5 Sets

5.1 The intuitive idea of a set

A **set** is a collection of things. The "things" can be concrete (such as "people") or abstract (such as "colours"). No reason for collecting the things into a set needs to be given but often there is a shared attribute. Examples:

- the set of undergraduate students at the University of Birmingham in the academic year 2021/22 (there are about 25,000);
- the seasons (there are four);
- the set of numbers that can be stored exactly as a `float` in Java (there are slightly fewer than 2^{32} many).

The "things" in the set are called the **elements** or the **members**. The following are important aspects of the idea of a set:

1. The elements of a set are identifiable and can be distinguished from each other.

   (So the "waves on the Atlantic" do not form a set.)

2. There is a clear criterion that defines when a "thing" is a member of the set and when it is not.

   (So we cannot form the set of "tall people in Britain.")

3. A member of a set is counted only once.

   (So the "set of staff members who serve on Staff-Student Committee or Teaching Committee" contains Mark Lee only once, although he satisfies both criteria.)

4. As far as the set is concerned, its elements are not ordered in any way.

   (So although it may appear natural to list the elements of the set of "books by Suzanne Collins" in the order

   - The Hunger Games
   - Catching Fire
   - Mockingjay

   we are not required to do so; listing them in any other order would define the same set.)

5. The set is defined by which members it has and nothing else. If two collections have the same elements, then they form the same set, independently of how the collections were defined.

   (So the "set of triangles which have three equal sides" is the same as the "set of triangles which have three equal angles.")

5.2 Notation

To indicate that x is an element of the set A, one writes

$$x \in A$$

In mathematics the symbol "$\in$" is used for this purpose and this purpose alone (it originates from the Greek letter *epsilon*, written "ε"). Principle 5 from the previous item can now be written in the form

31

Sets with few elements can often be written out explicitly, for example, the set of seasons:

$$\{\text{Spring}, \text{Summer}, \text{Autumn}, \text{Winter}\}$$

Infinite sets can be indicated by suggesting a formation principle:

$$\{1, 4, 9, 16, 25, \dots\}$$

Sets in Computer Science. Computer scientists use the language of sets for formal and precise descriptions, just like mathematicians, scientists, and engineers. They even developed it into a full-blown *specification language*, called Z.

Set theory has also influenced the design of systems and programming languages. A particularly striking example are "relational databases," suggested in 1970 by the British computer scientist Edgar F. "Ted" Codd. The principles of set theory, listed in Item 5.1, all have immediate relevance here, in particular 1, 3, and 4:

- 1: this is related to the issue of "keys" in a database table;
- 3: storing information about the same object more than once is wasteful and in many cases causes real errors;
- 4: giving the database management system the freedom to arrange the entries of a table in any order allows many optimizations in storage, processing, and retrieval.

Sets are also related to the programming language idea of a "type." Indeed, finite sets can be defined directly in C as "enumerated types," for example

```
enum Seasons {Spring, Summer, Autumn, Winter};
```

Java has supported enumerated types since version 1.5; the syntax is the same as in C. Underneath, however, `enum`'s are actually special `class` definitions and the Java compiler automatically creates several methods for it.

You should also know that Java supports the collection type `Set` (as an `interface`) and various concrete implementations, for example, the very useful `HashSet`.

5.3 Sets of numbers

Let us consider the various systems of numbers which we studied in the first two weeks. The **set of natural numbers**, as you may remember, is denoted by $\mathbb{N}$. You can imagine the elements of $\mathbb{N}$ to be given as strings of the form $s(s(\ldots s(0)\ldots))$, or as numbers in some place-value system such as the decimal one, or you can think of them as *abstract* concepts. In any case, for mathematicians there is only one set $\mathbb{N}$.

To describe the set of **integers**, we add to $\mathbb{N}$ the set of negative whole numbers. For this we use the operator $\cup$ which joins two sets into one. It is pronounced "**union**". Using it, we can write the set of integers as

32

where the elements of the new component can be thought of as written out in the usual decimal system with a minus sign in front.

Having both natural numbers and integers, we can define the set of **fractions** as **pairs** of an integer (the numerator) and a non-zero natural number (the denominator). In general, if A and B are sets (possibly the same one) we write $A \times B$ for the set whose elements are pairs (x,y) where x is an element of A and y is an element of B. If $A = B$ then we may abbreviate $A \times A$ to A^2. We call $A \times B$ the **product** of A and B. So the set of fractions could be written as the product $\mathbb{Z} \times P$ where P is the set of strictly positive natural numbers.

How do we define P? For this we use the **set difference** operator $\setminus$. In general, $A \setminus B$ is the set of all elements of A which do **not** belong to B. So:

33

All this is quite straightforward but how would we construct $\mathbb{Q}$, the set of rational numbers? Remember from Chapter 3 that the rationals are to be viewed as fractions *with a different equality*. At the moment our set theory does not provide a facility for imposing an alternate equality. We will deal with this systematically in Chapter 7. As a stop gap, we could restrict the set of fractions to those where numerator and denominator do not have a common factor greater than 1. In general, we say that A is a **subset** of a set B if all elements of A are also elements of B. In symbols, this is written as $A \subseteq B$. So our definition of $\mathbb{Q}$ is as a subset of the set of fractions. We'll have a lot more to say about subsets in the next chapter.

A similar issue arises with the finite rings $\mathbb{Z}_m$. On the one hand, each of these can be directly realised as the finite set $\{0,1,\ldots,m-1\}$, i.e., the possible remainders of dividing a natural number by m; on the other hand, it is much better to view $\mathbb{Z}_m$ as the set of integers but with the alternate equality "congruent modulo m".

The bit patterns that can be stored in a 32 bit register can be seen as 32-tuples, i.e., elements of the set B^{32} where $B = \{0,1\}$. This is a simple generalisation of the idea of a product. More interesting is the data type `String`. For this we introduce a new set formation operator: If A is a set then A^* is the set whose elements are finite sequences of elements of A. In computer science, the set A is often finite and called the "alphabet"; the elements of A^* are then the "words over A". The shortest element of A^* is the "empty word," or the "empty sequence," often denoted by ε (the real Greek letter epsilon). The alphabet for Java's strings is called **Unicode**; it has $2^{16} = 65,536$ characters.

In order to describe $\mathbb{R}$ as a set we also allow the construction of *infinite streams*. More precisely, if A is a set then A^ω is the set of infinite sequences of elements of A. With this device we can define a candidate for the set $\mathbb{R}$ of **real numbers**. Every element of $\mathbb{R}$ consists of an integer part and an infinite sequence of digits (the fractional part), so

34: The reals — not quite right

where $D = \{0,1,\ldots,9\}$. However, in doing so we are ignoring the fact that a decimal fraction that ends in an infinite sequence of nines is considered equal to one which ends in an infinite sequence of ones. For example, $0.999\ldots = 1.000\ldots$ We can remedy this by excluding from $\mathbb{Z} \times D^\omega$ those pairs where the second component ends in an infinite sequence of nines. However, as with the rational numbers and $\mathbb{Z}_m$, this does not reflect our day-to-day working with reals and we would much rather be able to postulate the equality between these special elements.

5.4 Counting the elements of a set

We all know how to do this, of course: If A is the set of seasons from Section 5.1 then we can say that it has four elements. This is written in set theory as

$$|A| = 4 \quad \text{or} \quad \mathrm{card}(A) = 4$$

where the function "card" is short for **cardinality**. We could also call it the "size" of a set.

Given two sets A and B, we don't need to count the elements in order to check that they have the same number of elements; instead we can line up the element of A and B in such a way that each element of A is paired with exactly one element of B (and vice versa):

35

which would persuade us that $\mathrm{card}(\mathrm{Seasons}) = \mathrm{card}(\mathrm{Compass})$, even if we had never heard of the number 4.

5.5 Cardinality of infinite sets

At first sight, the sets $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}$ all have infinitely many elements and we could say that their cardinality, therefore, is "infinity." It was Georg Cantor's great discovery that in fact not all infinite sets have the same cardinality. The argument is easy and does not use much from the theory of sets.

The first infinite set that comes to mind is that of the natural numbers, $\mathbb{N}$. Let us call a set **countable**, or more precisely, **countably infinite**, if it has the same cardinality as $\mathbb{N}$. There are more countable sets than you might think. We go through a sequence of examples.

- $A = \mathbb{N} \cup \{-1\}$ is countable:

36

$\mathbb{N}$

- $\mathbb{Z}$ is countable

37

$$\mathbb{N} = \{ \quad 0, \qquad 1, \qquad 2, \qquad 3, \qquad 4, \quad \ldots\}$$

- $\mathbb{N}^2$ is countable

38

$$\mathbb{N} = \{ \quad 0, \qquad 1, \qquad 2, \qquad 3, \qquad 4, \qquad 5, \quad \ldots\}$$

This enumeration trick is more clearly illustrated with a two-dimensional picture:

39

- If A is a finite set (an "alphabet") then A^* is countable. A good picture to illustrate this is the following (for $A = \{a,b\}$):

40

which is nothing other than a "breadth-first traversal" of the binary tree of all the words over $\{a,b\}$. You can see that it would work for any alphabet size (as long as the alphabet is finite).

- Java programs: These form a subset of A^* where A is the 65536 letter alphabet of Unicode. So it follows from the last example that:

  Theorem 2 *The set of valid* Java *programs is countable.*

  This is an important fact. Make sure you remember it!

5.6 Uncountability

We can think of the countable sets as being "infinite but still relatively small." A set that is not "small" in this sense is $\mathbb{R}$, the set of real numbers. The proof is by **contradiction**.[2] *Assume* we could list in the fashion of the previous item all the real numbers, that is, assume we had a listing that looks like the following:

	0	1	2	3	4	5
r_0	3.	1	4	1	5	9...
r_1	2.	7	1	8	2	8...
r_2	6.	0	2	2	1	4...
r_3	22.	7	1	0	9	8...
r_4	9.	8	0	6	6	5...
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Now consider the real number

$$a = a_0 \,.\, a_1\, a_2\, a_3\, \ldots$$

where

$$a_i = \begin{cases} 0 & \text{if } r_i[i] \neq 0 \\ 1 & \text{if } r_i[i] = 0 \end{cases}$$

where $r_i[i]$ means the digit in row i and column i in the assumed listing of all real numbers. For example, the number

$$r_0[0] \,.\, r_1[1]\, r_2[2]\, r_3[3]\, r_4[4]$$

derived from the example listing above starts with

$$3.7206\ldots$$

Consequently, the number a associated with this listing starts with

$$0.0010\ldots$$

As you can see, the number a is different from r_0 because it was constructed so that it differs in its integral part (the part before the decimal point); it differs from r_1 because it was made to differ from it in the first digit after the point; and so on.

[2] The proof idea is now truly famous. It was invented by Cantor in 1891, used in 1901 by Bertrand Russell in his "paradox," by Kurt Gödel in 1931 in his "Incompleteness Theorem," and by Alan Turing in 1937 in his proof of the "Undecidability of the *Entscheidungsproblem*."

In general, a differs from r_i in the i-th digit, so is certainly different from r_i. (Of course, it'll probably differ in many other places as well.)

This means that a is a real number that does not appear in the listing; in other words, our assumption that the listing is complete cannot be true. *Contradiction!* Since our construction of a was completely general and could be carried out for any listing of real numbers, we have to conclude that any such listing *by necessity* will miss out some real numbers. We therefore say:

Theorem 3 *The set of real numbers is* ***uncountable****.*

What does this mean for Computer Science? Since there are (only) countably many Java programs but uncountably many real numbers, there must exist real numbers that cannot be computed by a program. We can say:

Theorem 4 *There are* ***non-computable*** *real numbers.*

Comparing uncountable sets. The examples above illustrate how we can establish that two countably infinite sets have the same cardinality. To show that two *uncountable* sets have the same size requires a bit more ingenuity but the principle is the same: We establish a one-to-one relationship between the elements. For example, we can show that the elements of a circle (minus one point) can be lined up against the elements of the real line. The proof is geometric:

41

We see that every line through the "north pole" N links a point P on the circle and a point P' on the real line. This works for all points on the circle except the north pole itself.

Exercises

1. Let A and B be two countable sets. Argue that $A \cup B$ is also countable.
2. Find an arithmetic or a geometric argument that shows that the following two subsets of $\mathbb{R}$ have the same cardinality: $A = \{x \in \mathbb{R} \mid -1 < x < 1\}$ and $B = \{y \in \mathbb{R} \mid -2 < y < 2\}$.

Practical advice

In the exam I expect you to

- know the set-theoretic symbols $\in$, $\subseteq$, $\cup$, and $\setminus$ (more are to come);
- know the traditional symbol for sets of numbers $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}$;
- know the symbol for the set of strings A^* over an alphabet A, and likewise A^ω for the set of infinite streams;
- know what "countable" means for a set;
- be able to argue that $\mathbb{Z}$, $\mathbb{N}^2$, and A^* are countable;
- to be able to recall that $\mathbb{R}$ is uncountable, and give some hints as to the proof of this fact;
- demonstrate an understanding of Theorem 4 and its consequences.