Chapter 4

Constructions on sets

Forewarned by a problem first exposed by Bertrand Russell, we look to safe methods for constructing sets.

4.1 Russell's paradox

When set theory was being invented it was thought, first of all, that any property $P(x)$ determined a set

$$\{x \mid P(x)\} .$$

It came as a shock when Bertrand Russell realised that assuming the existence of certain sets described in this way gave rise to contradictions.[1]

Russell's paradox is really the demonstration that a contradiction arises from the liberal way of constructing sets above. His argument proceeds as follows. Consider the property

$$x \notin x$$

a way of writing "x is not an element of x." If we assume that properties determine sets, just as described, we can form the set

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

Either $R \in R$ or not. If so, *i.e.* $R \in R$, then in order for R to qualify as an element of R, from the definition of R, we deduce $R \notin R$. So we end up asserting both something and its negation—a contradiction. If, on the other hand, $R \notin R$ then from the definition of R we see $R \in R$—a contradiction again. Either $R \in R$ or $R \notin R$ lands us in trouble.

We need to have some way which stops us from considering a collection like R as a set, and so as a legitimate element. In general terms, the solution is to discipline the way in which sets are constructed, so that starting from certain given sets, new sets can only be formed when they are constructed by using particular, safe ways from old sets. We shall state those sets we assume to exist right from the start and methods we allow for constructing new sets. Provided these are followed we avoid trouble like Russell's paradox and at the same time have a rich enough world of sets to support most mathematics.[2]

4.2 Constructing sets

4.2.1 Basic sets

We take the existence of the empty set $\emptyset$ for granted, along with certain sets of basic elements such as

$$\mathbb{N}_0 = \{0, 1, 2, \cdots\} .$$

We shall also take sets of symbols like

$$\{\text{“a”}, \text{“b”}, \text{“c”}, \text{“d”}, \text{“e”}, \cdots, \text{“z”}\}$$

[1] The shock was not just to Russell and his collaborator Alfred North Whitehead. Gottlob Frege received the news as his book on the foundations of mathematics via sets was being printed—the paradox was devastating for his work. Some were delighted however. The great mathematician Henri Poincaré is reported as gleefully saying "Logic is not barren, it's brought forth a paradox!"

[2] Occasionally we consider collections which are not sets. For example, it can be useful to consider the collection of all sets. But such a collection is not itself a set, so cannot be made a proper element of any collection. The word 'class' which originally was synonymous with 'set' is now generally reserved for a collection which need not necessarily be a set.

for granted, although we could, alternatively have represented them as particular numbers for example. The equality relation on a set of symbols is that given by syntactic identity written =. Two symbols are equal iff they are literally the same.

4.2.2 Constructions

We shall take for granted certain operations on sets which enable us to construct sets from given sets.

Comprehension

If X is a set and $P(x)$ is a property, we can form the set

$$\{x \in X \mid P(x)\} \ ,$$

the subset of X consisting of all elements x of X which satisfy $P(x)$.

Sometimes we'll use a further abbreviation. Suppose $e(x_1, \ldots, x_n)$ is some expression which for particular elements $x_1 \in X_1, \cdots x_n \in X_n$ yields a particular element and $P(x_1, \ldots, x_n)$ is a property of such $x_1, \ldots, x_n$. We use

$$\{e(x_1, \ldots, x_n) \mid x_1 \in X_1 \ \& \ \cdots \& \ x_n \in X_n \ \& \ P(x_1, \ldots, x_n)\}$$

to abbreviate

$$\{y \mid \exists x_1 \in X_1, \cdots, x_n \in X_n.\ y = e(x_1, \ldots, x_n) \& \ P(x_1, \ldots, x_n)\} \ .$$

For example,

$$\{2m + 1 \mid m \in \mathbb{N} \ \& \ m > 1\}$$

is the set of odd numbers greater than 3.

Remark Note a consequence of comprehension. The collection $\{x \mid x \text{ is a set }\}$ is not itself a set. If it were then by using comprehension Russell's collection R would be a set, which is contradictory in the manner of Russell's original argument. As the collection of all sets is not a set we fortunately side-step having to consider whether it is a member of itself. □

Powerset

We can form a set consisting of the set of all subsets of a set, the so-called *powerset*:

$$\mathcal{P}(X) = \{Y \mid Y \subseteq X\}.$$

This is *the* important construction for building bigger sets. We shall see shortly that a powerset $\mathcal{P}(X)$ always has larger size than X.

Exercise 4.1 Let B be a fixed subset of the set A. Define the relation R on $\mathcal{P}(A)$ by

$$(X, Y) \in R \iff X \cap B = Y \cap B \ .$$

Show that R is an equivalence relation and describe a bijection between the set of R-equivalence classes and $\mathcal{P}(B)$. □

Unordered pairs

A seemingly modest but important way to produce sets is through forming unordered pairs. Given two objects x and y—they might be sets—we can form the set $\{x, y\}$ whose sole elements are x and y.

Indexed sets

Suppose I is a set and that for any $i \in I$ there is a unique object x_i, maybe a set itself. Then

$$\{x_i \mid i \in I\}$$

is a set. The elements x_i are said to be *indexed* by the elements $i \in I$. Any collection of objects indexed by a set is itself a set.

Union

As we've seen, the set consisting of the *union* of two sets has as elements those elements which are either elements of one or the other set:

$$X \cup Y = \{a \mid a \in X \text{ or } a \in Y\}.$$

This union is an instance of a more general construction, "big union," that we can perform on any set of sets.

Big union

Let X be a set of sets. Their *union*

$$\bigcup X = \{a \mid \exists x \in X.\ a \in x\}$$

is a set. Note that given two sets X and Y we can first form the set $\{X, Y\}$; taking its big union $\bigcup\{X, Y\}$ we obtain precisely $X \cup Y$. When $X = \{Z_i \mid i \in I\}$ for some indexing set I we often write $\bigcup X$ as $\bigcup_{i \in I} Z_i$.

The above operations are in fact enough for us to be able to define the remaining fundamental operations on sets, *viz.* intersection, product, disjoint union and set difference, operations which are useful in their own right.

Intersection

As we've seen, elements are in the *intersection* $X \cap Y$, of two sets X and Y, iff they are in both sets, *i.e.*

$$X \cap Y = \{a \mid a \in X \ \&\ a \in Y\}\ .$$

By the way, notice that one way to write $X \cap Y$ is as $\{a \in X \mid a \in Y\}$ comprising the subset of the set X which satisfy the property of also being in Y; so $X \cap Y$ is a set by Comprehension.

Big intersection

Let X be a nonempty collection of sets. Then

$$\bigcap X = \{a \mid \forall x \in X.\ a \in x\}$$

is a set called its *intersection*. Again, that such an intersection is a set follows by Comprehension.

When $X = \{Z_i \mid i \in I\}$ for a nonempty indexing set I we often write $\bigcap X$ as $\bigcap_{i \in I} Z_i$.[3]

Product

As we've seen, for sets X and Y, their *product* is the set

$$X \times Y = \{(a, b) \mid a \in X \ \&\ b \in Y\}\ ,$$

the set of ordered pairs of elements with the first from X and the second from Y.

More generally $X_1 \times X_2 \times \cdots \times X_n$ consists of the set of n-tuples $(x_1, x_2, \ldots, x_n)$. When all the components of a product are the same set X the n-ary product is written as X^n. By convention X^0, a zero-ary product, is generally understood to be a singleton set consisting just of the empty tuple $()$.

Exercise 4.2 Let X and Y be sets. Define the *projections*

$$\pi_1 : X \times Y \to X \text{ and } \pi_2 : X \times Y \to Y$$

by taking $\pi_1(a, b) = a$ and $\pi_2(a, b) = b$ for $(a, b) \in X \times Y$.

[3] In a context where all sets are understood to be subsets of a given universe U the empty intersection is taken to be U. In general though we can't assume there is a fixed set forming such a universe. We can't take the collection of all sets as the universe as this is not a set.

Let Z be a set and $f : Z \to X$ and $g : Z \to Y$. Show that there is a unique function $h : Z \to X \times Y$ such that $\pi_1 \circ h = f$ and $\pi_2 \circ h = g$.

$$\begin{tikzcd}
& X \times Y \arrow[dl, "\pi_1"'] \arrow[dr, "\pi_2"] & \\
X & & Y \\
& Z \arrow[ul, "f"] \arrow[ur, "g"'] \arrow[uu, dashed, "h"'] &
\end{tikzcd}$$

□

Disjoint union

Frequently we want to join sets together but, in a way which, unlike union, does not identify the same element when it comes from different sets. We do this by making copies of the elements so that when they are copies from different sets they are forced to be distinct:

$$X_1 \uplus X_2 \uplus \cdots \uplus X_n = (\{1\} \times X_1) \cup (\{2\} \times X_2) \cup \cdots \cup (\{n\} \times X_n).$$

In particular, for $X \uplus Y$ the copies $(\{1\} \times X)$ and $(\{2\} \times Y)$ have to be disjoint, in the sense that

$$(\{1\} \times X) \cap (\{2\} \times Y) = \emptyset,$$

because any common element would be a pair with first element both equal to 1 and 2, clearly impossible.

Exercise 4.3 Let X and Y be sets. Define the *injections*

$$inj_1 : X \to X \uplus Y \text{ and } inj_2 : Y \to X \uplus Y$$

by taking $inj_1(a) = (1, a)$ for $a \in X$, and $inj_2(b) = (2, b)$ for $b \in Y$.

Let Z be a set and $f : X \to Z$ and $g : Y \to Z$. Show that there is a unique function $h : X \uplus Y \to Z$ such that $h \circ inj_1 = f$ and $h \circ inj_2 = g$.

$$\begin{tikzcd}
& X \uplus Y \arrow[dd, dashed, "h"'] & \\
X \arrow[ur, "inj_1"] \arrow[dr, "f"'] & & Y \arrow[ul, "inj_2"'] \arrow[dl, "g"] \\
& Z &
\end{tikzcd}$$

□

Set difference

We can subtract one set Y from another X, an operation which removes all elements from X which are also in Y.

$$X \setminus Y = \{x \mid x \in X \;\&\; x \notin Y\} \,.$$

4.2.3 Axioms of set theory

The constructions we have described include most of the assumptions made in more axiomatic treatments of set theory based on the work of Zermelo and Frænkel. To spell out the connection, the existence of the set of natural numbers, powersets, unordered pairs, big unions and sets of indexed objects correspond to the axioms of *infinity*, *powerset*, *pairing*, *union* and *replacement* respectively; we have adopted the axiom of *comprehension* directly, and the axiom of *extensionality* amounts to saying a set is determined by its elements. For completeness we mention two remaining axioms, the axiom of *foundation* and the axiom of *choice*, which are generally assumed of sets. While sensible and safe axioms to assume of sets, they do not in fact follow from the constructions we have given so far.

The axiom of foundation

A set is built-up starting from basic sets by using the constructions described. We remark that a property of sets, called the *axiom of foundation*, follows from our informal understanding of sets and how we can construct them. Consider an element b_1 of a set b_0. It is either a basic element, like an integer or a symbol, or it is a set. If b_1 is a set then it must have been constructed from sets which have themselves been constructed earlier. Intuitively, we expect any chain of memberships

$$\cdots b_n \in \cdots \in b_1 \in b_0$$

to end in some b_n which is some basic element or the empty set. The statement that any such descending chain of memberships must be finite is called the axiom of foundation, and is an assumption generally made in set theory. Notice the axiom implies that no set X can be a member of itself as, if this were so, we'd get the infinite descending chain

$$\cdots X \in \cdots \in X \in X$$

—a contradiction.

General products and the axiom of choice

Occasionally it is important to have a general form of product in which instead of pairs with first and second coordinates, or finite tuples, we have tuples where the coordinates correspond to indices in a general set I. Let X_i be a set for each element i in a set I. By definition the *general product*

$$\prod_{i \in I} X_i = \{f : I \to \bigcup_{i \in I} X_i \mid \forall i \in I.\ f(i) \in X_i\} \ .$$

Given $f \in \prod_{i \in I} X_i$ we can get the ith coordinate as $f(i)$. Given $x_i \in X_i$ for each $i \in I$, we can form their tuple $f \in \prod_{i \in I} X_i$ by defining $f(i) = x_i$ for all $i \in I$. It's not too hard to construct the set $\prod_{i \in I} X_i$ out of the earlier constructions.

Is the set $\prod_{i \in I} X_i$ nonempty? It has to be empty if $X_i = \emptyset$ for any $i \in I$. But if X_i is nonempty for every $i \in I$ it would seem reasonable that one could make a function $f \in \prod_{i \in I} X_i$ by *choosing* some $f(i) \in X_i$ for all $i \in I$. When I is a finite set it is easy to make such a tuple, and so prove that $\prod_{i \in I} X_i$ is nonempty if each X_i is. For a while it was thought that this perfectly reasonable property was derivable from more basic axioms even when I is infinite. However, this turns out not to be so. Occasionally one must have recourse to the *axiom of choice* which says provided each X_i is nonempty for $i \in I$, then so is the product $\prod_{i \in I} X_i$.

4.3 Some consequences

4.3.1 Sets of functions

The set of all relations between sets X and Y is the set $\mathcal{P}(X \times Y)$. Using comprehension it is then easy to see that

$$(X \rightharpoonup Y) = \{f \in \mathcal{P}(X \times Y) \mid f \text{ is a partial function}\}$$

is a set; that of all partial functions from X to Y. Similarly,

$$(X \to Y) = \{f \in \mathcal{P}(X \times Y) \mid f \text{ is a function}\}$$

is also a set; that of all total functions from X to Y. Often this set is written Y^X.

Exercise 4.4 Let $A_2 = \{1, 2\}$ and $A_3 = \{a, b, c\}$. List the elements of the four sets $(A_i \to A_j)$ for $i, j \in \{2, 3\}$. Annotate those elements which are injections, surjections and bijections. □

Exercise 4.5 Let X and Y be sets. Show there is a bijection between the set of functions $(X \to \mathcal{P}(Y))$ and the set of relations $\mathcal{P}(X \times Y)$. □

When investigating the behaviour of a function $f \in (X \to Y)$ we apply it to arguments. Earlier in Proposition 3.3 we saw that equality of functions $f, f' \in (X \to Y)$ amounts to their giving the same result on an arbitrary argument x in X. We can treat functions as sets and so might introduce a function by describing the property satisfied by its input-output pairs. But this would ignore the fact that a function is most often introduced as an expression e describing its output in Y in terms of its input x in X. For this manner of description lambda notation (or λ-notation) is most suitable.

Lambda notation

Lambda notation provides a way to describe functions without having to name them. Suppose $f : X \to Y$ is a function which for any element x in X gives a value $f(x)$ described by an expression e, probably involving x. Sometimes we write

$$\lambda x \in X.\ e$$

for the function f. Thus

$$\lambda x \in X.\ e \ = \ \{(x, e) \mid x \in X\}\ .$$

So, $\lambda x \in X.\ e$ is an abbreviation for the set of input-output pairs determined by the expression e. For example, $\lambda x \in \mathbb{N}_0.\ x + 1$ is the successor function and we have $(\lambda x \in \mathbb{N}_0.\ x + 1) \in (\mathbb{N}_0 \to \mathbb{N}_0)$.

Exercise 4.6 Use lambda notation to describe bijections

$$[(A \times B) \to C] \cong [A \to (B \to C)]\ ,$$
$$[A \to (B \to C)] \cong [B \to (A \to C)]\ .$$

□

Exercise 4.7 Describe explicit bijections

$$[(A \uplus B) \to C] \cong (A \to C) \times (B \to C),$$
$$[A \to (B \times C)] \cong (A \to B) \times (A \to C)\ .$$

□

Characteristic functions

In set theory one starts with sets as primitive and builds functions. We built sets of functions $(X \to Y)$ with the help of powersets. There are alternative foundations of mathematics which work the other way round. They start with functions and "types" of functions $(X \to Y)$ and identify sets with special functions called *characteristic* functions to truth values. The correspondence between sets and characteristic functions is explored in the following exercise.[4]

Exercise 4.8 Let X be a set. The set $\{\mathtt{T}, \mathtt{F}\}$ consists of the truth values $\mathtt{T}$ and $\mathtt{F}$. Let $Y \subseteq X$. Define its *characteristic function* $\chi_Y : X \to \{\mathtt{T}, \mathtt{F}\}$ by taking

$$\chi_Y(x) = \begin{cases} \mathtt{T} & \text{if } x \in Y \\ \mathtt{F} & \text{if } x \notin Y \end{cases}$$

for all $x \in X$. Show the function taking Y to χ_Y is a bijection from $\mathcal{P}(X)$ to $(X \to \{\mathtt{T}, \mathtt{F}\})$. □

4.3.2 Sets of unlimited size

Cantor used a *diagonal argument* to show that X and $\mathcal{P}(X)$ are never in 1-1 correspondence for any set X. This fact is intuitively clear for finite sets but also holds for infinite sets. It implies that there is no limit to the size of sets.

Cantor's argument is an example of proof by contradiction. Suppose a set X is in 1-1 correspondence with its powerset $\mathcal{P}(X)$. Let $\theta : X \to \mathcal{P}(X)$ be the 1-1 correspondence. Form the set

$$Y = \{x \in X \mid x \notin \theta(x)\}$$

which is clearly a subset of X and therefore in correspondence with an element $y \in X$. That is $\theta(y) = Y$. Either $y \in Y$ or $y \notin Y$. But both possibilities are absurd. For, if $y \in Y$ then $y \in \theta(y)$ so $y \notin Y$, while, if $y \notin Y$ then $y \notin \theta(y)$ so $y \in Y$. We conclude that our first supposition must be false, so there is no set in 1-1 correspondence with its powerset.

[4] The seminal work on founding mathematics on functions is Alonzo Church's higher order logic. You can learn more on higher order logic and its automation in later courses of Mike Gordon and Larry Paulson.

Cantor's argument is reminiscent of Russell's paradox. But whereas the contradiction in Russell's paradox arises out of a fundamental, mistaken assumption about how to construct sets, the contradiction in Cantor's argument comes from denying the fact one wishes to prove.

As a reminder of why it is called a diagonal argument, imagine we draw a table to represent the 1-1 correspondence θ along the following lines. In the xth row and yth column is placed `T` if $y \in \theta(x)$ and `F` otherwise. The set Y which plays a key role in Cantor's argument is defined by running down the diagonal of the table interchanging `T`'s and `F`'s in the sense that x is put in the set iff the xth entry along the diagonal is `F`.

	$\cdots$	x	$\cdots$	y	$\cdots$
$\vdots$		$\vdots$		$\vdots$	
$\theta(x)$	$\cdots$	`T`	$\cdots$	`F`	$\cdots$
$\vdots$		$\vdots$		$\vdots$	
$\theta(y)$	$\cdots$	`F`	$\cdots$	`F`	$\cdots$
$\vdots$		$\vdots$		$\vdots$	

Exercise 4.9 This exercise guides you through to a proof that for any sets X and Y, with Y containing at least two elements, there cannot be an injection from the set of functions $(X \to Y)$ to X.

(i) Let X be a set. Prove there is no injection $f : \mathcal{P}(X) \to X$. [Hint: Consider the set $W =_{\text{def}} \{f(Z) \mid Z \subseteq X \ \& \ f(Z) \notin Z\}$.]

(ii) Suppose now that a set Y has at least two distinct elements. Define an injection $k : \mathcal{P}(X) \to (X \to Y)$, from the powerset of X to the set of functions from X to Y.

(iii) Prove that there is no injection from $(X \to Y)$ to X when the set Y has at least two distinct elements. [Hint: Recall that the composition of injections is an injection.]

□