A taste of set theory for philosophers

Jouko Väänänen*
Department of Mathematics and Statistics
University of Helsinki
and
Institute for Logic, Language and Computation
University of Amsterdam

November 17, 2010

Contents

1   Introduction

Originally set theory was a theory of infinity, an attempt to understand infinity in exact terms. Later it became a universal language for mathematics and an attempt to give a foundation for all of mathematics, and thereby to all sciences that are based on mathematics. So what is set theory?

Set theory is a very general but still entirely exact theory of objects called sets. It is useful in a number of fields of philosophy, like logic, semantics, philosophy of mathematics, philosophy of language and probably several others, but it is also useful in mathematics, computer science, cognitive science, linguistics, and even in the theory of music. It can be used anywhere where one needs an exact mathematical approach to objects that can be thought of as collections of something.

Even high school mathematics includes simple operations on sets, like union and intersection. College mathematics usually includes set theoretical concepts like ordered pair, cartesian product, relation, function, and so on. Elementary logic courses include such set theoretical concepts as finite sequence and relation. All the concepts mentioned so far are very useful for any philosophy student. Why? Because all these basic mathematical concepts can be given a uniform exact account. In this account any true properties of those concepts can be proved with a simple argument involving only a few lines.

The remarkable thing about set theory is that not only basic mathematics but indeed *all* mathematics can be represented as properties of sets. Thus we can define in set theory the natural numbers, the real numbers, the complex numbers, the Euclidean spaces $\mathbb{R}^n$, the Hilbert space, all the familiar Banach spaces, etc. Moreover, everything mathematicians prove about these objects can be proved from a few relatively simple axioms concerning sets. Therefore it is said that set theory can serve as a universal language of mathematics, indeed a foundation of mathematics. This gives set theory a special place in the philosophy of mathematics.

Of course, a representation of all mathematics in set theory is meant to be taken *only* as a representation. The fact that real numbers can be defined as sets does not mean that real numbers *are* sets. The point is that it is in principle possible to think of real numbers as sets. It is important to note that the goal of set theorists is not to convince other mathematicians that what mathematicians are doing is really set theory. The point of set theory as a universal language of mathematics is that set theory offers a common ground where any unclear argument can be scrutinized. If some argument in mathematics seems to use something that has not been stated, we can start the process of reducing the argument to the first principles in set theory. If this process is successful, then the argument can be considered valid without question. In this process it becomes clear whether, for example, the Axiom of Choice, a very powerful construction principle for abstract objects, is used. Also, some mathematical results depend on principles, such as the Continuum Hypothesis, that go beyond what is usually considered a priori true. Then the mathematical result can be stated as an implication: if the Continuum Hypothesis is assumed then this or that holds.

2   Elementary set theory

In this section sets are just collections of objects. We shall later define more exactly what this means. We use lower case letters $a, b, c, ...$ to denote sets. Since sets are collections, they have elements i.e. members of the collection in question. If a is a set and b is an element of a, then we write $b \in a$ and read this "b is an element of a" or "b is in a". Two sets are equal if they have the same elements. A set a is a *subset* of another set b, in symbols $a \subseteq b$ if all elements of a are also in b. The simplest sets the *singleton* set $\{a\}$ which has a as its only element, the *unordered pair* $\{a, b\}$ which has a and b and nothing else as its elements, and the *empty* set $\emptyset$, which has no elements whatsoever. Note that $\{a, b\} = \{b, a\}$ and that there is only one empty set, because any two sets without elements have the same elements and are therefore equal.

The most important non-trivial sets are: (1) The set $\{0, 1, 2, ...\}$ of natural numbers, denoted $\mathbb{N}$, (2) the set of rational numbers, denoted $\mathbb{Q}$, and (3) the set of real numbers, denoted $\mathbb{R}$. When we proceed deeper into set theory below we can actually *define* these sets, but let us take them for the moment as given.

We can form new sets from the ones we already know by means of set theoretic operations like the *union* $a \cup b$, *intersection* $a \cap b$, and *power set* $\mathcal{P}(a) = \{b : b \subseteq a\}$. There are a couple of others, and when one learns to use ordinals, there are the transfinite operations on sets.

Already with the simple operations $\cup$ and $\cap$ we get the following important concept: Let X be any set. Then obviously $\mathcal{P}(X)$ is closed under $\cup$ and $\cap$. Also, we can form the complement $-a = \{x \in X : x \notin a\}$ of any subset a of X. Finally, let us denote $\emptyset$ by 0 and X by 1. We have arrived at the structure

$$(\mathcal{P}(X), \cap, \cup, -, 0, 1)$$

which is a familiar algebraic structure, namely a *Boolean algebra*, because it satisfies the identities

$$
\begin{array}{rcll}
a \cap (b \cap c) & = & (a \cap b) \cap c & \text{Associativity law} \\
a \cup (b \cup c) & = & (a \cup b) \cup c & \\
a \cap (b \cup c) & = & (a \cap b) \cup (a \cap c) & \text{Distributivity law} \\
a \cup (b \cap c) & = & (a \cup b) \cap (a \cup c) & \\
a \cap b & = & b \cap a & \text{Commutativity law} \\
a \cup b & = & b \cup a & \\
a \cup -a & = & 1 & \text{Law of complements} \\
a \cap -a & = & 0 & \\
a \cup (a \cap b) & = & a & \text{Absorption law} \\
a \cap (a \cup b) & = & a & \\
-(a \cap b) & = & -a \cup -b & \text{De Morgan law} \\
-(a \cup b) & = & -a \cap -b &
\end{array}
$$

These are all easy to prove, even by just looking at a picture, as in Figure 1.

An important role in applications of set theory is played by the concept of a an ordered pair (a, b) of two sets a and b. The characteristic property of ordered pairs is: $(a_0, a_1) = (b_0, b_1)$ if and only if $a_0 = b_0$ and $a_1 = b_1$. The cartesian product of two
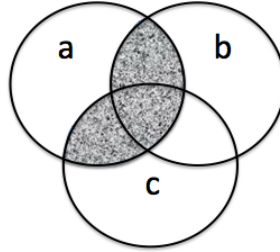
Figure 1: $a \cap (b \cup c) = (a \cap b) \cup (a \cap c)$.

sets a and b is $a \times b = \{(x, y) : x \in a, y \in b\}$. It is the idea of set theory that everything is defined in terms of the sole primitive symbol $\in$. This is by no means necessary but since it is possible it is tempting and is usually done. The most common definition for the ordered pair (x, y) in terms of $\in$ is $\{\{x\}, \{x, y\}\}$.

A function from a set a to another set b is any subset f of $a \times b$ such that for each $x \in a$ there is exactly one $y \in b$ such that $(x, y) \in f$. Then we write $f : a \to b$ and $y = f(x)$. In this definition of the concept of a function one notices a characteristic feature of set theory: the concept of a function is extremely general. We do not require that there is some "rule" which tells us how to compute $f(x)$ for a given x. All we require is that exactly one y such that $(x, y) \in f$ exists. Set theory uses classical logic so for a y such that $(x, y) \in f$ to exist it suffices that non-existence leads to a contradiction. There is also *constructive set theory* (see below) where intuitionistic logic is used and existence means more than deriving contradiction from non-existence.

A set a is *finite* if it is of the form $\{a_0, ..., a_{n-1}\}$ for some natural number n. This means that the set a has at most n elements. A set which is not finite is *infinite*. Finite sets have the following properties: $\emptyset$ is finite. If a and b are finite, then so is $a \cup b$. If a is finite and $b \subseteq a$, then also b is finite. If a and b are finite, then so is $a \times b$. If a is finite, then so is $\mathcal{P}(a)$.

With the above concepts one can already develop a lot of mathematics. One can define the integers as ordered pairs (n, m) of natural numbers with the intuitive meaning that (n, m) denotes the integer $n - m$. One can define the rationals as ordered pairs (r, q) of integers with the intuitive meaning that (r, q) denotes the rational r/q. One can define the reals as sets a of rationals, bounded from above, with the intuitive meaning that $a \subseteq \mathbb{Q}$ denotes the real $\sup(a)$.

3 Cardinal and ordinal numbers

A set is infinite if it is not of the form $\{a_1, ..., a_n\}$ for any natural number n. Set theory was developed to deal with problems of infinite sets and indeed there are some paradoxical phenomena related to infinite sets. A famous anecdotal example is *Hilbert's*

Hotel: Imagine a hotel the rooms of which are numbered by all natural numbers. Suppose the hotel is full but a tourist comes in and asks for a free room. The reception can ask the person in room 0 to move to room 1, the person in room 1 to move to room 2, ..., the person in room n to move to room $n + 1$, etc. This process leaves room 0 empty and the tourist can take it. There are many further variations of this anecdote. For example, one can fit infinitely many new tourists into a hotel which is already full. A vast extension of this idea, coupled with the so called Axiom of Choice, is the Banach-Tarski Paradox: The unit sphere in three-dimensional space can be split into five pieces so that if the pieces are rigidly moved and rotated they suddenly form two spheres of the original size (see Picture 2). The trick is that the splitting exists only in the abstract world of mathematics and can never actually materialize in the physical world. Conclusion: infinite abstract objects do not obey the rules we are used to among finite concrete objects. This is like the situation with sub-atomic elementary particles, where counter-intuitive phenomena, such as entanglement, occur.
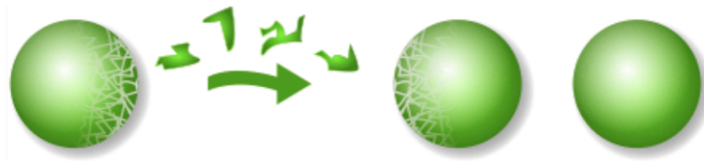
Picture due to Benjamin D. Esham for the Wikimedia Commons

Figure 2: The Banach-Tarski Paradox.

3.1 Equipollence

Equipollence of two sets means the existence of a bijection between the sets. A bijection is a mapping which is both one-to-one and onto. In other words, a bijection between two sets a and b is a function $f : a \to b$ so that for every $y \in b$ there is a unique $x \in a$ such that $f(a) = b$. Still in other words the equipollence of a and b means the existence of functions $f : a \to b$ and $g : b \to a$ such that for all $x \in a$ we have $g(f(x)) = x$ and for all $y \in b$ we have $f(g(y)) = y$. In set theory it is thought that if two sets are equipollent, then they have the same number of elements. Because the sets may be infinite, it is not a priori clear what it means to say that the sets have the same number of elements. However, if there is a bijection between the sets, it is quite credible to argue that whatever we mean by the number of elements of an infinite set, equipollent sets should get the same number.

For finite sets equipollence means indeed that the sets have the same number of elements. For infinite sets we have to give up the idea that *the part is smaller than the whole*, since for example the set of natural numbers $\{0, 1, 2, ...\}$ is equipollent with its

proper part $\{1, 2, 3, ...\}$, as the bijection $n \mapsto n + 1$ demonstrates. The part may not be smaller than the whole but at least it cannot be greater than the whole. And in some cases the part is smaller than the whole. Cantor proved that the set $\mathbb{N}$ of natural numbers is not equipollent with the set $\mathbb{R}$ of real numbers. This can be seen as follows: Suppose there were a bijection $f : \mathbb{N} \to \mathbb{R}$. Then there is an onto function $g : \mathbb{N} \to [0, 1]$. Let us construct a real number on $[0, 1]$ as follows. The number is $0.d_1 d_2 d_3...$ where $d_i = 1$ if the real number $g(i)$ has the decimal expansion $0.e_1 e_2 e_3...$, where $e_i \neq 1$. Otherwise $d_i = 0$. In this way we obtain a real number $r \in [0, 1]$. Since g is onto, there is $n \in \mathbb{N}$ such that $g(n) = r$. Let us look at d_n. We have $d_n = 1$ if and only if $d_n \neq 1$, a contradiction. Hence no such f can exist. So $\mathbb{N}$ is less than the whole $\mathbb{R}$ in harmony with our intuition. This result is due to Cantor. He went on to prove that the set $\mathbb{Q}$ of all rational numbers is equipollent with $\mathbb{N}$ and hence not equipollent with $\mathbb{R}$. Moreover, he showed that the set $\mathbb{A}$ of all algebraic numbers is also equipollent with $\mathbb{N}$ and hence not equipollent with $\mathbb{R}$. We get the surprising conclusion that there are fewer algebraic numbers than real numbers, hence many (if not most) of the real numbers must be transcendental. This was a remarkable conclusion by Cantor because at the time when the observation was made, very few transcendental numbers were known. Thus by purely abstract set theoretic methods Cantor had proved the existence of many many transcendental numbers.

Technically speaking, a bijection between two sets a and b is a function $f : a \to b$ which is *one-one* i.e. $\forall x \in a \forall y \in a(f(x) = f(y) \to x = y)$ and *onto* i.e. $\forall y \in b \exists x \in a(f(a) = b)$. With this definition, sets a and b are *equipollent*, $a \sim b$, if there is a bijection $f : a \to b$. Then $f^{-1} : b \to a$ is a bijection and $b \sim a$ follows. The composition of two bijections is a bijection, whence

$$a \sim b \sim c \Longrightarrow a \sim v.$$

Thus $\sim$ divides sets into equivalence classes. Each equivalence class has a canonical representative (a cardinal number, see Subsection "Cardinals" below) which is called the *cardinality* of (each of) the sets in the class. The cardinality of a is denoted by $|a|$ and accordingly $a \sim b$ is often written $|a| = |b|$. One of the basic properties of equipollence is that if

$$a \sim c, b \sim d \text{ and } a \cap b = c \cap d = \emptyset,$$

then

$$a \cup b \sim c \cup d.$$

Indeed, if $f : a \to c$ is a bijection and $g : b \to d$ is a bijection, then $f \cup g : a \cup b \to c \cup d$ is a bijection. If the assumption $a \cap b = c \cap d = \emptyset$ is dropped, the conclusion fails, of course, as we can have $a \cap b = \emptyset$ and $c = d$. It is also interesting to note that even if $a \cap b = c \cap d = \emptyset$, the assumption $a \cup b \sim c \cup d$ does not imply $b \sim d$ even if $a \sim c$ is assumed: Let $a = \mathbb{N}, b = \emptyset, c = \{2n : n \in \mathbb{N}\}$ and $d = \{2n + 1 : n \in \mathbb{N}\}$. However, for finite sets this holds: if $a \cup b$ is finite, $a \cup b \sim c \cup d, a \sim c, a \cap b = a \cap d = \emptyset$ then $b \sim d$. We can interpret this as follows: the cancellation law holds for finite numbers but does not hold for cardinal numbers of infinite sets.

A basic fact about equipollence, and indeed the starting point of all of set theory, is the result of Cantor that no set is equipollent with its power set. Let us see why this is

so. Suppose a set a is equipollent with $\mathcal{P}(a)$. Thus there is a bijection $f : a \to \mathcal{P}(a)$. Let $b = \{x \in a : x \notin f(x)\}$. Then $b \in \mathcal{P}(a)$ so there is some $x \in a$ such that $b = f(x)$. Is x in b or not? If $x \in b$, then $x \notin f(x)$, a contradiction, since $f(x) = b$. Therefore we must conclude $x \notin b$. But then $x \notin f(x)$, whence $x \in b$, a contradiction again. So no such f can exist. It is remarkable that with this simple short argument one can make the far-reaching conclusion that there are an unending sequence of greater and greater cardinalities, namely one needs only follow the sets $\mathbb{N}$, $\mathcal{P}(\mathbb{N})$, $\mathcal{P}(\mathcal{P}(\mathbb{N}))$, $\mathcal{P}(\mathcal{P}(\mathcal{P}(\mathbb{N})))$,...

There are many more interesting and non-trivial properties of equipollence that we cannot enter into here. For example the Schröder-Bernstein Theorem[1]: If $a \sim b$ and $b \subseteq c \subseteq a$, then $a \sim c$.

3.2 Countable sets

Countable sets are the most accessible infinite sets. They are the infinite sets that we can actually list, or rather, we can start listing a countable set and if we lived forever, we would list the entire set. So this is in sharp contrast to sets like $\mathbb{R}$, the set of all reals. Even if one lived forever, one could not list all real numbers. The quintessential example of a countable set is the set $\mathbb{N}$ of all natural numbers. Any set that is indexed by the natural numbers as $\{a_n : n \in \mathbb{N}\}$ is likewise called countable. And now we have already exhausted the class of countable sets! There are no others.

Countable sets already manifest the paradoxical feature of infinity that *the part need not be less than the whole*, for even the simplest countable set $\{0, 1, 2, ...\}$ is equipollent with its proper subset $\{1, 2, 3, ...\}$ via the bijection $n \mapsto n + 1$. By considering the bijection $n \mapsto 2n$ we can see that $\{0, 1, 2, ...\}$ is equipollent with the set of even numbers $\{0, 2, 4, 6, ...\}$. In fact, *all* infinite countable sets are equipollent: Suppose $A = \{a_n : n \in \mathbb{N}\}$ and $B = \{b_n : n \in \mathbb{N}\}$ are two infinite sets. Let $f(a_0) = b_0$. If $f(a_n) \in B$ has been defined, let $f(a_{n+1})$ be b_m with the smallest m such that $b_m \notin \{f(a_0), ..., f(a_n)\}$. Since B is infinite, such an m must always exist. Moreover, every b_m gets chosen at some point, for obviously $b_m \in \{f(a_0), ..., f(a_m)\}$.

Intuitively there are much more rational numbers than integers. Therefore it is a bit surprising that the set of all rational numbers is actually countable. Let us see how we can arrive at this conclusion. We can identify the rational number n/m (in lowest terms) with the ordered pair (n, m) of natural numbers. So let us first show that if a and b are countable, then so is $a \times b$. If either set is empty, the cartesian product is empty. So let us assume the sets are both non-empty. Suppose $a = \{a_0, a_1, ...\}$ and $b = \{b_0, b_1, ...\}$. Let

$$c_n = \begin{cases} (a_i, b_j), & \text{if } n = 2^i 3^j \\ (a_0, b_0), & \text{otherwise.} \end{cases}$$

Now $a \times b = \{c_n : n \in \mathbb{N}\}$, whence $a \times b$ is countable. So if we identify a rational number n/m (in lowest terms) with the pair (n, m), then there is some k such that $(n, m) = c_k$, and we have identified the set of all (non-negative) rational numbers with an infinite subset of $\mathbb{N}$, so in particular it is countable.

---

[1]The original formulation says: If there is a one-one function $a \to b$ and another $b \to a$ there is a bijection $a \to b$, see e.g. [12, p. 27].

We showed above that the cartesian product of two countable sets if countable. A similar, and very useful fact is the following: a countable union of countable sets is countable. The empty sets do not contribute anything to the union, so let us assume all the sets in our countable family are non-empty. Suppose A_n is countable for each $n \in \mathbb{N}$, say, $A_n = \{a_m^n : m \in \mathbb{N}\}$ (we use here the Axiom of Choice to choose an enumeration for each A_n). Let $B = \bigcup_n a_n$. We want to represent B in the form $\{b_n : n \in \mathbb{N}\}$. If n is given, we consider two cases: If n is $2^i 3^j$ for some i and j, we let $b_n = a_j^i$. Otherwise we let $b_n = a_0^0$. Now indeed $B = \{b_n : n \in \mathbb{N}\}$.

One of the reasons why countable sets are so important is that sets defined by induction are usually countable. Examples of such sets are abundant in logic, most notably the set of terms and and the set of formulas in a countable vocabulary. Any formal language based on a countable vocabulary generates a countable set of expressions. More generally, in a countable vocabulary the set of all strings of symbols of a fixed finite length is countable, and hence so is the set of all finite strings of symbols, as it is the union of a countable family of countable sets.

A powerful application of the above idea is the Löwenheim-Skolem Theorem of first order logic: Every countable first order theory has a countable model. There are reasons to believe—although this view is also contested[2]—that first order theories represent the best axiomatizations that we can ever get. Thus we are stuck with countable models whether we want it or not. For set theory this is called Skolem's Paradox. The paradox is that we can prove in set theory that the set of all reals is uncountable, but still set theory itself has countable models. That is the paradox. The solution of the paradox is that what seems countable from outside may not seem countable inside. More exactly, if we have a countable model of set theory, we can be sure that the mapping from the natural numbers onto the model is not an element of the model. This is a rough awakening to the reality that everything in set theory is relative. There are no signs that this would be the fault of set theory. It is even true of number theory vis a vis Gödel's Incompleteness Theorem.

3.3 Ordinals

The ordinal numbers introduced by Cantor are a marvelous general theory of measuring the *potentially infinite* on the one hand, and the actually infinity on the other hand. They are intimately related to inductive definitions and occur therefore widely in logic. It is easiest to understand ordinals in the context of games, although this was not Cantor's way. Suppose we have a game with two players I and II. It does not matter what the game is, but it could be something like chess. If II can force a win in n moves we say that the game has *rank* n. Suppose then II cannot force a win in n moves for any n, but after she has seen the first move of I, she can fix a number n and say that she can force a win in n moves. This situation is clearly different from being able to say in advance what n is. So we invent a symbol ω for the rank of this game. In a clear sense ω is greater than each n but there does not seem to be any possible rank between all the finite numbers n and ω. We can think of ω as an infinite number. However, there is nothing metaphysical about the infiniteness of ω. It just has infinitely many

---

[2]See e.g. [19].

predecessors. We can think of ω as a tree T_ω with a root and a separate branch of length n for each n above the root as in the tree on the left in Figure 3.

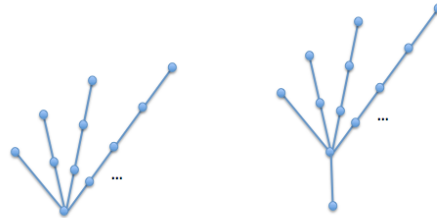Figure 3: T_ω and $T_{\omega+1}$.

Suppose then II is not able to declare after the first move how many moves she needs to beat I, but she knows how to play her first move in such a way that after I has played his second move, she can declare that she can win in n moves. We say that the game has rank $\omega + 1$ and agree that this is greater than ω but there is no rank between them. We can think of $\omega + 1$ as the tree which has a root and then above the root the tree T_ω, as in the tree on the right in Figure 3. We can go on like this and define the ranks $\omega + n$ for all n.

Suppose now the rank of the game is not any of the above ranks $\omega + n$, but still II can make an interesting declaration: she says that after the first move of I she can declare a number m so that after m moves she declares another number n and then in n moves she can force a win. We would say that the rank of the game is $\omega + \omega$. We can continue in this way defining ranks of games that are always finite but potentially infinite. These ranks are what set theorists call ordinals.

3.4 Cardinals

Historically cardinals (or more exactly cardinal numbers) are just representatives of equivalence classes of equipollence. Thus there is a cardinal number for countable sets, denoted $\aleph_0$, a cardinal number for the set of all reals, denoted $\mathfrak{c}$, and so on. There is some question as to what exactly are these cardinal numbers. The Axiom of Choice offers an easy answer, which is the prevailing one, as it says that every set can be well-ordered. Then we can let the cardinal number of a set be the order type of the smallest well-order equipollent with the set. Equivalently, the cardinal number of a set is the smallest ordinal equipollent with the set. If we leave aside the Axiom of Choice, some sets need not have have a cardinal number. However, as is customary in current set theory, let us indeed assume the Axiom of Choice. Then every set has a cardinal number and the cardinal numbers are ordinals, hence well-ordered. The αth infinite cardinal number is denoted $\aleph_\alpha$. Thus $\aleph_1$ is the next in order of magnitude from $\aleph_0$. The famous *Continuum Hypothesis* is the statement that $\aleph_1 = \mathfrak{c}$. Equivalently, for every set

A of reals, either A is countable of the cardinal number of A is $\mathfrak{c}$. For Borel[3] sets of real numbers it is true that there is no cardinality between $\aleph_1$ and $\mathfrak{c}$. If we assume large cardinals[4], it is even true that for sets of reals definable with real parameters there is no cardinality between $\aleph_1$ and $\mathfrak{c}$. So it is not *so* far-fetched to suggest that maybe the same holds for *all* sets of reals. On the other hand, the tenet of set theory is that properties of *definable* sets are different from the properties of *arbitrary* sets. So maybe indeed the "regular" sets of reals—for some sense of "regular"—obey the Continuum Hypothesis but when we enter the absurd and unintuitive world of totally undefinable—arbitrary—sets of reals, the Continuum Hypothesis fails.

4   Axiomatic set theory

After the above tour of basic concepts of set theory we can return to the beginning and ask what is it that we are doing. This is all the more important because, as we have indicated, a lot of mathematics can be developed in set theory, if not all of mathematics. So the philosophical question arises, what is set theory based on? The most commonly held view is that set theory is the most fundamental theory in mathematics and it is not possible to base set theory on anything even more primitive.

So how do we really know what is true of sets and what is not? This question is crucially important also because most of the sets we encounter in set theory are infinite and unquestionably abstract. They seem to exist only in their own abstract world which cannot be seen by eyes, binoculars or microscopes, cannot be touched by hand, and cannot be observed by listening, tasting or smelling. It is often said that we can observe sets only by thinking of them, but this seems an inadequate answer. The most commonly held view is that we simply accept certain simple facts about sets as axioms and then use rules of logic to derive more complicated facts. The axioms are accepted because of their intuitive appeal and because of their usefulness. From the axioms that we present below one can derive virtually all of mathematics, and that is ultimately the most important reason for accepting them. They simply seem to give a "house" for mathematics to live in.

Technically speaking, the axioms are first order sentences in a vocabulary which has just one binary predicate symbol $\in$ in addition to identity.

The simple idea of sets as collections of objects is too loose in a closer analysis. This can be seen from the many paradoxes it has led to. The most important is *Russell's Paradox*: Consider the set R of sets that are not elements of themselves. If $R \in R$, then $R \notin R$, and if $R \notin R$, then $R \in R$. This paradox shows that we cannot allow just any collection to be a set. Current thinking is that sets are in a sense "small" enough collections to be considered as sets. According to this thinking, arbitrary collections are called *classes*. A class that is not a set is called a *proper class*.

In the axiomatic approach paradoxes like Russell's Paradox are avoided because

---

[3]The class of Borel sets is the smallest class of sets containing the open sets and closed under complements and countable unions, see [12, p. 132].

[4]Large cardinals are "large" cardinals that have special properties that are used in proofs. Their existence cannot be proved, so they have to be just assumed. However, they seem quite necessary in modern set theory, see e.g. [12, p. 275]

sets and proper classes are kept away from each other. Technically speaking, objects in the axiomatic approach, that is, the range of all quantifiers, is sets. Classes are treated via formulas. A formula $\varphi(x)$, with perhaps parameters, is identified with the class $\{a : \varphi(a)\}$ of sets that satisfy $\varphi(x)$. So even if we think of our formulas as talking only about sets, we can talk about classes by talking about formulas defining the classes.

There is an intuitive model of set theory which goes beyond the simple idea that sets are "collections" of objects. According to this intuition sets have been created in stages. Elements of a set are, or have been, created before the set itself. This intuition does not mean that sets have really been created by someone, it is just a metaphor. The concept of an ordinal can be used to make the intuitive idea of stages more exact. The more exact version is called the *cumulative hierarchy* of sets. For this end, let $V_0 = \emptyset$, $V_{\alpha+1} = \mathcal{P}(V_\alpha)$ and $V_\nu = \bigcup_{\alpha<\nu} V_\alpha$ if ν is a limit ordinal. Finally, let $V = \bigcup_\alpha V_\alpha$. This is the intuitive model of set theory. Strictly speaking, it is not model in the sense of model theory because its domain is a proper class.

Now we present the axioms of set theory. They are called the Zermelo-Fraenkel axioms, denoted ZFC. When we discuss the axioms it is good to keep in mind the intuitive model offered by the cumulative hierarchy.

1. **Axiom of Extensionality:** Sets which have the same elements are equal i.e.

$$\forall x \forall y (\forall z (z \in x \leftrightarrow z \in y) \rightarrow x = y).$$

This axiom seems obvious but it is actually a deep axiom. It demonstrates that we do not want there to be anything else about sets than their elements. The elements form an aggregate we call a "set" but we do not *care* what it is that pulls these elements together. The opposite attitude would be to think that there is much more to a set than its elements, e.g. the *way*, whatever it means, how the elements are connected together into a set.

2. **Axiom of Pair:** From any two sets a and b we can form a new set $\{a, b\}$ which has exactly a and b as elements i.e.

$$\forall x \forall y \exists z \forall u (u \in z \leftrightarrow (u = x \lor u = y)).$$

Note that $\{a, b\}$ is *not* the union of a and b - however big sets a and b are the set $\{a, b\}$ has at most two elements, so in particular it is always finite. It is perfectly possible that $a = b$ and then $\{a, b\} = \{a\}$. We can form sets like $\{\mathbb{N}, \mathbb{Q}\}$, $\{\mathbb{Q}\}$ and $\{\mathbb{N}, \mathbb{Q}, \mathbb{R}\}$. Such sets are not particularly common or useful, but their existence in set theory is a manifestation of the basic tenet: whenever we have a set, we consider it as a "completed" totality, something we can use to build new sets.

3. **Axiom of Union:** For any set a we can form the *union* $\bigcup a$ of a, which consists of all sets which are elements of elements of a i.e

$$\forall x \exists y \forall z (z \in y \leftrightarrow \exists u (u \in x \land z \in u)).$$

Often sets are given in the form $a = \{a_i : i \in I\}$, that is, a is the range of the function $i \mapsto a_i$. Then $\bigcup a$ is the set $\bigcup_{i \in I} a_i$. This is a basic operation in mathematics and many applications of set theory.

4. **Axiom of Power set:** For any set a we can form the *power set* $\mathcal{P}(a)$ of a which consists of all sets which are subsets of a i.e

$$\forall x \exists y \forall z (z \in y \leftrightarrow \forall u(u \in z \rightarrow u \in x)).$$

One often hears criticism of this axiom but often also for a wrong reason. The problem with this axiom is *not* that it says that "all" subsets of a, whatever that means, exist. It says that those subsets which *do* exist can be collected together. The opposite of this axiom would be to think that some power sets are so large that they are proper classes. For example, we could think that, opposite to what the power set axioms says, the set of all reals, which is essentially the power set of $\mathbb{N}$, is a proper class. This is a coherent idea, but it does not mean that we have missed some subsets. We have all the subsets that we have, but we just cannot pull all of them together into a set. A smooth theory of the reals seems to require the power set axiom, but there are also alternative approaches.

5. **Axiom Schema of Subsets:** For any set a we can form a new set by taking the intersection of a and any class. In particular we can form new sets of the form $\{x \in a : \varphi(x)\}$ where $\varphi(x)$ is any formula. More exactly, for any formula $\varphi(x, \vec{y})$ we have the following axiom:

$$\forall x \forall x_1 ... \forall x_n \exists y \forall z (z \in y \leftrightarrow (z \in x \wedge \varphi(z, \vec{x}))).$$

Sets of the form $\{x \in a : \varphi(x)\}$ are very common in mathematics, for example $a \cap b = \{x \in a : x \in b\}$. Combined with the axioms of pair, union and power set, the Axiom of Subsets is very powerful indeed. This axiom has the impredicative element that the formula $\varphi(x)$ in $\{x \in a : \varphi(x)\}$ can have quantifiers and because these quantifiers range over the entire universe of sets the set $\{x \in a : \varphi(x)\}$ itself is also in the range of the quantifiers. We can remove this impredicativity by requiring that all quantifiers in $\varphi(x)$ are *bounded* i.e. of the form $\forall y \in z$ or $\exists y \in z$. However, this limits the applicability of the axiom seriously and leads to completely different kind of set theory, the so called Kripke-Platek set theory (see [4]).

6. **Axiom Schema of Replacement:** Suppose a is a set. If there is a way to associate to every element i of a a new set a_i, then we can form a new set $\{a_i : i \in a\}$, that is, a set which has all the a_i, where $i \in a$, as elements, and nothing else. More exactly, for any formula $\varphi(x, \vec{y})$ we have the following axiom:

$$\forall x \forall x_1 ... \forall x_n (\forall u \forall z \forall z'((u \in x \wedge \varphi(u, z, \vec{x}) \wedge \varphi(u, z', \vec{x})) \rightarrow z = z')$$
$$\rightarrow \exists y \forall z (z \in y \leftrightarrow \exists u(u \in x \wedge \varphi(u, z, \vec{x})))).$$

This axiom introduced by Fraenkel is needed e.g. in transfinite recursion.

7. **Axiom of Infinity:** This axiom simply says that there is an infinite set. More exactly,

$$\exists x (\exists y(y \in x \wedge \forall z \neg(z \in y)) \wedge \forall y(y \in x \rightarrow \exists z(y \in z \wedge z \in x))).$$

There are many ways to write this axiom, all equivalent, given the other axioms. The particular formulation here yields the set $A = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, ...\}$. It is easy to see on the basis of the Axiom of Extensionality that all elements of this set A are different.

8. **Axiom of Foundation:** This axiom says that every set has an element which is minimal with respect to $\in$, that is

$$\forall x \exists y (x \cap y = \emptyset).$$

This is the most useless axiom (of set theory) that anyone ever invented. In fact there are reasons to claim that no-one ever used this axiom! However, since the intuitive idea of sets is that they were "created" in stages, with elements of a set having been created before the set itself, then of course every set has an $\in$-minimal element, namely the one that was 'created first". Since we do not really think sets were created—creation being a mere metaphor—there is hardly any mathematical example where this axioms turns up. Set theorists count it in for their internal aesthetic reasons. Its usefulness is not based in what it gives but rather in that we can live without the circular sets it excludes.

5 Axiom of Choice

The Axiom of Choice is one of the axioms of set theory but we treat it here separately from the others because it is of a slightly different character. The Axiom of Choice states that if a set a of non-empty sets is given, then there is a function f such that $f(x) \in x$ for all $x \in a$. That is, the function f picks one element from each of those non-empty sets. There are so many equivalent formulations of this axiom that books have been written about it. The most notable is the Well-Ordering Principle: every set is equipollent with an ordinal (see e.g. [12, p. 45]).

The Axiom of Choice is the only axiom of ZFC which brings arbitrariness or abstractness into set theory, often with examples that can be justifiably called pathological, like the Banach-Tarski Paradox (see above). Every other axiom states the existence of some set and specifies what the set is. The Axiom of Union says the new set is the union $\bigcup_{i \in A} B_i$, the Axiom of Power Set says the new set is the powerset $\{B : B \subseteq A\}$, the Axiom of Subsets states that the new set is of the form $\{b \in a : \varphi(b)\}$.

Because of the abstractness brought about by the Axiom of Choice it has received criticism and some authors always mention explicitly if they use it in their work. The main problem in working without the Axiom of Choice is that there is no clear alternative and just leaving it out leaves many areas of mathematics, like measure theory, without proper foundation.

A basic problem with an axiom like the Axiom of Choice is that it has a formulations which are rather obvious, like the formulation above, and equivalent formulations which are completely unbelievable, like the Well-Ordering Principle. If one thinks of formulations that make it look obvious, one would like to accept it, but when one looks at the unbelievable consequences one would like to reject it. So which way to go?

It is sometimes wrongly believed that the problem of the Axiom of Choice is in that no-one knows which element to choose from each non-empty set. This is not the point. If a set a is non-empty, i.e. it is not the case that every set is not in a, then by the laws of logic there must be a set in a. This does not require the Axiom of Choice as it is simply a consequence of provability of $\neg\forall x\neg A \rightarrow \exists xA$. The problem is how to make *infinitely* many such choices.

6   Independence results

In set theory it is relatively easy to formulate questions that have turned out to be impossible to decide on the basis of the axioms. The most famous of these is the Continuum Hypothesis, already proposed by Cantor. The Continuum Hypothesis claims that every uncountable set of reals is equipollent with the entire set of reals $\mathbb{R}$ (see discussion on Continuum Hypothesis in Section 3.4).

The undecidability of a sentence on the basis of any axioms, set theory or not, can be proved by producing two models of the axioms, one where the sentence is true and another where it is false. In the case of the Continuum Hypothesis such two models have indeed been produced (see e.g. [12, chapters 13 and 14]). The two models, one due to Kurt Gödel and the other due to Paul Cohen, have led to an extensive study of models of set theory, and a profusion of different kinds of models have been uncovered. Most of these models are constructed by a method called *forcing*. This highly interesting method has turned out to be of relevance also outside set theory.

The basic idea of forcing is that instead of trying to build directly a model where something we are interested in is true, we settle with something less. We settle with contemplating what finite pieces of information, called *conditions*, "force" to be true, if ever a model based on them was constructed. For example, if we have a name $\dot{A}$ for a set of natural numbers, then the condition $\{0 \in \dot{A}, 1 \notin \dot{A}\}$ forces $\dot{A}$ to contain 0 but not 1, and this condition leaves it open whether e.g. 2 is in $\dot{A}$ or not. We form a particular infinite sequence of conditions called a generic sequence and build a model, called a generic model from that sequence. Remarkably, a sentence is true in the generic model if and only if some condition in the generic sequence forces it to be true. This can be done in such a manner that the Continuum Hypothesis is forced to be true or false in the generic model according to our will. If we want the Continuum Hypothesis to be true we use one kind of condition and if we want it to be false we use another kind of condition. For more on forcing see [12, Chapter 14] and [15, Chapter VII].

Forcing has turned out to have a connection to both modal and intuitionistic logic. This connection arises from the fact that we can think of the set of forcing conditions as the frame of a Kripke structure. For example, a condition p is said to force $\neg\varphi$ if and only if no extension of p forces φ. This is exactly the same as the definition of the truth of a negated sentence of intuitionistic logic at a node of a Kripke structure.

The philosophical importance of forcing is manifold. It represents a useful *weak* truth definition, and as such one which can be used in different parts of philosophical logic. It uncovers a huge gap in what the axioms of set theory decide leading to the philosophical question, whether there is ultimately any true universe of mathematical objects. Skeptics say that Gödel's Incompleteness Theorem casts a doubt on the ex-

istence of mathematical objects, and Cohen's forcing, especially the independence of the Continuum Hypothesis, was the last blow which to many people totally shattered the idea of a platonist reality of mathematics. The opposite view is that mathematical objects form a definite unique reality of their own and the results of Gödel and Cohen merely manifest an inherent underdetermination of the axioms of set theory in uncovering what is true in this invisible world and what is not.

7 Some recent work

7.1 Descriptive Set Theory

A set A is said to be *definable* if there is a formula $\varphi(x)$ such that A is the set of sets b that satisfy $\varphi(x)$. Since there are only countably many formulas there can be only countably many definable sets. However, if we allow parameters, we get more definable sets. Typical parameters that are sometimes allowed are on the one hand ordinal numbers and on the other hand real numbers. Descriptive Set Theory is an important sector of set theory which concentrates on sets that are definable with real parameters. The basic ideology is that the arbitrariness or pathology brought by the Axiom of Choice is only manifested in the realm of undefinable sets. The sets we actually work with are a fortiori definable—otherwise we could not talk about them! Seminal results of Martin-Steel-Woodin ([17]) show that assuming so called large cardinals, phenomena like the Banach-Tarski Paradox do not occur among definable sets. In other words, large cardinals remove the negative effect of arbitrariness that the Axiom of Choice brings to set theory. The abstract arbitrary sets are there, and are needed for the general theory, but they do not disturb the world of definable sets with their paradoxical counter-intuitive properties. Current work in Descriptive Set Theory further emphasizes this and at the same time brings set theory closer and closer to classical analysis, topology and measure theory (see e.g. [5]).

7.2 Non well-founded set theory

The Non-well-founded set theory of Peter Aczel ([2]) takes on the empirical fact that the Axiom of Foundation is not really a necessary axiom. So non-well-founded set theory replaces the Axiom of Foundation with its ultimate strongest possible denial: any combination of circularity in the $\in$-relation is manifested by some sets. Circularity comes up naturally in computer science: the state of a program may very well come back to itself. Of course, the common sense view is that then the program is in a loop and can be "dismissed" as a program with a bug. However, another common sense view is that most programs can enter a loop, and some programs, like operating systems, are even expected to come back to the same state time after time. It has turned out that non-well-founded set theory can be used to model conveniently processes in computer science (see e.g. [3]).

7.3 Constructive set theory

Constructive set theory drops classical logic from set theory. As a result, $\neg\forall x\neg\varphi(x)$ is not anymore a guarantee for $\exists x\varphi(x)$. For us to assert $\exists x\varphi(x)$ we have to have a construction of an x and a proof that $\varphi(x)$. At first sight this seems to have devastating consequences for set theory. However, if we just adopt constructive logic but do not change the axioms we do not gain much ([11]). To really make a difference in the direction of constructive mathematics, one has to rethink the axioms. One approach gaining popularity is the *Constructive Zermelo Fraenkel Set Theory CZF* (see [1]). The goal of CZF is to offer a simple intuitive foundation for constructive mathematics in the same way as ZFC offers one for classical mathematics.

8 Historical Remarks and Further Reading

Set theory was launched by Georg Cantor (see [6] and [7]) in 1874. There are many elementary books providing an introduction to set theory, for example [8], [9], [18], [16]. Textbooks covering a wide spectrum of modern set theory are [13] and [14]. A colossal recent source of advanced set theory is [10].

References

[1] Peter Aczel. The type theoretic interpretation of constructive set theory. In *Logic Colloquium '77 (Proc. Conf., Wrocław, 1977)*, volume 96 of *Stud. Logic Foundations Math.*, pages 55–66. North-Holland, Amsterdam, 1978.

[2] Peter Aczel. *Non-well-founded sets*, volume 14 of *CSLI Lecture Notes*. Stanford University Center for the Study of Language and Information, Stanford, CA, 1988. With a foreword by Jon Barwise [K. Jon Barwise].

[3] Peter Aczel. Final universes of processes. In *Mathematical foundations of programming semantics (New Orleans, LA, 1993)*, volume 802 of *Lecture Notes in Comput. Sci.*, pages 1–28. Springer, Berlin, 1994.

[4] Jon Barwise. *Admissible sets and structures*. Springer-Verlag, Berlin, 1975. An approach to definability theory, Perspectives in Mathematical Logic.

[5] Howard Becker and Alexander S. Kechris. *The descriptive set theory of Polish group actions*, volume 232 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1996.

[6] Georg Cantor. *Contributions to the founding of the theory of transfinite numbers*. Dover Publications Inc., New York, N. Y., 1952. Translated, and provided with an introduction and notes, by Philip E. B. Jourdain.

[7] Joseph Warren Dauben. *Georg Cantor*. Princeton University Press, Princeton, NJ, 1990. His mathematics and philosophy of the infinite.

[8] Keith Devlin. *The joy of sets*. Springer-Verlag, New York, second edition, 1993. Fundamentals of contemporary set theory.

[9] Herbert B. Enderton. *Elements of set theory*. Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1977.

[10] Matthew Foreman and Akihiro Kanamori (Eds.). *Handbook of Set Theory*. Springer-Verlag, Berlin, 2010.

[11] Harvey Friedman. The consistency of classical set theory relative to a set theory with intuitionistic logic. *J. Symbolic Logic*, 38:315–319, 1973.

[12] Thomas Jech. *Set theory*. Perspectives in Mathematical Logic. Springer-Verlag, Berlin, second edition, 1997.

[13] Thomas Jech. *Set theory*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2003. The third millennium edition, revised and expanded.

[14] Akihiro Kanamori. *The higher infinite*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, second edition, 2003. Large cardinals in set theory from their beginnings.

[15] Kenneth Kunen. *Set theory*. North-Holland Publishing Co., Amsterdam, 1980. An introduction to independence proofs.

[16] Kenneth Kunen. *Set theory*, volume 102 of *Studies in Logic and the Foundations of Mathematics*. North-Holland Publishing Co., Amsterdam, 1983. An introduction to independence proofs, Reprint of the 1980 original.

[17] Donald A. Martin and John R. Steel. A proof of projective determinacy. *J. Amer. Math. Soc.*, 2(1):71–125, 1989.

[18] B. Rotman and G. T. Kneebone. *The theory of sets and transfinite numbers*. Oldbourne, London, 1966.

[19] Stewart Shapiro. The "triumph" of first-order languages. In *Logic, meaning and computation*, volume 305 of *Synthese Lib.*, pages 219–259. Kluwer Acad. Publ., Dordrecht, 2001.