Axiomatic set theory

Bertrand Russell's paradox (#30 of §1.6) suggests that the intuitive handling of sets is not always sufficient for precise mathematical reasoning. Cautioned by his example, one is lead to limit the "collections of objects" that deserve the name *set*. To this end, in the early twentieth century, mathematicians began formalizing set theory axiomatically. A list of widely accepted rules for forming sets emerged – rules, which appeared to be absolutely necessary for conducting serious mathematics while (hopefully) being simple enough so as not to lead to any parodoxes.

Below are the nine axioms of set theory which are used today by (most) mathematicians. Very much like the postulates of geometry, these axioms are accepted on faith. All subsequent theorems, however, must be based on these axioms and strict rules of inference alone. (We will discuss these rules in §1.5)

First, we will list the axioms in plain English before we rephrase them in the language of first-order predicate logic.

(I) [**Equality**]
"Two sets are equal if and only if they have the same elements. Moreover, equal elements belong to the same sets."

(II) [**Empty set**]
"There exists a set with no elements."

> This (unique) set is called the *empty set* and will be denoted by $\emptyset$.

(III) [**Unordered pair**]
"Given two sets A and B, there is a set X which contains exactly A and B as its elements."

> We denote this (unique) set X by $\{A, B\}$.

(IV) [**Union of sets**]
"Given a set X whose elements are sets, there is a set U consisting of all those elements that are elements of some set belonging to X."

> We denote this (unique) set U by $\bigcup X$ and call it the *union* of the sets belonging to X. For example, if $X = \{\{1,2,3\},\{1,4,9\},\{2,10\}\}$, then $\bigcup X = \{1,2,3,4,9,10\}$. It is customary to write $A \cup B$ for $\bigcup\{A,B\}$ and to call it the *union of A and B*.

(V) [**Power set**]
"Given a set A, there is a set X whose elements are all the subsets of A."

> We call this (unique) set X the *power set* of A and denote it by $\mathcal{P}(A)$. Here, a set A is called a *subset* of a set B, denoted $A \subseteq B$, if every element of A also belongs to B. For example, if $A = \{1,2,3\}$, then $\mathcal{P}(A) = \{\emptyset,\{1\},\{2\},\{3\},\{1,2\},\{1,3\},\{2,3\},\{1,2,3\}\}$.

(VI) **[Forming subsets]**
"If $P(x)$ is a predicate and A is a set, then there is a set B whose elements are all those elements x of A for which $P(x)$ is true."

We denote this (unique) set B by $\{x \in A \mid P(x)\}$. Using this axiom, we can, for example, define the intersection $A \cap B$ of two sets A and B to be the set $\{x \in A \mid x \in B\}$.

(VII) **[Infinity]**
"There exists a set X with the following property: $\emptyset$ belongs to X and if A belongs to X, then so does $A \cup \{A\}$."

A set X that has this property is called an *inductive* set. The existence of an inductive set allows us, for example, to define the set $\mathbb{N}$ of natural numbers as the set of all those sets which must be contained in every inductive set. In other words,

$$\mathbb{N} = \{\emptyset,\ \{\emptyset\},\ \{\emptyset, \{\emptyset\}\},\ \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\},\ \cdots\}.$$

One then introduces the more familiar symbols $0, 1, 2, 3, \cdots$ to denote above sets. That is to say, we write 0 for $\emptyset$, 1 for $\{0\}$, 2 for $\{0, 1\}$, 3 for $\{0, 1, 2\}$, and so on. Based on this definition, one can also introduce the usual arithmetic into the set $\mathbb{N}$ of natural numbers – more about that later in the course.

(VIII) **[Replacement]**
"If the predicate $P(x, y)$ represents a function and A is a set, then there is a set B whose elements are all those y for which $P(x, y)$ is true for some element x of A."

Here, a predicate $P(x, y)$ is said to *represent a function* if for every x there is at most one y such that $P(x, y)$ is true. So, this axiom says that the range of a function, whose domain is a set, is a set.

(IX) **[Axiom of choice]**
"Given a set X whose elements are non-empty sets no two of which have elements in common, there is a set C which contains exactly one element of every set of X."

This last axiom, called the *axiom of choice*, sounds pretty harmless. For example, if $X = \{\{a, b\}, \{1, 2, 3\}, \{x\}\}$, we could take for C the set $\{b, 2, x\}$. It turns out that although this axiom is needed in proofs of many fundamental theorems of mathematics, it can also be used to prove rather mindboggling facts. This axiom was therefore very controversial for a long time. In fact, many mathematicians thought one should try to work without it. Over the years, however, the dust has settled and the mathematical community has (almost) unanimously come to adopt its use.

Let us now formulate the above axioms in the language of first-order predicate logic. Our universe of discourse shall be thought of as sets. We will use the predicate $Q(x, y)$ to mean "x is an element of y" and abbreviate the truth of this by $x \in y$. Similarly, $x \notin y$ will stand for $\neg\, Q(x, y)$. We will write $x \subseteq y$ for $\forall z(z \in x \to z \in y)$ and $x = y$ for $(x \subseteq y) \wedge (y \subseteq x)$. We will also use $x \neq y$ for $\neg\, (x = y)$. Notice that this definition already takes care of the first half of Axiom I above. Using the same notation as before, our nine axioms now read as follows:

(I) **[Equality]**
$\forall x \forall y (x = y \to \forall z[x \in z \leftrightarrow y \in z])$

(II) **[Empty set]**
$\exists x \forall y (y \notin x)$

(III) **[Unordered pair]**
$\forall x \forall y \exists z \forall u (u \in z \leftrightarrow [u = x \vee u = y])$

(IV) **[Union of sets]**
$\forall x \exists y \forall u (u \in y \leftrightarrow \exists v[u \in v \wedge v \in x])$

(V) **[Power set]**
$\forall x \exists y \forall u (u \in y \leftrightarrow u \subseteq x)$

(VI) **[Forming subsets]**
Given a predicate $P(u)$, $\forall x \exists y \forall u (u \in y \leftrightarrow [u \in x \wedge P(u)])$

(VII) **[Infinity]**
$\exists x (\emptyset \in x \wedge \forall z[z \in x \to z \cup \{z\} \in x])$

(VIII) **[Replacement]**
Given a predicate $P(u, v)$, $\forall x \forall u \forall v \Big(([P(x, u) \wedge P(x, v)] \to u = v)$
$\to \forall w \exists z \forall v (v \in z \leftrightarrow \exists u[u \in w \wedge P(u, v)]) \Big)$

(IX) **[Axiom of choice]**
$\forall x \Big(\forall u \Big(u \in x \to (u \neq \emptyset \wedge \forall v[v \in x \wedge (v \neq u \to v \cap u = \emptyset)\ ]\ )\ \Big)$
$\to \exists y \forall u \Big(u \in x \to \exists w[w \in u \cap y \wedge \forall z(z \in u \cap y \to z = w)] \Big) \Big)$