Introduction to Logic and Set Theory-2013-2014

General Course Notes

December 2, 2013

These notes were prepared as an aid to the student. They are not guaranteed to be comprehensive of the material covered in the course. These notes were prepared using notes from the course taught by Uri Avraham, Assaf Hasson, and of course, Matti Rubin. Many of the elegant proofs and examples are from their courses, though I am responsible for the mistakes and the choices in presentation. I was once told that it is the job of the student to keep the professor honest and I believe this to be true. Be critical of what you read as it is meant to be understood and not memorized. This is one of the few courses that is almost entirely self-contained, so everything you need will be right in front of you. There is no algorithm or specific way to write a proof, so what you write should be an expression of your thought processes and logic. We call proofs "arguments" and you should be convincing the reader that what you write is correct. The more you see your proofs in this light, the more enjoyable this course will be.

1 Truth Tables

The goal of this section is to understand both mathematical conventions and the basics of mathematical reasoning. When we discuss formulas later in the course, we will change notation slightly to distinguish between general mathematical argument and formal statements. In this section, we are learning how to think and write like a mathematician and we drop some formalism until we familiarize ourselves with convention. For a given statement P, we will look at the possible truth values of P. Namely, either P is true or P is False.

P
T
F

If we are analyzing two statements P and Q at the same time, we need to consider all possible combinations of truth values.

P	Q
T	T
T	F
F	T
F	F

The first row is the case when both are true. The second row is the case where P is true and Q is false. The third row is the case when P is false and Q is true. The fourth row is the case when both P and Q are false. Using truth tables, we can "define" what certain symbols and words mean in mathematics.

In mathematics, the terms "and", "or", "not" have precise meaning and are often written as symbols instead of words. We will use these symbols to construct more complicated statements from ones that we have.

$\neg$ = "NOT":

$\neg P$ is true when P is false and $\neg P$ is false when P is true.

P	$\neg P$
T	F
F	T

$\wedge$ = "AND":

In order for the statement $P \wedge Q$ to be true, both$P$ and Q must be true. Otherwise, it is false.

P	Q	$P \wedge Q$
T	T	T
T	F	F
F	T	F
F	F	F

Now let's move one step further and write the truth table for $\neg(P \wedge Q)$ which will be true exactly when $(P \wedge Q)$ is false.

P	Q	$P \wedge Q$	$\neg(P \wedge Q)$
T	T	T	F
T	F	F	T
F	T	F	T
F	F	F	T

It is crucial to note the placement of the parenthesis and how this is crucial to the meaning of the sentence. Without the parenthesis, we read from left to right. So $\neg P \wedge Q$ is understood to mean $(\neg P) \wedge (Q)$. Notice the difference in the truth tables:

P	Q	$\neg P$	$\neg P \wedge Q$
T	T	F	F
T	F	F	F
F	T	T	T
F	F	T	F

$\underline{\vee = \text{"OR"}}$:

In mathematics, we use an inclusive "or". This means that for $P \vee Q$ to be true, we require that one of them be true but allow for the possibility that both are true. Therefore, the truth table is as follows:

P	Q	$P \vee Q$
T	T	T
T	F	T
F	T	T
F	F	F

$\Rightarrow$ = "IMPLIES"

We will now discuss what it means for a statement to imply another statement. The symbol we use for this is $\Rightarrow$ and $P \Rightarrow Q$ is often read "P implies Q". Logically, this statement is equivalent to $\neg P \vee Q$ and to understand the mathematical implication consider the following: If your mother tells you "If I go to the candy store, I'll buy you a candy", when has she told the truth? Well, there is really only one way that she has lied - namely if she goes to the candy store and doesn't buy you a candy. Notice that if she doesn't go to the candy store, it doesn't really matter what happens later since her promise was conditional on her going to the candy store.

P	Q	$P \Rightarrow Q$
T	T	T
T	F	F
F	T	T
F	F	T

$\Longleftrightarrow$ = "IF AND ONLY IF":

We now discuss what it means for two statements to be equivalent. The symbol we use for this is $\Longleftrightarrow$ and $P \Longleftrightarrow Q$ is often read "P if and only if Q". Logically, this statement is equivalent to $(P \Rightarrow Q) \wedge (Q \Rightarrow P)$ and means that the two statements have exactly the same truth values or truth tables. Let us look at the truth table for $P \Longleftrightarrow Q$.

P	Q	$P \Rightarrow Q$	$Q \Rightarrow P$	$P \Longleftrightarrow Q$
T	T	T	T	T
T	F	F	T	F
F	T	T	F	F
F	F	T	T	T

TAUTOLOGIES and CONTRADICTIONS:

In mathematics, a *tautology* is a statement that is always true regardless of the "circumstances". Here is a simple example:

P	$\neg P$	$P \vee \neg P$
T	F	T
F	T	T

The next example is very common in mathematics. We consider the implication $P \Rightarrow Q$ and the *contrapositive* $\neg Q \Rightarrow \neg P$. It turns out that these two are equivalent, i.e. the statement $(P \Rightarrow Q) \iff (\neg Q \Rightarrow \neg P)$ is a tautology.

P	Q	$\neg P$	$\neg Q$	$P \Rightarrow Q$	$\neg Q \Rightarrow \neg P$	$(P \Rightarrow Q) \iff (\neg Q \Rightarrow \neg P)$
T	T	F	F	T	T	T
T	F	F	T	F	F	T
F	T	T	F	T	T	T
F	F	T	T	T	T	T

In mathematics, a *contradiction* is the assertion of a statement and its negation, or equivalently, a statement that can never be true. A contradiction is equivalent to the negation of a tautology.

P	$\neg P$	$P \wedge \neg P$
T	F	F
F	T	F

EXERCISES

1. Write the truth table for $(P \vee Q) \Rightarrow R$. (Hint: There will be 8 possible truth combinations for P, Q, and R.)
2. Show that $\neg(P \vee Q)$ is equivalent to $\neg P \wedge \neg Q$.
3. Show that $\neg(P \wedge Q)$ is equivalent to $\neg P \vee \neg Q$.
4. Show that $P \wedge (Q \wedge R)$ is equivalent to $(P \wedge Q) \wedge R$.
5. Show that $P \Rightarrow (Q \Rightarrow R)$ is equivalent to $(P \wedge Q) \Rightarrow R$.
6. Show that $P \Rightarrow Q$ is not equivalent to the *converse* $Q \Rightarrow P$.

2 Sets

We will introduce the notion of a set without formally defining it. For our purposes, a set is a collection of objects or symbols. The objects in a set will be called elements of the set. Sets are usually described using "{}" and inside these curly brackets a list of the elements or a description of the elements of the set. If a is an element of a set A, we use the notation $a \in A$ and often say "a in A" instead of "a an element of A". The notation $a \notin A$ indicates that a is not an element of A and is often read "a is not in A".

NOTE: When we use letters as elements of sets we do NOT consider the letters as variables unless specified in the context. It is just the letter. For example, $\{a, b, c\}$ is the set containing three elements, the letters a, b, and c since we have not described them as variables in this context. When we describe the set, we will need to be clear what letters are variables and what letters are actual elements. You will see that this doesn't cause much confusion in reality.

Examples:

1. $\{a, b, c\}$. $b \in \{a, b, c\}$.
2. The natural numbers: $\mathbb{N} = \{0, 1, 2, 3, ...\}$. $5 \in \mathbb{N}$, $-1 \notin \mathbb{N}$.
3. The integers: $\mathbb{Z} = \{\cdots -2, -1, 0, 1, 2, \dots\}$
4. The nonnegative even integers: $S = \{a \in \mathbb{N} : \text{a is even }\}$. $0 \in S$, $5 \notin S$.

In the last example, the colon is often read "such that" and sometimes replaced with a straight line |.

**The set with no elements is called the empty set and is denoted by $\emptyset$.

Definition 1. Let A and B be sets. Then $A \subseteq B$ (or equivalently $B \supseteq A$) if $a \in A \Rightarrow a \in B$.

In this case, we say A is a *subset* of B or equivalently that A *is contained* in B. To prove that a given set A is contained in B, one needs to show that $x \in A \Rightarrow x \in B$. Since this implication is logically true if $x \notin A$ as we saw before, we begin with the assumption that $x \in A$ and proceed to show that $x \in B$. See examples below.

Example: Consider the nonnegative even integers: $S = \{a \in \mathbb{N} : \text{a is even }\}$. This is a subset of $\mathbb{N}$. Notice that by definition, $\mathbb{N}$ is a subset of $\mathbb{N}$ as well.

A set is completely determined by the elements and we define equality on sets as follows:

Definition 2. Let A and B be sets. Then $A = B$ if they contain exactly the same elements, that is $a \in A \iff a \in B$.

To prove that two sets A and B are equal, we need to show that for all $a \in A$ we have $a \in B$ and for all $a \in B$, we have $a \in A$.

Claim 3. *Let A and B be sets. Then $A = B$ if both of the following conditions are satisfied:*

- $A \subseteq B$.
- $B \subseteq A$.

The proof is left as an exercise.

We use the notation $A \subset B$ or $A \subsetneq B$ to mean that $A \subseteq B$ and $A \neq B$.

Set Operations

- *Union*: $A \cup B = \{x : x \in A \text{ or } x \in B\}$
- *Intersection*: $A \cap B = \{x : x \in A \text{ and } x \in B\}$
- *Set minus*: $A \setminus B = \{x \in A : x \notin B\}$.
- *Symmetric Difference*: $A \triangle B = \{x \in (A \cup B) : x \notin A \cap B\} = A \cup B \setminus A \cap B$.

  Given an ambient set U which we call the universe, we can discuss the complement of $A \subseteq U$.

- *Complement*: $A^c = \{x \in U : x \notin A\}$

  We must fix a universe to discuss complements and when we fix a universe, all sets we consider in the given context are considered as subsets of this universe.

We are now ready to look at some basic facts and get a feeling for how to write proofs. I want to emphasize here that there is NO trick to writing proofs. A proof must be logical and each line must follow from the previous line. Throughout these statements, A and B are sets.

1. Claim: $A \subseteq A \cup B$.

   proof:

   Suppose $a \in A$. Then $a \in A$ or $a \in B$. So $a \in A \cup B$.

2. Claim: $A \cap B \subseteq A$.

   proof:

   Suppose $a \in A \cap B$. Then, by definition of the intersection, $a \in A$ and $a \in B$. So $a \in A$.

   We now move onto De Morgan's Laws: Let A and B be subsets of some ambient universe U.

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

   proof: We divide this proof into two parts.

   - $(A \cup B)^c \subseteq A^c \cap B^c$ Suppose $a \in (A \cup B)^c$. Then $a \notin A \cup B$ by definition of the complement. By definition of the union, we know that a is not in the set given by $\{x : x \in A \text{ or } x \in B\}$. Notice the following: Let P be the statement "$x \in A$ or $x \in B$". The negation of this statement is equivalent to $x \notin A$ and $x \notin B$ as we proved in the exercises above. So $a \notin A$ and $a \notin B$. Therefore $a \in A^c$ and $a \in B^c$ by definition of the complement and thus, $a \in A^c \cap B^c$ by definition of the intersection.

- $A^c \cap B^c \subseteq (A \cup B)^c$

  Suppose $a \in A^c \cap B^c$. Then $a \in A^c$ and $a \in B^c$ by definition of the intersection. So $a \notin A$ and $a \notin B$ by definition of the complement. In the exercises, you proved that $\neg P \wedge \neg Q$ is equivalent to $\neg(P \vee Q)$. So we know that $a \notin A$ and $a \notin B$ is equivalent to $\neg(a \in A$ or $a \in B)$. So $a \notin A \cup B$ by the definition of the union and thus $a \in (A \cup B)^c$ by the definition of the complement.

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

   proof: This is left as an exercise.

Definition 4. Let A be a set. The *power set* of A, $\mathcal{P}(A)$, is the set of all subsets of A, i.e. $\mathcal{P}(A) = \{B : B \subseteq A\}$

EXAMPLE: Let $A = \{1, 2, 3\}$. Then

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

Let us now prove some basic facts:

Claim 5.
- *For any set A, $\emptyset \in \mathcal{P}(A)$. In particular, the powerset of any set cannot be empty.*
- *For any set A, $A \in \mathcal{P}(A)$.*

Proof.
- Our claim is that $\emptyset \subseteq A$ for any set A. To see this consider the definition of subset. We need to show that $x \in \emptyset \Rightarrow x \in A$. Since there is no $x \in \emptyset$, this implication is always true!
- For any set A, $a \in A \Rightarrow a \in A$. So $A \subseteq A$. This is a bit silly, but it is worth seeing once.

□

Russel's Paradox:

Throughout this course, we will be using a more intuitive notion of set to develop basic mathematical concepts. Pretty much any collection of objects

or symbols we can describe we are calling a set. This is often called "Naive Set Theory". At this point in our mathematical adventure, we will take a step back to realize that defining what is a set is not so simple. Consider the following "set":

$$X = \{x : x \notin x\}$$

This seems to be the set of all sets that don't contain themselves. However this leads to a contradiction:

$$X \in X \iff X \notin X$$

For those that take axiomatic set theory, you will learn about something called "bounded comprehension" or "restricted comprehension". Essentially, to avoid this paradox, we can only describe subsets of those collections we know to be sets. This is well beyond the scope of this course, but it is worth noting how essential and yet difficult this most basic of notions is.

EXERCISES

1. Let $A = \{a, a, b, c\}$ and $B = \{a, b, c\}$. Prove that $A = B$. Thus, the number of times an element is listed does not matter.

2. Let $A = \{a, b, c\}$ and $B = \{b, c, a\}$. Prove that $A = B$. Thus, the order in which the elements are listed does not matter.

3. Let $A = \{a\}$ and $B = \{\{a\}\}$. Prove that in general $A \neq B$. Thus, the presence of curly brackets is crucial.

4. Prove that if $A \subseteq \emptyset$ then $A = \emptyset$.

5. Prove Claim 3.

6. Let A be a subset of some universe U. Prove that $A \cap A^c = \emptyset$.

7. Let A and B be two subsets of some universe U. Prove that $(A \cap B)^c = A^c \cup B^c$.

8. Let $A = \{5, 6, 7\}$. What is $\mathcal{P}(A)$? Do not merely describe the set, but rather write down all the elements.

3 Ordered Pairs and Relations

An *ordered pair* (a, b) is a mathematical object comprised of two objects a and b belonging to sets A and B respectively. The order in which the element is written is crucial and we define equality on ordered pairs as follows:

$$(a, b) = (c, d) \iff a = c \wedge b = d$$

Definition 6. Let A and B be sets. Then, the *Cartesian product* of A and B, written as $A \times B$, is the set of ordered pairs

$$\{(a, b) : a \in A \text{ and } b \in B\}$$

We will now prove a few basic facts about the Cartesian product.

Proposition 7. *Let $A, B,$ and C be sets.*

- $A \times (B \cup C) = (A \times B) \cup (A \times C)$.
- $A \times (B \cap C) = (A \times B) \cap (A \times C)$.
- $A \times (B \setminus C) = (A \times B) \setminus (A \times C)$.

Proof. We will prove the first bullet and leave the second and third as exercises. Again, we split this proof into two parts.

1. $A \times (B \cup C) \subseteq (A \times B) \cup (A \times C)$:

   Let $(a, d) \in A \times (B \cup C)$. Then $a \in A$ and $d \in B \cup C$. By the definition of union, we know that $d \in B$ or $d \in C$.

   Case 1: $d \in B$. Then $(a, d) \in A \times B$ and we have that $(a, d) \in (A \times B) \cup (A \times C)$.

   Case 2: $d \in C$. Then $(a, d) \in A \times C$ and we have that $(a, d) \in (A \times B) \cup (A \times C)$.

2. $(A \times B) \cup (A \times C) \subseteq A \times (B \cup C)$:

   Let $(a, d) \in (A \times B) \cup (A \times C)$. Then $(a, d) \in (A \times B)$ or $(a, d) \in (A \times C)$.
   Case 1: $(a, d) \in (A \times B)$. Then $a \in A$ and $d \in B$ by definition of the Cartesian product. So $d \in B \cup C$ and $(a, d) \in A \times (B \cup C)$.

   Case 2: $(a, d) \in (A \times C)$. Then $a \in A$ and $d \in C$ by definition of the Cartesian product. So $d \in B \cup C$ and $(a, d) \in A \times (B \cup C)$.

□

NOTATION: At this point, it is worthwhile to introduce new notation. The symbols below will be used informally in arguments as a way of simplifying notation and making our proofs and definitions neater. At some point, we will introduce more formal logic and certain symbols will have a more precise meaning at that time.

- The universal quantifier: $\forall$ = "for all".
- The existential quantifier: $\exists$ = "there is" or "exists".

Let us take a moment to consider what these symbols mean. Let $P(x)$ be a statement about x and consider $\forall x P(x)$. If something is true of all x, then you cannot find a counterexample. The negation of this statement is that there is an x such that $\neg P(x)$ and intuitively that you have a counterexample. Therefore $\neg(\forall x P(x)) \iff \exists x(\neg P(x))$.

Let us now consider the statement $\exists x P(x)$. Then, this statement is true if there is an x such that $P(x)$ is true. So the negation of this statement is intuitively that for every possible x, $P(x)$ is not true. Therefore, $\neg(\exists x P(x)) \iff \forall x \neg P(x)$.

We often use the notation $\forall x \in A(P(x))$ where A is some set. The meaning of this is the obvious one, namely that $P(x)$ holds for $x \in A$. It is equivalent to $\forall x(x \in A \Rightarrow P(x))$.

We now move on to the definition of a relation.

Definition 8. A *relation* R is just a set of ordered pairs where the following are assumed to exist: The *domain* of a relation R is the set

$$\text{dom}(R) = \{x : \exists y((x, y) \in R)\}$$

The *image* of a relation R is the set

$$\text{image}(R) = \{y : \exists x((x, y) \in R)\}$$

If for some set A and $R \subseteq A \times A$, then we say that R *is a relation on* A. We also use the following notation when convenient: We write aRb to mean $(a, b) \in R$.

EXAMPLE: Consider the classic relation on $\mathbb{N}$: $\leq$. In this example, both notations are clear. We can view it as the set of ordered pairs $\{(a, b) : a \leq b\}$ and we normally write $a \leq b$ instead of $(a, b) \in \leq$. However, both notations are acceptable.

Operations and Properties of relations:

- If R and S are relations, we say $R = S$ if they are equal as sets.
- If R and S are relations, then so are the sets $R \cup S$, $R \cap S$, and $R \setminus S$.
- If R is a relations, then we can look at the inverse relation $R^{-1} = \{(y, x) : (x, y) \in R\}$.
- If $R \subseteq X \times Y$ is a relation, you can draw the relation in the Cartesian plane where the horizontal axis is X and the vertical axis is Y. Therefore, the $\text{dom}(R) \subseteq X$ is also called the *projection of R onto the first coordinate* or equivalently, the *projection of R onto the X axis.* Similarly, the $\text{image}(R) \subseteq Y$ is also called the *projection of R onto the second coordinate* or equivalently, the *projection of R onto the Y axis.*

How relations are related to graphs:

Definition 9. Let A be a set and $R \subseteq A \times A$. Then $\langle A, R \rangle$ is a directed graph.

A graph is comprised of vertices and edges which join these vertices. A directed graph has the property that given two points a and b, we distinguish between an edge between a and b and one between b and a. Since in general a relation is just a set of ordered pairs, it is possible that $(a, b) \in R$ and $(b, a) \notin R$, so a relation is a directed graph.

Definition 10. Let A be a set and $R \subseteq A \times A$ be a relation on A. Then R is a *symmetric relation* if it satisfies $\forall a, b \in A((a, b) \in R \iff (b, a) \in R$.

If R is a symmetric relation on A, then $\langle A, R \rangle$ is a graph.

EXERCISES

1. Let A, B, and C be sets. Prove $A \times (B \cap C) = (A \times B) \cap (A \times C)$.
2. Let A, B, and C be sets. Prove $A \times (B \setminus C) = (A \times B) \setminus (A \times C)$.

3.
   - Let X and Y be sets and $R, S \subseteq X \times Y$ relations. For a set A, let $\pi_X(A)$ be the projection of A onto the X axis. Prove that $\pi_X(R \cup S) = \pi_X(R) \cup \pi_X(S)$.
   - Let R and S be as above. Is it true in general that $\pi_X(R \cap S) = \pi_X(R) \cap \pi_X(S)$? Justify your answer.

4 Partial Orders

In this section we consider a special type of relation called a partial order:

Definition 11. Let R be a relation on a set A.
R is *reflexive* if for all $a \in A$, aRa.
R is *symmetric* if for all $a, b \in A$, $(aRb) \iff (bRa)$.
R is *transitive* if for all $a, b, c \in A$, we have that $(aRb \wedge bRc) \Rightarrow (aRc)$.

Let us consider a few examples:

- Consider the set $\mathbb{N} = \{0, 1, 2, 3 \dots\}$ and the relation $\leq$. Then this relation is reflexive and transitive, but NOT symmetric. In fact, this is an example of a relation that is *anti-symmetric* which means that $(a \leq b) \wedge (b \leq a) \Rightarrow a = b$.
- Consider the same set $\mathbb{N}$ with the relation $<$. This relation is transitive, but not reflexive or symmetric.

Definition 12. A reflexive, transitive, antisymmetric relation R on a set A is called a *partial order* on A. The two pieces of information together $\langle A, R \rangle$ is called a *partially ordered set*.

When it is clear which partial order is being referred to, we sometimes call A a partially ordered set with the given partial order implied.

Definition 13. If A is a set and R is a transitive, antisymmetric and antireflexive relation, i.e. $\forall x(\neg xRx)$, then A is called a *strict partial order*.

Notation: We use the notation $\leq$ very often to refer to a partial order. When we fix a partial order $\leq$, we will use the notation $<$ to refer to the strict partial order $(a, b) : a \leq b \wedge a \neq b$.

When we wish for a definition or statement to apply to both partial orders and strict partial orders, we will state it for "(strict) partial orders". Some claims are stated for partial orders, but a similar statement could be proved for strict orders as well. Consider this when reading.

Examples:

- $\mathbb{N}$ with the usual $\leq$ is a partial order. $\mathbb{N}$ with the usual $<$ is a strict order.
- Let A be any set. Then consider the relation $\subseteq$. This is a partial order on $\mathcal{P}(A)$. The relation $\subsetneq$ is a strict order.

The proofs of these are left as exercises.

Definition 14. Let $\langle A, R\rangle$ be a (strict) partial order and $a, b \in A$ such that aRb and aRb. Then we say b is the the *immediate successor* to a if there is no $x \in A$ satisfying $x \neq a$,$x \neq b$ and $aRx \wedge xRb$.

Example: In $\langle \mathbb{N}, \leq\rangle$ where $\leq$ is the usual less than, 1 is the immediate successor of 0.

Definition 15. Let $\langle A, R\rangle$ be a (strict) partially ordered set and $a, b \in A$.

We say that a and b are *comparable* if either aRb, bRa, or $a = b$.

We say that a subset of a partial order $A' \subseteq A$ is an *antichain* if no two nonequal elements in A' are comparable.

We say that $\langle A, R\rangle$ is a *linear order* if any two elements of A are comparable.

Example: Let A be a set with at least two elements. Then $\langle \mathcal{P}(A), \subseteq\rangle$ is <u>not</u> a linear order. $\langle \mathbb{N}, <\rangle$ is a linear order.

Proof: Let $a, b \in A$ where $a \neq b$. Then $\{a\}$ is not a subset of $\{b\}$ and $\{b\}$ is not a subset of $\{a\}$. But you will prove in the exercises that $\langle \mathcal{P}(A), \subseteq\rangle$ is a partial order.

Definition 16. Let $\langle A, R\rangle$ be a (strict) partially ordered. Then $a \in A$ is a *minimal element* if there is no $x \in A$ such that $xRa \wedge x \neq a$.

If a satisfies that aRx for all $x \in A$, then a is a *minimum*.

We will now state some basic facts about partial and linear orders. Some of these we will prove and some are left as exercises.

- FACT 1: Let $\langle A, \leq_1\rangle$ and $\langle B, \leq_2\rangle$ be two partial orders. Then $\langle A \cap B, \leq_1 \cap \leq_2$ is a partial order where $\leq_1 \cap \leq_2$ are all the ordered pairs that are in both $\leq_1$ and $\leq_2$.

  Proof:

  It is an easy exercise to show that $\leq_1 \cap \leq_2$ is indeed a relation on $A \cap B$ as defined.

  Reflexivity If $a \in A \cap B$, then $a \leq_1 a$ and $a \leq_2 a$, so $\leq_1 \cap \leq_2$ is still reflexive.

  Antisymmetry If $(a, b) \in \leq_1 \cap \leq_2$, then in particular, $(a, b) \in \leq_1$, so either $a = b$ or $(b, a) \notin \leq_1$, so if $a \neq b$, then $(b, a) \notin \leq_1 \cap \leq_2$. Thus the intersection is antisymmetric.

  Transitivity Now suppose that $(a, b), (b, c) \in \leq_1 \cap \leq_2$. Then $(a, b), (b, c) \in \leq_1$ so $(a, c) \in \leq_1$ since $\leq_1$ is a partial order. Similarly, $(a, b), (b, c) \in \leq_2$ so $(a, c) \in \leq_2$. So (a, c) is in the intersection and the intersection is transitive.

- FACT 2: If $\langle A, \leq\rangle$ is a partially ordered set, then $\leq^{-1}$ is also a partial order on A.

  Proof: Exercise.

- FACT 3: Suppose $\langle A, \leq_1\rangle$ and $\langle B, \leq_2\rangle$ are partial orders. Then the following relation $\leq_3$ is a partial order on the cartesian product $A \times B$:

$$(a_1, b_1) \leq_3 (a_2, b_2) \iff a_1 \leq_1 a_2 \wedge b_1 \leq_2 b_2$$

  Proof:

Reflexivity Since $a \leq_1 a$ and $b \leq_2 b$, we have that $(a,b) \leq_3 (a,b)$ and $\leq_3$ is reflexive.

Antisymmetry Suppose $(a_1,b_1) \leq_3 (a_2,b_2)$ and $(a_1,b_1) \neq (a_2,b_2)$. Then, since $\leq_1$ is antisymmetric, we know that either $a_1 = a_2$ or $(a_2,a_1) \notin \leq_1$. If $(a_2,a_1) \notin \leq_1$ then $(a_2,b_2) \not\leq_3 (a_1,b_1)$ and $\leq_3$ is antisymmetric. If $a_1 = a_2$, then we know that $b_1 \neq b_2$ since $(a_1,b_1) \neq (a_2,b_2)$. Since $\leq_2$ is antisymmetric, we have that $b_2 \not\leq_2 b_1$ and thus $(a_2,b_2) \not\leq_3 (a_1,b_1)$.

Transitivity Now suppose $(a_1,b_1) \leq_3 (a_2,b_2)$ and $(a_2,b_2) \leq_3 (a_3,b_3)$. Then, as before, $a_1 \leq_1 a_2$ and $a_2 \leq_1 a_3$ so by transitivity of a_1, we have $a_1 \leq_1 a_3$. Similarly $b_1 \leq_2 b_2$ and $b_2 \leq_2 b_3$ so by transitivity of $\leq_2$, we have $b_1 \leq_2 b_3$. Thus $(a_1,b_1) \leq_3 (a_3,b_3)$ and $\leq_3$ is transitive.

- FACT 4: Suppose $\langle A, \leq_1 \rangle$ and $\langle B, \leq_2 \rangle$ are partial orders. Then the following relation $\leq_4$, which we call the *lexicographic order*, is a partial order on the cartesian product $A \times B$:

$$(a_1,b_1) \leq_4 (a_2,b_2) \iff (a_1 \neq a_2 \wedge a_1 \leq_4 a_2) \vee (a_1 = a_2 \wedge b_1 \leq b_2)$$

Proof: Exercise.

EXERCISES:

1. Prove that the relation $\subseteq$ is a partial order on $\mathcal{P}(A)$.
2. Let A and B be two subsets of some universe U. Prove that there is a maximal element C (with respect to the partial order on $\mathcal{P}(U)$ given by $\subseteq$) such that $C \subseteq A$ and $C \subseteq B$, i.e any other set D satisfying this is a subset of C. Prove that there is a minimal element E such that $A \subseteq E$ and $B \subseteq E$.
3. Suppose that A is a set and $<$ is a strict partial order on A. Prove that $\leq = \{(a,b) : a < b \vee a = b\}$ is a partial order.
4. Prove that $\leq_1 \cap \leq_2$ is indeed a relation on $A \cap B$ as defined above.
5. Suppose $\langle A, \leq \rangle$ is a partially ordered set. Prove $\leq^{-1}$ is also a partial order on A.

6. Prove that the lexicographic order is a partial order on the cartesian product of two partially ordered sets. Prove that the lexicographic order on the cartesian product of two linear orders is a linear order.

5 Equivalence Relations

We will again consider the following properties that a relation can satisfy:

Definition 17. Let R be a relation on a set A.
R is *reflexive* if for all $a \in A$, aRa.
R is *symmetric* if for all $a, b \in A$, $(aRb) \iff (bRa)$.
R is *transitive* if for all $a, b, c \in A$, we have that $(aRb \wedge bRc) \Rightarrow (aRc)$.

Definition 18. Let A be a set. We say that a relation R is an equivalence relation if R is reflexive, transitive and symmetric.

Let us consider the following example which we will carry out throughout the section. Let E be the following relation on $\mathbb{N}$: $R_1 = \{(a, b) : a \text{ and } b \text{ are both even or both odd }\}$. We claim that this is an equivalence relation.

Proof:
Reflexivity: For any $a \in \mathbb{N}$, aR_1a since if a is even then so is a and if a is odd, then so is a. This may sound absurd, but we are not ready yet to use phrases like "clearly".
Symmetry: For any $a, b \in \mathbb{N}$, if a and b are both even, then b and a are both even. If a and b are both odd, then b and a are both odd. So $aR_1b \iff bR_1a$.
Transitivity: Let $a, b, c \in \mathbb{N}$ such that aR_1b and bR_1c. Either b is even or b is odd.

Case 1 - b is even: Then a is even since aR_1b and c is even since bR_1c. So aR_1c.
Case 2 - b is odd: Then a is odd since aR_1b and c is odd since bR_1c. So aR_1c.

Therefore this is an equivalence relation. Let us now prove an easy fact about equivalence relations.

Proposition 19. *If E is an equivalence relation on A, then* $\text{dom}(E) = \text{image}(E) = A$.

Proof. Let $a \in A$. Then $(a, a) \in E$ since E is reflexive, and therefore $a \in \text{dom}(E)$ and $a \in \text{image}(E)$.

□

Definition 20. Given a set A, an equivalence relation E on A, and an element $a \in A$, we call the set $[a] = \{x \in A : aEx\}$ the equivalence class of a.

It is clear that $a \in [a]$ since E is reflexive. Therefore, if $[a] = [b]$ then since $b \in [b]$, $b \in [a]$ and we have that aEb. So we have that $[a] = [b] \Rightarrow aEb$. We now prove the converse.

Claim 21. *If aEb, then $[a] = [b]$. Thus $aEb \iff [a] = [b]$.*

Proof. $\subseteq$ Suppose $c \in [a]$. Then cEa. Since aEb and E is transitive, we have the cEb. Since E is symmetric, we have bEc and $c \in [b]$.

$\supseteq$Suppose $c \in [b]$. Then cEb. Since aEb and E is symmetric, we have bEa. Since E is transitive, we have cEa. Since E is symmetric, we have aEc and $c \in [a]$. □

What are the equivalence classes in our example R_1 on $\mathbb{N}$? Intuitively, we've split $\mathbb{N}$ into two sets- the even numbers and the odd numbers. So there are two equivalence classes: $[0], [1]$.

Definition 22. Let A be a set and P a set of nonempty subsets of A, i.e. every element of P is a nonempty subset of A. Then we say that P is a *partition* of A if for ever every element of a there is exactly one element $p \in P$ such that $a \in p$.

We now connect the notion of an equivalence relation and a partition.

Claim 23. *If E is an equivalence relation on A, the the set of equivalence classes is a partition of A.*

Proof. We need to show that every element is in exactly one equivalence class. It is clear that $a \in [a]$, so we know that every element is in at least one equivalence class. Now we need to show it is in at most one equivalence class. To do this, we will suppose that $c \in [a]$ and $c \in [b]$ and show that $[b] = [a]$. Suppose $c \in [a]$ and $c \in [b]$. Then we know that aEc and bEc. Since E is symmetric, we know that cEb and since E is transitive, we know that aEb, so as we proved before, $[a] = [b]$. □

For this reason, there are many possible ways to represent an equivalence class - we simply choose a *representative* from the class, say a, and write $[a]$ to mean the class to which a belongs. This choice is not unique, and we will deal with the implications of this shortly.

Claim 24. *If P is a partition on a set A, then the relation $R = \{(a, b) : \exists p \in P$ such that $a \in p$ and $b \in p\}$ is an equivalence relation.*

Proof. Reflexive: Let $a \in P$. Since $(a, a) \in R$ since there is a $p \in P$ such that $a \in p$.

Symmetric: Suppose there is a $p \in P$ such that $a \in p$ and $b \in p$, i.e. $(a, b) \in R$. Then $b \in p$ and $a \in p$ so $(b, a) \in R$.

Transitive: Suppose $(a, b) \in R$, i.e. there is a $p_1 \in P$ such that $a \in p_1$ and $b \in p_1$,and suppose $(b, c) \in R$, i.e. there is a $p_2 \in P$ such that $b \in p_2$ and $c \in p_2$. Since there is a unique $p \in P$ such that $b \in p$, $p_1 = p_2$ and $c \in p_2$. Therefore, $(a, c) \in R$. □

Now that we have all the definitions, we move onto the notion of quotient spaces.

Definition 25. Suppose that E is an equivalence relation on A. Then the *quotient space* is the set of equivalence classes $A/E = \{[a] : a \in A\}$ and A/E is often read "A mod E".

Example: $\mathbb{N}/R_1 = \{[0], [1]\}$.

When considering a quotient space, we may wish to define a relation on the quotient space. To do this, we must ensure that the definition doesn't rely in an essential way on the choice of representative of a class.

Definition 26. Let E be an equivalence relation on A. A relation on the set A is *well-defined* on A/E if for any $a, b, c, d \in A$, if $[a] = [b]$ and $[c] = [d]$ then $(a, c) \in R$ if and only if $(b, d) \in R$.

Crucial point: If you define a relation on A, it can be thought of as a relation on A/E as well provided it is well-defined. It is often necessary to check that what we say makes sense as a notion before determining whether what we have said is correct.

Let us look at two examples to illustrate this point.

- Let us consider our equivalence relation R_1 on $\mathbb{N}$ and consider the relation T on $\mathbb{N}/R_1$ where $T = \{([a], [b]) : a - b \text{ is even } \}$. Note that $a - b$ could be a negative whole number. We claim that this relation is well-defined. To prove this, we need to prove that if $a - b$ is even, then for any $c \in [a]$ and any $d \in [b]$, $c - d$ is even as well. We will use the following subclaim.

  Subclaim: For $a, b \in \mathbb{N}$, $a - b$ is even if and only if a and b are both even or a and b are both odd.

  Proof:

  ($\Leftarrow$) Suppose a and b are both even. Then $a = 2r$ for some natural number r and $b = 2s$ for some natural number s. So $a - b = 2r - 2s = 2(r - s)$ and $a - b$ is even. Suppose a and b are both odd. Then $a = 2r + 1$ and $b = 2s + 1$ for some natural numbers r and s. So $a - b = 2r + 1 - (2s + 1) = 2r - 2s = 2(r - s)$ and $a - b$ is even.

  ($\Rightarrow$) To prove this direction, we will prove the contrapositive. Namely, we will show that if a is even and b is odd (or similarly if a is odd and b is even), then $a - b$ is odd. Suppose $a = 2r + 1$ and $b = 2s$ for some natural numbers r and s. Then $a - b = 2r + 1 - 2s = 2(r - s) + 1$ and $a - b$ is odd.

  We now return to our original claim. Suppose that $([a], [b]) \in T$, i.e. $a - b$ is even, and suppose that $c \in [a]$ and $d \in [b]$. By the subclaim, either a and b are both even in which case $[a] = [b]$, or a and b are both odd and still $[a] = [b]$. More than this we showed that if $[a] = [b]$ then $a - b$ is even. If $c \in [a]$, then $[c] = [a]$ and if $d \in [b]$, then $[d] = [b]$, so if $[a] = [b]$, then $[c] = [d]$.

- We say two non-zero natural numbers are relatively prime if their greatest common divisor is 1. Let us consider our equivalence relation R_1 on $\mathbb{N}$, and consider the relation S on $\mathbb{N}/R_1$ where $S = \{([a],[b]) : a \text{ and } b \text{ are relatively prime.}\}$. This relation is not well-defined. To see this consider that fact that as we've defined it, $([3],[10]) \in S$ since 3 and 10 are relatively prime and $([5],[10]) \notin S$. However, $[3] = [5]$.

We will extend the notion of "well-defined" to functions on the quotient space in the next section.

EXERCISES:

1. We say two non-zero natural numbers are relatively prime if their greatest common divisor is 1. Consider the set of natural number $\mathbb{N}$ and the relation $\{(a,b) : a \text{ and } b \text{ are relatively prime.}\}$. Is this relation reflexive? Is this relation transitive? Is this relation symmetric? Justify your answers with proofs.

2. Let $a, b \in \mathbb{N}$. We say that $a \equiv b \bmod 3$ if $a - b$ is divisible by 3. (0 is divisible by 3). Show that $\equiv \bmod 3$ is an equivalence relation. (Notice that the relation E in our example above is $\equiv \bmod 2$.) What are the equivalence classes? Justify your answer.

3. Prove the general fact that $\equiv \bmod n$ is an equivalence relation.

6 Functions

We will present functions in two ways. We will view functions as a map between sets as well as a set of ordered pairs.

Definition 27. Let X and Y be sets. A relation $F \subseteq X \times Y$ is a *function* if for any $x \in X$ there is at most one $y \in Y$ such that $(x,y) \in F$.

There are many ways to represent a function. Let us consider some examples:

- $F \subseteq \mathbb{N} \times \mathbb{N}$, where $F = \{(a,b) : b = a\}$
- $F \subseteq \mathbb{N} \times \mathbb{N}$, where $F = \{(1,2)(3,4)\}$.

The following is not a function.

- $F \subseteq \mathbb{Z} \times \mathbb{Z}$, where $F = \{(a, b) : a = b^2\}$. Then $(4, 2)$ and $(4, -2)$ are both in F.

**The relation F is not a function, but F^{-1} is the familiar function $f(x) = x^2$.

Recall the following definitions which will remain the same for relations which are also functions:

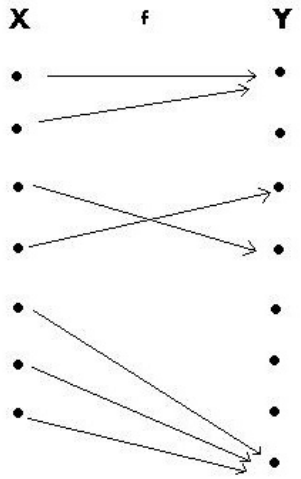
Definition 28. The *domain* of a function F is the set

$$\text{dom}(F) = \{x : \exists y((x, y) \in F)\}$$

If $\text{dom}(F) \neq X$, then we say F is a *partial function* on X. When $X = \text{dom}(F)$ the function is sometimes called a *total function*. Whether a function is total or not depends on how we choose to represent it. IN THIS CLASS WE ASSUME FUNCTIONS ARE TOTAL UNLESS WE ASSERT OTHERWISE. In other words, we assume that we present F as a subset of $\text{dom}(F) \times Y$.

Definition 29. The *range* of a function F is the set Y. The range depends on how the function is presented.

Now let us look at functions as maps: A function can be viewed as a map from X to Y and we often write $F : X \to Y$. Intuitively, we see F as assigning to each value in X a unique value in Y. We use the notation $F(x) = y$ for $x \in X$ and $y \in Y$ to say that F assigns the value y to x, or when viewing it as a relation, that $(x, y) \in F$.



Viewing it this way, it becomes very clear that a function is given by THREE pieces of information: The domain, the range, and the assignment between the two. We will go back and forth between the two ways of viewing a function so that both become intuitive.

Recall the following definition:

Definition 30. .

The *image* of a function F is the set

$$\text{image}(F) = \{y : \exists x((x, y) \in F)\}$$

If we write $F : X \to Y$ then, the *image* of F is $\{y : \exists x(F(x) = y)\}$. If $A \subseteq X$, the *image of A under F*, $F[A] = \{y : \exists a \in A(F(a) = y)\}$.

Claim 31. *Let $F : X \to Y$ be a function and $A, B \subseteq X$. Then $F[A \cup B] = F[A] \cup F[B]$.*

Proof. $F[A \cup B] \subseteq F[A] \cup F[B]$: Let $x \in F[A \cup B]$. Then there is an $a \in A \cup B$ such that $F(a) = x$. Then $a \in A$ or $a \in B$. If $a \in A$, then $x = F(a) \in F[A]$. If $a \in B$ then $x = F(a) \in F[B]$. Either way, $x \in F[A] \cup F[B]$.

$F[A] \cup F[B] \subseteq F[A \cup B]$: Let $x \in F[A] \cup F[B]$. Then either $x \in F[A]$ or $x \in F[B]$. If $x \in F[A]$, there is an $a \in A$ such that $x = F(a)$. In this case $a \in A \cup B$ and $x = F(a) \in F[A \cup B]$. If $x \in F[B]$, there is an $a \in B$ such that $x = F(a)$. In this case $a \in A \cup B$ and $x = F(a) \in F[A \cup B]$.

□

We now define a number of properties a function can have. We consider $F : X \to Y$:

Definition 32.

- A function F is *surjective* if $\text{range}(F) = \text{image}(F)$.
- A function F is *injective* if for every $y \in Y$ there is at most one $x \in X$ such that $(x, y) \in F$, or equivalently, if $F(x) = F(y)$ then $x = y$.
- A function F is *bijective* if it is both surjective and injective.

Intuitively, a bijection pairs up the two sets perfectly, i.e. nothing in Y is used twice and everything in Y is used.

Let us look at some examples:

- $F : \mathbb{N} \to \mathbb{N}$, where $F = \{(a, b) : b = a\}$. This function is bijective.

  proof: Firstly, notice that we could have defined this assignment as follows: $F(x) = x$. This is the conventional notation. First we will prove that F is injective. Suppose $F(a) = F(b)$. Since $F(a) = a$ and $F(b) = b$ we have that $a = b$. So F is injective. Now we need to prove that F is surjective. To do this, consider $a \in \mathbb{N}$. Then, $F(a) = a$ so a is in the image. Thus the function is bijective.

- $F \subseteq \mathbb{Z} \times \mathbb{Z}$, where $F = \{(a, b) : a^2 = b\}$. This function is neither injective nor surjective.

  proof: Since $F(2) = 4$ and $F(-2) = 4$, this function is not injective. There is no integer a such that $a^2 = -1$. So F is not surjective.

- $F : \mathbb{N} \to \mathbb{N}$ where $F(n) = n + 1$. This function is injective but not surjective.

  proof: 0 is not in the image, so the function is not surjective. $n + 1 = m + 1 \Rightarrow n = m$ so the function is injective.

- $F : \mathbb{N} \to \{0\}$ where $F(n) = 0$ for all $n \in \mathbb{N}$. This function is surjective, but not injective.

  proof: $F(0) = 0$ so the function is surjective, but $F(0) = F(1)$ so the function is not injective.

Claim 33. *If $F : A \to B$ is a bijection, then $F^{-1} : B \to A$ is also a bijection.*

Proof. We need to prove 3 things: the the relation F^{-1} is a function with domain B, that this function is injective, and that this function is surjective.

F^{-1} is a function with domain B: Since F is surjective, we have that for every $b \in B$ there is an $a \in A$ such that $(a, b) \in F$. Thus $(b, a) \in F^{-1}$ and $b \in \text{dom}(F^{-1}$. Since F is injective, there is a single A such that $(a, b) \in A$. Thus there is only one a such that $(b, a) \in F^{-1}$. So $F^{-}1$ is a function.

F^{-1} is injective: Since F is a function, we have that for every $a \in A$ there is a unique $b \in B$ such that $(a, b) \in F$. Therefore, for every $a \in A$, there is a unique b such that $(b, a) \in F^{-1}$.

F^{-1} is surjective:Let $a \in A$. Since F is a function with domain A, there is some $b \in B$ such that $(a, b) \in A$. Thus, there is some $b \in B$ such that $(b, a) \in F^{-1}$. □

Definition 34. Let A be a set. A *binary function on A* is a function $F : A \times A \to A$.

Examples: Both $+$ and $\cdot$ are binary functions on $\mathbb{Z}$ and $\mathbb{N}$. These satisfy that $a + b = b + a$ and $a \cdot b = b \cdot a$. This is not true of all binary functions! Consider $-$ as a binary function on $\mathbb{Z}$.

Definition 35. Consider functions $F : X \to Y$ and $G : Y \to Z$. The composition function $G \circ F : X \to Z$ is the function given by $G \circ F(x) = z$ if there is a $y \in Y$ such that $F(x) = y$ and $G(y) = z$.

Notice that in order for the composition $G \circ F$ to be defined, the image of F must be contained in the domain of G.

We will not prove a few properties of composition functions.

Claim 36. *Suppose we are given functions $F : X \to Y$ and $G : Y \to Z$.*

1. *If $G \circ F$ is injective, then F is injective. However $G \circ F$ is injective does not imply that G is injective.*

2. *If $G \circ F$ is surjective, then G is surjective. However, the converse is not true in general.*

Proof. 1. Suppose F is not injective. Then there is $a, b \in F$ where $a \neq b$ and $F(a) = F(b)$. Then, $G(F(a)) = G(F(b))$ and $G \circ F$ is not injective. To see that $G \circ F$ is injective does not imply that G is injective, consider the following functions from $\mathbb{N} \to \mathbb{N}$. $F(n) = 2n$ and G the function defined as follows: $G(n) = n$ for n even, and $G(n) = n - 1$ for n odd. Here, G is not injective, but $F \circ G$ is injective. The proof of this is left to the reader.

2. Suppose $z \in Z$. Then there is a $a \in X$ such that $G \circ F(a) = z$. Let $y = F(a)$. Then, $G(y) = z$ and z is in the image of G. To see that the converse is not true in general, consider the following functions on $\mathbb{N}$. Let $F(n) = 2n$ and let $G(n) = n$. G is clearly surjective, but $G \circ F$ has image only the even natural numbers.

□

Definition 37. Let $F : A \to B$ be a function and let $C \subseteq A$. Then the restricted function $F \restriction C : C \to B$ sometimes written $F|_C$ is the function given by $\{(c, b) : c \in C \text{ and } F(c) = b\}$ where c is considered as an element of A.

Claim 38. *Suppose $F : A \to B$ is a function that is not injective. Then, there is some subset $A' \subseteq A$ such that $F \restriction A'$ is injective. Furthermore, this can be done so that the image of $F \restriction A'$ is the same as the image of F.*

Proof. If you just want an injective function, you can always choose any element $a \in A$ and let $A' = \{a\}$. A function on a single element must be injective. However, this function will not generally have the same image.

Consider the following equivalence relation: aEb if $F(a) = F(b)$. Now consider the quotient space A/E. Each element of the quotient space $[a]$ is a subset of A. Now let A' be a set (and there are many options) containing precisely one representative from each class. Then $F \restriction A'$ is injective and the image is the same as the image of F.

□

A word of caution: This proof uses what is known as the axiom of choice. An intuitive notion of what this axiom states is that if you give me a collection of nonempty sets, I can choose one element from each set to form a new set.

Well-Defined Function on Quotient Spaces

We are now going to use equivalence relations to define interesting quotient spaces. In the notes, we will define $\mathbb{Z}_3$ and prove that the induced addition and multiplication are well-defined. In the exercises, you will construct the rational numbers with addition and multiplication and prove that these are well- defined.

Definition 39. Let E be an equivalence relation on A. Then a function $F : A \to A$ is *well-defined* on A/E for every $a \in A$, if $F(a) = b$ and cEa, then $F(c)Eb$. A binary function F on A is well-defined on A/E if for any $a, b \in A$, if $F((a, b)) = c$ and $a'Ea$,$b'Eb$, then $F((a', b'))Ec$.

What are the elements of $\mathbb{Z}_3$?

Consider the set of integers $\mathbb{Z} = \{\cdots -2, -1, 0, 1, 2, \dots\}$ and the equivalence relation $\equiv$ mod 3, i.e. aEb iff $a - b$ is divisible by 3. We saw this example in the exercises above and you proved that this is an equivalence relation. $\mathbb{Z}_3$ is the quotient space $\mathbb{Z}/E$. So $\mathbb{Z}_3$ is comprised of 3 equivalence classes: $[0], [1], [2]$. The functions $+$ and $\cdot$ refer to the ordinary addition and multiplication of integers which satisfy all of the usual properties- commutativity, associativity, and the distributive property. We use the concatenation notation ab instead of $a \cdot b$ to make it easier to write.

$+$ is well defined on $\mathbb{Z}_3$.

We need to show the following: If $a'Ea$ and $b'Eb$ then $(a + b)E(a' + b')$:

$a \equiv a'$ mod $3 \Rightarrow a' - a = 3r$ for some integer r.

$b \equiv b'$ mod $3 \Rightarrow b' - b = 3s$ for some integer s.

So $(a'+b')-(a+b) = a'-a+b'-b = 3r-3s = 3(r-s)$ so $(a+b) \equiv (a'+b')$ mod 3.

$\cdot$ is well defined on $\mathbb{Z}_3$.

We need to show the following: If $a'Ea$ and $b'Eb$ then $(ab)E(a'b')$:

$a \equiv a' \bmod 3 \Rightarrow a' - a = 3r$ for some integer r.
$b \equiv b' \bmod 3 \Rightarrow b' - b = 3s$ for some integer s.
So $a'b' - ab = a'b' - ab' + ab' - ab$
$= (a' - a)b' + (b' - b)a = 3rb' + 3sa = 3(rb' + sa)$ so $(ab) \equiv (a'b') \bmod 3$.

EXERCISES:

1. For any set A, the function $id : A \to A$ is the function $id(x) = x$. Show that for any set A, id is a bijection.

2. Recall that if R is a relation, then we can look at the inverse relation $R^{-1} = \{(y, x) : (x, y) \in R\}$. Show that for a function $F : X \to Y$, F^{-1} is a (total) function if and only if F is bijective.

3. Show that if $F : X \to Y$ is a bijection, then $F^{-1} : Y \to X$ is a bijection.

4. Consider functions $F : X \to Y$ and $G : Y \to Z$. The composition function $G \circ F : X \to Z$ is the function given by $G \circ F(x) = z$ if there is a $y \in Y$ such that $F(x) = y$ and $G(y) = z$. Show that image$(G \circ F) \subseteq$ image(G).

5. Consider functions $F : X \to Y$ and $G : Y \to Z$. The composition function $G \circ F : X \to Z$ is the function given by $G \circ F(x) = z$ if there is a $y \in Y$ such that $F(x) = y$ and $G(y) = z$. Prove that if F and G are injective, then so is $G \circ F$. Show that if F and G are surjective, then so is $G \circ F$. Conclude that the composition of two bijections is a bijection.

6. Consider functions $F : X \to Y$ and $G : Y \to Z$. The composition function $G \circ F : X \to Z$ is the function given by $G \circ F(x) = z$ if there is a $y \in Y$ such that $F(x) = y$ and $G(y) = z$. Is the image of $G \circ F$ necessarily the image of G? Justify your answer.

7. CONSTRUCTING $\mathbb{Q}$: There are a number of ways to represent $\mathbb{Q}$. We can write $\mathbb{Q} = \{\frac{a}{b} : a, b \in \mathbb{Z}, b \neq 0\}$. This notation has a lot of repeated elements. Namely $\frac{2}{4} = \frac{1}{2}$. In this exercise, you will construct $\mathbb{Q}$ as a quotient space modulo the classic equivalence on fractions.

   Consider the set $\mathbb{Z} \times \mathbb{Z}^*$ where $\mathbb{Z}^* = \mathbb{Z} \setminus \{0\}$.

(a) Show that the relation $(a,b)E(c,d)$ if and only if $ad = bc$ is an equivalence relation.

(b) Consider the binary function on $\mathbb{Z} \times \mathbb{Z}^*$ defined by $(a,b)+(c,d) = (ad+bc, bd)$. Show that this is well-defined on $\mathbb{Z} \times \mathbb{Z}^*/E$.

(c) Consider the binary function on $\mathbb{Z} \times \mathbb{Z}^*$ defined by $(a,b) \cdot (c,d) = (ac, bd)$. Show that this is well-defined on $\mathbb{Z} \times \mathbb{Z}^*/E$.

7 Induction

Let us begin with a few basic definitions that we will use in the next few sections.

- $\sim$: Let A and B be sets. We write $A \sim B$ if there is a bijection from A to B. This is an equivalence relation on subsets of a given set as you will prove in the exercises.
- $\mathbb{N}^{<n}$: Let n be a natural number. Then $\mathbb{N}^{<n} = \{0, 1, ..., n-1\}$. It follows that $\mathbb{N}^{<0} = \emptyset$.

Definition 40. Let $S \subseteq \mathbb{N}$. S is *bounded* if for some $n \in \mathbb{N}$, $S \subseteq \mathbb{N}^{<n}$.

Example: The set containing all the factors of 12, $\{1, 2, 3, 4, 6, 12\}$ is a bounded subset of $\mathbb{N}$ since it is a subset of $\mathbb{N}< 13$. The set of even natural numbers $\{0, 2, 4, 6, ...\}$ is unbounded.

We will take the following principle as fact:

PIGEONHOLE PRINCIPLE: Let $n > m$ two natural numbers. A function from $\mathbb{N}^{<n} \to \mathbb{N}^{<m}$ cannot be injective. In particular, if m and n are two natural numbers and $m \neq n$, then either $m > n$ or $n > m$ and so there is no bijection from $\mathbb{N}^{<n} \to \mathbb{N}^{<m}$.

Definition 41. Let A be any set. A is *finite* if for some $n \in \mathbb{N}$ there is a bijection from A to $\mathbb{N}^{<n}$. In this case, we write $|A| = n$ and say "A has size n" or "A has n elements".

NOTE: The empty set is a bijection between the empty set and the empty set. (Think about whether it satisfies the conditions... hopefully you realize that the conditions are all satisfied vacuously.) Therefore, the empty set has size 0.

If there was a bijection $f : A \to \mathbb{N}^{<n}$ and a bijection from $g : A \to \mathbb{N}^{<m}$, then $g(f^{-1})$ is a bijection from $\mathbb{N}^{<n}$ to $\mathbb{N}^{<m}$ which can only exist if $n = m$ by the pigeonhole principle. Therefore a finite set can have only one size. And yes- we have now used quite a bit of math to prove an obvious fact, but we are making sure that our definitions give us the notion we want.

In general, the notation $|A|$ refers to the size of A for any set A. We will see more of this in the next section.

Claim 42. *If A and B are finite sets and $A \sim B$ then $|A| = |B|$.*

Proof. Suppose $|A| = n$. We will show that $|B| = n$. Since $|A| = n$ there is a bijection $f : A \to \mathbb{N}^{<n}$. Since $A \sim B$ there is a bijection $g : A \to B$. Then consider $f \circ g^{-1} : B \to \mathbb{N}^{<n}$. Since g is a bijection, then so is g^{-1}, and you proved in the exercises that the composition of two bijections is a bijection. So $f \circ g^{-1}$ is a bijection.

□

Example: The set $\{a, b, c\}$ is finite and has 3 elements since the function $\{(a, 0), (b, 1), (c, 2)\}$ is a bijection from the set to $\mathbb{N}^{<3}$.

NOTE: This is an example of an intuitive concept that in mathematics needs a rigorous definition. You should understand that when you say a set has 3 elements, you mean that there is a bijection (which you normally determine by pointing and counting aloud) from the set to $\mathbb{N}^{<3}$. The only unnatural part of this is the fact that we start counting from 0 - *ma laasot.*

Claim 43. *Let A and B be finite sets. Then if $A \cap B = \emptyset$, then $|A \cup B| = |A| + |B|$.*

Proof. Suppose A has size $n \in \mathbb{N}$ and B has size $m \in \mathbb{N}$. Then there exist $f : A \to \mathbb{N}^{<n}$ and $g : B \to \mathbb{N}^{<m}$ bijections. Now consider $h : A \cup B \to \mathbb{N}^{<m+n}$ as follows:

$$h(x) = \begin{cases} f(x) & \text{if } x \in A \\ g(x) + n & \text{if } x \in B \end{cases}$$

So, for $b \in B$, $h(b) \geq n$. We will prove that h ia a bijective function.

$\underline{h \text{ is a function:}}$ Let $a \in A \cup B$. Then $a \in A$ or $a \in B$ but not both and $h(a)$ has one value as defined.

$\underline{h \text{ is surjective:}}$ Let $s \in \mathbb{N}^{<n+m}$. Then either $s < n$ or $n \leq s < n+m$. If $s < n$, then since f is surjective, there is an $a \in A$ such that $f(a) = h(a) = n$. If $n \leq s < n+m$, then $s - n < m$ and there is $b \in B$ such that $g(b) = s - n$. Then $h(b) = g(b) + n = s$.

$\underline{h \text{ is injective:}}$ Suppose $h(a_1) = h(a_2)$. If $h(a_1), h(a_2) \in \mathbb{N}^{<n}$, then by definition of h, $a_1, a_2 \in A$. Then $h(a_1) = f(a_1)$ and $h(a_2) = f(a_2)$ and since f is injective, if $f(a_1) = f(a_2)$ then $a_1 = a_2$. Similarly, if $h(a_1), h(a_2) > n$, then $a_1, a_2 \in B$. $h(a_1) = h(a_2) \Rightarrow h(a_1) - n = h(a_2) - n$ and thus $g(a_1) = g(a_2)$. Since g is injective, this implies that $a_1 = a_2$. □

Now consider a fable: Once upon a time there was a mathematician who wanted to prove that a statement was true about all the natural numbers. So he proved it was true about $0, 1, 2, ..., 999$ but when he got to 1000, he died. His student began to prove the statement about 1000 when suddenly it dawned on him: he would also die one day and the work will never be done! So he went back to the drawing board and decided there should be a way to prove that something is true about all the natural numbers without actually proving it for each number separately - and induction was born. (If you are rolling your eyes, chuckling, or snorting, this is good- it means you are still reading!)

To understand the induction principle, we first consider how the natural numbers are constructed. We start with 0 and then we begin adding 1. Every time we add 1, we get a new element- the successor to the previous one. Every element is obtained in just this way. The induction principle plays on this construction to give us a new way to prove statements about the natural numbers.

We will now state a principle that we will accept without proof and then move onto other principles that follow from it:

INDUCTION PRINCIPLE 1 (Weak Induction): Suppose $A \subseteq \mathbb{N}$ and A satisfies the following two conditions:

- $0 \in A$.

- If $a \in A$ then $a + 1 \in A$.

Then $A = \mathbb{N}$.

This principle is both reasonable and intuitive given the construction of $\mathbb{N}$ that we saw above. We can use this to prove an easy fact about $\mathbb{N}$ - indeed this was the construction itself, but it is useful to see it as an example of how to apply the induction principle. Let $\leq$ be the usual order on $\mathbb{N}$. Recall the following definition:

Definition 44. Let $\langle A, \leq \rangle$ be a partial order and $a, b \in A$ such that $a \leq b$ and $a \neq b$. Then we say b is the *successor* or *immediate successor* to a if there is no $x \in A$ satisfying $x \neq a$,$x \neq b$ and $a \leq x \leq b$.

Now it is clear that for $\langle \mathbb{N}, \leq \rangle$ the immediate successor of a natural number n is $n + 1$. Therefore we have the following claim:

Claim 45. *Every element of* $\mathbb{N}$ *is either* 0 *or the successor of another number.*

Proof. Let $A = \{n \in \mathbb{N} : n = 0$ or n is a successor.$\}$ Then $0 \in A$ and if $a \in A$ then $a + 1$ is the successor of a, so $a + 1 \in A$. Thus by the induction principle, $A = \mathbb{N}$.

□

We now move onto the second induction principle and we will show it follows from the first:

INDUCTION PRINCIPLE 2 (Strong Induction): Suppose $A \subseteq \mathbb{N}$ and for all $n \in \mathbb{N}$, $\mathbb{N}^{<n} \subseteq A \Rightarrow \mathbb{N}^{<n+1} \subseteq A$, then $A = \mathbb{N}$.

Proof: Suppose A satisfies the condition above. Let $X = \{n \in A : \mathbb{N}^{<n} \subseteq A\}$. We will show that $X = \mathbb{N}$, and then that A satisfies the conditions for weak induction.

Since the empty set is a subset of every set, $\mathbb{N}^{<0} \subseteq A$ and so we know that since A satisfies the condition strong induction, $\{0\} = \mathbb{N}^{<1} \subseteq A$ and $0 \in X$. Now suppose $n \in X$. Then $\mathbb{N}^{<n} \subseteq A$ by definition and therefore $\mathbb{N}^{<n+1} \subseteq A$ and so is $\mathbb{N}^{<n+2}$. Thus $n+1 \in A$ and $\mathbb{N}^{<n+1} \subseteq A$, so $n+1 \in X$. Thus $X = \mathbb{N}$ by the weak induction principle. Therefore, since $X \subseteq A$, $A = \mathbb{N}$ and we get the conclusion of the strong induction principle.

We now recall a definition and move onto another principle which we will prove using the strong induction principle:

Definition 46. Let $\langle A, \leq \rangle$ be a partially ordered set. Then $a \in A$ is a *minimal element* if there is no $x \in A$ such that $x \leq a \wedge x \neq a$.

If a satisfies that $a \leq x$ for all $x \in A$, then a is a *minimum*.

Note that a minimum element is always minimal and that if A is a linear order (i.e. for all $x, y \in A$, either $x \leq y$ or $y \leq x$), then a minimal element is a minimum.

WELL-ORDERING PRINCIPLE: If $A \subseteq \mathbb{N}$ is nonempty, then A has a minimal element.

Let us see how this follows from the strong induction principle.

Proof: We will prove that for all $n \in \mathbb{N}$, if $X \subseteq \mathbb{N}$ and $n \in X$, then X has minimal element. Since every nonempty subset of $\mathbb{N}$ has an element by definition, this will imply the well-ordering principle. Consider

$$A = \{n \in \mathbb{N} : \forall X \subseteq \mathbb{N}(n \in X \Rightarrow \text{ X has a minimal element.})\}$$

We will use the strong induction principle to prove that $A = \mathbb{N}$. Suppose $n \in \mathbb{N}$ such that $\mathbb{N}^{<n} \subseteq A$. We need to show that $n \in A$. Let $X \subseteq \mathbb{N}$ such that $n \in X$. Then we have two cases:

- Case 1: There is an $m < n$ such that $m \in X$. Then, since $m \in X$ and we assumed that $m \in A$, we know that X has a minimal element.
- Case 2: There is no $m < n$ such that $m \in X$. Then n is a minimal element by definition.

Thus, in either case, X has a minimal element and $n \in A$. Therefore, by strong induction, $A = \mathbb{N}$. Note that this minimal element must be the minimum since the order on $\mathbb{N}$ is linear. Thus, every nonempty subset of $\mathbb{N}$ has a minimum.

What made weak induction so reasonable? We pointed out that it was really the construction of $\mathbb{N}$ that made this idea so natural. What we need is

a base case- that is 0 - and step by step process in building $\mathbb{N}$. We will now show that if a set is defined *recursively*, these notions are just as natural. This is *not* the most general notion of recursively defined sets, but the others (especially those involving what is known as transfinite induction) are beyond the scope of this course.

Definition 47. Suppose V is a set. Then we say that $A \subseteq V$ is *inductively defined* or equivalently, *recursively defined* by G if G is a function $G : \mathbb{N} \to V$ and A satisfies:

- $G(0) \in A$.
- $G(n) \in A \Rightarrow G(n+1) \in A$.
- A is minimal such. (i.e. A is minimal as an element of the set $\mathcal{P}(V)$ with the order $\subseteq$ satisfying the above two conditions.)

This definition doesn't say much. The first condition states that we have a base case. The second condition guarantees that there is a reasonable way to move from one element of A to another. The third condition just states that all elements of A can be gotten to using this step by step process starting from the base case. This definition should make it clear how to use induction on an inductively defined set. Let us look at some examples where viewing the set as inductively defined instead of resorting to putting it in the context of traditional induction is useful.

Proposition 48. *Suppose $A \in V$ is an inductively defined set and $G : \mathbb{N} \to V$ is an injective function witnessing this. Suppose $B \subseteq A$ satisfies the following two conditions:*

- $G(0) \in B$.
- *If $G(a) \in B$ then $G(a+1) \in B$.*

Then $B = A$.

Proof. Use induction principle 1 on the set $\{n \in \mathbb{N} : G(n) \in B\}$. It is an easy exercise to see that this is all of $\mathbb{N}$. Thus, B is a subset of A satisfying the first two conditions of the definition of A being inductively defined set. Since A is minimal, we conclude that $B = A$. □

Example: Consider the set $A \subseteq \mathbb{N}$ where $A = \{5, 6, 7, \dots\}$. Consider $G : \mathbb{N} \to \mathbb{N}$ where $G(n) = n + 5$. Now let's check that A satisfies the conditions.

- $G(0) \in A$: $G(0) = 5 \in A$
- By definition of A, if $G(n) \in A$, then so is $G(n) + 1$. Since $G(n) + 1 = n + 5 + 1 = n + 1 + 5 = G(n + 1)$, we satisfy the second condition.
- Suppose $B \subsetneq A$. Let n be the minimum of the set $\{m \in A : m \notin B\}$. If $n = 5$, then B does not satisfy the first bullet. So we may assume that $n \geq 6$ and $n - 1 \in A$. Since n is the minimum of the set $A \setminus B$, $n-1 \in A \Rightarrow n-1 \in B$. $n-1 = G(n-6) \in B$, and $G(n-5) = n \notin B$. Thus, B does not satisfy the second bullet.

Example: Sometimes we actually want to define a set using an inductive definition. Namely, we give a function G and tell you how to construct $G(0)$ and then how to construct $G(n + 1)$ given $G(n)$. This sort of definition actually sets us up nicely for induction. For example, we could define the set of even numbers as follows: $G(0) = 0$ and $G(n + 1) = G(n) + 2$. Then if you want to apply induction, it's quite clear how to use the information at the n^{th} stage to proceed to the $(n + 1)^{\text{th}}$ stage. This will be very useful in examples in the next two sections.

Example: Let us look at a slightly more complicated example. Suppose I wanted to prove something about all finite subsets of $\mathbb{N}$. Well, consider the equivalence relation on $\mathbb{N}$, given by $\sim$. Then, two finite subsets of $\mathbb{N}$ are equivalent if they have the same size. The finite subsets of $\mathbb{N}$ are inductively defined by G where G is the function $G : \mathbb{N} \to \mathcal{P}(\mathbb{N})/\sim$ where $G(n)$ is the equivalence class of subsets of size n. This is incredibly useful if we are trying to prove something about the finite subsets of $\mathbb{N}$ based only on their size.

The definition of an inductively defined set (namely, the minimality condition) immediately gives us the following conclusion. Aside from the intuition, notice that we would obtain the same conclusion if we applied induction to the domain of G instead of the image which is A, i.e. apply induction to the subset $X \subseteq \mathbb{N}$ where $X = \{x \in \mathbb{N} : G(x) \in B\}$ where B is a subset of A. If you conclude that $X = \mathbb{N}$, then $B = A$.

We will now go through some examples of how to use induction in a proof. Proofs by weak or strong induction are not usually written in the format of constructing a subset of $\mathbb{N}$ or a subset of an inductively defined set and then proving it is the entire set. If we are proving that $\mathbb{N}$ or an inductively defined set has a certain property, the implied subset is the one containing all the elements that satisfy the property. We first prove the base case, i.e. that 0 or $G(0)$ satisfies the property. We then assume for purposes of induction that the property is true on the elements either up to $n+1$ or $G(n+1)$ if we are using strong induction or simply true of n and $G(n)$ for weak induction, and we show the property to be true of $n+1$ or $G(n+1)$. Then, having satisfied the conditions, we conclude that $\mathbb{N}$ or the inductive set has the property.

Claim 49. *For every $m \in \mathbb{N}$,*

$$\sum_{i=0}^{m} i = \frac{m(m+1)}{2}$$

Proof. We will use weak induction to prove this claim.

Base case: 0. Then we have

$$\sum_{i=1}^{0} i = 0 = \frac{0(0+1)}{2}$$

and the claim is true.

Inductive step: Suppose the claim is true for some m.

Now we need to show the claim is true for m+1,i.e. that $\sum_{i=0}^{m+1} i = \frac{(m+1)(m+2)}{2}$:

$$\sum_{i=0}^{m+1} i = (\sum_{i=0}^{m} i) + (m+1)$$

We assumed for purposes of induction that

$$\sum_{i=0}^{m} i = \frac{m(m+1)}{2}$$

, so we have that

$$\sum_{i=0}^{m+1} i = \frac{m(m+1)}{2} + (m+1)$$

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

and by induction, the claim is true for all natural numbers.

□

Proposition 50. *Suppose $A \subseteq \mathbb{N}$ is a bounded set. Then A is finite. Thus, $A \subseteq \mathbb{N}$ is finite if and only if it is bounded.*

Proof. We will induct on n such that $A \subseteq \mathbb{N}^{<n}$. In this case, notice that the set of sets $\{\mathbb{N}^{<m} : m \in \mathbb{N}\}$ is an inductively defined set. So by inducting on the index, we are in fact using the proposition that induction applies to inductively defined sets by the same rationale as it applied to $\mathbb{N}$.

n=0: Suppose $A \subseteq \mathbb{N}^{<0}$. Then $A \subseteq \emptyset$ and thus $A = \emptyset$ and is finite.

Assume the claim for all subsets of $\mathbb{N}^{<k}$.

n=k+1: Now suppose that $A \subseteq \mathbb{N}^{<k}$. Then, if $A = \mathbb{N}^{<k}$, then A is finite. If $k-1 \notin A$, then $A \subseteq \mathbb{N}^{<k-1}$ and it is finite. So assume that $k-1 \in A$. Then consider $A \setminus \{k-1\}$. $A \setminus \{k-1\} \subseteq \mathbb{N}^{<k-1}$ and is thus finite. Therefore $A \setminus \{k-1\}$ is in bijection with $\mathbb{N}^m$ for some $m \in \mathbb{N}$ and it is easy to see that if we take this bijection union $(k-1, m)$ this will be a bijection from A to $\mathbb{N}^{m+1}$.

□

Theorem 51 (Fundamental Theorem of Arithmetic)**.** *Every natural number greater than 1 is a product of prime numbers, i.e. if $n > 1$, then there are $p_1, ..., p_k$ prime numbers such that $p_1 p_2, ..., p_k = n$. (This factorization is also unique, though we will not prove this.)*

Proof. We will prove this by strong induction, but we will need a definition first: Recall that a natural number is prime if it has exactly 2 divisors, 1 and itself. (1 is therefore *not* considered prime.)

Assume for purposes of induction that the claim is true for all natural numbers $< n$.

Now we need to prove the claim for n. We have 2 cases. Either n is prime, or $n = k \cdot l$ where $k, l < n$.

Case 1: If n is prime, then n is a product of prime numbers, namely $n = n$.

Case 2: If $n = k \cdot l$, then by induction, k and l have prime factorizations, namely $k = p_1, ..., p_s$ and $l = q_1, ..., q_t$ where p_i and q_i are

prime numbers. Therefore $n = p_1...p_s q_1....q_t$ and n is also a product of prime numbers. □

We will need the following lemma to prove a theorem:

Lemma 52. *Suppose A is a finite set and $a \in A$. Then $|\{B \in \mathcal{P}(A) : a \in B\}| = |\mathcal{P}(A \setminus \{a\})|$. Namely, there are the same number of subsets of A which contain a as those which do not.*

Proof. It is enough to give a bijection $F : \mathcal{P}(A \setminus \{a\}) \to \{B \in \mathcal{P}(A) : a \in B\}$. For $B \in \mathcal{P}(A \setminus \{a\})$, let $F(B) = B \cup \{a\}$.

F is injective: Suppose $C, D \in \mathcal{P}(A \setminus \{a\}$. Suppose $C \neq D$. Then either there is a $b \neq a$ such that $b \in C$ and $b \notin D$ or there is a $b \neq a$ such that $b \in D$ and $b \notin C$. Then $C \cup \{a\} \neq D \cup \{a\}$ since b is in one and not the other and $F(C) \neq F(D)$.

F is surjective: Suppose B is such that $a \in B$. Then $a \notin B \setminus \{a\}$ and it is clear that $F(B \setminus \{a\}) = B$.

□

Theorem 53. *Let A be a finite set. Then $\mathcal{P}(A)$ has $2^{|A|}$ elements.*

Proof. We will prove this by induction on the size of A.

Base case: A has 0 elements, i.e. $A = \emptyset$. Then $\mathcal{P}(A) = \{\emptyset\}$, since the only subset of $\emptyset$ is $\emptyset$. So $|\mathcal{P}(A)| = 1 = 2^0$.

Assume for purposes of induction that for all finite sets of size n

the claim is true, i.e. if $|A| = n$ then $|\mathcal{P}(A)| = 2^n$.

Now consider a set A where $|A| = n + 1$. Let $a \in A$, and consider $A \setminus \{a\}$. Then $|A \setminus \{a\}|$ has size n. (This is an exercise.)

Let B be a subset of A. Then either $a \in B$ or $B \subseteq A \setminus \{a\}$ and not both. Thus $\mathcal{P}(A) = \{B \in \mathcal{P}(A) : a \in B\} \cup \{B \in \mathcal{P}(A) : B \subseteq A \setminus \{a\}$. and it is clear that these two sets are disjoint. Thus, by the claims earlier in this

section, the size of the union is the sum of the sizes of each of the sets and each of these sets have the same size. So

$$|\mathcal{P}(A)| = |\mathcal{P}(A \setminus \{a\})| + |\{B \cup \{a\} : B \in \mathcal{P}(A \setminus \{a\})\}|$$
$$= 2|\mathcal{P}(A \setminus \{a\})| = 2 \cdot 2^n = 2^{n+1}$$

.

□

Definition 54. For A and B sets, A^B is the set of all functions from B to A.

Traditionally we write 2^A instead of $\{0,1\}^A$.

Proposition 55. *Let A be a set. Then $\mathcal{P}(A) \sim 2^A$.*

Proof. Let A be a set. For every $B \subseteq A$, let $\chi_B : A \to \{0,1\}$ be the map

$$\chi_B(x) = \begin{cases} 1 & \text{if } x \in B \\ 0 & \text{if } x \in A \setminus B \end{cases}$$

This function is called the *characteristic function* of B. Now consider the map $h : \mathcal{P}(A) \to 2^A$ where for $B \subseteq A$, $h(B) = \chi_B$.

We claim this function is a bijection.

h is injective: To see this suppose $B_1, B_2 \subseteq A$ and $B_1 \neq B_2$. Then either there is $a \in A$ such that $a \in B_1$ and $a \notin B_2$ or there is an $a \in A$ such that $a \in B_2$ and $a \notin B_1$. Without loss of generality, assume there is $a \in A$ such that $a \in B_1$ and $a \notin B_2$. Then, since $\chi_{B_1}(a) = 1$ and $\chi_{B_2}(a) = 0$, $h(B_1) \neq h(B_2)$. Therefore h is injective.

h is surjective: Suppose $f : A \to \{0,1\}$. Then consider the set $B = \{a \in A : f(a) = 1\}$. This is a subset of A and $f = \chi_B$. □

To see the beauty of this notation, recall we proved that for A finite, $|\mathcal{P}(A)| = 2^{|A|}$ so $|2^A| = 2^{|A|}$.

EQUIVALENT STATEMENTS REVISITED:

The following will be left as an exercise in truth tables: Show that $[(P \Rightarrow Q) \wedge (Q \Rightarrow R)] \Rightarrow (P \Rightarrow R)$.

We previously discussed what it means for two mathematical statements to be equivalent. Namely, P and Q are equivalent if $P \iff Q$. First consider three statements (or as many as you like) P, Q, and R. Then P, Q and R are equivalent if they are pairwise equivalent, namely $P \iff Q$, $Q \iff R$, and $R \iff P$. However, we don't actually need to prove all of this. It is enough to show that we can create a "circle of implication". Namely, $P_1, \ldots, P_n$ are equivalent if $P_1 \Rightarrow P_2, \ldots, P_{n-1} \Rightarrow P_n$, and $P_n \Rightarrow P_1$.

Using this logic, we will prove that the weak induction principle, strong induction principle, and the well ordering principle are all equivalent. We've already shown that the weak induction principle implies the strong induction principle, and the strong induction principle implies the well ordering principle. All that remains is to show that the well ordering principle implies the weak induction principle.

Claim 56. *Assume that if $A \subseteq \mathbb{N}$ is nonempty, then A has a minimal element. Now suppose that $B \subseteq \mathbb{N}$ and B satisfies the following two conditions:*

- $0 \in B$.
- *If $a \in B$ then $a + 1 \in B$.*

Then $B = \mathbb{N}$.

Proof. We will show that the well ordering principle implies the contrapositive of the weak induction principle. That is, if we assume the well ordering principle, then $B \neq \mathbb{N}$ implies that B cannot satisfy both of the conditions.

Suppose $B \neq \mathbb{N}$ and assume that $0 \in B$. (If we assume that $0 \notin B$ then B doesn't satisfy the first condition and we are done in this case.) Then, there is a $m \in \mathbb{N} \setminus B$ minimal by the well-ordering principle, and $m > 0$ since $0 \in B$. Then $m \notin B$ but $m - 1 \in B$ and B does not satisfy the second condition.

□

EXERCISES:

1. Let $\mathcal{C}$ be a set of sets. Show that $\sim$ is an equivalence relation on $\mathcal{C}$.

2. Let A be a finite set and $a \in A$. Then $|A \setminus \{a\}| = |A| - 1$.

3. Prove the pigeonhole principle. (Hint: use induction on n)

4. Prove using the pigeonhole principle that any injective function from $\mathbb{N}^{<n} \to \mathbb{N}^{<n}$ is surjective.

5. Prove that the sets of multiples of 6, i.e. $\{6, 12, 18, ...\}$ is an inductively defined set. (Give the function G and show it satisfies the conditions.) Prove by induction that the sum of the first m multiples of 6 is $3n(n+1)$.

6. Let $\langle A, \leq \rangle$ be a finite partial order. Show that for every a in A there is a minimal $x \leq a$. (Hint: Use induction on the size of A.)

7. Prove that if A and B are finite, then $A \times B$ is finite.

8. Using truth tables, show that $[(P \Rightarrow Q) \wedge (Q \Rightarrow R)] \Rightarrow (P \Rightarrow R)$. Show that the circle of implication $(P \Rightarrow Q)$ and $(Q \Rightarrow R)$ and $(R \Rightarrow P)$ implies that P, Q, R are pairwise equivalent.

9. In the proof that the well-ordering principle implies weak induction, we made use of a subtle point. That is that $(P \wedge Q) \Rightarrow R$ is equivalent to $(\neg R \wedge P) \Rightarrow \neg Q$. Prove this equivalence.

10. BONUS: Prove that the function $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ where $f((a, b)) = \frac{(a+b)(a+b+1)}{2} + a$ is a bijection.

11. BONUS:

    Recall that every element of the real numbers can be uniquely given in decimal form $n + \sum a_i 10^{-i}$ where $n \in \mathbb{Z}, a_i \in \{0, 1, ..., 9\}$ and for all $m \in \mathbb{N}$,there is $k > m$ such that $a_k \neq 9$. (The sequences that are eventually 9 are equivalent to ones which are not eventually 9.) We know that $\mathbb{Q}$ is dense in $\mathbb{R}$, i.e., for any r_1, r_2 in $\mathbb{R}$ where $r_1 < r_2$, there is $q \in \mathbb{Q}$ such that $r_1 < q < r_2$.

    - Show that for any real number r, the set of rational numbers less than r is unique to r, namely that if $r_1 \neq r_2$, then $\{q \in \mathbb{Q} : q < r_1\} \neq \{q \in \mathbb{Q} : q < r_2\}$.
    - Give an injective function from $\mathbb{R} \to 2^{\mathbb{Q}}$. Give an injective function from $2^{\mathbb{N}} \to \mathbb{R}$. This will be very useful later.

8 Counting Principles

In the previous section, we learned how to determine the size of finite sets using bijections. In this section, we will learn that not all infinite sets have the same size. At this point, it is worth mentioning again that we assume the Axiom of Choice throughout this section. These concepts are difficult to absorb, so we will avoid complicating the discussion with formal discussions on set theory. However, the reader ought to keep in mind that these concepts are not so simple and that the effects of choice on mathematics is a topic studied to this day.

Let us begin with the following notation for two sets A and B:

$$A \sim B : \text{There is a bijection from } A \text{ to } B$$
$$A \precsim B : \text{There is an injection from } A \text{ to } B$$

It is clear that $A \sim B \Rightarrow A \precsim B$. As the notation suggests, we would hope that if $A \precsim B$ and $B \precsim A$ then $A \sim B$.

Theorem 57 (Cantor-Berenstein Theorem)**.** *Let A and B be sets. If there are injections $g : A \to B$ and $h : B \to A$, then there is a bijection $f : A \to B$.*

First we will be begin with a subclaim.

Subclaim 58. *Suppose $X \subseteq Y$ and there is a function $f : Y \to X$ that is injective. Then $X \sim Y$.*

Proof. We will inductively define sets E_n.

$E_0 = Y \setminus X$. Given E_i, we define

$$E_{i+1} = f[E_i] = \{f(a) : a \in E_i\}$$

.

Let $E = \bigcup_{i=0}^{\infty} E_i$. We define the function $h : Y \to X$ as follows:

$$h(x) = \begin{cases} f(y) & \text{if } y \in E \\ y & \text{if } y \in Y \setminus E \end{cases}$$

We claim that h is a bijection which will finish this subclaim. Because and element of Y is either in E or it isn't, it is clear that h is a function from Y to Y.

- We will show first that the range is actually a subset of X. If $y \in E$, then $f(y) \in \text{image}(f) \subseteq X$. If $y \in Y \setminus E$, then since $E_0 = Y \setminus X$, $Y \setminus E \subseteq X$.

- h is injective: Suppose $y_1 \neq y_2 \in Y$. Then there are four options: either $y_1, y_2 \in E$, $y_1, y_2 \in Y \setminus E$, $y_1 \in E$ and $y_2 \in Y \setminus E$, or $y_1 \in Y \setminus E$ and $y_2 \in E$. Clearly the last two cases are similar.

  1. $y_1, y_2 \in E$: Then $h(y_1) = f(y_1)$ and $h(y_2) = f(y_2)$. Since f is injective, $y_1 \neq y_2$ implies that $f(y_1) \neq f(y_2)$ and thus $h(y_1) \neq h(y_2)$.
  2. $y_1, y_2 \in Y \setminus E$: Then $h(y_1) = y_1$ and $h(y_2) = y_2$. So $y_1 \neq y_2$ implies that $h(y_1) \neq h(y_2)$.
  3. $y_1 \in E$ and $y_2 \in Y \setminus E$. So $h(y_1) = f(y_1)$ and $h(y_2) = y_2$. $y_1 \in E$ so there is an $n \in \mathbb{N}$ such that $y_1 \in E_n$. Then $f(y_1) \in E_{n+1}$ and $f(y_1) \in E$. Therefore, since $y_2 \notin E$, we conclude that $f(y_1) \neq y_2$ and thus $h(y_1) \neq h(y_2)$.

- h is surjective: Suppose $x \in X$. Then if $x \in Y \setminus E$, then $h(x) = x$ and $x \in \text{image}(h)$. Now suppose $x \in E$. Then $x \in E_n$ for some $n \in \mathbb{N}$. Since $E_0 = Y \setminus X$, $x \notin E_0$, so $n > 0$. $x \in E_n$ for $n > 0$ implies, by definition of E_n for $n > 0$, that there is a $y \in E_{n-1}$ such that $x = f(y)$. Since $y \in E_{n-1}$, $y \in E$ and therefore $h(y) = f(y) = x$ and $x \in \text{image}(h)$.

This concludes the proof of the subclaim. We now return to the proof of the theorem.

Suppose A and B are sets and there are injections $g : A \to B$ and $h : B \to A$. Let $Y = A$ and $X = h(B) \subseteq Y$. Then $h \circ g$ is a composition of two injective functions and is thus injective. (This was an exercise.) Since $X \subseteq Y$ and $h \circ g : Y \to X$ is injective, we conclude that $X \sim Y$,i.e. that $A \sim h(B)$. Now consider the function $h' : B \to h[B]$ where $h'(x) = h(x)$ for all $x \in B$. (We are just restricting the range to the image.) Then h' is

surjective and $B \sim h(B)$. Thus, since $\sim$ is both symmetric and transitive (this was also an exercise), $A \sim B$.

□

The following useful observation is left as an exercise:

Claim 59. *Suppose A and B are sets and there is a surjection from A to B. Then $B \precsim A$.*

Proof. Exercise

□

At this point in our mathematical excursions, we need to contemplate how we view sizes of sets. In the previous section, we showed that when we say a set has size 3, the precise meaning of this is that there is a bijection from the set to $\mathbb{N}^{<3}$. Now we will start analyzing how we view sizes of infinite sets.

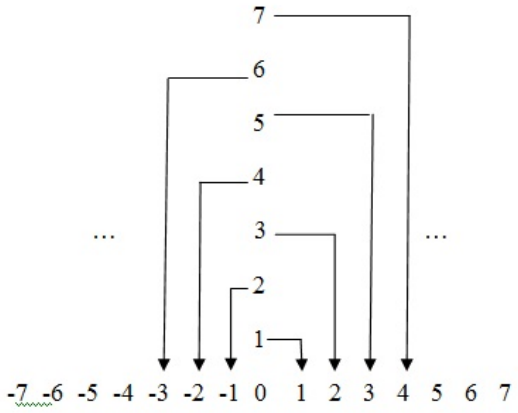
Definition 60. Let A be a set. We say A is *countable* if A is finite or $A \sim \mathbb{N}$. If $A \sim \mathbb{N}$ we say that A is *countably infinite* or $|A| = \aleph_0$ and say that the size of A is $\aleph_0$.

EXAMPLES:

- $\mathbb{N}$ is clearly countable since the identity map is a bijection.
- $2\mathbb{N} = \{0, 2, 4, 6, ...\}$ is countable. To see this consider the map $f : \mathbb{N} \to 2\mathbb{N}$ where $f(x) = 2x$. It is easy to prove this map is bijective.
- $\mathbb{Z}$ is countable! Consider the map $f : \mathbb{N} \to \mathbb{Z}$ as follows:

$$f(x) = \begin{cases} \frac{x+1}{2} & \text{if } x \text{ is odd} \\ \frac{-x}{2} & \text{if } x \text{ is even} \end{cases}$$

  Here is a picture of the map:



We now want to verify that an infinite subset of $\mathbb{N}$ is the same size as $\mathbb{N}$ - that is that $\aleph_0$ is the smallest size of infinite sets.

Proposition 61. *Let A be a set. Then A is countable if and only if $A \precsim \mathbb{N}$.*

Proof. ($\Rightarrow$) If A is countable then A is finite or $A \sim \mathbb{N}$. If A is finite, then there is a bijection $f : A \to \mathbb{N}^{<n}$ for some $n \in \mathbb{N}$. So the map $g : A \to \mathbb{N}$ where $g(a) = f(a)$ for all $a \in A$ is an injection from A to $\mathbb{N}$ and $A \precsim \mathbb{N}$. If $A \sim \mathbb{N}$ then as we saw earlier $A \precsim \mathbb{N}$.

($\Leftarrow$) If $A \precsim \mathbb{N}$, then there is an injection $f : A \to \mathbb{N}$. We now divide the proof into two cases. image(f) is bounded and image(f) is unbounded.

- image(f) is bounded. Then, as we proved earlier, the image(f) is finite and there is an $n \in \mathbb{N}$ and a bijection $g : \text{image}(f) \to \mathbb{N}^{<n}$. Then $g \circ f$ is a bijection from $A \to \mathbb{N}^{<n}$ and A is finite. Thus A is countable.

- image(f) is unbounded. Let $B = \text{image}(f)$ and consider the following function $g \subset \mathbb{N} \times B$:

  $$g(i) = \min(B \setminus \{g(0), ..., g(i-1)\})$$

  .

  That this map is a function is clear as the minimum is unique. Furthermore, this function is defined inductively, so if $n \in \text{dom}(g)$, then

$m \in \text{dom}(g)$ for all $m < n$. If domain of g is $\mathbb{N}^{<n}$, then this would imply that B is contained in $\mathbb{N}^{\leq g(n-1)}$ and is thus bounded. Therefore the domain of g is $\mathbb{N}$ because B is unbounded. Thus $g : \mathbb{N} \to B$. We now prove that it is a bijection.

g is injective: Notice that $g(i) < g(j)$ for $i < j$ since $g(i)$ is the minimum of a set including $g(j)$ and $g(j)$ is a minimum of a set not including $g(i)$. This in particular implies that g is injective.

g is surjective: Suppose $n \in B$. Then, if $n \notin \{g(0), ..., g(n)\}$, then $g(0), ..., g(n) < n$ since these are minima of sets including n. That would imply that there are $n+1$ elements of $\mathbb{N}$ less that n. Contradiction. Thus $n \in \{g(0), ..., g(n)\}$. So g is surjective.

Thus g is a bijection from $\mathbb{N}$ to image(f) and $g^{-1} \circ f : A \to \mathbb{N}$ is also a bijection and $A \sim \mathbb{N}$.

□

First we will see many examples of countable sets.

Proposition 62. *Suppose A is countable and $A' \subseteq A$. Then A' is countable.*

Proof. We showed that A is countable if and only if $A \precsim \mathbb{N}$. Let $f : A \to \mathbb{N}$ be the injection witnessing this. Now take the restricted function $f|_{A'} : A' \to \mathbb{N}$. This is still an injection and therefore $A' \precsim \mathbb{N}$ and A' is countable.

□

Lemma 63. *Suppose $A_1, ..., A_n$ are finite sets. Then $\bigcup_i^n A_i$ is finite. In other words, the finite union of finite sets is finite.*

Proof. Exercise. □

Lemma 64. *Suppose A is finite and B is countable. Then $A \cup B$ is countable.*

Proof. Exercise. □

Proposition 65. *The finite union of countable sets is countable.*

Proof. Suppose $A_0, ..., A_{n-1}$ are countable. Since the finite union of finite sets is finite and the union of a finite set and a countable set is countable, without loss of generality, we may assume that $A_0, ..., A_{n-1}$ are countably infinite. We will first prove the special case where $A_1, ..., A_n$ are pairwise disjoint.

Suppose $A_0, ..., A_{n-1}$ are pairwise disjoint and countably infinite. Then there are $f_0, ..., f_{n-1}$ bijections where $f_i : A_i \to \mathbb{N}$. The picture for this will look similar to the picture we used to count $\mathbb{Z}$. Now consider $\mathbb{N} \equiv_{mod n}$. This is the set of equivalence classes $[0], ..., [n-1]$. Each of these classes is an unbounded subset of $\mathbb{N}$. Therefore, they are all countably infinite and $\mathbb{N} \sim [i]$ for each $0 \leq i \leq n-1$. Let $g_i : \mathbb{N} \to [i]$ be bijections that witness this. Then since $\cup_{i=1}^{n-1}[i] = \mathbb{N}$ we consider the following map:

$$h = \bigcup_{i=1}^{n-1}(g_i \circ f_i) : \bigcup_{i=1}^{n-1} A_i \to \mathbb{N}$$

For each i, $g_i \circ f_i$ is a bijection from $A_i \to [i]$, and since the domains are DISJOINT and the images are DISJOINT, and the union of the images is $\mathbb{N}$, h is a bijection.

(WORD OF WARNING: the arbitrary unions of bijections are not necessarily bijections. Think about this. Think of some examples. Now you know why the word "disjoint" is highlighted.)

Now we will deal with the case that the A_i are not necessarily disjoint. Consider $B_0, ..., B_{n-1}$ defined as follows:

$$B_0 = A_0$$

$$B_i = A_i \setminus \bigcup_{j \in \{0,...,i-1\}} B_j$$

Subclaim: $\bigcup A_i = \bigcup B_i$

Proof: If $a \in \bigcup A_i$, then either $a \in A_m$ for some m and therefore $a \in B_m$ or $a \in B_k$ for some $k < m$. Thus $a \in \bigcup B_i$ and $\bigcup A_i \subseteq \bigcup B_i$. $B_i \subseteq A_i$ so $\bigcup B_i \subseteq \bigcup A_i$. Thus $\bigcup A_i = \bigcup B_i$.

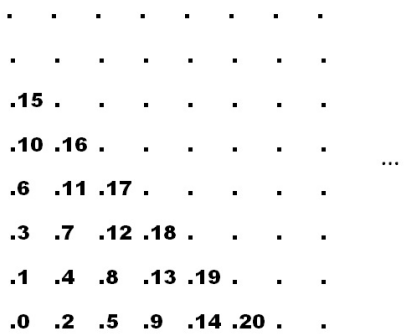
Subclaim: $B_i \cap B_j = \emptyset$ for $i \neq j$.

Proof: If $i \neq j$ then without loss of generality, we may assume $i > j$. Then, $a \in B_i$ implies that $a \notin B_j$.

If some of the B_i are finite, we may take $C = \bigcup\{b \in B_i : B_i \text{ is finite and } 0 \leq i \leq n-1\}$. Then, since C is a finite union of finite sets, C is finite. If we take $D = \bigcup\{b \in B_i : B_i \text{ is infinite and } 0 \leq i \leq n-1\}$, then D is the finite union of countably infinite disjoint sets and it is thus countable. $\bigcup B_i = C \cup D$ and since a union of a countable set and a finite set is countable, $\bigcup B_i$ is countable and thus $\bigcup A_i$ is countable. □

Theorem 66. $\mathbb{N} \times \mathbb{N}$ *is countable. Therefore, the Cartesian product of two countable sets is countable.*

Proof. You actually already proved the first part in the bonus question! You showed that the function $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ where $f((a,b)) = \frac{(a+b)(a+b+1)}{2} + a$ is a bijection. To understand this intuitively, consider the following picture:



For the second part, consider the following fact which you will prove as an exercise: If $A \precsim B$ and $C \precsim D$ then $A \times C \precsim B \times D$.

Therefore, if A and C are countable, then $A \times C \precsim \mathbb{N} \times \mathbb{N} \sim \mathbb{N}$.

□

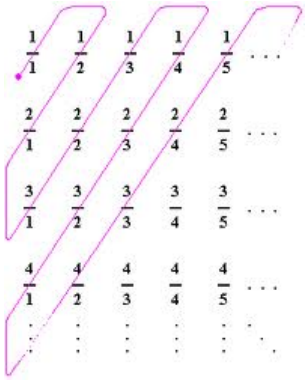
We will now prove that various sets are countable by using the following useful observation: We don't always need to find a bijection to $\mathbb{N}$ to prove that an infinite set is countable. It is enough to find an injection from the set to any other set we know is countable since compositions of injections are injections. It also suffices to find a surjection from a countable set to the set

in question as this is also equivalent to the set being countable as you will prove in the exercises.

Corollary 67. $\mathbb{Q}$ *is countable!*

Proof. For any fraction $\frac{a}{b}$, let a' in $\mathbb{Z}$ and b' in $\mathbb{N} \setminus \{0\}$ such that $\frac{a}{b} = \frac{a'}{b'}$ and a', b' are relatively prime. a' and b' are unique to any fraction. (Recall that every natural number divides zero, so the reduced form of 0 is $\frac{0}{1}$). So we will take $\mathbb{Q}$ to be the set of reduced fractions. (This is equivalent to the construction you did at the end of section 6 where we are taking specific representatives of the equivalence classes.) Since $\mathbb{Z} \times \mathbb{N} \setminus \{0\}$ is the product of two countable sets, it is countable. Now consider the function $f : \mathbb{Q} \to \mathbb{Z} \times \mathbb{N} \setminus \{0\}$, where $f(\frac{a'}{b'}) = (a, b)$. This function is an injection, and thus $\mathbb{Q}$ is countable.

Since there is an injection from $\mathbb{Q}$ to a countable set, it is countable. Here is a nice picture that might help you to imagine a surjective (but not injective) function from $\mathbb{N}$ to the positive rational numbers.



□

Corollary 68. *The countable union of countable sets is countable. Namely, if we have a countable set of sets*

$$A = \{A_n | n \in \mathbb{N}\}$$

and each $A_i \in A$ *is countable, then the union* $\bigcup_{n \in \mathbb{N}} A_n$ *is countable.*

Proof. We will construct $h : \mathbb{N} \times \mathbb{N} \to \bigcup_{n\in\mathbb{N}} A_n$ that is surjective. Each A_i is countable, so fix a $g_i : \mathbb{N} \to A_i$ surjective. (Our ability to choose these functions follows from the axioms of choice.) Now define h as follows:

$$h((a, b)) = g_a(b)$$

We claim that this is surjective. Let $a \in \bigcup_{n\in\mathbb{N}} A_n$. Then $a \in A_i$ for some $i \in \mathbb{N}$. Therefore, $a \in \text{image}(g_i)$ and thus there is some $m \in \mathbb{N}$ such that $a = g_i(m)$. So $h((i, m)) = a$. □

Definition 69. Let $A_1, ..., A_n$ be nonempty sets. Then we may extend the definition of Cartesian product to $A_1 \times \cdots \times A_n$ to be the set

$$\{(a_1, ..., a_n) : a_1 \in A_1, \ldots a_n \in A_n\}$$

The elements of these sets are called n-tuples. A 2-tuple is just an ordered pair. The Cartesian product of the empty set with any other set is clearly empty. So if A_j is empty for any $1 \leq j \leq n$, then $A_1 \times ... \times A_n = \emptyset$.

Theorem 70. *If $A_1, \ldots, A_n$ are countable sets, then $A_1 \times \cdots \times A_n$ is countable.*

Proof. Exercise. □

We now have an easy corollary to theorem 70 and corollary 68.

Corollary 71. *Let $Fin(\mathbb{N})$ be the set of finite sequences of natural numbers. Then $Fin(\mathbb{N})$ is countable.*

Proof.

$$Fin(\mathbb{N}) = \bigcup_{m\in\mathbb{N}} \underbrace{\mathbb{N} \times \cdots \times \mathbb{N}}_{m-times}$$

and it is therefore a countable union of countable sets and is thus countable. □

We now move on to prove that not all sets are countable!

Theorem 72. $2^{\mathbb{N}}$ *is NOT countable.*

Proof. We will show first an illustration of Cantor's beautiful diagonalization proof: Every element of $2^{\mathbb{N}}$ can be represented as a sequence of 0's and 1's. We will show that for any attempt to enumerate them, we can find a sequence that is not on our list!

Let the following be any enumeration of elements $2^{\mathbb{N}}$:

$$(\mathbf{0}, 1, 1, 0, 0, 1, 0, 0, 1, 1, 1, \dots)$$

$$(0, \mathbf{1}, 1, 1, 0, 0, 0, 0, 1, 1, 1, \dots)$$

$$(1, 1, \mathbf{1}, 0, 0, 1, 0, 0, 1, 0, 1, \dots)$$

$$(1, 1, 0, \mathbf{0}, 0, 1, 0, 0, 1, 1, 1, \dots)$$

$$(0, 0, 1, 0, \mathbf{1}, 1, 0, 0, 1, 1, 1, \dots)$$

$$(1, 1, 1, 1, 1, \mathbf{1}, 1, 1, 1, 1, 1, \dots)$$

$$(0, 1, 1, 0, 0, 1, \mathbf{1}, 0, 1, 1, 0, \dots)$$

$$(1, 1, 1, 0, 1, 1, 0, \mathbf{0}, 1, 1, 1, \dots)$$

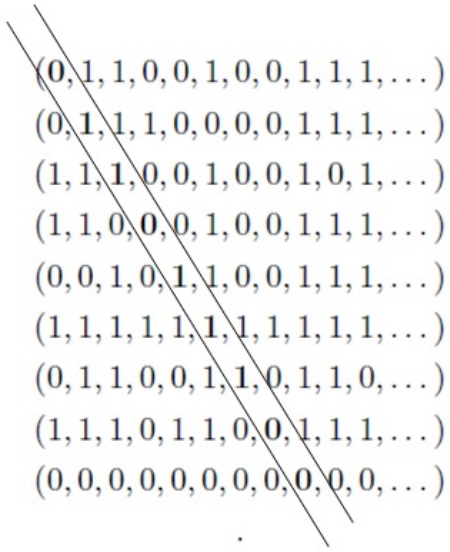
$$(0, 0, 0, 0, 0, 0, 0, 0, \mathbf{0}, 0, 0, \dots)$$

$$.$$

$$.$$

$$.$$

We now isolate the diagonal sequence:



Now swap the 0's and 1's of this sequence. This new sequence of 0's and 1's disagrees with the first sequence at the first element, the second sequence on the second element, and so on. So it is NOT on my original list and the enumeration was not surjective! Now let's prove this formally:

Suppose $f : \mathbb{N} \to 2^{\mathbb{N}}$. Then consider the following $g : \mathbb{N} \to \{0, 1\}$ defined as follows:

$$g(n) = \begin{cases} 0 & \text{if } (f(n))(n) = 1 \\ 1 & \text{if } (f(n))(n) = 0 \end{cases}$$

Then we claim that $g \neq f(n)$ for any n. This is clear since $(f(n))(n) \neq g(n)$. Thus f is not surjective. □

Corollary 73. $\mathbb{R}$ *is not countable.*

Proof. You already showed that $\mathbb{R} \precsim 2^{Q}$ and that $2^{\mathbb{N}} \precsim \mathbb{R}$. Since $Q \sim \mathbb{N}$, we have that $\mathbb{R} \precsim 2^{\mathbb{N}}$ and that $2^{\mathbb{N}} \precsim \mathbb{R}$. (This is an exercise.) Thus, by Cantor-Berenstein, $\mathbb{R} \sim 2^{\mathbb{N}}$. Since $2^{\mathbb{N}} \nsim \mathbb{N}$, $\mathbb{R} \nsim \mathbb{N}$, and since $\mathbb{R}$ is not finite, we conclude that $\mathbb{R}$ is not countable. □

Cantor generalized his diagonalization proof to show that we can always make larger sets:

Theorem 74 (Cantor's Theorem)**.** *Let A be any set. Then $A \precnsim \mathcal{P}(A)$.*

Proof. Let $f : A \to \mathcal{P}(A)$ be any function. Then we claim that f is not surjective. Consider the following subset of A:

$$\{x \in A | x \notin f(x)\}$$

Now suppose this set is $f(y)$ for some $y \in A$. Then $y \in f(y) \iff y \notin f(y)$. Thus, we obtain a paradox and this set cannot be in the image of f. $\square$

Cardinals and the Continuum Hypothesis

We have made a brief introduction to *cardinals* which are the possible sizes of sets- and since we use bijections to measure the relative size of a set, $A \sim B$ implies that they have the same *cardinality* or size. We use the notation $|A|$ to refer to the size or cardinality of A. Therefore, when a set is finite and of size n, we say it has *cardinality* n. A set is countable if it is finite or has cardinality $\aleph_0$ which is the size of $\mathbb{N}$. We call the cardinality of $2^{\mathbb{N}}$, $2^{\aleph_0}$. Cantor asked the following question:

Continuum Hypothesis: Is there a set A such that $\aleph_0 < |A| < 2^{\aleph_0}$? In other words, is $\aleph_1 = 2^{\aleph_0}$?

In 1963, Paul Cohen showed that this question is independent of the most widely accepted axioms of set theory known as ZFC.

EXERCISES:

1. Prove that if $A \precsim B$ and $C \precsim D$ then $A \times C \precsim B \times D$
2. Prove claim 59. Now prove that a set A is countable if there is a surjection from a countable set to A.
3. Prove that if $A \sim B$, then $\mathcal{P}(A) \sim \mathcal{P}(B)$ and $2^A \sim 2^B$.
4. Prove lemma 63.
5. Prove lemma 64.
6. Prove that $\mathbb{C}$ is not countable. Prove that $\mathbb{C} \setminus \mathbb{R}$ is not countable. Prove that $\mathbb{R} \setminus \mathbb{N}$ is not countable.
7. We say an element $r \in \mathbb{C}$ is *algebraic* if there is an $n \in \mathbb{N}$ and $a_0, ..., a_n \in \mathbb{Z}$, not all zero, such that $a_n r^n + a_{n-1} r^{n-1} + ... + a_1 r + a_0 = 0$. For example, $\sqrt{2}$ is algebraic since it is a solution to $x^2 - 2 = 0$. Show that the set of algebraic numbers is countable.

8. Prove theorem 70. (Hint: find a bijection from $A_1 \times ... \times A_n$ to $(A_1 \times ... \times A_{n-1}) \times A_n$.)

9. BONUS: Prove that $\mathbb{R} \sim \mathbb{R} \setminus \mathbb{N}$. Prove that $\mathbb{R} \sim (-1, 1)$.

9 First Order Logic

This section is based on "Model Theory: an Introduction" by David Marker. The main ideas that are covered are languages, structures, formulas, sentences, embeddings, and isomorphisms. The approach taken is one aimed at examples and applications that the students will see in future mathematics coursework. There are many approaches to such a system and we forgo a more abstract and formal approach to one aimed at modern model theory.

Definition 75. A language $\mathcal{L}$ is given by specifying the following data:

1. a set of function symbols $\mathcal{F}$ and positive integers n_f for each $f \in \mathcal{F}$;

2. a set of relation symbols $\mathcal{R}$ and positive integers n_R for each $R \in \mathcal{R}$;

3. a set of constant symbols $\mathcal{C}$.

The numbers n_f and n_R tell us that f is a function of n_f variables and R is an n_R-ary relation. Any or all of the sets $\mathcal{F}, \mathcal{R}$, and $\mathcal{C}$ may be empty.

We assume that all languages contain the symbol $=$ as well as the Boolean connectives and quantifiers as we will discuss soon. It is also worth noting that just as we extended the notion of function to the notion of binary function, we can extend it to n-ary by simply taking a function $f : A \times ... \times A \to A$. Furthermore, we can extend the notion of a relation to the notion of an n-ary relation where we allow it to be a subset of $A \times ... \times A$, and the elements are now n-tuples. A relation as we defined it is often called a binary relation.

Examples:

- $\mathcal{L} = \emptyset$: This is the language of pure sets.

- $\mathcal{L} = \{R\}$ where $n_R = 2$: This is the language of graphs or equivalence relations.

- $\mathcal{L} = \{+, \cdot, 0, 1\}$ where $+, \cdot$ are binary $(n., n_+ = 2)$, and $0, 1$ are constants: This is the language of rings and fields.

Definition 76. An $\mathcal{L}$-structure $\mathcal{M}$ is given by the following data:

1. a nonempty set M called the universe, domain, or underlying set of $\mathcal{M}$;
2. a function $f^{\mathcal{M}} : M^{n_f} \to M$ for each $f \in \mathcal{F}$;
3. a set $R^{\mathcal{M}} \subseteq M^{n_R}$ for each $R \in \mathcal{R}$;
4. an element $c^{\mathcal{M}} \in M$ for each $c \in \mathcal{C}$.

For a finite language, we will usually write the structure in angular brackets $\langle M, ... \rangle$ where we give the interpretations of the function symbols, relation symbols, and constant symbols in the same order they are listed in the language.

Let us look at some examples:

- Let $\mathcal{L} = \emptyset$. Then if M is any nonempty set, $\mathcal{M}$ is an $\mathcal{L}$-structure.
- $\mathcal{L} = \{R\}$ where $n_R = 2$. Then we can take $\langle \mathbb{Z}, \equiv_{\bmod n} \rangle$ as an $\mathcal{L}$-structure.
- $\mathcal{L} = \{+, \cdot, 0, 1\}$ where $+, \cdot$ are binary $(n., n_+ = 2)$, and $0, 1$ are constants. Then $\langle \mathbb{R}, +, \cdot, 0, 1 \rangle$ is an $\mathcal{L}$-structure, and so is $\langle \mathbb{R}, +, +, 0, 1 \rangle$, and $\langle \mathbb{R}, +, \cdot, 2, 2 \rangle$.

We now move onto the definitions of $\mathcal{L}$-terms, formulas, and sentences. However, we need to make clear symbols that are present in every language we consider: parentheses $(,)$, the Boolean connectives $\wedge, \vee, \neg$, the symbol $=$, the quantifiers $\exists, \forall$, and a set of variables $V = \{v_1, v_2, ...\}$.

Definition 77. The set of $\mathcal{L}$-terms T is the smallest set satisfying the following:

1. $c \in T$ for all $c \in \mathcal{C}$.

2. $v_i \in T$ for all $v_i \in V$.

3. If $t_1, ..., t_{n_f} \in T$, then $f(t_1, ..., t_{n_f}) \in T$.

Let's see some examples: Let $\mathcal{L} = \{+, 0\}$ where $+$ is a binary function and 0 is a constant. Then the following are terms: v_1, v_2, 0, $+(v_1, v_2)$ which is usually written $v_1 + v_2$ for notation sake, $((v_1 + v_2) + v_2) + 0$. You can make these terms very complicated. Intuitively, these terms don't assert anything. In a given structure, they will only refer to elements and not say anything about them.

Definition 78. We say that φ is an atomic $\mathcal{L}$-formula if φ is either

1. $t_1 = t_2$, where t_1 and t_2 are $\mathcal{L}$-terms, or

2. $R(t_1, ..., t_{n_R})$, where $R \in \mathcal{R}$ and $t_1, ..., t_{n_R}$ are terms.

Definition 79. The set of $\mathcal{L}$-formulas is the smallest set W containing all the atomic formulas such that
i) if φ is in W, then $\neg\varphi$ is in W,
ii) if φ and ψ are in W , then $(\varphi \wedge \psi)$ and $(\varphi \vee \psi)$ are in W, and
iii) if φ is in W, then $\exists v_i \varphi$and $\forall v_i \varphi$ are in W.

Examples: Let $\mathcal{L} = \{+, 0\}$ where $+$ is a binary function and 0 is a constant. Then a formula could assert the equality of two terms, say $v_1 + v_2 = 0$. It could exert existence of an element satisfying a property. For example $\exists v_1(v_1 + 0 = 0)$. We will need the following important concept:

Definition 80. We say that an occurrence of a variable v in a formula φ is *free* if it is not inside a $\exists v$ or $\forall v$ quantifier. Otherwise we say that it is *bound*.

Examples: In the formula $\forall v_1(v_2 \geq 0 \vee \exists v_3(v_2 \cdot v_3 = v_2))$. In this example, v_1 and v_3 are bound and v_2 is free. Now consider the more complicated example $v_1 > 0 \wedge \exists v_1(v_1 + v_2 = 0)$. In this example, the first occurrence of v_1 is free and the second is bound. We will attempt to write formulas in the clearest manner possible so that confusion doesn't arise, but we must make sure that on a formal level this is precise and clear.

Definition 81. We say a formula φ is *quantifier-free* if there are no quantifiers appearing in φ. Alternatively, φ is quantifier free if all variables occurring in φ are free.

Let $\mathcal{L} = \{+, 0\}$ as before. Given an $\mathcal{L}$-structure, we know how to interpret the symbols + and = and 0. But we can't determine whether a formula is true in the model without concretizing what elements the terms refer to and what the formula is asserting. To make this formula concrete, we will use assignments.

Definition 82. Let $\mathcal{L}$ be a language and $\mathcal{M}$ and $\mathcal{L}$-structure. Then an *assignment* is any function $\sigma : V \to M$.

We can now determine what the value of a term is given an assignment.

Definition 83. Let $\mathcal{L}$ be a language and $\mathcal{M}$ an $\mathcal{L}$-structure. Let $\sigma : V \to \mathcal{M}$ be an assignment. We inductively define $t^{\mathcal{M}}[\sigma] \in M$ for any term $t \in T$.

1. If t is a constant symbol $c \in \mathcal{C}$, then $t^{\mathcal{M}}[\sigma] = c^{\mathcal{M}}$.
2. If t is a variable symbol $v_i \in V$, then $t^{\mathcal{M}}[\sigma] = \sigma(v_i)$.
3. If $t_1, ..., t_{n_f}$ are terms, and $t = f(t_1, ..., t_{n_f})$ then $t^{\mathcal{M}}[\sigma] = f^{\mathcal{M}}(t_1^{\mathcal{M}}[\sigma], ..., t_1^{\mathcal{M}}[\sigma])$

Given an assignment $\sigma : V \to M$ and $v \in V$ and $a \in M$, we can define a new assignment $\sigma\left[\frac{a}{v}\right]$ to be the assignment

$$\sigma\left[\frac{a}{v}\right](v_i) = \begin{cases} \sigma(v_i) & \text{if } v_i \neq v \\ a & \text{if } v_i = v \end{cases}$$

Example:

$\mathcal{L} = \{+, 0\}$ as before. Let t_1 be $v_1 + 0$, t_2 is v_2 and t is $t_1 + t_2$. Then if $\mathcal{M} = \langle \mathbb{Z}, +, 5 \rangle$, and $\sigma(v_i) =$, we get $t_1^{\mathcal{M}}[\sigma] = 1 + 5 = 6$ and $t_2^{\mathcal{M}}[\sigma] = 2$. Therefore, $t^{\mathcal{M}}[\sigma] = 6 + 2 = 8$. $t^{\mathcal{M}}[\sigma[\frac{5}{v_2}]] = 6 + 5 = 11$.

Now that we know how to interpret these terms, we can decide whether a formula is true in a structure with a given assignment. Namely, it's silly to ask whether $v_1 + v_2 = v_3$ in a structure until I know what v_1,v_2, and v_3 are since the truth value will often change if you plug in different values.

Definition 84. Let $\mathcal{M}$ be and $\mathcal{L}$-structure and σ and assignment. We inductively define $\mathcal{M} \models_\sigma \varphi$, said $\mathcal{M}$ *models* φ with assignment σ, for all $\mathcal{L}$-formulas φ as follows:

1. If φ is $t_1 = t_2$ then $\mathcal{M} \models_\sigma \varphi$ if $t_1^{\mathcal{M}}[\sigma] = t_2^{\mathcal{M}}[\sigma]$.
2. If φ is $R(t_1, ..., t_{n_R})$, then $\mathcal{M} \models_\sigma \varphi$ if $(t_1^{\mathcal{M}}[\sigma], ..., t_{n_R}^{\mathcal{M}}[\sigma]) \in R^{\mathcal{M}}$.
3. If φ is $\neg\psi$, then $\mathcal{M} \models_\sigma \varphi$ if $\mathcal{M} \not\models_\sigma \psi$.
4. If φ is $(\psi \wedge \theta)$, then $\mathcal{M} \models_\sigma \varphi$ if $\mathcal{M} \models_\sigma \psi$ and $\mathcal{M} \models_\sigma \theta$.
5. If φ is $(\psi \vee \theta)$, then $\mathcal{M} \models_\sigma \varphi$ if $\mathcal{M} \models_\sigma \psi$ or $\mathcal{M} \models_\sigma \theta$.
6. If φ is $\exists v_j \psi$ then $\mathcal{M} \models_\sigma \varphi$ if there is an $a \in M$ such that $\mathcal{M} \models_{\sigma[\frac{a}{v_j}]} \psi$.
7. IF φ is $\forall v_j \psi$, then $\mathcal{M} \models_\sigma \varphi$ if for any $a \in M$ we have that $\mathcal{M} \models_{\sigma[\frac{a}{v_j}]} \psi$.

NOTE: We will often use abbreviations and convenient notation to help us.

- $\varphi \to \psi$ is an abbreviation for the formula $\neg\varphi \vee \psi$. This is very intuitive from the first section! In the same spirit, we use $\varphi \leftrightarrow \psi$ to mean $\varphi \to \psi$ and $\psi \to \varphi$.
- We didn't even need $\vee$ or $\forall$!! We could have take $\neg(\neg\varphi \wedge \neg\psi)$ instead of $\varphi \vee \psi$. We included them to make notation simpler, but we will use the fact that they are abbreviations when we prove things about formulas.
- We will also use
$$\bigwedge_{i=1}^{n} \psi_i = \psi_1 \wedge \cdots \wedge \psi_n$$
and
$$\bigvee_{i=1}^{n} \psi_i = \psi_1 \vee \cdots \vee \psi_n$$
.
- In addition to the symbols $v_1, v_2, ...$ we will use $w, x, y, z, ...$ as variable symbols. It will be clear what the variables are in the context.

- We can only quantify over the elements of the model, not subsets, or other unrelated sets! This is part of what makes this first-order logic.

We will now prove that to determine the truth value of a formula, we only need to know what the assignment does to the free variables in the formula.

Lemma 85 (Coincidence Lemma)**.** *Suppose $\mathcal{M}$ is and $\mathcal{L}$-structure.*

1. *Suppose t is an $\mathcal{L}$-term and $\sigma, \tau : V \to M$ are assignments that agree on all variables in t. Then $t^{\mathcal{M}}[\sigma] = t^{\mathcal{M}}[\tau]$.*
2. *Suppose φ is an $\mathcal{L}$-formula and $\sigma, \tau : V \to M$ are assignments which agree on all free variables occurring in φ. Then $\mathcal{M} \models_\sigma \varphi$ if and only if $\mathcal{M} \models_\tau \varphi$.*

Proof. 1. We constructed terms inductively, so we will prove this by induction on terms.

If $t = c \in \mathcal{C}$ is a constant, then

$$t^{\mathcal{M}}[\sigma] = c^{\mathcal{M}} = t^{\mathcal{M}}[\tau]$$

If $t = v_i$ is a variable, then

$$t^{\mathcal{M}}[\sigma] = \sigma(v_i) = \tau(v_i) = t^{\mathcal{M}}[\tau]$$

Suppose the lemma is true for $t_1, ..., t_{n_f}$. Let $t = f(t_1, ..., t_{n_f})$. Then

$$t^{\mathcal{M}}[\sigma] = f^{\mathcal{M}}(t_1^{\mathcal{M}}[\sigma], ..., t_{n_f}^{\mathcal{M}}[\sigma]) = f^{\mathcal{M}}(t_1^{\mathcal{M}}[\tau], ..., t_{n_f}^{\mathcal{M}}[\tau]) = t^{\mathcal{M}}[\tau]$$

2. We prove this by induction on formulas.

If φ is $t_1 = t_2$ where t_1, t_2 are $\mathcal{L}$-terms. Then

$$\begin{aligned} \mathcal{M} \models_\sigma \varphi &\iff t_1^{\mathcal{M}}[\sigma] = t_2^{\mathcal{M}}[\sigma] \\ &\iff t_1^{\mathcal{M}}[\tau] = t_2^{\mathcal{M}}[\tau] \\ &\iff \mathcal{M} \models_\tau \varphi \end{aligned}$$

.

If R is a relation symbol, $t_1, ..., t_{n_R}$ $\mathcal{L}$-terms, and φ is $R(t_1, ..., t_{n_R})$, then

$$\mathcal{M} \models_\sigma \varphi \iff (t_1^{\mathcal{M}}[\sigma], ..., t_{n_R}^{\mathcal{M}}[\sigma]) \in R^{\mathcal{M}}$$

$$\iff (t_1^{\mathcal{M}}[\tau], ..., t_{n_R}^{\mathcal{M}}[\tau]) \in R^{\mathcal{M}}$$

$$\iff \mathcal{M} \models_\tau \varphi$$

.

Suppose the claim is true for ψ and φ is $\neg\psi$. Then

$$\mathcal{M} \models_\sigma \varphi \iff \mathcal{M} \not\models_\sigma \psi$$

$$\iff \mathcal{M} \not\models_\tau \psi$$

$$\iff \mathcal{M} \models_\tau \varphi$$

Suppose the claim is true from ψ and θ and φ is $\psi \wedge \theta$. Then

$$\mathcal{M} \models_\sigma \varphi \iff \mathcal{M} \models_\sigma \psi \text{ and } \mathcal{M} \models_\sigma \theta$$

$$\iff \mathcal{M} \models_\tau \psi \text{ and } \mathcal{M} \models_\tau \theta$$

$$\iff \mathcal{M} \models_\tau \varphi$$

Now suppose the claim is true for ψ and φ is $\exists v_i \psi$. Then

$$\mathcal{M} \models_\sigma \varphi \iff \mathcal{M} \models_{\sigma[\frac{a}{v_i}]} \psi \text{ for some } a \in M$$

. The assignments σ, τ agree on all free variables in φ , so the assignments $\sigma[\frac{a}{v_i}], \tau[\frac{a}{v_i}]$ also agree on all free variables in ψ. So by induction

$$\mathcal{M} \models_{\sigma[\frac{a}{v_i}]} \psi \iff \mathcal{M} \models_{\tau[\frac{a}{v_i}]} \psi$$

$$\Rightarrow \mathcal{M} \models_\tau \varphi$$

The other direction is similar.

Thus, by induction, for any formula φ,

$$\mathcal{M} \models_\sigma \varphi \iff \mathcal{M} \models_\tau \varphi$$

□

Definition 86. An $\mathcal{L}$-formula φ is a *sentence* if it has no free variables.

Consider $\mathcal{L} = \{+, 0\}$ where $+$ is a binary function symbol as usual and 0 is a constant symbol. Then $v_1 + v_2 = 0$ is not a sentence, but $\exists v_1 \exists v_2 (v_1 + v_2 = 0)$ is a sentence. The intuition is that sentences are either true or not in a structure regardless of the assignment.

Corollary 87. *Suppose that φ is an $\mathcal{L}$-sentence and $\mathcal{M}$ is an $\mathcal{L}$-structure. Then $\mathcal{M} \models_\sigma \varphi$ for some assignment σ if and only if $\mathcal{M} \models_\sigma \varphi$ for all assignments σ.*

This immediate corollary allows us to define the following:

Definition 88. If φ is a sentence, we write $\mathcal{M} \models \varphi$ if $\mathcal{M} \models_\sigma \varphi$ for all assignments $\sigma : V \to M$.

Disjunctive Normal Form:

Definition 89. Let $\mathcal{L}$ be a language and φ and ψ $\mathcal{L}$-formulas. We say φ is *equivalent* to ψ if for any $\mathcal{L}$-structure $\mathcal{M}$ and σ an assignment, $\mathcal{M} \models_\sigma \varphi \iff \mathcal{M} \models_\sigma \psi$, or equivalently if every $\mathcal{L}$-structure $\mathcal{M}$ satisfies $\mathcal{M} \models_\sigma \varphi \leftrightarrow \psi$ for any assignment σ.

We will state and not prove the following:

Theorem 90. *Let $\mathcal{L}$ be a language and φ a quantifier free formula. Then φ is equivalent to*

$$\bigvee_{i=1}^{n} \bigwedge_{j=1}^{m_i} \psi_{i,j}$$

where $\psi_{i,j}$ are atomic or negated atomic formulas.

Let $\mathcal{L} = \{+, 0\}$ where $+$ is a binary function symbol and 0 is a constant symbol. Then let us look at the following:

$$((v_1 + v_2 = 0) \vee (v_1 + 0 = v_3)) \wedge \neg(0 + 0 = v_2)$$

.

This is equivalent to

$$((v_1 + v_2) = 0 \wedge \neg(0 + 0 = v_2)) \vee ((v_1 + 0 = v_3) \wedge \neg(0 + 0 = v_2)).$$

Let's look at a more abstract example. Suppose $\varphi_1, \varphi_2, \varphi_3$ are atomic $\mathcal{L}$ formulas for some language $\mathcal{L}$. Then, it is not hard to see (you can use truth tables if you like), that

$$(\neg(\varphi_1 \wedge \varphi_2) \wedge \varphi_3) \wedge \varphi_2$$

is equivalent to

$$((\neg\varphi_1 \vee \neg\varphi_2) \wedge \varphi_3) \wedge \varphi_2$$

which is equivalent to

$$((\neg\varphi_1 \wedge \varphi_3) \vee (\neg\varphi_1 \wedge \varphi_3)) \wedge \varphi_2$$

which is equivalent to

$$((\neg\varphi_1 \wedge \varphi_3) \wedge \varphi_2) \vee ((\neg\varphi_2 \wedge \varphi_3) \wedge \varphi_2)$$

which is equivalent to

$$(\neg\varphi_1 \wedge \varphi_3 \wedge \varphi_2) \vee (\neg\varphi_2 \wedge \varphi_3 \wedge \varphi_2)$$

Isomorphism and Elementary Equivalence:

Definition 91. Suppose $\mathcal{M}$ and $\mathcal{N}$ are $\mathcal{L}$-structures with universes M and N respectively. Then and $\mathcal{L}$-*embedding* $\eta : \mathcal{M} \to \mathcal{N}$ is an injective function $\eta : M \to N$ that preserves the interpretations of the symbols in $\mathcal{L}$, namely

(i) $\eta(f^{\mathcal{M}}(a_1, ..., a_{n_f})) = f^{\mathcal{N}}(\eta(a_1), ..., \eta(a_{n_f}))$ for all $f \in \mathcal{F}$ and all $a_1, ..., a_{n_f} \in M$.

(ii) $(a_1, ..., a_{n_R}) \in R^{\mathcal{M}} \iff (\eta(a_1), ..., \eta(a_{n_R})) \in R^{\mathcal{N}}$ for all $R \in \mathcal{R}$ and $a_1, ..., a_{n_R} \in N$.

(iii) $\eta(c^{\mathcal{M}}) = c^{\mathcal{N}}$ for all $c \in \mathcal{C}$.

A bijective $\mathcal{L}$-embedding is called an *$\mathcal{L}$-isomorphism*. If $\mathcal{M}$ and $\mathcal{N}$ are $\mathcal{L}$-structures, $M \subseteq N$ and $\mathrm{Id} : M \to N$ is an $\mathcal{L}$-embedding, then $\mathcal{M}$ is a *substructure* of $\mathcal{N}$.

Notice that if $\mathcal{M}$ is a substructure of $\mathcal{N}$, then any assignment $\sigma : V \to M$ can be considered an assignment $\sigma : V \to N$.

Proposition 92. *Suppose that $\mathcal{M}$ and $\mathcal{N}$ are $\mathcal{L}$-structures and $\mathcal{M}$ is a substructure of $\mathcal{M}$. Then for any quantifier-free $\mathcal{L}$-formula φ and assignment $\sigma : V \to M$, $\mathcal{M} \models_\sigma \varphi \iff \mathcal{N} \models_\sigma \varphi$.*

Proof. Exercise.

□

Definition 93. We say that two $\mathcal{L}$-structures $\mathcal{M}$ and $\mathcal{N}$ are *elementarily equivalent* and write $\mathcal{M} \equiv \mathcal{N}$ if for all $\mathcal{L}$-sentences φ,

$$\mathcal{M} \models \varphi \iff \mathcal{N} \models \varphi$$

We will end the course with the following theorem which shows us that the same sentences are true in isomorphic structures. This will allow us to show that two structures are not isomorphic if there is a sentence which is true in one and not true in the other.

Theorem 94. *Suppose that $\mathcal{M}$ and $\mathcal{N}$ are $\mathcal{L}$-structures and $j : \mathcal{M} \to \mathcal{N}$ is an isomorphism. Then $\mathcal{M} \equiv \mathcal{N}$.*

Proof. First note that if $j : \mathcal{M} \to \mathcal{N}$ is an $\mathcal{L}$-isomorphism and $\sigma : V \to M$ is an assignment, then $j \circ \sigma : V \to N$ is an assignment.

We will prove by induction on formulas that for any formula φ,

$$\mathcal{M} \models_\sigma \varphi \iff \mathcal{N} \models_{j\circ\sigma} \varphi$$

First we must prove that terms behave well, that is that $j(t^{\mathcal{M}}[\sigma]) = t^{\mathcal{N}}[j \circ \sigma]$

We will prove this by induction on terms:

- t is $c \in \mathcal{C}$. Then

$$j(t^{\mathcal{M}}[\sigma]) = j(c^{\mathcal{M}}) = c^{\mathcal{N}} = t^{\mathcal{N}}[j \circ \sigma]$$

- t is $v \in V$. Then

$$j(t^{\mathcal{M}}[\sigma]) = j(\sigma(v)) = t^{\mathcal{N}}[j \circ \sigma]$$

- Suppose the claim is true on $t_1, ..., t_{n_f}$ and t is $f(t_1, ..., t_{n_f})$. Then

$$j(t^{\mathcal{M}}[\sigma]) = j(f^{\mathcal{M}}(t_1^{\mathcal{M}}[\sigma], ..., t_{n_f}^{\mathcal{M}}[\sigma]))$$

$$= f^{\mathcal{N}}(j(t_1^{\mathcal{M}}[\sigma]), ..., j(t_{n_f}^{\mathcal{M}}[\sigma]))$$

$$(f^{\mathcal{N}}(t_1^{\mathcal{N}}[j \circ \sigma], ..., t_{n_f}^{\mathcal{N}}[j \circ \sigma]))$$

$$= t^{\mathcal{N}}[j \circ \sigma]$$

We now proceed with our original claim by inducting on the complexity of φ.

- φ is $t_1 = t_2$ where t_1, t_2 are $\mathcal{L}$-terms:

$$\mathcal{M} \models_\sigma \varphi \iff t_1^{\mathcal{M}}[\sigma] = t_2^{\mathcal{M}}[\sigma]$$

$$\iff j(t_1^{\mathcal{M}}[\sigma]) = j(t_2^{\mathcal{M}}[\sigma]) \text{ since } j \text{ is injective}$$

$$\iff t_1^{\mathcal{N}}[j \circ \sigma] = t_2^{\mathcal{N}}[j \circ \sigma]$$

$$\mathcal{N} \models_{j \circ \sigma} \varphi$$

- φ is $R(t_1, ..., t_{n_R})$ where $R \in \mathcal{R}$ and $t_1, ..., t_{n_R}$ are $\mathcal{L}$-terms:

$$\mathcal{M} \models_\sigma \varphi \iff (t_1^{\mathcal{M}}[\sigma], ..., t_{n_R}^{\mathcal{M}}[\sigma]) \in R^{\mathcal{M}}$$

$$\iff (j(t_1^{\mathcal{M}}[\sigma]), ..., j(t_{n_R}^{\mathcal{M}}[\sigma])) \in R^{\mathcal{N}}$$

$$\iff (t_1^{\mathcal{N}}[j \circ \sigma], ..., t_{n_R}^{\mathcal{N}}[j \circ \sigma]) \in R^{\mathcal{N}} \iff \mathcal{N} \models_{j \circ \sigma} \varphi$$

- φ is $\neg\psi$ and we assume the claim for ψ for induction:

$$\mathcal{M} \models_\sigma \varphi \iff \mathcal{M} \not\models_\sigma \psi \iff \mathcal{N} \not\models_{j \circ \sigma} \psi \iff \mathcal{N} \models_{j \circ \sigma} \varphi$$

- φ is $\psi \wedge \theta$ and we assume the claim for ψ and θ for induction:

$$\mathcal{M} \models_\sigma \varphi \iff \mathcal{M} \models_\sigma \pi \text{ and } \mathcal{M} \models_\sigma \theta$$
$$\iff \mathcal{N} \models_{j\circ\sigma} \psi \text{ and } \mathcal{N} \models_{j\circ\sigma} \theta$$
$$\iff \mathcal{N} \models_{j\circ\sigma} \varphi$$

- φ is $\exists v\psi$ and we assume the claim for ψ for induction:

$$\mathcal{M} \models_\sigma \varphi \iff \mathcal{M} \models_{\sigma[\frac{a}{v}]} \psi \text{ for some } a \in M$$
$$\Rightarrow \mathcal{N} \models_{j\circ(\sigma[\frac{a}{v}])} \psi$$
$$\Rightarrow \mathcal{N} \models_{(j\circ\sigma)[\frac{j(a)}{v}]} \psi$$
$$\Rightarrow \mathcal{N} \models_{j\circ\sigma} \varphi$$

For the other direction, we will use that j is surjective.

$$\mathcal{N} \models_{j\circ\sigma} \varphi \iff \mathcal{N} \models_{(j\circ\sigma)[\frac{b}{v}]} \psi \text{ for some } b \in N$$
$$\Rightarrow \mathcal{N} \models_{(j\circ\sigma)[\frac{j(a)}{v}]} \psi \text{ for some } a \in M \text{ since } j \text{ is surjective}$$
$$\Rightarrow \mathcal{N} \models_{j\circ(\sigma[\frac{a}{v}])} \psi$$
$$\Rightarrow \mathcal{M} \models_{\sigma[\frac{a}{v}]} \psi$$
$$\Rightarrow \mathcal{M} \models_\sigma \varphi$$

□

EXERCISES:

1. Suppose $\mathcal{L}$ is a language and $\mathcal{L}^* \subseteq \mathcal{L}$. (All the symbols appearing in $\mathcal{L}^*$ appear in $\mathcal{L}$.) Prove that if $\eta : \mathcal{M} \to \mathcal{N}$ is an $\mathcal{L}$-embedding, then $\eta : \mathcal{M} \to \mathcal{N}$ is an $\mathcal{L}^*$-embedding. Give an example showing the converse is not true.

2. Let $\mathcal{L} = \{\leq\}$. Write a sentence φ such that if $\mathcal{M} \models \varphi$ then $\leq^{\mathcal{M}}$ is a partial order on M. Write a sentence ψ such that if $\mathcal{M} \models \psi$, then $\leq^{\mathcal{M}}$ is linear order. Write a sentence θ such that if $\mathcal{M} \models \theta$ then $\leq^{\mathcal{M}}$ is a linear order on M with a maximal element.

3. Let $\mathcal{L} = \emptyset$.

   - Write a sentence φ_n such that if $\mathcal{M} \models \varphi$ then M has size exactly n.
   - Write sentences $\varphi_1, \varphi_2, ...$, such that if $\mathcal{M} \models \varphi_i$ for all $i \in \mathbb{N}$ then M is infinite.
   - Show that there is a language $\mathcal{L}^*$ and a set Φ of $\mathcal{L}^*$-sentences such that if $\mathcal{M} \models \varphi$ for all $\varphi \in \Phi$, then M is uncountable. It is worth noting that there is no language or set of sentences which forces a model to be countably infinite.

4. Prove Proposition 92

5. Determine whether the following are isomorphic or not. Justify your answer.

   - $\mathcal{L} = \{<\}$, $<$ a binary relation symbol. $\mathcal{M} = \langle \mathbb{Z}, < \rangle$, $\mathcal{N} = \langle \mathbb{N}, < \rangle$.
   - $\mathcal{L} = \{<\}$. $<$ a binary relation symbol. $\mathcal{M} = \langle \mathbb{Z}, < \rangle$, $\mathcal{N} = \langle \mathbb{Z}, > \rangle$.
   - $\mathcal{L} = \{<\}$. $<$ a binary relation symbol. $\mathcal{M} = \langle \mathbb{Z}, < \rangle$, $\mathcal{N} = \langle \mathbb{Z}, \leq \rangle$.
   - $\mathcal{L} = \{+, \cdot, 0, 1\}$. $+, \cdot$ are binary function symbols, $0, 1$ constant symbols. $\mathcal{M} = \langle \mathbb{Q}, +, \cdot, 0, 1 \rangle$, $\mathcal{N} = \langle \mathbb{R}, +, \cdot, 0, 1 \rangle$.
   - $\mathcal{L} = \{+, \cdot, 0, 1\}$. $+, \cdot$ are binary function symbols, $0, 1$ constant symbols. $\mathcal{M} = \langle \mathbb{C}, +, \cdot, 0, 1 \rangle$, $\mathcal{N} = \langle \mathbb{R}, +, \cdot, 0, 1 \rangle$.
   - $\mathcal{L} = \{+\}$, $+$ a binary function symbol. $\mathcal{M} = \langle \{f | f : \mathbb{N} \to \{0, 1\}\}, \circ \rangle$, $\mathcal{N} = \langle \{f | f : \mathbb{R} \to \{0, 1\}\}, \circ \rangle$.