Three kinds of set: Finite, Countably infinite and Uncountable

1 Finite and infinite sets

Every set is either finite or infinite. Every finite set has a *cardinality* (size), which is a natural number. Examples of finite sets:

- The empty set. This has cardinality 0.
- The set of all modules in the CS first year. This has cardinality 6.
- The set $\{2, 3, 5, 7, 11, 13, 17.19\}$ of all prime numbers less than 20. This has cardinality 8.
- Any subset of a finite set.

Examples of infinite sets:

- The set $\mathbb{N}$ of all natural numbers.
- The set $\mathbb{Z}$ of all integers.
- The set $\mathbb{Q}$ of all rational numbers.
- The set $\mathbb{R}$ of all real numbers.
- The set of all prime numbers.
- The set $\{0, 1, 2\}^*$ of all words in the alphabet $\{0, 1, 2\}$.
- The set $\mathcal{P}\mathbb{N}$ of all sets of natural numbers.
- The set $\mathcal{P}(\{0, 1, 2\}^*)$ of all languages in the alphabet $\{0, 1, 2\}$.

2 Countably infinite sets

A set A is *countably infinite* when there's a bijection (one-to-one correspondence) between $\mathbb{N}$ and A. Here are some examples.

- The set of all prime numbers is countably infinite:

$$\begin{array}{ccccccc} 0 & 1 & 2 & 3 & 4 & 5 & \cdots \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \\ 2 & 3 & 5 & 7 & 11 & 13 & \cdots \end{array}$$

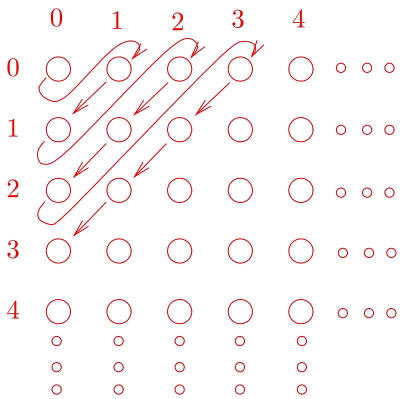
- More generally, *any* infinite subset of $\mathbb{N}$ is countably infinite.
- The set $\mathbb{Z}$ of all integers is countably infinite:

$$\begin{array}{ccccccc} 0 & 1 & 2 & 3 & 4 & 5 & \cdots \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \\ 0 & 1 & -1 & 2 & -2 & 3 & \cdots \end{array}$$

- The set $\mathbb{N} \times \mathbb{N}$ of all ordered pairs of natural numbers is countably infinite:

$$\begin{array}{ccccccc} 0 & 1 & 2 & 3 & 4 & 5 & \cdots \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \\ (0,0) & (1,0) & (0,1) & (2,0) & (1,1) & (0,2) & \cdots \end{array}$$

  This can be explained with a picture:



- The set $\{0,1\}^*$ is countably infinite, because these words can be enumerated as follows:

$$\begin{array}{l} \varepsilon, \\ 0, 1, \\ 00, 01, 10, 11, \\ 000, 001, 010, 011, 100, 101, 110, 111, \\ 0000, 0001, \ldots, 1111, \\ 00000, 00001, \ldots, 11111, \\ \ldots \end{array}$$

- Likewise A^* is countably infinite, for any finite set A with at least one element.
- Any infinite subset of a countably infinite set is countably infinite.
- The set of all Java programs is an infinite subset of the set of all ASCII strings. So it is countably infinite.
- The set of all regular expressions on the alphabet $\{a, b\}$ (or any finite alphabet) is an infinite subset of the set of all ASCII strings. So it is countably infinite.

3 Countable sets

A *countable* set is one that is either finite or countably infinite. Every subset of a countable set is countable.

4 The set of all bitstreams is uncountable

A *bitstream* is an infinite sequence of 0's and 1's. (Imagine a program that prints 0's and 1's forever.) Here are some examples:

$$\begin{array}{l} 010101010101\ldots \\ 10100100010000100001\ldots \\ 1111111111\ldots \end{array}$$

The set of all bitstreams is written $\{0,1\}^\omega$. Clearly this is infinite. Let's show that it's uncountable. The proof method is called "diagonalization".

If it's countable, there's an enumeration:

$$\begin{array}{rcl} x_0 & = & y_{0,0}\ y_{0,1}\ y_{0,2}\ y_{0,3}\cdots \\ x_1 & = & y_{1,0}\ y_{1,1}\ y_{1,2}\ y_{1,3}\cdots \\ x_2 & = & y_{2,0}\ y_{2,1}\ y_{2,2}\ y_{2,3}\cdots \\ x_3 & = & y_{3,0}\ y_{3,1}\ y_{3,2}\ y_{3,3}\cdots \\ \vdots & & \end{array}$$

Each $y_{i,j}$ is either 0 or 1. Our assumption is that every bitstream appears in this list.

Let's write $\overline{0} \stackrel{\text{def}}{=} 1$ and $\overline{1} \stackrel{\text{def}}{=} 0$. Consider the bitstream

$$\overline{y_{0,0}}\ \overline{y_{1,1}}\ \overline{y_{2,2}}\ \overline{y_{3,3}}\cdots$$

This bitstream is not equal to x_0, nor to x_1, nor to x_2, nor to x_n for any n. This contradicts our assumption that every bitstream appears in the list. Thus the set $\{0,1\}^*$ is uncountable.

Application We can consider Java programs that print 0's and 1's forever. The set of all such programs is countably infinite. But the set of all bitstreams is uncountable. So there are some bitstreams that are not Java-computable.

5 More examples of uncountable sets

From this, we can obtain other examples of uncountable sets:

- The set $\mathbb{R}$ of all real numbers is uncountable. That's because a real number in the interval (0.1) can be represented in binary, as a bitstream. (Since some real numbers have two distinct binary representations, the argument is slightly more complicated.)
- A set of natural numbers corresponds to a bitstream. For example, the set of all even numbers corresponds to $101010101\ldots$, and the set of all prime numbers to $00110101000101\ldots$ Therefore the set $\mathcal{P}\mathbb{N}$ of all sets of natural numbers is uncountable.
- In the same way, for any infinite set B, the set $\mathcal{P}B$ is uncountable.
- In particular, the set of languages $\mathcal{P}(\{0,1\}^*)$ is uncountable.

Application For the alphabet $\{0,1\}$, there are uncountably many languages but only countably many regular expressions. Therefore some languages are irregular.

6 Test your understanding

Are the following sets finite, countably infinite or uncountable?

- The set of all ordered pairs of primes.
- The set of all sets of ordered pairs of primes.
- The set of all ordered pairs of primes less than a million.
- The set of all functions from $\{5,6,7\}$ to $\{8,9,10,11\}$.
- The set of all functions from $\mathbb{N}$ to $\{5,6,7\}$.
- The set of all functions from $\{5,6,7\}$ to $\mathbb{N}$.
- The set of all English sentences.
- The set of all English sentences that appear on the Web.