Introducing sets

1 What is a set?

A *set* is a collection of things. Here is a set of cities:

$$\{\text{Ankara}, \text{Birmingham}, \text{Chicago}, \text{Dubai}\}$$

Here is a set of numbers:

$$\{4, 7, 9, 571, 572, 573\}$$

Here is a set of people:

$$\{\text{Paul}, \text{Rajesh}, \text{Usman}, \text{Vincent}\}$$

The things that belong to a set are called its *elements* (or members). We sometimes describe a set by enumerating its elements, as in the above examples.

The notation $a \in B$ means that a is an element of B, and $a \notin B$ means that a is not an element of B. For example:

$$\begin{array}{rcl} 573 & \in & \{4, 7, 9, 571, 572, 573\} \\ 574 & \notin & \{4, 7, 9, 571, 572, 573\} \end{array}$$

2 Equality of sets

Two sets A and B are equal when they have the same elements. In symbols:

$$A = B \quad \text{if and only if} \quad \forall x.\, x \in A \iff x \in B$$

For example:

$$\{4, 7, 9, 9\} \quad = \quad \{9, 4, 7, 4, 7, 4, 7\}$$

because

- any element of $\{4, 7, 9, 9\}$ is either 4 or 7 or 9 or 9, and each of these belongs to $\{9, 4, 7, 4, 7, 4, 7\}$
- any element of $\{9, 4, 7, 4, 7, 4, 7\}$ is either 9 or 4 or 7 or 4 or 7 or 4 or 7, and each of these belongs to $\{4, 7, 9, 9\}$.

As this example illustrates, *order and repetition don't matter* when we present a set. By contrast with lists.

3 Cardinality

Each set is either *finite* or *infinite*. The examples above are finite, but $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ are infinite sets. A finite set can be enumerated—at least in principle. For any infinite set A, there's an infinite sequence of elements $a_0, a_1, a_2, \ldots$ with no repetitions. An example, in the case of the set $\mathbb{R}$, is the sequence $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$. (Of course, this doesn't contain every real number.)

Are the following sets finite or infinite?

$$\begin{array}{l} [4 \mathinner{..} 8) \\ [4 \mathinner{..} 10^{10^{100}}) \\ [4 \mathinner{..} \infty) \\ [4 \mathinner{..} 4) \\ [4, 8) \\ [4, 10^{10^{100}}) \\ [4, \infty) \\ [4, 4) \end{array}$$

Each finite set B has a *cardinality* (or size), written $|B|$, which is a natural number. For example:

$$|\{9,4,7,4,7,4,7\}| \quad = \quad ?$$

There is just one set whose cardinality is 0. It is called the *empty set*, written $\{\,\}$ or $\emptyset$.

A set whose cardinality is 1 is called a *singleton*. For example, $\{5\}$ and $\{\text{Birmingham}\}$ and $\{\text{Paul}\}$ are all singletons. Please do not confuse $\{a\}$ with a. For example, $\{\emptyset\}$ has one element but $\emptyset$ has none.

4 Subsets

A *subset* of a set B is a set A with the following property: every element of A is an element of B. The notation $A \subseteq B$ means that A is a subset of B. For example:

$$\begin{array}{rcl} \{7,9,9\} & \subseteq & \{9,4,7,4,7,4,7\} \\ \{4,7,9,9\} & \subseteq & \{9,4,7,4,7,4,7\} \\ \emptyset & \subseteq & \{9,4,7,4,7,4,7\} \\ \{4,7,9,9,83\} & \not\subseteq & \{9,4,7,4,7,4,7\} \end{array}$$

How many subsets does

$$\{4,7,9\}$$

have? More generally, for a finite set A of cardinality n, how many subsets does it have?

We can now formulate equality of sets more concisely.

$$A = B \text{ if and only if } A \subseteq B \text{ and } B \subseteq A.$$

5 Union, intersection and difference

Given two sets A and B, we obtain some more sets.

1. $A \cup B$ is the *union*. It consists of all things that are in either A or B.
2. $A \cap B$ is the *intersection*. It consists of all things that are in both A and B.
3. $A \setminus B$ is the *set-difference*. It consists of all things that are in A but not B.

Examples:

$$\begin{array}{rcl} \{3,5,7,10\} \cup \{5,7,8,9\} & = & ? \\ \{3,5,7,10\} \cap \{5,7,8,9\} & = & ? \\ \{3,5,7,10\} \setminus \{5,7,8,9\} & = & ? \end{array}$$

6 Set-builder notation

Set-builder notation is a convenient way of constructing subsets, by taking those elements that have a particular property. For example:

$$\{n \in \mathbb{N} \mid n^3 \geqslant 100\}$$

This is the set of all natural numbers such that $n^3 \geqslant 100$. It is the integer interval $[5\,..\,\infty)$.

A fancier version of this notation uses an operation. For example:

$$\{n^2 \mid n \in \mathbb{N},\ n^3 \geqslant 100\}$$

This is the set of squares of all natural numbers such that $n^3 \geqslant 100$. It is $\{25, 36, 49, \ldots\}$.

7 Tuples

We often talk about *ordered pairs* such as $(\text{Paul}, 7)$. This might be an record in a database, meaning that my score is 7. The basic property of ordered pairs is that $(a, b) = (x, y)$ iff $a = x$ and $b = y$. (The word "iff" is short for "if and only if".)

Likewise, we can speak of *ordered triples* such as $(\text{Paul}, \text{Birmingham}, 7)$. The basic property of ordered triples is that $(a, b, c) = (x, y, z)$ iff $a = x$ and $b = y$ and $c = z$.

Likewise we can speak of there are ordered quadruples, ordered quintuples, etc. The word "ordered" is usually dropped. Generally, for any natural number n, an n*-tuple* has n components. There is just one 0-tuple, written $()$.

Given two sets A and B, their *product* is defined to be

$$A \times B \stackrel{\text{def}}{=} \{(x, y) \mid x \in A,\ y \in B\}$$

(The symbol $\stackrel{\text{def}}{=}$ means "is defined to be".) For example.

$$\begin{aligned} \{\text{Paul}, \text{Rajesh}, \text{Usman}\} \times \{4, 7\} \quad = \quad & \{(\text{Paul}, 4), (\text{Paul}, 7) \\ & (\text{Rajesh}, 4), \text{Rajesh}, 7), \\ & (\text{Usman}, 4), (\text{Usman}, 7)\} \end{aligned}$$

What about multiplying three sets A, B, C? We define

$$A \times B \times C \stackrel{\text{def}}{=} \{(x, y, z) \mid x \in A,\ y \in B,\ z \in C\}$$

In the same way, any list of sets has a product. We define 1 to be the singleton set $\{()\}$, and this is the product of the empty list.

In programming, a type that corresponds to a product of sets is often called a "record type".

8 Applications

Sets are used throughout computer science. Here are just a few applications.

- *Types* appear in many programming languages, such as Java. They can be "modelled" using sets.
- *Databases.* I'll say more in the coming weeks.
- *Algorithmic problems.* For example: "write an efficient program that takes two lists of ID numbers, each with repetitions, and determines the number of IDs that appear in the first list but not the second." Sophisticated data structures, such as HashSet, are used to solve problems of this kind.
- *Automata.* An automaton is a set of states, with additional structure. This notion will feature in the "Theories of Computation" module.

9 Subtleties

Set theory has given rise to deep and subtle questions that you might like to ponder (though they're not part of this module). For example

- We've seen that finite sets have a cardinality. What about infinite sets?
- We've seen that multiplication of natural numbers is associative. What about multiplication of sets?
- Is there a set that is *universal*, i.e., contains absolutely everything?