1.10 Basics Concepts of Set Theory

Having learned some fundamental notions of logic, it is now a good place before proceeding to more interesting things, such as functions and relations, to go through a very quick review of some basic concepts of set theory.

This section will take the very "naive" point of view that a set is a collection of objects, the collection being regarded as a single object.

Having first-order logic at our disposal, we could formalize set theory very rigorously in terms of axioms.

This was done by Zermelo first (1908) and in a more satisfactory form by Zermelo and Fraenkel in 1921, in a theory known as the "Zermelo-Fraenkel" (ZF) axioms.

Another axiomatization was given by John von Neumann in 1925 and later improved by Bernays in 1937.

A modification of Bernay's axioms was used by Kurt Gödel in 1940. This approach is now known as "von Neumann-Bernays" (VNB) or "Gödel-Bernays" (GB) set theory.

However, it must be said that set theory was first created by Georg Cantor (1845-1918) between 1871 and 1879.

However, Cantor's work was not unanimously well received by all mathematicians. Cantor regarded infinite objects as objects to be treated in much the same way as finite sets, a point of view that was shocking to a number of very prominent mathematicians who bitterly attacked him (among them, the powerful Kronecker).

Also, it turns out that some paradoxes in set theory popped up in the early 1900, in particular, *Russell's paradox*.

Russell's paradox (found by Russell in 1902) has to to with the

"set of all sets that are not members of themselves"

which we denote by

$$R = \{x \mid x \notin x\}.$$

(In general, the notation $\{x \mid P\}$ stand for the set of all objects satisfying the property P.)

Now, classically, either $R \in R$ or $R \notin R$. However, if $R \in R$, then the definition of R says that $R \notin R$; if $R \notin R$, then again, the definition of R says that $R \in R$!

So, we have a contradiction and the existence of such a set is a paradox.

The problem is that we are allowing a property (here, $P(x) = x \notin x$), which is "too wild" and circular in nature.

As we will see, the way out, as found by Zermelo, is to place a restriction on the property P and to also make sure that P picks out elements from some already given set (see the Subset Axioms below).

The apparition of these paradoxes prompted mathematicians, with Hilbert among its leaders, to put set theory on firmer grounds. This was achieved by Zermelo, Fraenkel, von Neumann, Bernays and Gödel, to only name the major players.

In what follows, we are assuming that we are working in classical logic.

We will introduce various operations on sets using defintion involving the logical connectives $\wedge$, $\vee$, $\neg$, $\forall$ and $\exists$.

In order to ensure the existence of some of these sets requires some of the axioms of set theory, but we will be rather casual about that.

Given a set, A, we write that some object, a, is an element of (belongs to) the set A as

$$a \in A$$

and that a is not an element of A (does not belong to A) as

$$a \notin A.$$

When are two sets A and B equal? This corresponds to the first axiom of set theory, called

Extensionality Axiom

Two sets A and B are equal iff they have exactly the same elements, that is

$$\forall x(x \in A \Rightarrow x \in B) \wedge \forall x(x \in B \Rightarrow x \in A).$$

The above says: Every element of A is an element of B and conversely.

There is a special set having no elements at all, the *empty set*, denoted $\emptyset$. This is the

Empty Set Axiom

There is a set having no members. This set is denoted $\emptyset$ and it is characterized by the property

$$\forall x(x \notin \emptyset).$$

Remark: Beginners often wonder whether there is more than one empty set. For example, is the empty set of professors distinct from the empty set of potatoes?

The answer is, by the extensionality axiom, there is only *one* empty set!

Given any two objects a and b, we can form the set $\{a, b\}$ containing exactly these two objects. Amazingly enough, this must also be an axiom:

Pairing Axiom

Given any two objects a and b (think sets), there is a set, $\{a, b\}$, having as members just a and b.

Observe that if a and b are identical, then we have the set $\{a, a\}$, which is denoted by $\{a\}$ and is called a *singleton set* (this set has a as its only element).

To form bigger sets, we use the union operation. This too requires an axiom.

Union Axiom (Version 1)

For any two sets A and B, there is a set, $A \cup B$, called the *union of A and B* defined by

$$x \in A \cup B \quad \text{iff} \quad (x \in A) \vee (x \in B).$$

This reads, x is a member of $A \cup B$ if either x belongs to A or x belongs to B (or both).

We also write

$$A \cup B = \{x \mid x \in A \quad \text{or} \quad x \in B\}.$$

Using the union operation, we can form bigger sets by taking unions with singletons. For example, we can form

$$\{a, b, c\} = \{a, b\} \cup \{c\}.$$

Remark: We can systematically construct bigger and bigger sets by the following method: Given any set, A, let

$$A^{+} = A \cup \{A\}.$$

If we start from the empty set, we obtain sets that can be used to define the natural numbers and the $+$ operation corresponds to the successor function on the natural numbers, i.e., $n \mapsto n + 1$.

Another operation is the *power set formation*.

It is indeed a "powerful" operation, in the sense that it allows us to form very big sets.

For this, it is helpful to define the notion of inclusion between sets. Given any two sets, A and B, we say that *A is a subset of B* (or that *A is included in B*), denoted $A \subseteq B$, iff every element of A is also an element of B, i.e.

$$\forall x(x \in A \Rightarrow x \in B).$$

We say that *A is a proper subset of B* iff $A \subseteq B$ and $A \neq B$. This implies that that there is some $b \in B$ with $b \notin A$. We usually write $A \subset B$.

Observe that the equality of two sets can be expressed by

$$A = B \quad \text{iff} \quad A \subseteq B \quad \text{and} \quad B \subseteq A.$$

Power Set Axiom

Given any set, A, there is a set, $\mathcal{P}(A)$, (also denoted 2^A) called the *power set of A* whose members are exactly the subsets of A, i.e.,

$$X \in \mathcal{P}(A) \quad \text{iff} \quad X \subseteq A.$$

For example, if $A = \{a, b, c\}$, then

$$\mathcal{P}(A) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\},$$

a set containing 8 elements. Note that the empty set and A itself are always members of $\mathcal{P}(A)$.

Remark: If A has n elements, it is not hard to show that $\mathcal{P}(A)$ has 2^n elements. For this reason, many people, including me, prefer the notation 2^A for the power set of A.

At this stage, we would like to define intersection and complementation.

For this, given any set, A, and given a property, P, (specified by a first-order formula) we need to be able to define the subset of A consisting of those elements satisfying P. This subset is denoted by

$$\{x \in A \mid P\}.$$

Unfortunately, there are problems with this construction. If the formula, P, is somehow a circular definition and refers to the subset that we are trying to define, then some paradoxes may arise!

The way out is to place a restriction on the formula used to define our subsets, and this leads to the subset axioms, first formulated by Zermelo.

These axioms are also called *comprehension axioms* or *axioms of separation*.

Subset Axioms

For every first-order formula, P, we have the axiom:

$$\forall A \exists X \forall x (x \in X \quad \text{iff} \quad (x \in A) \land P),$$

where P does *not* contain X as a free variable. (However, P may contain x free.)

The subset axiom says that for every set, A, there is a set, X, consisting exactly of those elements of A so that P holds. For short, we usually write

$$X = \{x \in A \mid P\}.$$

As an example, consider the formula

$$P(B, x) = x \in B.$$

Then, the subset axiom says

$$\forall A \exists X \forall x (x \in A \wedge x \in B),$$

which means that X is the set of elements that belong both to A and B.

This is called the *intersection of A and B*, denoted by $A \cap B$. Note that

$$A \cap B = \{x \mid x \in A \quad \text{and} \quad x \in B\}.$$

We can also define the *relative complement of B in A*, denoted $A - B$, given by the formula $P(B, x) = x \notin B$, so that

$$A - B = \{x \mid x \in A \quad \text{and} \quad x \notin B\}.$$

In particular, if A is any given set and B is any subset of A, the set $A - B$ is also denoted $\overline{B}$ and is called the *complement of B*.

Because $\wedge, \vee$ and $\neg$ satisfy the de Morgan laws (remember, we are dealing with classical logic), for any set X, the operations of union, intersection and complementation on subsets of X satisfy various identities, in particular the de Morgan laws

$$\begin{aligned}\overline{A \cap B} &= \overline{A} \cup \overline{B}\\ \overline{A \cup B} &= \overline{A} \cap \overline{B}\\ \overline{\overline{A}} &= A,\end{aligned}$$

and various associativity, commutativity and distributivity laws.

So far, the union axiom only applies to two sets but later on we will need to form infinite unions. Thus, it is necessary to generalize our union axiom as follows:

Union Axiom (Final Version)

Given any set X (think of X as a set of sets), there is a set, $\bigcup X$, defined so that

$$x \in \bigcup X \quad \text{iff} \quad \exists B(B \in X \wedge x \in B).$$

This says that $\bigcup X$ consists of all elements that belong to some member of X.

If we take $X = \{A, B\}$, where A and B are two sets, we see that

$$\bigcup \{A, B\} = A \cup B,$$

and so, our final version of the union axiom subsumes our previous union axiom which we now discard in favor of the more general version.

Observe that

$$\bigcup\{A\} = A, \quad \bigcup\{A_1, \ldots, A_n\} = A_1 \cup \cdots \cup A_n.$$

and in particular, $\bigcup \emptyset = \emptyset$.

Using the subset axiom, we can also define infinite intersections. For every nonempty set, X, there is a set, $\bigcap X$, defined by

$$x \in \bigcap X \quad \text{iff} \quad \forall B(B \in X \Rightarrow x \in B).$$

Observe that

$$\bigcap\{A, B\} = A \cap B, \quad \bigcap\{A_1, \ldots, A_n\} = A_1 \cap \cdots \cap A_n.$$

Note that $\bigcap \emptyset$ is not defined. Intuitively, it would have to be the set of all sets, but such a set does not exist, as we now show. This is basically a version of Russell's paradox.

Theorem 1.10.1 *(Russell) There is no set of all sets, i.e., there is no set to which every other set belongs.*

Proof. Let A be any set. We construct a set, B, that does not belong to A. If the set of all sets existed, then we could produce a set that does not belong to it, a contradiction. Let

$$B = \{a \in A \mid a \notin a\}.$$

We claim that $B \notin A$. We proceed by contradiction, so assume $B \in A$. However, by the definition of B, we have

$$B \in B \quad \text{iff} \quad B \in A \quad \text{and} \quad B \notin B.$$

Since $B \in A$, the above is equivalent to

$$B \in B \quad \text{iff} \quad B \notin B,$$

which is a contradiction. Therefore, $B \notin A$ and we deduce that there is no set of all sets. $\square$

Remarks:

(1) We should justify why the equivalence $B \in B$ iff $B \notin B$ is a contradiction.

What we mean by "a contradiction" is that if the above equivalence holds, then we can derive $\perp$ (falsity) and thus, all propositions become provable.

This is because we can show that for any proposition, P, if $P \equiv \neg P$ is provable, then $\perp$ is also provable.

We leave the proof of this fact as an easy exercise for the reader. By the way, this holds classically as well as intuitionistically.

(2) We said that in the subset axiom, the variable X is not allowed to occur free in P.

A slight modification of Russell's paradox shows that allowing X to be free in P leads to paradoxical sets. For example, pick A to be any nonempty set and set $P(X, x) = x \notin X$. Then, look at the (alleged) set

$$X = \{x \in A \mid x \notin X\}.$$

As an exercise, the reader should show that X is empty iff X is nonempty!

This is as far as we can go with the elementary notions of set theory that we have introduced so far. In order to proceed further, we need to define relations and functions, which is the object of the next Chapter.

The reader may also wonder why we have not yet discussed infinite sets.

This is because we don't know how to show that they exist!

Again, perhaps surprisingly, this takes another axiom, the *axiom of infinity*. We also have to define when a set is infinite. However, we will not go into this right now.

Instead, we will accept that the set of natural numbers, $\mathbb{N}$, exists and is infinite. Once we have the notion of a function, we will be able to show that other sets are infinite by comparing their "size" with that of $\mathbb{N}$ (This is the purpose of *cardinal numbers*, but this would lead us too far afield).

Remark: In an axiomatic presentation of set theory, the natural numbers can be defined from the empty set using the operation $A \mapsto A^+ = A \cup \{A\}$ introduced just after the union axiom.

The idea due to von Neumann is that the natural numbers, $0, 1, 2, 3, \ldots$, can be viewed as concise notations for the following sets:

$$
\begin{aligned}
0 &= \emptyset \\
1 &= 0^+ = \{\emptyset\} = \{0\} \\
2 &= 1^+ = \{\emptyset, \{\emptyset\}\} = \{0, 1\} \\
3 &= 2^+ = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\} = \{0, 1, 2\} \\
&\vdots \\
n+1 &= n^+ = \{0, 1, 2, \ldots, n\} \\
&\vdots
\end{aligned}
$$

However, the above subsumes induction! Thus, we have to proceed in a different way to avoid circularities.

Definition 1.10.2 We say that a set, X, is *inductive* iff

(1) $\emptyset \in X$;

(2) For every $A \in X$, we have $A^+ \in X$.

Axiom of Infinity

There is some inductive set.

Having done this, we make the

Definition 1.10.3 A *natural number* is a set that belongs to every inductive set.

Using the subset axioms, we can show that there is a set whose members are exactly the natural numbers. The argument is very similar to the one used to prove that arbitrary intersections exist.

Therefore, the set of all natural numbers, $\mathbb{N}$, does exist. The set $\mathbb{N}$ is also denoted ω.

We can now easily show

Theorem 1.10.4 *The set $\mathbb{N}$ is inductive and it is a subset of every inductive set.*

It would be tempting to view $\mathbb{N}$ as the intersection of the family of inductive sets, but unfortunately this family is not a set; it is too "big" to be a set.

As a consequence of the above fact, we obtain the

Induction Principle for $\mathbb{N}$ (Version 1): Any inductive subset of $\mathbb{N}$ is equal to $\mathbb{N}$ itself.

Now, in our setting, $0 = \emptyset$ and $n^+ = n + 1$, so the above principle can be restated as follows:

Induction Principle for $\mathbb{N}$ (Version 2): For any subset, $S \subseteq \mathbb{N}$, if $0 \in S$ and $n + 1 \in S$ whenever $n \in S$, then $S = \mathbb{N}$.

We will see how to rephrase this induction principle a little more conveniently in terms of the notion of function in the next chapter.

Remarks:

1. We still don't know what an infinite set is or, for that matter, that $\mathbb{N}$ is infinite! This will be shown in the next Chapter (see Corollary 2.9.7).

2. Zermelo-Fraenkel set theory (+ Choice) has three more axioms that we did not discuss: The *Axiom of Choice*, the *Replacement Axioms* and the *Regularity Axiom*. For our purposes, only the Axiom of Choice will be needed and we will introduce it in Chapter 2.

   Let us just say that the Replacement Axioms are needed to deal with ordinals and cardinals and that the Regularity Axiom is needed to show that every set is grounded.

   The Regularity Axiom also implies that no set can be a member of itself, an eventuality that is not ruled out by our current set of axioms!

As we said at the beginning of this section, set theory can be axiomatized in first-order logic.

To illustrate the generality and expressiveness of first-order logic, we conclude this section by stating the axioms of *Zermelo-Fraenkel set theory* (for short, *ZF*) as first-order formulae.

The language of Zermelo-Fraenkel set theory consists of the constant, $\emptyset$ (for the empty set), the equality symbol, and of the binary predicate symbol, $\in$, for set membership.

We will abbreviate $\neg(x = y)$ as $x \neq y$ and $\neg(x \in y)$ as $x \notin y$.

The axioms are the equality axioms plus the following seven axioms:

$$
\begin{array}{l}
\forall A \forall B(\forall x(x \in A \equiv x \in B) \Rightarrow A = B) \\
\forall x(x \notin \emptyset) \\
\forall a \forall b \exists Z \forall x(x \in Z \equiv (x = a \vee x = b)) \\
\forall X \exists Y \forall x(x \in Y \equiv \exists B(B \in X \wedge x \in B)) \\
\forall A \exists Y \forall X(X \in Y \equiv \forall z(z \in X \Rightarrow z \in A)) \\
\forall A \exists X \forall x(x \in X \equiv (x \in A) \wedge P) \\
\exists X(\emptyset \in X \wedge \forall y(y \in X \Rightarrow y \cup \{y\} \in X)),
\end{array}
$$

where P is any first order formula that does not contain X free.

- Axiom (1) is the extensionality axiom.
- Axiom (2) is the emptyset axiom.
- Axiom (3) asserts the existence of a set, Y, whose only members are a and b. By extensionality, this set is unique and it is denoted $\{a, b\}$. We also denote $\{a, a\}$ by $\{a\}$.

- Axiom (4) asserts the existence of set, Y, which is the union of all the set that belong to X. By extensionality, this set is unique and it is denoted $\bigcup X$. When $X = \{A, B\}$, we write $\bigcup\{A, B\} = A \cup B$.
- Axiom (5) asserts the existence of set, Y, which is the set of all subsets of A (the power set of A). By extensionality, this set is unique and it is denoted $\mathcal{P}(A)$ or 2^A.
- Axioms (6) are the subset axioms (or axioms of separation).
- Axiom (7) is the infinity axiom, stated using the abbreviations introduced above.