{i=1}^{m} i.$$

Now we state some theorems (mathematical statements whose truth we will work to establish).

THEOREM 1.1. *For every natural number n,*

$$H(n) = G(n - 1).$$

THEOREM 1.2. *For every natural number n,*

$$H(n) = \frac{1}{2} n \cdot (n - 1).$$

THEOREM 1.3. *For every natural number m,*

$$G(m) = \frac{1}{2}m \cdot (m+1).$$

First off, let us observe that Theorem 1.2 gives a formula for $H(n)$, Theorem 1.3 gives a formula for $G(m)$ and Theorem 1.1 expresses a relationship between H and G. Theorem 1.1 is a translation between two problems: a combinatorial problem of counting handshakes is related to an algebraic problem of adding a sequence of natural numbers.

We observe the following logical implications:

(1) Theorem 1.1 + Theorem 1.2 $\Longrightarrow$ Theorem 1.3;
(2) Theorem 1.1 + Theorem 1.3 $\Longrightarrow$ Theorem 1.2;
(3) Theorem 1.2 + Theorem 1.3 $\Longrightarrow$ Theorem 1.1.

In other words, if any two of those three theorems are true, then the third one is also true. Another way to think of the logical relationship between these three statements is that **if one of them is true, then the remaining two become equivalent**. We will revisit this fact and use it in the groupwork.

2. Groupwork

Refer to the quantities $G(n), H(n)$ defined in Section 1.3. As we said, before embarking on proving anything, we should test our statements to make sure we can't immediately show that they are false.

PROBLEM 1.1. Compute $H(n)$ for $n \leq 4$ and $G(m)$ for $m \leq 4$, and show that the statements of the three theorems are verified for these small values of n and m.

Next, we are going to establish the truth of Theorem 1.1.

PROBLEM 1.2. Prove Theorem 1.1 by organizing the count of the handshakes $H(n)$ in a way that makes it clear that the total number of handshakes is $G(m-1)$.

Here are two examples of how one might write a solution to this problem. In black is what constitutes a proof. In blue are my comments to explain what is going on.

(A)

First we organize the number of shakes in a way that will be useful.

We organize the counting of handshakes as follows: we number the n persons 1 to n. We let the persons in the room one at a time, in the order corresponding to their numbering. When a person enters the room, they shake hands with all the persons which are already in the room (equivalently, with all persons with lower number). After person n has entered the room and shaken the hands of the people in the room, all handshakes will have occured.

We see that such organization expresses the number of handshakes as a sum,

For $j = 1 \dots n$, when the person labeled j enters the room, there are $j-1$ other persons in the room. Person j then makes $j-1$ handshakes. The total number of handshakes is equal to the sum of the handshakes that each person does when entering the room. Therefore:

$$H(n) = \sum_{j=1}^{n} (j-1)$$

Algebra now shows that the right hand side of the above equation is what we want.
By applying the substitution $i = j - 1$, we have:

$$H(n) = \sum_{j=1}^{n} (j-1) = \sum_{i=0}^{n-1} i = \sum_{i=1}^{n-1} i = G(n-1)$$

The second equality is obtained by applying the substitution and reindexing appropriately the summation bounds. The third equality follow from the fact that the $i = 0$ summand in the summation is 0 and it therefore can be omitted. The last equality is the definition of G.
(B)
Prove by induction. Base case.
$H(2) = 1$, since two persons will shake hands, and that will be the only handshake.
$G(1) = \sum_{i=1}^{1} i = 1$.
Therefore $H(2) = G(1)$ and the base case $(n = 2)$ has been established.
Stating the inductive step.
Assume $H(N) = G(N-1)$ and let us analyze $H(N+1)$. We want to show that it is equal to $G(N)$.
Breaking up $H(N+1)$ into two parts.
We can subdivide the number of handshakes in two groups: those that do not involve the $(N+1)$-st person, and those that do. This leads to the equation:

$$H(N+1) = H(N) + \text{number of handshakes involving person } N+1.$$

Now we must use the inductive hypothesis for the first summand, and we must compute the second part.
By the inductive hypothesis $H(N) = G(N-1)$. For the second term, person $N + 1$ shakes hands with N other people. Therefore the equation above becomes:

$$H(N+1) = G(N-1) + N = \sum_{i=1}^{N-1} i + N = \sum_{i=1}^{N} i = G(N),$$

where the second equality is the definition of G, the third equality is an algebraic manipulation, and the fourth equality is again the definition of G.

After proving Theorem 1.1, the remaining two Theorems are now equivalent in the sense that any proof of either of the remaining theorems will imply that all three theorems hold true. We will study a few different proofs, just to become familiar with different ways one may go about proving the same statement. For each of the proofs we present, ask yourselves what are the things you like and dislike about such proof (this is a subjective question, so there isn't a correct answer).

Problem 1.3 (Baby Gauss trick). Prove Theorem 1.3 by using a trick attributed (probably an urban legend) to Gauss in kindergarden. Write the numbers 1 to m in a $m \times 2$ array. In the first row write them in increasing order, in the second row write them in decreasing order:

$$\begin{array}{c|c|c|c|c|c|c} 1 & 2 & 3 & \dots & m-2 & m-1 & m \\ m & m-1 & m-2 & \dots & 3 & 2 & 1 \end{array}$$

What is the relationship between the sum of all the numbers in this array and $G(n)$? Find a way to compute the sum of all the numbers in this array that allows to give a formula for this sum. Organize these observations into a string of sentences that coherently and completely proves Theorem 1.3.

We now give a rather complicated proof of Theorem 1.3. We will see however that this proof has the advantage of giving the possibility of generalizing the statement of Theorem 1.3. Start from the following identity, which is well known to hold true for any value of i:

$$(i+1)^2 = i^2 + 2i + 1$$

We manipulate it as follows:

$$\begin{aligned} -i^2 + (i+1)^2 &= 2i + 1 \\ -i^2 + (i+1)^2 - 1 &= 2i \\ \frac{1}{2}\left(-i^2 + (i+1)^2 - 1\right) &= i \end{aligned}$$

This seems a bizarre thing to do, let us ponder for a second what we just did: we expressed an integer i as an expression that contains the difference of the squares of two consecutive numbers. This doesn't seem all that convenient at first, but let us remember that our task is to add integers. So let us add summation signs on either sides of the above identity:

$$\sum_{i=1}^{m} \frac{1}{2}\left(-i^2 + (i+1)^2 - 1\right) = \sum_{i=1}^{m} i$$

Since addition is commutative and distributive, we can rewrite this as:

$$\frac{1}{2}\left(\sum_{i=1}^{m}[-i^2 + (i+1)^2] - \sum_{i=1}^{m} 1\right) = \sum_{i=1}^{m} i = G(m). \tag{3}$$

Problem 1.4. Consider the summation:

$$\sum_{i=1}^{m}[-i^2 + (i+1)^2]$$

Compute the value of this summation for small values of m. Guess a formula for a general value of m. Provide an argument that proves your formula.

This summation is what is called a **telescoping sum**: the second term of the first summand cancels with the first term of the second summand, the second term of the second summand cancels with the first term of the third summand, and so on; so, in the end, out of this whole big summation, the only surviving terms are the very first and the very last.

Problem 1.5. Use the formula from Problem 1.4 to compute (3), and show that this gives a proof of Theorem 1.3.

3. Homework

Each regular homework question will be graded on a 3 *points scale. In order to get the full* 3 *points, the solution must be* **correct, complete, well-organized and tidily written-up**. *Extra-credit homework questions are worth* 2 *points each.*

EXERCISE 1.1. What do you think are are the pros and cons of proof by induction?

EXERCISE 1.2. Give a proof by induction of Theorem 1.3.

EXERCISE 1.3. Give a geometric proof of Theorem 1.3, again by thinking of the integers as counting the number of beads, and realizing a strategy of positioning the beads on the plane in such a way that the quantity $G(m)$ can be counted. **Hint:** look for arranging $1+2+3+\ldots+m$ beads inside a square $m \times m$ grid.

EXERCISE 1.4. Use Theorem 1.3 to give another proof of the statement:

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

EXERCISE 1.5 (challenge - extra credit). Use an argument similar to the last proof in the groupwork to find a formula for the sum of the first n squares:

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

EXERCISE 1.6 (challenge - extra credit). Prove that the formula for the sum of the first n k-powers is a polynomial in n of degree $k+1$.

$$\sum_{i=1}^{n} i^k = 1^k + 2^k + \ldots + (n-1)^k + n^k = a_{k+1}n^{k+1} + \ldots + a_0$$

Note: I am not asking you to find the formula for such a polynomial!!

UNIT 2

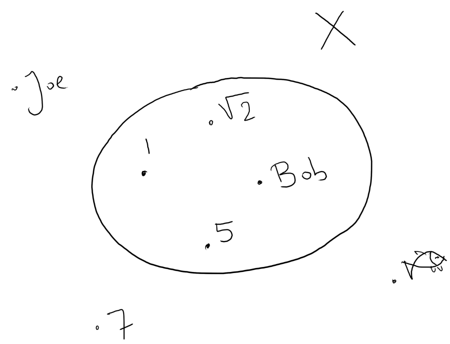
Sets

1. Key Ideas

As we mentioned on our first day of class, we are not going to properly define what a set is, since doing it honestly would require way too much work. We all have this intuitive notion of a set being a collection of "things", called the elements of the set. Such intuitive notion will suffice for us to be able to work with sets, and set operations.

Here are some examples of sets:

(1) The set of students of this class.
(2) The set $A = \{1, 3, d, f, banana\}$.
(3) The set of integers $\mathbb{Z}$.
(4) The set X:



Observations:

elements: elements of a set can be anything you want. Set A containing banana as an element is just as mathematical as the set of integers.

how to describe sets: there are many ways that we can describe sets:
(1) by giving properties that identify the elements of the set (Example 1).
(2) by listing the elements of the set (Example 2). The order does not matter!
(3) certain common mathematical sets have their own name (Example 3).

(4) by drawing an oval, and putting the elements of the set inside the oval: this is called the Venn diagram representation of a set (Example 4).

equality of sets: given that sets may be described in different ways, it is important to clarify what it means that two sets are equal. Two sets are equal if they have the same elements.

cardinality of a set: the number of elements of a set is called the cardinality of the set. If a set does not contain a finite number of elements, we say its cardinality is infinite.

DEFINITION 2.1. A set A is a **subset** of a set B if every element of A belongs to B. We write

$$A \subseteq B$$

to say that A is a subset of B.

To prove that $A \subseteq B$ one must show that if $x \in A$ (read x is an element of A, or x belongs to A, or x in A) , then $x \in B$.

QUESTION 2.1. Is a set A a subset of itself?

DEFINITION 2.2. A set A is a **proper subset** of a set B if $A \subseteq B$ and $A \neq B$. We write

$$A \subset B$$

to say that A is a proper subset of B.

To prove that $A \subset B$ one must show that if $x \in A$, then $x \in B$, plus there exists an element $y \in B$, $y \notin A$.

DEFINITION 2.3. Given two sets A, B we define:

Intersection: The intersection $A \cap B$ is the set of all common elements of A and B. In other words, $x \in A \cap B$ if and only if $x \in A$ **and** $x \in B$.

Union: The union $A \cup B$ is the set of all elements that belong to A or to B (and yes, they can also belong to both A and B). In "mathematese", $x \in A \cup B$ if and only if $x \in A$ **or** $x \in B$.

Observe that the set $A \cap B$ is a subset of A and it is a subset of B. On the other hand A is a subset of $A \cup B$, and B is a subset of $A \cup B$.

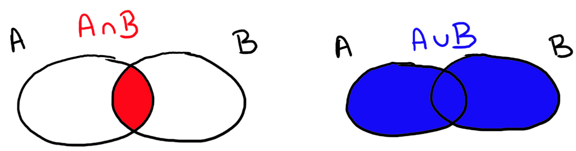


FIGURE 2.1. The Venn representation of the intersection and union of two sets.

Here is an example of the kind of simple mathematical statements you can make about sets.

THEOREM 2.1. *For any three sets* A, B, C,

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C).$$

The basic trick is that in order to prove an equality of sets $X = Y$, one typically proves two inclusions: that X is a subset of Y and Y is a subset of X. *I also recommend, when proving statments like this, to draw Venn diagrams to see what is going on.* The proof of this statement follows in black. In blue I am commenting how each step of the proof comes about.

PROOF:

STEP 1 We show that $(A \cup B) \cap C \subseteq (A \cap C) \cup (B \cap C)$.
The following sentence is a rephrasing of what we need to prove, in a way that makes it easier to show the statement is true.
Consider an element $x \in (A \cup B) \cap C$; we want to show that $x \in (A \cap C) \cup (B \cap C)$.
We use the definition of intersection to make the next statement:
We know that $x \in C$ and $x \in A \cup B$.
We use the definition of union to make the next statement:
There are two possibilities:

(1) $x \in C$ **and** $x \in A$, **or**
(2) $x \in C$ **and** $x \in B$.

There are two possible cases, so we must analyze each case separately.
In the first case $x \in (A \cap C)$, *using the definition of intersection*
and therefore $x \in (A \cap C) \cup (B \cap C)$ *using the definition of union*.
In the second case $x \in (B \cap C)$, *using the definition of intersection*
and therefore $x \in (A \cap C) \cup (B \cap C)$ *using the definition of union*
. *We have thus established that in all possible cases, when x is an element of $(A \cup B) \cap C$, it follows that x is also an element of $(A \cap C) \cup (B \cap C)$.*
It follows that $(A \cup B) \cap C \subseteq (A \cap C) \cup (B \cap C)$.

STEP 2 *We must show that* $(A \cup B) \cap C \supseteq (A \cap C) \cup (B \cap C)$.
Consider an element $x \in (A \cap C) \cup (B \cap C)$.
There are two possibilities:

(1) $x \in A$ **and** $x \in C$, **or**
(2) $x \in B$ **and** $x \in C$.

In both cases, $x \in C$, so we conclude that $x \in C$.
Since in the first case $x \in A$ and in the second case $x \in B$, we conclude that $x \in A \cup B$.
Therefore $x \in (A \cup B) \cap C$.
We have shown that $(A \cap C) \cup (B \cap C) \subseteq (A \cup B) \cap C$.

Since $(A \cup B) \cap C \subseteq (A \cap C) \cup (B \cap C)$ and $(A \cap C) \cup (B \cap C) \subseteq (A \cup B) \cap C$, it must be that $(A \cap C) \cup (B \cap C) = (A \cup B) \cap C$. This concludes the proof.

□

Let us conclude this part with a couple more definitions.

DEFINITION 2.4. The **universe** $\mathcal{U}$ is the set of all things that we wish to consider as potential elements of sets. Note that the universe set depends on what we are trying to do with it...in mathematics, the universe might be the set of all numbers that we want to work with. In probability, it may be the set of all events that we wish to consider. In any case, we just want to think of the universe as containing "everything".

By this definition, any set that we are wishing to consider is a subset of the universe. I like to represent the universe in Venn diagrams as a large rectangle. Then I will fit any set that I want to consider inside such rectangle.

QUESTION 2.2.
For any set A, what is $A \cup \mathcal{U}$? What is $A \cap \mathcal{U}$?

DEFINITION 2.5. A set that contains no elements is called the **empty set** and denoted ϕ.

By the definition of empty set, the empty set is a subset of any other set!

QUESTION 2.3.
For any set A, what is $A \cup \phi$? What is $A \cap \phi$?

2. Groupwork

DEFINITION 2.6. Let us introduce some other operations on sets.

Complement: Given a set A, we define the **complement** of A, denoted A^c, by the following property: an element y belongs to A^c if and only if y does not belong to A.

Difference: Given two sets A, B, the **difference** $A \setminus B$ consists of all elements of A which do not belong to B.

Symmetric difference: Given two sets A, B, the **symmetric difference** $A \Delta B$ is defined to be:

$$A \Delta B = (A \setminus B) \cup (B \setminus A).$$

PROBLEM 2.1. Draw Venn diagrams that represent the three definitions above.

PROBLEM 2.2. Convince yourselves that for any set A,

$$A^c = \mathcal{U} \setminus A.$$

PROBLEM 2.3. When is $A \setminus B = B \setminus A$?

PROBLEM 2.4. Prove, with the same level of detail as the proof of Theorem 2.1, that for every two sets A, B,

$$(A \cap B)^c = A^c \cup B^c.$$

PROBLEM 2.5. Decide, by drawing appropriate Venn diagrams, whether the operations of difference and symmetric difference are associative. More precisely, this means to check whether, for sets A, B, C:

(1)
$$(A \setminus B) \setminus C = A \setminus (B \setminus C)?$$

(2)
$$(A \Delta B) \Delta C = A \Delta (B \Delta C)?$$

3. Homework

Each regular homework question will be graded on a 3 *points scale. In order to get the full* 3 *points, the solution must be* **correct, complete, well-organized and tidily written-up**. *Extra-credit homework questions are worth* 2 *points each.*

EXERCISE 2.1. Prove, with the same level of detail as the proof of Theorem 2.1, that given two sets A, B,

$$A \subseteq B \iff B^c \subseteq A^c.$$

(Read: A is a subset of B if and only if B^c is a subset of A^c).

EXERCISE 2.2. Prove (with the same level of detail as the proof of Theorem 2.1) that

$$A \Delta B = (A \cup B) \smallsetminus (A \cap B).$$

EXERCISE 2.3. Draw Venn diagrams representing each of the sets below:

(1)
$$(A \cup B) \Delta C.$$

(2)
$$A \cup (B \Delta C).$$

(3)
$$(A \Delta B \Delta C)^c \smallsetminus (A \cup B \cup C).$$

(4)
$$((A \cap B) \Delta (B \cap C)) \Delta (A \cap C).$$

(5)
$$((A \cap B) \cup (B \cap C) \cup (A \cap C)) \smallsetminus (A \cap B \cap C).$$

DEFINITION 2.7 (Cartesian product). Given two sets X, Y the **cartesian product** $X \times Y$ is the set of ordered pairs

$$X \times Y := \{(x, y) | x \in X, y \in Y\}.$$

EXERCISE 2.4. Let us get familiar with the cartesian product.

- If $X = \{1, 2\}$ and $Y = \{a, b, c\}$, write down the elements of $X \times Y$ and $Y \times X$. Notice that $X \times Y \neq Y \times X$.
- If X is a set with n elements, and Y a set with m elements, what is the cardinality of $X \times Y$?
- What is $X \times \phi$?

EXERCISE 2.5 (challenge - extra credit). More questions on the cartesian product:

- If $A \subseteq X$ and $B \subseteq Y$, prove that $A \times B \subseteq X \times Y$.
- True or false:

$$(X \times Y) \smallsetminus (A \times B) = (X \smallsetminus A) \times (Y \smallsetminus B)?$$

EXERCISE 2.6 (challenge - extra credit). A finite list of sets $U_1, U_2, \ldots, U_n$ is called **nested** if

$$U_1 \supseteq U_2 \supseteq \ldots \supseteq U_n.$$

In this case, what is the intersection

$$U_1 \cap U_2 \cap \ldots \cap U_n?$$

An infinite list of sets $U_1, U_2, \ldots$ is called **nested** if

$$U_1 \supseteq U_2 \supseteq \ldots$$

Can you construct an example of an infinite nested list of sets, all different from the empty set, but such that

$$\bigcap_{i=1}^{\infty} U_i = \phi.$$

UNIT 3

Definitions, Quantifiers, Implications

1. Key Ideas

In real life, we seldom use definitions. For example, you never worried about what is the definition of a *table*. Over the course of your life you were shown a bunch of examples of tables and you abstracted the notion of *table*, so that now you recognize a *table* when you see one. In a sense, in elementary mathematics, we also do that. We don't ever really define what numbers **are**, we just become familiar with them by using them.

In more advanced mathematics, definitions are the starting point of everything. A **definition** is a sentence that identifies a mathematical concept, or equivalently , a rule that allows you to decide, given any mathematical object that comes up to you, whether that object is or isn't that mathematical concept that is being defined. Given any mathematical statement, in order to stand a chance to prove it, you have to be familiar with the definitions of all concepts involved in the statement. Let us start with some simple examples.

DEFINITION 3.1. A natural number n is:

even: if $n = 2 \cdot q$, where q is a natural number;
prime: if it is different from 1, and its only divisors are 1 and itself;
foofy: if it is even, prime, and greater than 7.

THEOREM 3.1. *If a natural number is divisible by* 4*, then it is even.*

PROOF:
Let n be a natural number divisible by 4.
This means that we can write $n = 4 \cdot k$, where k is also a natural number.
Since $4 = (2 \cdot 2)$, we have $n = (2 \cdot 2) \cdot k$.
Multiplication is associative, therefore we can rearrange the parenthesis:

$$n = 2 \cdot (2 \cdot k).$$

Since $2 \cdot k$ is a natural number, this shows that n satisfies the definition of being even by letting $q = 2k$.

□

THEOREM 3.2. *There are no foofy numbers.*

PROOF: The only even prime number is 2, since, by definition, a prime number is divisible only 1 and itself. So 2 is the only candidate to be a foofy number. Since 2 is less than 7, 2 is not foofy. Therefore there are no foofy numbers.

□

Note: defining foofy numbers was kind of a silly thing, since such definition does not apply to any number. Nonetheless, sometimes one is given a definition and has to work with it before they know whether it may be silly or not.

Quantifiers are the sentence fragments *there is* and *for every*. They are extremely important in mathematical sentences. Notice the big difference between the following two statements:

(1) **There exists** a prime divisor p equal to 2 in the prime factorization of n.
(2) **For every** prime divisor p in the prime factorization of n, $p = 2$.

The first statement is telling us that n is an **even** number, the second statement is telling us that $n = 2^k$.

Suppose that a number n walks up to you and you want to assess whether it satisfies statements (1) or (2).
For statement (1).

- If you want to establish (1) to be true, you must look at elements of the prime factorization of n until you find the factor 2. If you find it, you can stop.
- If you want to establish (1) to be false, you must look at ALL the prime factorization of n, and check that all prime divisors are different from 2.

For statement (2).

- If you want to establish (2) to be true, you must look at ALL the prime factorization of n, and check that all prime divisors are equal to 2.
- If you want to establish (2) to be false, you must look at the prime factorization of n until you find a factor different from 2. If you find it, you can stop.

In other words, the negative of a statement with a "there exists" contains a "for every" and viceversa. In this specific case, let us write the negatives of these statements. Oh, and since we are talking about definitions, let us define the negative.

Definition 3.2. The **negative** of a mathematical statement S is another mathematical statement *not* S with the property that S is true when *not* S is false and S is false when *not* S is true.

With this definition in place, we have:

Not (1): For every prime divisor p in the prime factorization of n, $p \neq 2$
Not (2): There exists a prime divisor different from 2 in the prime factorization of n.

An **implication** is a mathematical statement of the form *if A, then B*, also denoted as $A \implies B$, where A and B are themselves mathematical statements. If the implication $A \implies B$ applies to a set of events X, then the implication is true if for every element $x \in X$:

(1) $A(x)$ and $B(x)$ are both true, **or**
(2) $A(x)$ is false.

Alternatively, we can think of $A \subset X$ as the set of events $x \in X$ for which $A(x)$ is true, and of $B \subset X$ as the set of events $x \in X$ for which $B(x)$ is true. Then the implication $A \implies B$ corresponds to the fact that $A \subseteq B$.

Let us explore this in a simple example. Consider the statement:

(S) *For every day of February 2021, if it is raining when I am about to go to work, then I take the umbrella.*

In this case we have:

- the set X of events I want to consider has 28 elements, corresponding to the days of February 2021;
- the statement A is *it is raining when I am about to go to work*;
- the statement B is *I take the umbrella*

For the statement (S) to be a true mathematical statement, I have to look at all 28 days of February, and make sure that those days when it was raining when I was about to go to work, I actually did take the umbrella with me. On the sunny days, it does not matter what I do: if some days I did choose to take the umbrella anyway, maybe out of worry that it may rain later, that is allright, it does not make (S) false, since (S) is not giving me any constrain on days when it does not rain.

On the other hand, to make (S) false I need to find one day, just one, when it was raining and I did not take the umbrella. It does not matter if it rained all February, and I forgot to take the umbrella only once out of the whole months. That one time makes (S) false.

We define what it means for two mathematical statements A and B to be **equivalent** (we say *"A if and only if B"* and write $A \iff B$). If the statements A and B apply to a set of events X, then $A \iff B$ when for every element $x \in X$, either $A(x)$ and $B(x)$ are both true, or they are both false.

Returning to our example, now let us say:

(S') *For every day of February 2021, it is raining when I am about to go to work if and only if I take the umbrella.*

For this statement (S') to be true, meaning that A is equivalent to B, every rainy day of February I must have the umbrella and every not rainy day of February I must not have the umbrella.

Note that $A \iff B$ corresponds to the simultaneous application of the two implications $A \implies B$ and $B \implies A$.

2. Groupwork

We are going to train our skills by making up some random definitions, and running with them.

For this groupwork, we must first of all fix a universe $\mathcal{U}$ consisting of a set of words. Any set we are going to consider now is going to be a subset of this universe. You may choose $\mathcal{U}$ to be the set of all words in the English language, but you may also want to choose $\mathcal{U}$ to be a smaller set of words, for example a finite set, if you want to make things simpler.

DEFINITION 3.3. An **asfl** is a set A such that every element of A starts with the same letter.

DEFINITION 3.4. A **tesfall** is a set T such that there exists an element of T whose first and last letter are the same.

The first thing to do is getting comfortable with these definitions by giving some examples and counterexamples.

PROBLEM 3.1. Give an example of an asfl, a tesfall, of something which is an asfl and a tesfall, an asfl but not a tesfall, a tesfall but not an asfl, and neither a asfl nor a tesfall.

Then we start making some statements about our newly introduced concepts.

PROBLEM 3.2. For each of the following statements, decide whether they are true or false. If they are true, give a proof. If they are false, provide a counterexample showing it.

(1) If $\mathcal{U}$ is an asfl, then every subset of $\mathcal{U}$ is an asfl.
(2) If $\mathcal{U}$ is a tesfall, then every subset of $\mathcal{U}$ is a tesfall.
(3) If there exists an asfl A, then $\mathcal{U}$ is an asfl.
(4) If there exists a tesfall T, then $\mathcal{U}$ is a tesfall.
(5) The empty set is an asfl.
(6) The empty set is a tesfall.

We now introduce some attributes of tesfalls and asfls.

DEFINITION 3.5. An asfl A is **closed** if A is the empty set or if there is no A' properly containing A which is also an asfl.

DEFINITION 3.6. A tesfall T is **full** if every proper subset $\phi \neq S' \subset S$ is also a tesfall.

PROBLEM 3.3. Give some examples of closed asfl's, not closed ones, full tesfall's and not full ones. Notice that being a closed asfl doesn't only depend on B, but it depends essentially on $\mathcal{U}$. Give an example of a set B and two choices of $\mathcal{U}$ such that in one case B is a closed asfl, in the other case it isn't.

PROBLEM 3.4. Can you figure out equivalent statements for Definitions 3.5 and 3.6, that may be more intuitive? Formulate these statements as Theorems of the form: " A asfl is closed if and only if...", "A tesfall is full if and only if...".

DEFINITION 3.7. Given an asfl A, the **closure** of A, denoted $\overline{A}$, is the unique closed asfl containing A.

PROBLEM 3.5. Show that Definition 3.7 makes sense: for any asfl B, there exists a unique closed asfl containing B.

DEFINITION 3.8. Given a tesfall T, the **core** of T, denoted T°, is the largest subset of T which is a full tesfall.

PROBLEM 3.6. Show that Definition 3.8 makes sense: for any tesfall T, there exist a unique largest full tesfall contained in T.

Problem 3.4 can come in handy to solve the last two questions.

3. Homework

Each regular homework question will be graded on a 3 *points scale. In order to get the full* 3 *points, the solution must be* **correct, complete, well-organized and tidily written-up**. *Extra-credit homework questions are worth* 2 *points each.*

EXERCISE 3.1. Refer back to Statements (1) and (2) at the end of the Key Ideas section. Decide, for every $n \leq 20$ the truth or falsehood of the statements in the leftmost column of the table:

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
(1)																				
Not (1)																				
(2)																				
Not (2)																				

Observe that the negative of each statement has indeed always opposite answer than the statement.

EXERCISE 3.2. Here is another fun definition to play with.

DEFINITION 3.9. A set B is **babadabubi** if all of its elements are italian words.

In mathemathese, this is written:

B is **babadabubi** if for every $b \in B$, b is an italian word.

Answer the following questions about babadabubi sets:

(1) Write down the definition of not being babadabubi.
(2) Give an example of a babadabubi set.
(3) Give an example of a set which is not babadabubi.
(4) Is the empty set babadabubi?
(5) TRUE OR FALSE: the union of two babadabubi sets is babadabubi.
(6) TRUE OR FALSE: the intersection of two babadabubi sets is babadabubi.
(7) TRUE OR FALSE: the union of two not babadabubi sets is not babadabubi.
(8) TRUE OR FALSE: the intersection of two not babadabubi sets is not babadabubi.

EXERCISE 3.3 (contrapositive). We introduce the notion of the **contrapositive** of a mathematical statement.

DEFINITION 3.10. Let S_1, S_2 be two sentences and consider a statement (S) of the form: "If S_1, then S_2". We define the **contrapositive** statement to (S) to be:

If not S_2, then not S_1.

Write down the contrapositive statement to each of the following:

(1) If it rains, then I carry an umbrella.
(2) If it doesn't rain, then I don't carry an umbrella
(3) If a natural number n is a multiple of m, then $n \geq m$.

(4) If $x = y$, then $x^2 = y^2$.

EXERCISE 3.4. The interesting thing is that the contrapositive of a statement (S) is equivalent to (S), in the sense that they are true or false together (there cannot be any situation in which a statement is true and its contrapositive is false, or viceversa). Observe this by filling in the table below. Put T if a scenario supports a statement, F if a scenario shows that the statement is false, and X if the scenario doesn't allow you to decide. Let (1) and (2) refer to the first two statements in the previous exercise.

	(1)	contrapos. to (1)	(2)	contrapos. to (2)
It rains and I have my umbrella				
It rains and I don't have my umbrella				
It doesn't rain and I have my umbrella				
It doesn't rain and I don't have my umbrella				
It doesn't rain and I tip-tap dance				
I tip-tap dance with an umbrella				

EXERCISE 3.5 (challenge - extra credit). Sometimes in order to prove a mathematical statement to be true, it is easier to prove the contrapositive statement. This proof technique is called **proof by contradiction**. Fill in the blanks to complete the proof.

THEOREM 3.3. *There are infinitely many prime numbers.*

Comment: This is another case where all the hypothesis are hidden: they consist of knowing the definition of prime numbers, and the usual axioms of arithmetics. So you should think of this theorem as the statement: "If the hidden hypothesis hold, then there are infinitely many prime numbers". Therefore the contrapositive statement is: " If there are finitely many prime numbers, then something goes wrong in arithmetics". What goes wrong is the following: if we assume there are finitely many prime numbers, and we take them all, then we can always produce one more prime number. Now let us write down this idea formally.

PROOF: Assume there are ____________ prime numbers, in particular we can assume there are n of them and we can list them all: $p_1, \dots, p_n$. Consider the number $q = p_1 \cdot p_2 \cdot \dots \cdot p_n + 1$, obtained by multiplying all the prime numbers together, and then adding 1. We make the claim that q is not divisible by ____________ of the p_i's. If the claim is true, then there are two possibilities: either q is a new prime number, or q has a prime factorization with primes that are different from the p_i's. In either case, we are contradicting the statement that the list $p_1, \dots, p_n$ contains ____________ prime numbers. Assuming that there are ____________ many primes led us to a contradiction, which shows that it must be the case that there are ____________ many prime numbers.

We now prove the claim, also by contradiction. Assume that q ____________ by one of the p_i's, without loss of generality let us assume q ____________ by p_1.

This means $q = p_1 \cdot k$, with k a natural number.

Then

$$q - p_1 \cdot p_2 \cdot \dots \cdot p_n = \_\_\_\_\_\_\_\_\_\_,$$

which shows that $q - p_1 \cdot p_2 \cdot \ldots \cdot p_n$ is divisible by ______________. But $q - p_1 \cdot p_2 \cdot \ldots \cdot p_n =$ ______________, which is not divisible by any prime. Assuming that q is divisible by one of the p_i's led to a contradiction, which allows us to conclude the claim: q is not divisible by ______________ of the p_i's.

□

EXERCISE 3.6 (challenge - extra credit). A number x is a **rational number** if $x = \frac{p}{q}$, where both p and q are integers. Prove (by contradiction) that $\sqrt{2}$ is not a rational number.

UNIT 4

Bijections and Infinities

1. Key Ideas

This week we are going to focus on the notion of **bijection**. This is a way we have to compare two different sets, and say that they have "the same number of elements". Of course, when sets are finite everything goes smoothly, but when they are not... things get spiced up! We start, as usual, with some definitions.

DEFINITION 4.1. A **bijection** between a set X and a set Y is a procedure matching elements of X with elements of Y in such a way that every element of X has a unique partner, and every element of Y has a unique partner.

Note: the word **unique** in mathematics has a specific technial meaning: it means "only one".

EXAMPLE 4.1. Let $X = \{1,2,3\}$ and $Y = \{a,b,c\}$. An example of a bijection between X and Y consists of matching 1 with a, 2 with b and 3 with c.

A NOT example would match 1 and 2 with a and 3 with c. In this case, two things fail: a has two partners, and b has no partner.

Note: if you are familiar with these concepts, you may have recognized a bijection as a function between X and Y which is both injective (1:1) and surjective (onto). We will revisit the notion of bijection in this language in a few weeks. For now, we content ourselves of a slightly lower level approach. However, you may be a little unhappy with the use of the word "matching" in a mathematical definition. Here is a more formal definition for the same concept.

DEFINITION 4.2. A **bijection** between a set X and a set Y is a subset $B \subseteq X \times Y$ such that every element of X is the first term of exactly one element of B and every element of Y is the second term of exactly one element of B.

Caution: with Definition 4.1 it may be confusing to think about what a bijection between a set X and itself is. This becomes transparent using Definition 4.2: in $X \times X$, the left and right factors of X play different roles (the pairs are ordered).

So for example, if $X = \{a,b,c\}$, B is a bijection between X and itself:

$$B = \{(a,a),(b,c),(c,b)\}.$$

Bijections are intimately related with the concept of cardinality.

THEOREM 4.1. *Assume X and Y are two finite sets. There exists a bijection between X and Y if and only if X and Y have the same cardinality (i.e. the same number of elements).*

We make the following definition, regardless of whether the sets are finite or not.

DEFINITION 4.3. Two sets X and Y have the same cardinality if there exists a bijection between them.

The first infinite set that most people encounter is the set of natural numbers $\mathbb{N}$. In a sense that I don't want to make too precise at this point, the natural numbers are the "smallest" type of infinity there is. We make the following definition.

DEFINITION 4.4. A set X is **countable** if it has the same cardinality as $\mathbb{N}$. An infinite set which is not countable is called **uncountable**.

The reason that historically the word countable was used is that giving a bijection between $\mathbb{N}$ and X is the same as choosing an element of X to be number 1, another to be number 2 and so on... in other words it is like *counting* the elements of X. Yet another way of thinking of a bijection between a set X and the natural numbers consists in writing down the elements of X as an ordered list:

$$X = \{x_1, x_2, x_3, \dots\}.$$

The bijection determined by this notation matches the natural number n with the element x_n.

THEOREM 4.2. *A subset of a countable set is either finite or countable.*

PROOF. Let X be a countable set and $Y \subseteq X$.
By assumption, X is countable which means that we can write the elements of X in a (not finite) ordered list:

$$X = \{x_1, x_2, x_3, \dots\}.$$

We call L the ordered list above.
We want to show that we can write the elements of Y also in an ordered list. Define y_1 to be the first element of Y appearing in L, y_2 to be the second element of Y appearing in L, and so on.
This procedure creates an ordered list containing all elements of Y. If this list contains a finite number of elements, Y is finite. Otherwise, Y is countable. □

THEOREM 4.3. *Let X and Y be two countable sets. Then $X \cup Y$ is also countable.*

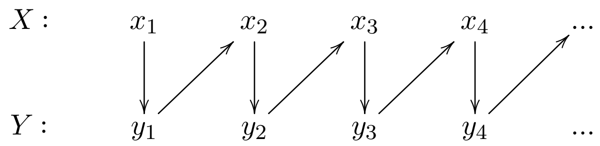
PROOF. Assume that we fix a bijection between X and $\mathbb{N}$, and let us call x_n the element of X which is matched with the natural number $n \in \mathbb{N}$. Similarly we fix a bijection between Y and $\mathbb{N}$, and call y_n the element of Y which is matched with the natural number $n \in \mathbb{N}$.

We now define a bijection between $X \cup Y$ and $\mathbb{N}$ as follows: match each x_n with the natural number $2n-1$, and each y_n with the natural number $2n$.

Via this bijection we can arrange the elements of $X \cup Y$ in the following ordered list:

$$X \cup Y = \{x_1, y_1, x_2, y_2, x_3, y_3, \ldots\}.$$

Another way to visualize this is by noting that the the bijection we described between $X \cup Y$ and $\mathbb{N}$ is obtained by following the arrows in the diagram below.



□

QUESTION 4.1. **True or false?** given any non-empty set X, there exists a bijection between X and itself.

QUESTION 4.2. **True or false?** If X is a finite set and $A \subset X$ is a proper subset of X, then there can be a bijection between A and X.

2. Groupwork

In this groupwork we examine some of the counterintuitive behavior of infinite sets. In particular, we see the analogous for infinite sets to question 4.2 has a surprising answer.

PROBLEM 4.1 (Hilbert's hotel). Hilbert's hotel has a countable number of rooms, indexed by the natural numbers. The hotel is currently full.

(1) A traveler shows up at the reception. Can the receptionist accomodate the traveler, without asking anyone to shack up with somebody else, and without leaving any of the current clients without a room?
(2) An infinite bus with a countable number of passengers shows up. Can the receptionist accomodate the passengers of the bus, without asking anyone to shack up with somebody else, and without leaving any of the current clients without a room?

QUESTION 4.3. **True or false?** If X is a set and $A \subset X$ is a proper subset of X, then there can be a bijection between A and X.

PROBLEM 4.2. Show that if X and Y are countable, then the cartesian product $X \times Y$ is countable.

Hint: Assume that you have given bijections for X and Y, and organize the elements of the cartesian product $X \times Y$ in a square, infinite, array. Then try and find a pattern for counting the elements of $X \times Y$, similar to what was done in the proof of Theorem 4.3.

Now we prove a fun fact: even though there seem to be way more rational numbers than natural numbers, in fact $\mathbb{Q}$ is a countable set.

PROBLEM 4.3. The set of rational numbers $\mathbb{Q}$ is countable. Use Problem 4.2, and Theorems 4.2, 4.3 to prove this fact.

At this point we should have at least one example of a non-countable set. Here we go.

PROBLEM 4.4. The set of infinite sequences of 0's and 1's is not countable.

Hint: We prove this fact by contradiction. Assume that the set of such sequences is countable, and imagine you can write all infinite sequences in a list. Now you need to devise a procedure to construct a sequence of 0's and 1's that cannot possibly be in that list.

Now try to adapt the same argument to prove the following statement.

PROBLEM 4.5. The set of real numbers $\mathbb{R}$ is not countable.

3. Homework

Each regular homework question will be graded on a 3 points scale. In order to get the full 3 points, the solution must be **correct, complete, well-organized and tidily written-up**. *Extra-credit homework questions are worth 2 points each.*

EXERCISE 4.1. Explain how Problem 4.1 and Question 4.3 are related.

EXERCISE 4.2. Generalize Theorem 4.3 as follows: for $n \in \mathbb{N}$, let X_n be a countable set. Show that

$$\bigcup_{n \in \mathbb{N}} X_n = X_1 \cup X_2 \cup X_3 \cup \dots$$

is a countable set.

EXERCISE 4.3. Given a set X, the **power set** of X, denoted $\mathcal{P}(X)$, is the set of all subsets of X. Prove that if X has n elements, then $\mathcal{P}(X)$ has 2^n elements, in two different ways:

(1) Do a proof by induction.
(2) Establish a bijection between $\mathcal{P}(X)$ and the set of ordered lists of 0's and 1's of length n.

EXERCISE 4.4. Prove that if X is countable, then $\mathcal{P}(X)$ is uncountable.

EXERCISE 4.5 (challenge - extra credit). Let X be an uncountable set, and $Y \subset X$ be a countable subset of X. Prove that $X \setminus Y$ is uncountable.

EXERCISE 4.6 (challenge - extra credit). Let us get comfortable with weird bijections.

- Establish a bijection between the set $(-1,1)$ and all of $\mathbb{R}$.
- Establish a bijection between $\mathbb{R} \cup \{\star\}$ and $\mathbb{R}$.
- Establish a bijection between the set $[-1,1]$ and $\mathbb{R}$.

UNIT 5

Combinatorics and Newton's theorem

1. Key Ideas

This week we are going to explore Newton's binomial expansion theorem. This is a very useful tool in analysis, but it also offers us an opportunity to explore an interesting connection with the combinatorics of finite sets. Newton's theorem is a formula for the expansion of an expression of the form $(x+y)^n$, for any natural number n. Once one has the formula, then a proof by induction can show that the formula is true. However, the connection to combinatorics actually explains why the formula holds. So, as usual, we will "overproof" our statement.

Let us start with some definitions.

DEFINITION 5.1. For any natural number n, we define $n!$ (read n **factorial**) to be the product of all natural numbers from 1 to n (including n).

For example, $1! = 1, 2! = 2, 3! = 6, 4! = 24, 5! = 120, \ldots$
There are good reasons to define $0! = 1$, even though we will not get into them right now.

DEFINITION 5.2. For any two natural numbers n, k, with $k \le n$, we define $\binom{n}{k}$ (read n **choose** k) to be:

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

So for example $\binom{4}{2} = 6$, $\binom{4}{1} = \binom{4}{3} = 4$.

Note that, since we defined $0! = 1$, we have $\binom{n}{0} = \binom{n}{n} = 1$.

With these two definitions in place, we are ready to state Newton's binomial expansion theorem.

THEOREM 5.1. *For any natural number n, we have:*

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}.$$

Let us see this theorem in action. If you want to expand $(x+y)^6$, you can immediately write:

$$(x+y)^6 = x^6 + 6x^5y + 15x^4y^2 + 20x^3y^3 + 15x^2y^4 + 6xy^5 + y^6,$$

because, in this case, we have: $\binom{6}{0} = \binom{6}{6} = 1$, $\binom{6}{1} = \binom{6}{5} = 6$, $\binom{6}{2} = \binom{6}{4} = 15$, $\binom{6}{3} = 20$.

Let us take a side-step for a second, and observe that the number $n!$ counts the number of ways you can order the elements of a set with n elements.

THEOREM 5.2. *Let $X = \{x_1, \dots, x_n\}$ be a set with n elements. There are $n!$ distinct ways of ordering the elements of X.*

PROOF. We have n choices for what element to put in the first position. Once we have chosen an element for first position, we have $n - 1$ choices for what element to put in second position.
Once we have chosen the second element, we have $n - 2$ choices for what element to put in third position.
This pattern continues until we have only one element left, which must be put in the last position.
All together we have $n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot 2 \cdot 1$ choices, which is precisely the definition of $n!$. □

Theorem 5.2 has an interpretation in terms of bijections.

THEOREM 5.3. *Let $X = \{x_1, \dots, x_n\}$ be a set with n elements. The cardinality of the set of bijections from X to itself is $n!$.*

PROOF. We show that an ordering of X is equivalent to a bijection of X with itself; then the statement follows from Theorem 5.2.
Consider an ordering of the elements of X:

$$X = \{x_{r_1}, x_{r_2}, \dots, x_{r_n}\}.$$

Then one may associate to such ordering the bijection:

$$B = \{(x_i, x_{r_i})\}_{x_i \in X}.$$

Vice-versa, given a bijection from X to itself, you may create an ordering of X by declaring the i-th element of the ordering to be the element which via the bijection is matched with x_i. □

2. Groupwork

Let give the number $\binom{n}{k}$ a combinatorial meaning in a way similar to Theorems 5.2, 5.3.

PROBLEM 5.1. If X is a set with n elements, show that $\binom{n}{k}$ is the number of distinct k-element subsets of X.

QUESTION 5.1. After solving Problem 5.1, answer the following:

(1) Why is $\binom{n}{k} = \binom{n}{n-k}$? (can you use the combinatorial interpretation given in Problem 5.1 together with the notion of bijection to prove this?)
(2) Are the definitions $\binom{n}{0} = \binom{n}{n} = 1$ consistent with Problem 5.1?

PROBLEM 5.2. Show that the following identity holds:

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

Show it in two different ways:

(1) Use the algebraic definition 5.2 and do some algebra.
(2) Try to subdivide the set of *subsets of X with k elements* into two subsets, one of cardinality $\binom{n-1}{k}$, and the other of cardinality $\binom{n-1}{k-1}$.

Now let us tackle the proof of Newton's binomial theorem. To warm up, let us do the following exercise.

PROBLEM 5.3. Write down $(x+y)(x+y)(x+y)$ in three different colors, one for each binomial. Now expand the above expression keeping track of the colors. For example, instead of writing x^3, do write xxx. Subdivide the terms into four groups, corresponding to the four monomials x^3, x^2y, xy^2, y^3 that you obtain when you forget the colors. Can you relate the number of elements in each group with a number that has to do with the set of colors $X = \{black, blue, red\}$?

Here comes the main problem in this worksheet, so make sure you spend a fair amount of time on this and understand it very well.

PROBLEM 5.4. Generalize the idea from the previous problem to give a proof of Theorem 5.1. In particular use n-distinct colors to color the binomial $(x+y)$, and establish a bijection between the terms of the colored expansion of the binomial and subsets of the set of colors.

Let us conclude this groupwork with some fun exploration.

PROBLEM 5.5 (Pascal's triangle). Pascal's triangle is an infinite triangle of number that contains all numbers of the form $\binom{n}{k}$, as shown in the following figure:

$$\binom{0}{0}$$

$$\binom{1}{0} \quad \binom{1}{1}$$

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

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

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

$$\binom{5}{0} \quad \binom{5}{1} \quad \binom{5}{2} \quad \binom{5}{3} \quad \binom{5}{4} \quad \binom{5}{5}$$

$$\binom{6}{0} \quad \binom{6}{1} \quad \binom{6}{2} \quad \binom{6}{3} \quad \binom{6}{4} \quad \binom{6}{5} \quad \binom{6}{6}$$

Plugging in the numbers, you get

$$
\begin{array}{ccccccccccccc}
 & & & & & & 1 & & & & & & \\
 & & & & & 1 & & 1 & & & & & \\
 & & & & 1 & & 2 & & 1 & & & & \\
 & & & 1 & & 3 & & 3 & & 1 & & & \\
 & & 1 & & 4 & & 6 & & 4 & & 1 & & \\
 & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & \\
1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1
\end{array}
$$

Spot as many interesting numerical patterns as you can in Pascal's triangle. Then, try to explain them!

3. Homework

Each regular homework question will be graded on a 3 *points scale. In order to get the full* 3 *points, the solution must be* **correct, complete, well-organized and tidily written-up**. *Extra-credit homework questions are worth* 2 *points each.*

EXERCISE 5.1. Give a proof by induction of Newton's Theorem (Theorem 5.1).

EXERCISE 5.2. Use Newton's theorem to show the following facts:

- The sum of the numbers in each row of Pascal's triangle is a power of 2.
- The alternating sum of the numbers in each row of Pascal's triangle is equal to 0.

EXERCISE 5.3. A *hockeystick* in Pascal's triangle is obtained by starting on the side of the triangle (at one of the 1's), moving diagonally down for as long as you want, and then making one ninety degree downward turn and stopping. We call all the numbers that you traverse in the first direction the *rod* of the hockeystick, and the last number the *tip*. For example, if we start at $\binom{3}{0} = 1$, we could do $\{1, 4, 10, 15\}$. In this case the rod is $\{1, 4, 10\}$ and the tip is 15. Show that for any hockey stick the sum of the numbers in the rod equals the tip.

EXERCISE 5.4. Given a natural number n, and three numbers k_1, k_2, k_3 such that $k_1 + k_2 + k_3 = n$, what is the number of ways you can subdivide X into three disjoint subsets labeled U_1, U_2, U_3, where the cardinality of U_1 is k_1, the cardinality of U_2 is k_2 and the cardinality of U_3 is k_3? Prove your formula.

EXERCISE 5.5 (challenge - extra credit). Use the previous exercise to develop a formula for the expansion of a trinomial $(x + y + z)^n$. Prove that the formula holds in a similar way to what you did in Problem 5.4. Can you conjecture a general statement for Newton's theorem, in the case that you have m-distinct variables? (i.e. a combinatorial description for the coefficients of the expansion of $(x_1 + x_2 + \ldots + x_m)^n$?

EXERCISE 5.6 (challenge - extra credit). Observe that the first five rows of Pascal's triangle correspond to powers of 11: $11^0 = 1, 11^1 = 11, 11^2 = 121, 11^3 = 1331, 11^4 = 14641$. Can you explain why? Why doesn't the pattern continue after the fifth row?

UNIT 6

Functions: part I

1. Key Ideas

The concept of function is central to mathematics. Functions make different mathematical objects "communicate" with one another. In order to have a function, you need the following ingredients:

- a set X called the **domain** (or **input set**, or **starting set**), and
- a set Y called **codomain** (or **output set**, or **arrival set**).

We then talk about a function **from** X **to** Y.

DEFINITION 6.1. A function $f : X \to Y$ is *a prescription* that assigns, to each element x of X, one element of y that we call the **image** of x and denote $f(x)$.

When sets are given their Venn representation, functions can be described by drawing arrows between elements of the two sets, as shown in Figure 6.1. A few important observations:

(1) You can think of f as some machine with an input hole and an output hole. You can put elements of x into the input hole, and the machine will spit out elements of y. You can put in any elements of X, and the machine will always work to give you elements of Y. If you put the same element of X in the machine a bunch of times, the machine will always produce the same output.
(2) It is possible that the machine will produce the same output for different inputs.
(3) It is possible that not all elements of Y actually appear as outputs for some inputs.

EXAMPLE 6.1. Let $X = \{1, 2, 3\}$ and $Y = \{a, b, c, d\}$. A couple examples of functions from X to Y:

- $f_1 : X \to Y$ defined by $f_1(1) = a, f_1(2) = b, f_1(3) = c$;
- $f_2 : X \to Y$ defined by $f_2(1) = c, f_2(2) = c, f_2(3) = b$.

And a couple examples of things that are NOT functions from X to Y:

- $f : X \to Y$ defined by $f(1) = a, f(2) = b$;
- $f : X \to Y$ defined by $f(1) = c, f(1) = b, f(2) = c, f(3) = b$.

As we did for bijections, we can give a slightly more formal definition of function, as a subset of the cartesian product $X \times Y$.

DEFINITION 6.2. A **function** $f : X \to Y$ is equivalent information to a subset $F \subseteq X \times Y$ with the property that for any element $x \in X$, there is a unique element of F whose first term is x. We also call F the **graph** of f.

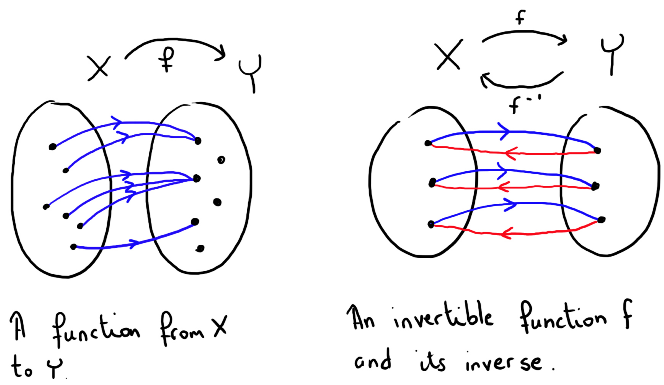


FIGURE 6.1. An example of a function, and of an invertible function. Tails of arrows represent input elements and heads of arrows the corresponding outputs.

For every element of F, the first term is the input, and the second term the corresponding output.

EXAMPLE 6.2. For the two functions f_1 and f_2 in Example 6.1, we have:

- $F_1 = \{(1,a),(2,b),(3,c)\}$;
- $F_2 = \{(1,c),(2,c),(3,b)\}$.

DEFINITION 6.3. Given any set X, there is always a special function from X to X, called the **identity function**, which we will denote Id_X. This function is defined by sending each element $x \in X$ to itself: $Id_X(x) = x$.

One important feature of functions is that they can be composed. In the analogy of the machines, suppose that you have a machine f whose input hole accepts elements of X and produces outputs in a set Y, and then another machine g that accepts inputs belonging to Y and produces outputs in some set Z. Then you can connect the output hole of the first machine with the input hole of the second, to create a new machine that "eats" elements of X and produces elements of Z.

DEFINITION 6.4. Given $f : X \to Y$ and $g : Y \to Z$, we define their **composition** $g \circ f : X \to Z$ to be the function

$$g \circ f(x) = g(f(x)).$$

EXAMPLE 6.3. Let $X = \{1,2,3\}$, $Y = \{a,b,c,d\}$, $Z = \{\star,\bullet\}$. Let

- $f : X \to Y$ defined by $f(1) = a, f(2) = b, f(3) = b$;
- $g : Y \to Z$ defined by $g(a) = \star, g(b) = \star, g(c) = \bullet, g(d) = \star$.

Then the function $g \circ f : X \to Z$ is defined to be:

$$g \circ f(1) = \star, g \circ f(2) = \star, g \circ f(3) = \star.$$

QUESTION 6.1. What happens when you compose a function with the identity function? More precisely, let $f : X \to Y$ be a function. What are the functions $f \circ Id_X$ and $Id_Y \circ f$?

Some "machines" are special in the sense that you can run them also in the opposite direction: feed in an element $y \in Y$, turn the crank backwards, and produce the element of X that y was the output of. Such special machines are called **invertible functions**.

DEFINITION 6.5. A function $f : X \to Y$ is called **invertible** if there exists a function $g : Y \to X$ such that:

$$g \circ f = Id_X, \qquad f \circ g = Id_Y.$$

The function g is called the **inverse** of f and denoted f^{-1}.

EXAMPLE 6.4. Let $X = \{1,2,3\}$ and $Y = \{a,b,c\}$. If $f : X \to Y$ is given by

$$f(1) = b, \quad f(2) = a, \quad f(3) = c,$$

then $g : Y \to X$ defined by

$$g(a) = 2, \quad g(b) = 1, \quad g(c) = 3,$$

is the inverse of f.

THEOREM 6.1. *If a function $f : X \to Y$ is invertible, then its inverse function is unique.*

PROOF. To prove this theorem, we assume that we have two inverse functions g_1 and g_2, and we show that it must be the case that $g_1 = g_2$.
The important thing to be aware of is what it means to be an inverse: it means satisfying Definition 6.5. So when we say that g_1 is an invers of f, we mean that

$$g_1 \circ f = Id_X, \qquad f \circ g_1 = Id_Y. \tag{4}$$

And when we say that g_2 is an inverse of f, we mean that

$$g_2 \circ f = Id_X, \qquad f \circ g_2 = Id_Y. \tag{5}$$

Consider the composition:

$$g_2 \circ f \circ g_1.$$

Since composition of functions is associative, we can place parenthesis however we want and we don't change the result.
This is one possible way:

$$g_2 \circ f \circ g_1 = (g_2 \circ f) \circ g_1 = Id_X \circ g_1 = g_1,$$

where for the second equality we have used (4). But by placing parentheses differently we obtain:

$$g_2 \circ f \circ g_1 = g_2 \circ (f \circ g_1) = g_1 \circ Id_Y = g_2,$$

where for the second equality we have used (5). By comparing these two computations we have shown that $g_1 = g_2$, which is what we wanted. □

2. Groupwork

In this groupwork we define and become familiar with the concept of inverse image. It is often the case that inverse image and inverse functions are confused. Partly, because the standard notation for the two concepts is the same. So please pay close attention!

DEFINITION 6.6. Given a function $f : X \to Y$ and a subset $U \subseteq Y$, the **inverse image** of U, denoted $f^{-1}(U)$, is the subset of X consisting of **all** elements of X whose output via f belongs to U. In symbols:

$$f^{-1}(U) = \{x \in X \text{ such that } f(x) \in U\}.$$

Here are the things to pay attention to: first of all, in order to talk about the inverse image, you must also specify a subset of Y. Secondly, the inverse image is a subset of X: it may be empty, it may be a single element, or it may contain multiple elements. The reason that it is acceptable to use the same notation for inverse image and inverse function, is that, when f is an invertible function and $U = \{y\}$ is a singleton, then the two concepts coincide, in the sense that the inverse image of the one element subset U is the same as the output of the inverse function for the element y:

$$f^{-1}(\{y\}) = \{f^{-1}(y)\}.$$

PROBLEM 6.1. Let $X = 1, 2, 3, 4, 5$, $Y = \{a, b, c\}$ and $f : X \to Y$ be defined by:

$$f(1) = b, \quad f(2) = b, \quad f(3) = c, \quad f(4) = c, \quad f(5) = b.$$

For every subset $U \in \mathcal{P}(Y)$, write down $f^{-1}(U)$.

PROBLEM 6.2. Given any function $f : X \to Y$, what are $f^{-1}(\phi)$, $f^{-1}(Y)$?

PROBLEM 6.3. Give an example of a function $f : X \to Y$ and a subset $\phi \neq U \subseteq Y$ such that $f^{-1}(U) = \phi$.
Give an example of a function $f : X \to Y$ and a proper subset $\mathcal{U} \subset Y$ such that $f^{-1}(U) = X$.

PROBLEM 6.4. Consider a function $f : X \to Y$. Give a characterization of the subsets of Y whose inverse image is the empty set, and of the subsets of Y whose inverse image is all of X. This consists of a statement of the form: $f^{-1}(U) = \phi$ (respectively X) if and only if...

PROBLEM 6.5. Prove that if $f : X \to Y$, $g : Y \to Z$, and $U \subseteq Z$:

$$(g \circ f)^{-1}(U) = f^{-1}(g^{-1}(U)).$$

3. Homework

Each regular homework question will be graded on a 3 points scale. In order to get the full 3 points, the solution must be **correct, complete, well-organized and tidily written-up**. *Extra-credit homework questions are worth 2 points each.*

Exercise 6.1. Consider the following prescriptions that assign to a human, another human. Decide which of them is a function from the set of humans to the set of humans.

birthmom: assigns to each human their birth mother;
son: assigns to each human their son;
president: assigns to each human the president of the country they are a citizen of;
birthpresident: assigns to each human the president of the country they were born in, at the time they were born.

Exercise 6.2. Consider the function "birthmom" defined in the previous exercise. What is the function birthmom∘birthmom?
Now define a function "biodad: Humans → Humans", which assigns to each human their biological father. Are the compositions birthmom∘biodad and biodad∘birthmom the same or different functions? Conclude whether the operation of composition of functions is commutative or not.

Exercise 6.3. Consider the following functions from $\mathbb{Z}$ to $\mathbb{Z}$. For each of them decide if they are invertible. If they are, write down the inverse function.

- $f(n) = n + 3$;
- $g(n) = 2n$;
- $h(n) = n^2$;
- $l(n) = \begin{cases} n & \text{if } n \text{ is even;} \\ n+2 & \text{if } n \text{ is odd.} \end{cases}$

Exercise 6.4. In each of the cases below, compute the inverse image $f^{-1}(U)$. All functions are from $\mathbb{R}$ to $\mathbb{R}$

- $f(x) = 3x, U = (-1, 2]$;
- $f(x) = x^2, U = (-1, 2]$;
- $f(x) = e^x, U = (-1, 2]$;
- $f(x) = sin(x), U = (0, 1/2]$;
- $f(x) = tan(x), U = (-1, 1)$.

Exercise 6.5 (challenge - extra credit). Use Definition 6.2 to show that a function is invertible if and only if it is a bijection.

Exercise 6.6 (challenge - extra credit). Given $f : X \to Y$, show that the procedure of inverse image defines a function

$$f^{-1} : \mathcal{P}(Y) \to \mathcal{P}(X).$$

Prove that f is a bijection between X and Y if and only if f^{-1} is a bijection between $\mathcal{P}(Y)$ and $\mathcal{P}(X)$.

UNIT 7

Functions: part 2

1. Key Ideas

We now introduce some attributes of functions that are extremely important in mathematics.

Definition 7.1. A function $f : X \to Y$ is called **injective** (or **1 : 1**) if f maps distinct inputs to distinct outputs.

In other words, a function is injective if no two different inputs are sent to the same output.

There are two standard ways to prove that a function is injective:

(1) One can show that $x_1 \neq x_2 \implies f(x_1) \neq f(x_2)$,
(2) or that $f(x_1) = f(x_2) \implies x_1 = x_2$.

Example 7.1. Let $X = \{1, 2, 3\}$, and $Y = \{a, b, c, d\}$. Then $f : X \to Y$ defined by

$$f(1) = b, \quad f(2) = c, \quad f(3) = a$$

is injective. But $g : X \to Y$ defined by

$$g(1) = b, \quad g(2) = b, \quad g(3) = c$$

is NOT injective.

The intuition is that injective functions use elements of Y as identifications for elements of X. For example, think of the function that assigns to each CSU student, their CSU-id number. The use of such a function is that numbers are more easily entered into a database, for example. Not every number corresponds to a student (I believe CSU-id's are 9 digits long or so,

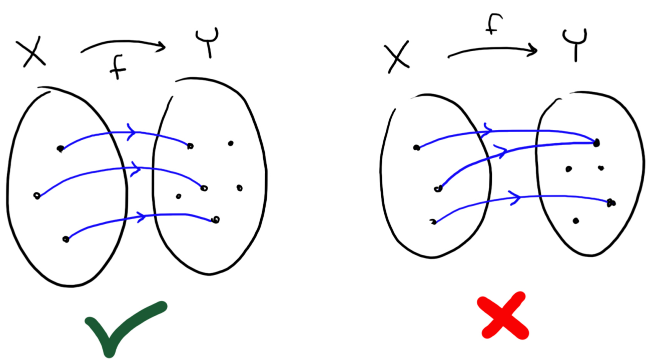


Figure 7.1. An example of an injective, and a not injective functions.

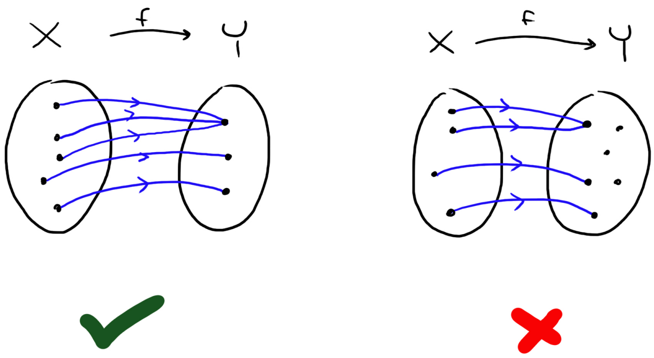


FIGURE 7.2. An example of an surjective, and a not surjective functions.

and we definitely don't have a billion students at CSU), but it is essential that no two students have the same CSU-id. This is the prototypical example of an injective function!

DEFINITION 7.2. A function $f : X \to Y$ is called **surjective** (or **onto**) if every element of Y is the output via f for some input element of X.

In other words, a function is surjective if for every $y \in Y$, there exists $x \in X$ such that $f(x) = y$.

EXAMPLE 7.2. Let $X = \{1, 2, 3\}$, and $Y = \{a, b\}$. Then $f : X \to Y$ defined by

$$f(1) = b, \quad f(2) = a, \quad f(3) = a$$

is surjective. But $g : X \to Y$ defined by

$$g(1) = b, \quad g(2) = b, \quad g(3) = b$$

is NOT surjective.

Surjective functions come into play when you only want to remember certain information about elements of X. For example, you might need to perform a task that depends only on the nationality of a person (say decide the color of their passport). Then, instead of working with the set of all living humans (which has about 6 billion elements), you may want to work with the set of nations (which has about 200 elements). Then the function that assigns to each person their nationality (assume that there is no double citizenship, just for the sake of this actually being a function) is a surjective function. It is surjective because there is no nation which has no inhabitant.

DEFINITION 7.3. Given a function $f : X \to Y$, the **image** (or **range**) of f is the set $f(X) \subseteq Y$ of all elements of Y that are outputs for some element of X,i.e.

$$Im(f) = f(X) = \{y \in Y \text{ such that there exists some } x \in X \text{ with } f(x) = y\}.$$

QUESTION 7.1. What does f being surjective mean in terms of the image of f?

DEFINITION 7.4. A function $f : X \to Y$ is called **bijective** if it is both injective and surjective.

Note that bijective functions are precisely what we have called bijections, and, after Exercise 6.5, they are also the invertible functions.

2. Groupwork

PROBLEM 7.1. Give an example of a function which is both injective and surjective, of one which is injective but not surjective, of one which is surjective but not injective and of one which is neither injective nor surjective.

PROBLEM 7.2.

(1) Suppose X and Y are finite sets, and $f : X \to Y$ is an injective function. What can you say about the relationship between $|X|$ and $|Y|$?
(2) Suppose X and Y are finite sets, and $f : X \to Y$ is a surjective function. What can you say about the relationship between $|X|$ and $|Y|$?
(3) Let $f : X \to Y$ be an injective function. Prove that if Y is countable, then X is either finite or countable.
(4) Let $f : X \to Y$ be a surjective function. Prove that if X is countable, then Y is either finite or countable.

PROBLEM 7.3.

(1) Let $f : X \to Y$ be an injective function. What can you say about the inverse images of subsets of Y? In particular, if $U = \{y\}$ is a singleton, what can $f^{-1}(U)$ be?
(2) Let $f : X \to Y$ be a surjective function. What can you say about the inverse images of subsets of Y? In particular, if $U = \{y\}$ is a singleton, what can $f^{-1}(U)$ be?

PROBLEM 7.4. Prove that the composition of two injective functions is injective and the composition of two surjective functions is surjective.

PROBLEM 7.5. Given a subset $U \subseteq Y$, there is a natural function $i_U : U \to Y$ defined by $i_U(u) = u$. Convince yourself that this function is injective. Viceversa, any injective function $f : X \to Y$ realizes a bijection between X and the image of f. Make sure this makes sense to you. This means that the notion of injective functions with output set Y is equivalent to the notion of subsets of the set Y. Spend a little time to understand this statement, and discuss it with the other members of the group.

PROBLEM 7.6. Let $f : X \to Y$ be a surjective function, $|Y| = n$ and for every $y \in Y$ suppose you have $|f^{-1}(\{y\})| = m$. Then what is $|X|$?

3. Homework

Each regular homework question will be graded on a 3 points scale. In order to get the full 3 points, the solution must be **correct, complete, well-organized and tidily written-up**. *Extra-credit homework questions are worth 2 points each.*

EXERCISE 7.1. For each of the following situations, provide an example if it is possible, or explain why it is not possible.

(1) $f : X \to Y$ not injective, $g : Y \to Z$ a function and $g \circ f$ injective;
(2) $f : X \to Y$ a function, $g : Y \to Z$ not injective, and $g \circ f$ injective;
(3) $f : X \to Y$ not surjective, $g : Y \to Z$ a function and $g \circ f$ surjective;
(4) $f : X \to Y$ a function, $g : Y \to Z$ not surjective, and $g \circ f$ surjective

EXERCISE 7.2. Use Problem 7.6 to count (yet again) the number of handshakes among n persons. Consider the following sets:

- X = the set of ordered pairs of distinct persons in the group.
- Y = the set of handshakes among those persons.

Define a natural, surjective function $f : X \to Y$. What is the number of elements of X? For each element $y \in Y$, what is $|f^{-1}(y)|$? Use the conclusion of Problem 7.6 to deduce $|Y|$.

EXERCISE 7.3. Given $f : X \to Y$, prove the following:

(1) If $y_1 \neq y_2$, then $f^{-1}(\{y_1\}) \cap f^{-1}(\{y_2\}) = \phi$;
(2)
$$\bigcup_{y \in Y} f^{-1}(\{y\}) = X$$

EXERCISE 7.4. Use the previous exercise to interpret the equality

$$2^n = \sum_{k=0}^{n} \binom{n}{k} \tag{6}$$

as follows. Let X be a set with n elements, and consider its power set $\mathcal{P}(X)$. Define a surjective function $f : \mathcal{P}(X) \to \{0, 1, 2, \ldots, n\}$ by $f(U) = |U|$. What is $|\mathcal{P}(X)|$? For any element $i \in \{0, 1, 2, \ldots, n\}$, what is $|f^{-1}(i)|$? Now use Exercise 7.3 to prove (6).

EXERCISE 7.5 (challenge - extra credit). Given the cartesian product $X \times Y$, we have two natural functions, called projections, defined as follows:

- $p_1 : X \times Y \to X$ is defined by $p_1((x, y)) = x$.
- $p_2 : X \times Y \to Y$ is defined by $p_2((x, y)) = y$.

Show that if X and Y are non empty, then the two projections are surjective.

EXERCISE 7.6 (challenge - extra credit). Use the concept of **graph of a function** (see Definition 6.2) to show that any function $f : X \to Y$ may be written as the composition $s \circ i$, where i is an injective function and s is a surjective function.

UNIT 8

Equivalence relations and quotient sets

1. Key Ideas

The notion of a quotient set is one of the really confusing things that students struggle with in classes like M366, M369 and M317. So try to really pay attention to what happens in this unit, as it will be coming back over and over in your future classes.

Let us start by saying something that should feel pretty natural:

- **(i1):** Given a subset $X \subset Y$, we can think of it as an injective function $i_X : X \to Y$.
- **(i2):** Viceversa, given any injective function $f : X \to Y$, f realizes a bijection between X and the subset $f(X) \subseteq Y$.

These two statements together mean that the concept of injective function contains the same amount of information as the concept of subset, in the sense that there is a procedure that, given an injective function, produces a subset, and an inverse procedure that, given a subset, produces an injective function.

Now we would like to do the same for the concept of surjective function. Surprisingly, things get quite a bit more complicated and confusing. Let us start by writing right away the analogous statements to (i1) and (i2), and then we will spend a fair amount of time making sense of these statements.

- **(s1):** Given a **partition** of a set X, or alternatively an **equivalence relation** on X with quotient set $X/\sim$, we can get a surjective function $\pi : X \to X/\sim$.
- **(s2):** Viceversa, given any surjective function $f : X \to Y$, f induces an equivalence relation on X and a natural bijection between $X/\sim$ and Y.

These two procedures are inverse of each other, showing that the notion of surjective function $f : X \to Y$ is equivalent to the notion of equivalence relation on (or partition of) X.

In order to make sense of (s1) and (s2) we must introduce a bunch of definitions.

Definition 8.1. A **partition** of a set X is a collection of disjoint (non-empty) subsets of X whose union is all of X:

$$X = \bigcup_{\alpha \in X/\sim} U_\alpha \quad \textit{and} \quad U_{\alpha_1} \cap U_{\alpha_2} = \phi \text{ (when } \alpha_1 \neq \alpha_2).$$

An example of this concept is illustrated in Figure 8.1. The indexing set for the subsets of the partition is denoted $X/\sim$.

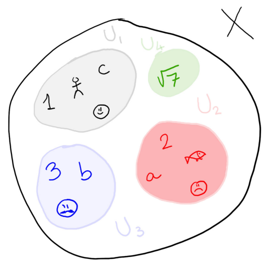


FIGURE 8.1. A partition of the set X.

The notation $X/\sim$, for now, seems a little bizarre; the reason is that it comes from the concept of equivalence relation, which is equivalent (no pun intended) to the notion of partition.

DEFINITION 8.2. Given a set X, an **equivalence relation** $\sim$ on X is a procedure that decides when two elements x_1 and x_2 are equivalent (in which case we write $x_1 \sim x_2$), which satisfies the following three requirements:

r: for every x, $x \sim x$;
s: if $x_1 \sim x_2$, then $x_2 \sim x_1$;
t: if $x_1 \sim x_2$ and $x_2 \sim x_3$, then $x_1 \sim x_3$.

See Figure 8.2 for an example.

The first way to define an equivalence relation is more slick, and ultimately, more useful. The second one however explains better why we use the name.

DEFINITION 8.3. Given an equivalence relation $\sim$ on a set X, a non-empty subset consisting of elements that are all equivalent to each other, and maximal with respect to this property, is called an **equivalence class** for $\sim$. The set of equivalence classes is called the **quotient set** and denoted $X/\sim$.

If $x \in X$, then we denote the equivalence class of x by $[x]$.

So the quotient set is a set whose elements are subsets of the set X. There are two very confusing things going on here:

(1) sometimes we want to think of an equivalence class as a subset of X, sometimes we want to think of it as an element of the quotient set. Unfortunately, the notation $[x]$ leaves the distinction to the context of the statement.
(2) the symbol $[x]$ is just one possible name for the equivalence class of x. For any other element $y \in [x]$, we have $[x] = [y]$. Therefore one equivalence class has as many different names as it has elements that it contains.

Finally we can define the projection function.

DEFINITION 8.4. The **projection function** $\pi : X \to X/\sim$ is defined by

$$\pi(x) = [x].$$

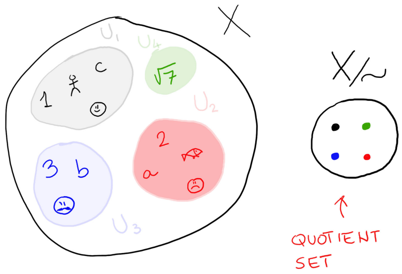


FIGURE 8.2. An equivalence relation on X may be given by the rule $x \sim y$ *if they have the same color*. The resulting equivalence classes and the quotient set.

In other words, π assigns to each element of X the equivalence class it belongs to.

Now that we have all relevant definitions in place, let us get to work to understand things!

2. Groupwork

We begin by understanding why the two definitions 8.1 and 8.2 are equivalent, by looking at a couple examples:

(1) Consider the set $X = \{a, b, c, d, e\}$, and the partition $X = \{a, c, e\} \cup \{b, d\}$. To such a partition, we can assign the following notion of equivalence of elements: $a \sim c \sim e$ and $b \sim d$.
(2) Let $X = \mathbb{Z}$ and let $\sim$ be defined as follows: $m \sim n$ when their difference is even. Then $\mathbb{Z} = E \cup O$, where E is the set of even numbers, and O is the set of odd numbers.

PROBLEM 8.1. Check that the prescription $\sim$ in 2. satisfies the three requirements **r,s,t**.

PROBLEM 8.2. By observing the two examples (1) and (2), come up with general procedures to go from a partition of a set to a notion of $\sim$ and viceversa.

PROBLEM 8.3. Let $X = \{1, 2, 3, 4, 5, 6, 7, 8\}$, $Y\{a, b, c\}$ and $f : X \to Y$ be defined by:

$$f(1) = a \;\; f(2) = a \;\; f(3) = c \;\; f(4) = b \;\; f(5) = a \;\; f(6) = b \;\; f(7) = c \;\; f(8) = a.$$

(1) Write down $f^{-1}(\{a\})$, $f^{-1}(\{b\})$, $f^{-1}(\{c\})$. Observe that they realize a partition of X.
(2) Define $x_1 \sim x_2$ when $f(x_1) = f(x_2)$. Check that $\sim$ satisfies **r,s,m**. What are the equivalence classes?
(3) We have seen now that we have constructed an equivalence relation on X. Write down the projection function $\pi : X \to X/\sim$.
(4) Consider the function $[f] : X/\sim \to Y$, defined by $[f]([x]) = f(x)$. Show that $[f]$ is a well defined function and that it is a bijection.

(5) Show that $f = [f] \circ \pi$.

PROBLEM 8.4. Abstract the example in the previous problem, and show that any surjective function $f : X \to Y$ gives rise to an equivalence relation on X and a natural bijection between $X/\sim$ and Y.

It looks like in the definition of $[f]$ nothing is really happening...we seem to only be rearranging brackets and then using f. This is a quite subtle point however. The next problem illustrates what could go wrong.

PROBLEM 8.5. Consider $X = \{1,2,3,4,5,6,7,8\}$ with the equivalence relation $X = \{1,2,4\} \cup \{3,7\} \cup \{6,8\} \cup \{5\}$. Then consider the function $F : X \to \{a,b\}$ that sends all even numbers to a and odd numbers to b. Does the prescription

$$[F]([x]) = F(x)$$

define a function $[F] : X/\sim \to \{a,b\}$?

PROBLEM 8.6. Let $\sim$ be an equivalence relation on a set X and $f : X \to Y$ a function. When is there a function $[f] : X/\sim \to Y$ such that $f = [f] \circ \pi$?

3. Homework

Each regular homework question will be graded on a 3 points scale. In order to get the full 3 points, the solution must be **correct, complete, well-organized and tidily written-up**. *Extra-credit homework questions are worth 2 points each.*

Exercise 8.1. Consider the set $X = \{a, b, c, d, e, g, h\}$ and the equivalence relation given by the partition $X = \{a, c, e\} \cup \{b, h\} \cup \{d, g\}$. Answer the following questions:

(1) Is $a \sim b$?
(2) How many elements does the quotient set have?
(3) Is $[a] = [e]$?
(4) Write down the projection function $\pi : X \to X/\!\sim$.
(5) Is $\pi(b) = [h]$?

Exercise 8.2. Consider the set $X = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ and the equivalence relation prescribed by $m \sim n$ when $m - n$ is a multiple of 4.

(1) How many elements does $X/\!\sim$ have?
(2) Write down the equivalence classes.
(3) Write down the projection function.

Exercise 8.3. Consider the equivalence relation on $\mathbb{Z}$ defined by the prescription that all positive numbers are equivalent, all negative numbers are equivalent, and 0 is only equivalent to itself. Let $f : \mathbb{Z} \to \{a, b\}$ be the function that maps all negative numbers to a and all non-negative numbers to b. Does there exist a function $F : X/\!\sim \to \{a, b\}$ such that $f = F \circ \pi$? If so, describe it.

Exercise 8.4. Consider the equivalence relation on $\mathbb{Z}$ defined by the prescription that all positive numbers are equivalent, all negative numbers are equivalent, and 0 is only equivalent to itself. Show that you can find $n_1 \sim n_2$ and $m_1 \sim m_2$ such that $n_1 + m_1 \not\sim n_2 + m_2$.

Exercise 8.5 (challenge - extra credit). Let $\sim$ be the equivalence relation on $\mathbb{Z}$ defined by $n \sim m$ when $n - m$ is a multiple of 5. Show that if $n_1 \sim n_2$ and $m_1 \sim m_2$, then $n_1 + m_1 \sim n_2 + m_2$.

Exercise 8.6 (challenge - extra credit). An equivalence relation on a set X may be described as a subset $R \subseteq X \times X$ that satisfies some properties corresponding to **r, s, t**. Assume X is a finite set and that you may describe the Cartesian product as an array of elements in the plane. Describe these three axioms in terms of the geometry of this representation of the cartesian product $X \times X$.

UNIT 9

Quotients and Operations

1. Key Ideas

In Unit 8 we discussed quotient sets and equivalence relations. Let us recap now the important points:

(1) An **equivalence relation** $\sim$ on a set X produces a partition of X into disjoint subsets, called **equivalence classes**.
(2) The set whose elements are the equivalence classes is called **quotient set** and denoted $X/\sim$.
(3) Given an element $x \in X$, the equivalence class that x belongs to is denoted $[x]$ (note that $[x] = [y]$ if and only if $x \sim y$).
(4) The **projection function** $\pi : X \to X/\sim$ assigns to each element x the equivalence class that x belongs to. ($\pi(x) = [x]$).
(5) A function $F : X \to Y$ determines a function $[F] : X/\sim \to Y$ via the prescription $[F]([x]) = f(x)$ if and only if the following condition holds:

$$x_1 \sim x_2 \implies f(x_1) = f(x_2).$$

The last point is of particular interest to us: in order for a function $F : X \to Y$ to naturally define a function $[F]$ from the quotient set, F must satisfy some special property, which we think of as a compatibility condition between F and the equivalence relation $\sim$. Now we want to boost this idea. Let us consider an equivalence relation $\sim$ on a set X and a function:

$$F : X \to X$$

We would like to define a function $[F] : X/\sim \to X/\sim$ by the prescription:

$$[F]([x]) = [F(x)]. \tag{7}$$

Let us see in an example when this works. Let

$$X = \{boy, cat, dog, bat, chloe, doe, tic, tac, toe\},$$

and $F : X \to X$ be defined by:

$$F(boy) = bat, \quad F(cat) = dog, \quad F(dog) = bat$$

$$F(bat) = boy, \quad F(chloe) = doe, \quad F(doe) = bat$$

$$F(tic) = tac, \quad F(tac) = toe, \quad F(toe) = toe,$$

Define an equivalence relation on X by saying that two elements of X are equivalent if they begin with the same letter. Then $X/\sim = \{b, c, d, t\}$ has four elements. Then the prescription given in equation (7) defines the function $[F] : X/\sim \to X/\sim$ as:

$$[F](b) = b, \quad F(c) = d, \quad F(d) = b, \quad F(t) = t.$$

Now define an equivalence relation on X by saying that two elements of X are equivalent if they end with the same letter. The quotient set $X/\sim = \{c, e, g, t, y\}$ has five elements. The prescription given in equation (7) does NOT define a function:

$$[chloe] = [doe] = e,$$

but

$$[F(chloe)] = [doe] = e \neq [F(doe)] = [bat] = t,$$

which means that $[F]$ is trying to sent the input e to two different outputs.

Definition 9.1. Given an equivalence relation $\sim$ on a set X, we say that a function $F : X \to X$ is **compatible with** $\sim$ if

$$x_1 \sim x_2 \implies F(x_1) \sim F(x_2).$$

If this is the case, the prescription in equation (7) defines a function $[F] : X/\sim \to X/\sim$; we also equivalently say that the function $[F]$ is **well-defined**.

Our goal this week is to generalize this idea even further: if X is a set of numbers and $\sim$ an equivalence relation, when can we define operations of addition and multiplication on the quotient set $X/\sim$? We will use this idea to give a precise and formal proof of the following rule, that probably a lot of you remember from elementary school.

Theorem 9.1. *An integer n is divisible by* 3 *if and only if the sum of its digits is divisible by* 3*.*

We introduce a couple definitions that will get us started for the groupwork.

Definition 9.2. A set X is a **number system** if there are two operations on the elements of X (addition and multiplication) that satisfy all the usual rules of arithmetics that the common addition and multiplication of ordinary numbers satisfy.

Question 9.1. Given an equivalence relation on a number system, when is the quotient set also a number system in a natural way?

Definition 9.3. An equivalence relation $\sim$ on a number system X is **compatible with addition** if

$$x_1 \sim x_2 \text{ and } y_1 \sim y_2 \implies x_1 + y_1 \sim x_2 + y_2.$$

An equivalence relation $\sim$ on a number system X is **compatible with multiplication** if

$$x_1 \sim x_2 \text{ and } y_1 \sim y_2 \implies x_1 \cdot y_1 \sim x_2 \cdot y_2.$$

When an equivalence relation on a number system is compatible with addition and multiplication, the quotient set $X/\sim$ becomes a number system with the following definitions. We write here $\boxplus$ and $\boxdot$ to define the operations on the quotient set, so as to not confuse them with the operations $+, \cdot$ on the original number system.

DEFINITION 9.4.
$$[x] \boxplus [y] := [x+y]$$
and
$$[x] \boxdot [y] := [x \cdot y]$$

In other words, we are using the operations of addition and multiplication on X to define operations on $X/\sim$. The dangerous thing is that elements of the quotient set have multiple names. So what may happen is that when two different names are used, the same operation on the same elements may produce two different outcomes.

2. Groupwork

PROBLEM 9.1. Let X be a set with exactly two elements denoted E, O. Let addition and multiplication be defined by the following tables:

$+$	E	O
E	E	O
O	O	E

$\cdot$	E	O
E	E	E
O	E	O

In order to prove that X is a number system we should check that such addition and multiplication are commutative, associative and they respect the distributive laws. One really should check that these properties hold in **all possible instances!** This is of course very boring, so let us instead check, for each of the properties, one particular instance.

commutativity: Check that $E+O=O+E$ and $E \cdot O = O \cdot E$.
associativity of +: Check that $(E+O)+O = E+(O+O)$.
associativity of ·: Check that $(E \cdot O) \cdot O = E \cdot (O \cdot O)$.
distibution laws: Check that $O \cdot (E+O) = (O \cdot E) + (O \cdot O)$.

PROBLEM 9.2. Give an example of an equivalence relation on a number system which is not compatible with addition, nor multiplication.

PROBLEM 9.3. Let $X = \mathbb{Z}$ be the integers, and $\sim$ the equivalence relation defined as follows:
$$m \sim n \iff m - n = 9k,$$
for some $k \in \mathbb{Z}$. Show that $\sim$ is compatible with addition and multiplication.

PROBLEM 9.4. Let $X = \mathbb{Z}$ and $\sim$ the equivalence relation that declares all negative integers to be equivalent, and all non-negative integers to be equivalent. Apply Definition 9.4 above to $[4] \boxplus [-3]$ and $[2] \boxplus [-3]$. What is the problem?

The condition of being compatible with addition and multiplication are precisely what is needed to make sure the problem we just witnessed does not happen. Finally, the number systems X and $X/\sim$ are very closely related, in this sense.

THEOREM 9.2. *Given any sequence of arithmetic operations on elements of X, one obtains the same result in the following two ways:*

(1) *Perform all operations in X, and at the end, put square brackets around the result.*

(2) *Perform all operations in $X/\sim$, i.e., put square brackets around all numbers and boxes around all operation signs.*

Formal proofs of all these statements are a bit tedious and complicated, so we will skip them for now (raincheck till your MATH 366 class). But let us see one example of the above theorem. Consider the equivalence relation from Problem 9.3, and denote by $\mathbb{Z}/9\mathbb{Z}$ the quotient set. Then the following computations give the same result:

$$[12 + 3 \cdot (11 + 7)] = [66] = [3]$$

or

$$[12] \boxplus [3] \boxdot ([11] \boxplus [7]) = [3] \boxplus [3] \boxdot ([2] \boxplus [7]) = [3] \boxplus [3] \boxdot [0] = [3].$$

Notice that in the second case, since we get to choose different names for the same element (i.e. we can replace $[12]$ with $[3]$, since $[12] = [3]$), we can choose to work with smaller numbers. This is the idea at the base of the test for divisibility by 3. Here are all the important steps.

Problem 9.5. Let $\mathbb{Z}/9\mathbb{Z}$ be defined as above.

(1) Prove that an integer n is divisible by 3 if an only if $[n] = [0], [3]$ or $[6]$.
(2) Prove that for any nonnegative integer n, $[10^n] = [1]$.
(3) Prove that for any integer n, $[n]$ is equal to the equivalence class of the sum of the digits of n.

3. Homework

Each regular homework question will be graded on a 3 points scale. In order to get the full 3 points, the solution must be **correct, complete, well-organized and tidily written-up**. *Extra-credit homework questions are worth 2 points each.*

Exercise 9.1.

(1) Give an example of an equivalence relation $\sim$ on a set X, and a function $F : X \to X$ which is compatible with $\sim$.
(2) Give an example of an equivalence relation $\sim$ on a set X, and a function $F : X \to X$ which is NOT compatible with $\sim$.

Exercise 9.2. Let $\sim$ be an equivalence relation on the set X. Decide if the following statements are true or false:

(1) If $F : X \to X$ is a constant function, then $[F]$ is a function.
(2) If $Id_X : X \to X$ denotes the identity function of X, then $[Id_X]$ is a function.
(3) If for every $x \in X$, $F(x) \sim x$, then $[F] = Id_{X/\sim}$ is well-defined.
(4) If for every $x \in X$, $F(x) \sim x$, then $[F] = Id_{X/\sim}$ is the identity function of the quotient set.
(5) If $[F] = Id_{X/\sim}$ is the identity function of the quotient set, then $F = Id_X$.

Exercise 9.3. We make some definitions about number systems:

zero: An element z of a number system is a **zero** if, for every element x in the number system, we have $x + z = z + x = x$, and $x \cdot z = z \cdot x = z$.
one: An element u of a number system is a **one** if, for every element x in the number system, we have $x \cdot u = u \cdot x = x$.
subtraction: The operation of **subtraction** is defined as follows: $x - y$ is defined to be an element w such that $x = y + w$.

Consider the number system $X = \{E, O\}$ from Problem 9.1.

(1) Show that E is a zero for X.
(2) Show that O is a one for X.
(3) What is $E - O$?

Exercise 9.4. Use the concepts from Problem 9.5 to outline a formal proof of Theorem 9.1. An outline means that for the details you can just refer to the work you have done in the problems. But you need to write a clear and logical sequence of steps that prove the thesis of the theorem.

Exercise 9.5 (challenge - extra credit). Devise a test to check divisibility by the number 11.

Exercise 9.6 (challenge - extra credit). For $n \in \mathbb{Z}$, denote by $\mathbb{Z}/n\mathbb{Z}$ the number system obtained as the quotient set for the equivalence relation on the integers:

$$x \sim y \iff x - y = nk.$$

The operation of division should be defined as follows: for $y \neq [0]$, x/y is the element w such that $x = y \cdot w$. Note however that the operation of division is not defined for all choices of n. Why? What goes wrong? What are the values of n for which it is defined?