A Short Introduction to Set Theory

Moti Ben-Ari

http://www.weizmann.ac.il/sci-tea/benari/

Contents

1 Finite and infinite sets **2**

2 Subset **3**

3 Proving inclusion and equality of sets **4**

4 Union, intersection and difference **4**

5 Sequences **6**

6 Cartesian product **7**

7 Relations **8**

8 Functions **10**

9 Cardinality **12**

10 Powersets **15**

11 Induction **15**

12 The Well-ordering Principle **18**

13 References **19**

This document is based upon Appendix A of [1] and used by permission of Springer.

The presentation of mathematical logic in [1] is based upon an informal use of set theory whose definitions and theorems are summarized here. For an elementary, but detailed, development of set theory, see [4].

I would like to express my thanks to Jørgen Villadsen for his helpful suggestions.

1 Finite and infinite sets

The concept of an *element* is undefined, but informally the concept is clear: an element is any identifiable object like a number, color or node of a graph. Sets are built from elements.

Definition 1 *A* set *is composed of* elements*. $a \in S$ denotes that a is an element of set S and $a \notin S$ denotes that a is* not *an element of S. $\emptyset$, the* empty set*, is the set with no elements. Capital letters like S, T and U are used for sets. The elements of sets are written within braces* { }.

There are two ways to define a set: (a) We can explicitly write the elements comprising the set. If a set is large and if it is clearly understood what its elements are, an ellipsis '…' is used to indicate the elements that are not listed. (b) A set may be defined by set comprehension, where the set is specified to be composed of all elements that satisfy a condition.

Example

- The set of colors of a traffic light is $\{red, yellow, green\}$.
- The set of atomic elements is $\{hydrogen, helium, lithium, \ldots\}$.
- $\mathbb{Z}$, the set of *integers*, is $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$.
- $\mathbb{N}$, the set of *natural numbers*, is $\{0, 1, 2, \ldots\}$. $\mathbb{N}$ can also be defined by *set comprehension*: $\mathbb{N} = \{n \in \mathbb{Z} \mid n \geq 0\}$. Read this as "$\mathbb{N}$ is the set of all integers n such that n is greater than or equal to zero."
- $\mathbb{R}$, the set of *real numbers*, can be informally defined as all numbers that can be expressed in decimal notation: 3.14159 or 0.31459×10^1.
- E, the set of even natural numbers, is $\{n \in \mathbb{N} \mid n \bmod 2 = 0\}$.
- P, the set of prime numbers, is:

$$\{n \in \mathbb{N} \mid n \geq 2 \text{ and for all } m\,(n \bmod m = 0 \text{ implies } (m = 1 \text{ or } m = n))\}.$$

There is no meaning to the order of the elements in a set nor to repetition of elements: $\{3, 2, 1, 1, 2, 3\} = \{1, 2, 3\} = \{3, 1, 2\}$. A set containing a single element (a *singleton set*) and the element itself are not the same: $5 \in \{5\}$.

2 Subset

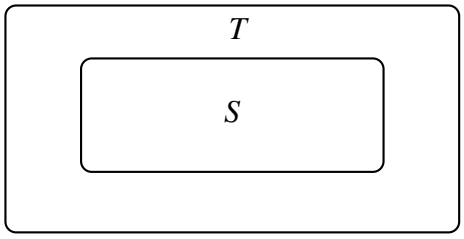
Definition 2 *Let S and T be sets. S is a* subset *of T, denoted $S \subseteq T$, iff every element of S is an element of T, that is, $x \in S$ implies $x \in T$. S is a* proper subset *of T, denoted $S \subset T$, iff $S \subseteq T$ and $S \neq T$.*

Example $\mathbb{N} \subset \mathbb{Z}$, $E \subset \mathbb{N}$, $\{red, green\} \subset \{red, yellow, green\}$, $\{red, green\} \subseteq \{green, red\}$.

Theorem 3 $\emptyset \subseteq T$.

Proof We have to show that $x \in T$ holds for all $x \in \emptyset$. But there are no elements in $\emptyset$, so the statement is vacuously true. ∎

The relationships among sets can be shown graphically by the use of *Venn diagrams*. These are closed curves drawn in the plane and labeled with the name of a set. A point is in the set iff it is within the interior of the curve. In the following diagram, S is a subset of T since every point within S is also within T.



Theorem 4 *The subset property is transitive:*

1. *If $S \subseteq T$ and $T \subseteq U$ then $S \subseteq U$.*
2. *If $S \subset T$ and $T \subseteq U$ then $S \subset U$.*
3. *If $S \subseteq T$ and $T \subset U$ then $S \subset U$.*
4. *If $S \subset T$ and $T \subset U$ then $S \subset U$.*

Proof Let us prove (2). By the assumption $S \subset T$: if $x \in S$ then $x \in T$. By the assumption $T \subseteq U$ if $x \in T$ then $x \in U$. Therefore, $S \subseteq U$. We have to show that $S \neq U$. By assumption, $S \subset T$ so $S \neq T$ and there is an element $x \in T$ and $x \notin S$. Since $T \subseteq U$, $x \in U$ and by assumption $x \notin S$, so $S \neq U$. ∎

3 Proving inclusion and equality of sets

To prove $S \subseteq T$ choose an *arbitrary* element $x \in S$ and show $x \in T$.

Example Prove that every prime number greater than 2 is odd. Formally:

$$\begin{aligned} S &= \{n \in \mathbb{N} \mid n > 2 \text{ and } n \in P\} \\ T &= \{n \in \mathbb{N} \mid n = 2k+1 \text{ for some } k \in \mathbb{N}\} \\ S &\stackrel{?}{\subseteq} T. \end{aligned}$$

Let n be an *arbitrary* element of S. Suppose that $n \notin T$. Then $n = 2k$ for some $k \in \mathbb{N}$. If $k = 0$ or $k = 1$ so that $n = 1$ or $n = 2$, then $n \not> 2$ and $n \notin S$. Otherwise, $n > 2$ has (at least) the factors $2, k$ that are different from 1 and n, so $n \notin S$. Since n was an arbitrary element of S, $S \subseteq T$.

To prove that two sets are equal, use the following theorem whose proof is left to the reader:

Theorem 5 *$S = T$ iff $S \subseteq T$ and $T \subseteq S$.*

Example Prove that E, the set of even natural numbers, is equal to the set of natural numbers of the form $2k$ for $k \in \mathbb{N}$. Formally:

$$\begin{aligned} S &= \{n \in \mathbb{N} \mid n \bmod 2 = 0\} \\ T &= \{n \in \mathbb{N} \mid \text{ for some } k \in \mathbb{N}, n = 2k\} \\ S &\stackrel{?}{=} T. \end{aligned}$$

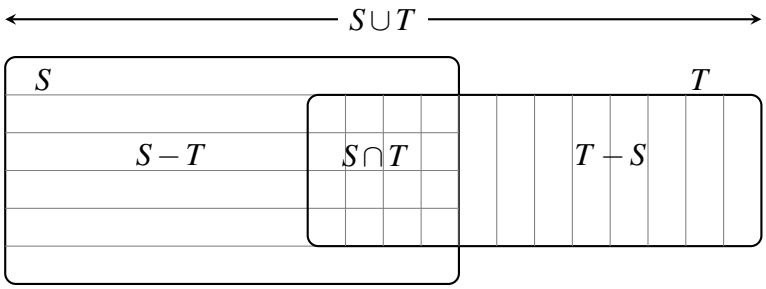
Let $m \in S$. Then $m \bmod 2 = 0$, which by definition means there is a $k \in \mathbb{N}$ such that $m = 2k+0 = 2k$. Therefore, $m \in T$ and $S \subseteq T$. If $m \in T$, then there is a k such that $m = 2k = 2k+0$ so $m \bmod 2 = 0$. Therefore, $T \subseteq S$. By Theorem 5 $S = T$.

4 Union, intersection and difference

Definition 6

1. $S \cup T = \{x \mid x \in S \text{ or } x \in T\}$, *the* union *of S and T, is the set consisting of those elements which are elements of* either S or T *or both.*
2. $S \cap T = \{x \mid x \in S \text{ and } x \in T\}$, *the* intersection *of S and T, is the set consisting of those elements which are elements of* both S and T. *If $S \cap T = \emptyset$ then S and T are* disjoint.
3. $S - T = \{x \mid x \in S \text{ and } x \notin T\}$, *the* difference *of S and T, is the set of elements of S that are not elements of T.*
4. *Let U be understood as a universal set; then $\overline{T}$, the* complement *of T, is $U - T$.*

The following Venn diagram illustrates these concepts.



Example Here are some examples of the operations on sets:

$$\begin{array}{lcl} \{red, yellow\} \cup \{red, green\} & = & \{red, yellow, green\}, \\ \{red, yellow\} \cap \{red, green\} & = & \{red\}, \\ \{red, yellow\} - \{red, green\} & = & \{yellow\}, \\ P \cap E & = & \{2\}. \end{array}$$

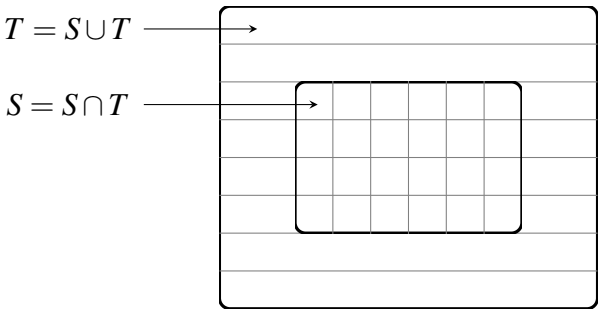
The following theorem states some properties of the set operators.

Theorem 7

1. $T = (T - S) \cup (S \cap T)$.
2. *If* $S \subseteq T$ *then:* $S \cap T = S$, $S \cup T = T$, $S - T = \emptyset$.
3. *If* S *and* T *are disjoint then* $S - T = S$.
4. $S \cup \emptyset = S$, $S \cap \emptyset = \emptyset$, $S - \emptyset = S$.

Proof

1. See the Venn diagram above.
2. See the following Venn diagram.



For example, $P \cap \mathbb{N} = P$, $P \cup \mathbb{N} = \mathbb{N}$.

3. $S - T = \{x \mid x \in S \text{ and } x \notin T\}$. By definition, S and T are disjoint means $S \cap T = \emptyset$, that is, $x \in S$ implies $x \notin T$, so every element in $S - T$ is in S.

4. Let us prove $S \cap \emptyset = \emptyset$. $S \cap \emptyset = \{x \mid x \in S \text{ and } x \in \emptyset\}$, but there are no x such that $x \in \emptyset$ so there are no elements in $S \cap \emptyset$.

∎

The operators $\cup$ and $\cap$ are commutative, associative and distributive.

Theorem 8

1. $S \cup T = T \cup S$.
2. $S \cap T = T \cap S$.
3. $(S \cup T) \cup U = S \cup (T \cup U)$.
4. $(S \cap T) \cap U = S \cap (T \cap U)$.
5. $S \cup (T \cap U) = (S \cup T) \cap (S \cup U)$.
6. $S \cap (T \cup U) = (S \cap T) \cup (S \cap U)$.

The reader is invited to give informal proofs of these properties by drawing Venn diagrams.

5 Sequences

Definition 9 *Let S be a set. A* finite sequence of elements of *S is a function:*[1]

$$f : \{0, \ldots, n-1\} \mapsto S.$$

The length *of the sequence is n. An* infinite sequence of elements of *S is a function:*

$$f : \mathbb{N} \mapsto S.$$

Example Let S be the set of three colors $\{red, yellow, green\}$. Suppose that you at are a traffic light and see a green light, but before you can cross the street, the light changes and you have to wait until the light is green again. L_1, the finite sequence of colors that you will see until you cross, is:

$$f_{L_1}(0) = green,\ f(1)_{L_1} = yellow,\ f(2)_{L_1} = red,\ f(3)_{L_1} = green.$$

L_2, the infinite sequence of colors that the light shows (assuming that it is always on), is:

$$f_{L_2}(0) = green,\ f_{L_2}(1) = yellow,\ f_{L_2}(2) = red,\ f_{L_2}(3) = green,\ f_{L_2}(4) = yellow,\ \ldots,$$

[1] *Functions* and the symbol $\mapsto$ are formally defined in Section 8.

where the ellipsis ... indicates that we know how to continue constructing the sequence. Alternatively, we could formally define the infinite sequence L_2 as:

$$\begin{aligned} f_{L_2}(i) &= \mathit{green} \quad \text{if } i \bmod 3 = 0 \\ f_{L_2}(i) &= \mathit{yellow} \quad \text{if } i \bmod 3 = 1 \\ f_{L_2}(i) &= \mathit{red} \quad \text{if } i \bmod 3 = 2\,. \end{aligned}$$

The elements of a sequence are listed within parentheses ( ) to differentiate them from the elements of sets which are written within braces { }. In a sequence $(s_0, s_1, s_2, \ldots)$, $s_i = f(i)$:

$$L_2 = (\mathit{green},\ \mathit{yellow},\ \mathit{red},\ \mathit{green},\ \mathit{yellow},\ \mathit{red},\ \ldots)\,.$$

Definition 10 *A finite sequence of length n is called an* n-tuple*. The following terms are also used: a* 2*-tuple is a* pair*, a* 3*-tuple is a* triple*, a* 4*-tuple is a* quadruple*.*

Example Examples of sequences:

- A 1-tuple: (red).
- A pair: $(5, 25)$.
- A triple: $(\mathit{red}, \mathit{yellow}, \mathit{green})$.
- A different triple: $(\mathit{red}, \mathit{green}, \mathit{yellow})$.
- A triple with repeated elements: $(\mathit{red}, \mathit{green}, \mathit{green})$.
- An infinite sequence: $(1, 2, 2, 3, 3, 3, 4, 4, 4, 4, \ldots)$.

It is important to understand that a sequence can have multiple occurrences of the same element, whereas a set has only one occurrence of each element:

$$\begin{aligned} (\mathit{green},\ \mathit{yellow},\ \mathit{red},\ \mathit{green}) &\neq (\mathit{green},\ \mathit{yellow},\ \mathit{red}) \\ \{\mathit{green},\ \mathit{yellow},\ \mathit{red},\ \mathit{green}\} &= \{\mathit{green},\ \mathit{yellow},\ \mathit{red}\}\,. \end{aligned}$$

6 Cartesian product

Definition 11 *Let $S_1, \ldots, S_n$ be sets. $S_1 \times \cdots \times S_n$, their* Cartesian product*, is the set of n-tuples $(s_1, \ldots, s_n)$, such that $s_i \in S_i$. If all the sets S_i are the same set S, the notation S^n is used for $S \times \cdots \times S$.*

The case $n = 2$ is very common. Let S and T be sets. $S \times T$ is the set of all pairs (s, t) such that $s \in S$ and $t \in T$.

Example

- $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$ is the set of all pairs of real numbers. This set can be used to represent coordinates in the plane. The term Cartesian plane or Cartesian coordinates is often used for this set.
- $\mathbb{N} \times \{red, yellow, green\}$ is the set of all pairs whose first element is a number and whose second element is a color. This set could be used to represent the color of a traffic light at different points of time: $(28, red)$.
- $\mathbb{N}^3$ can be used to represent dates as (day,month,year): $(11, 12, 1948)$.
- $\mathbb{N}^3$ can also be used to represent times as (hours,minutes,seconds): $(16, 38, 52)$. Alternatively, $\mathbb{N} \times \mathbb{N} \times \mathbb{R}$ or $\mathbb{N}^2 \times \mathbb{R}$ can represent times with fractions of a second: $(16, 38, 52.83)$.

We leave it to the reader to prove the distributive laws for Cartesian products:

Theorem 12

1. $S \times (T \cap U) = (S \times T) \cap (S \times U)$.
2. $S \times (T \cup U) = (S \times T) \cup (S \times U)$.
3. $(S \times T) \cap (U \times V) = (S \cap U) \times (T \cap V)$.
4. $(S \times T) \cup (U \times V) = (S \cup U) \times (T \cup V)$.

7 Relations

Definition 13 *Let $S_1, S_2, \ldots$ be sets. An n-*ary relation *$\mathbb{R}$ is a subset of $S_1 \times \cdots \times S_n$. $\mathbb{R}$ is said to be a relation* on *$S_1 \times \cdots \times S_n$.*
For $n = 1$, a subset of S_1 is called a *unary relation*. For $n = 2$, a subset of $S_1 \times S_2$ is called a *binary relation*.[2]

Example Here are some relations over $\mathbb{N}^k$ for various $k \geq 1$:

- The set of prime numbers P is a relation on $\mathbb{N}^1 = \mathbb{N}$.
- $SQ = \{(n_1, n_2) \in \mathbb{N}^2 \mid n_2 = n_1^2\}$ is a relation on $\mathbb{N}^2$; it is the set of pairs of numbers and their squares: $(4, 16) \in SQ$, $(7, 49) \in SQ$, $(10, 99) \notin SQ$.
- RP is the set of relatively prime numbers, that is, pairs of natural numbers that have no common divisor except for 1:

  $$RP = \{(n, m) \in \mathbb{N}^2 \mid \text{ for all } k\,(n \bmod k = 0 \text{ and } m \bmod k = 0 \text{ imply } k = 1)\}.$$

  Examples are: $(4, 9) \in RP$, $(15, 28) \in RP$, $(14, 63) \notin RP$.

[2] Velleman [3, Definition 4.2.1, p. 171] uses the term *relation* for what we call a *binary relation*. In mathematical logic, n-ary relations are used to define interpretations for n-ary predicates [2, Definition 7.16, p. 136], so we use the more general definition.

- PT, Pythagorean triples, are triples of values that can be the lengths of the sides of a right triangle:
$$\{(x,y,z) \in \mathbb{N}^3 \mid x^2 + y^2 = z^2\}.$$
Examples are: $(3,4,5) \in PT$, $(8,15,17) \in PT$, $(6,8,9) \notin PT$.

- Let F be the set of quadruples $\{(x,y,z,n) \in \mathbb{N}^4 \mid x,y,z > 0, n > 2 \text{ and } x^n + y^n = z^n\}$. Fermat's Last Theorem (which was recently proved) states that $F = \emptyset$, the empty set.

Definition 14 *Let S be a set and R a binary relation on S^2.*

1. *R is* reflexive *iff $R(x,x)$ for all $x \in S$.*
2. *R is* symmetric *iff $R(x_1,x_2)$ implies $R(x_2,x_1)$.*
3. *R is* transitive *iff $R(x_1,x_2)$ and $R(x_2,x_3)$ imply $R(x_1,x_3)$.*

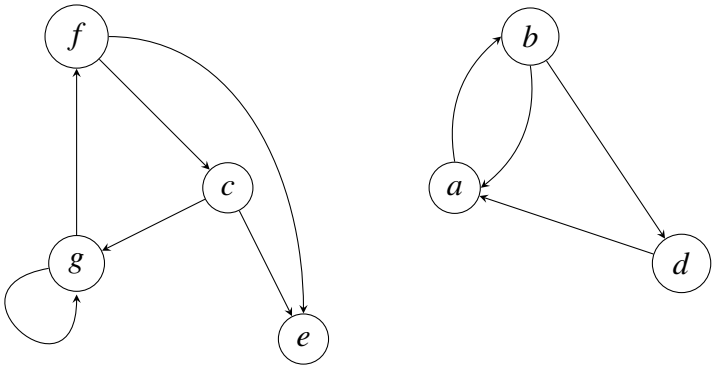
Definition 15 *R^*, the* reflexive transitive closure of a binary relation $\mathbb{R} \subseteq S \times S$, *is the smallest relation that satisfies:*

1. *$R \subseteq R^*$.*
2. *$R^*(x,x)$ for any $x \in S$.*
3. *For any $x_1,x_2,x_3 \in S$, if $R^*(x_1,x_2)$ and $R^*(x_2,x_3)$ then $R^*(x_1,x_3)$.*

Example Consider the set $G = \{a,b,c,d,e,f,g\}$ and the relation ρ defined by:

$$\begin{array}{lllll} \rho(a,b), & \rho(b,a), & \rho(b,d), & \rho(d,a), & \rho(c,e), \\ \rho(f,e), & \rho(f,c), & \rho(c,g), & \rho(g,f), & \rho(g,g). \end{array}$$

It is easy to visualize the relation by drawing a graph with a node for each element of G and a directed arrow from node n_1 to node n_2 if $\rho(n_1,n_2)$:



The relation ρ is not reflexive: although the relation is reflexive at node g since $\rho(g,g)$ is in the relation, the relation does not include $\rho(e,e)$ and similarly for the other nodes. The relation is also not transitive: although the relation is transitive for the nodes f,c,e since $\rho(f,c),\rho(c,e),\rho(f,e)$ are in the relation, the relation does not include $\rho(g,e)$ even though it does include $\rho(g,f),\rho(f,e)$.

ρ^*, the reflexive transitive closure of ρ, includes $\rho^*(n,n)$ for every $n \in G$, as well as $\rho^*(n_1,n_2)$ if there exists a *path* $n_1,\ldots,n_2$ in the graph from n_1 to n_2.

There are no elements in ρ such that $\rho(x,y)$ for $x \in \{a,b,d\}$ and $y \in \{c,e,f,g\}$, nor are there such elements in ρ^*. If you look again at (3) in the definition, we have to ensure $\rho^*(x_1,x_3)$ only if $R^*(x_1,x_2)$ and $R^*(x_2,x_3)$, but there are no such elements in the relation or its reflexive transitive closure, so the requirement is vacuous.

8 Functions

Definition 16 *A relation F on $S_1 \times \cdots \times S_n$ is a* function *iff for every n–1-tuple $(x_1,\ldots,x_{n-1}) \in S_1 \times \cdots \times S_{n-1}$, there is at most one $x_n \in S_n$, such that $F(x_1,\ldots,x_n)$.*

Example The relation $SQ = \{(n_1,n_2), n_1,n_2 \in \mathbb{Z} \mid n_2 = n_1^2\}$ is a function because for each integer, there is exactly one integer that is its square. The relation $SQRT = \{(n_1,n_2), n_1,n_2 \in \mathbb{Z} \mid n_2^2 = n_1\}$ is *not* a function because there are integers that have two (integer) square roots, one positive and one negative, for example, $(7)^2 = 49$ and $(-7)^2 = 49$.

Definition 17 *Here are some terms used with functions:*

1. *The* domain *of F is the set of all $(x_1,\ldots,x_{n-1}) \in S_1 \times \cdots \times S_{n-1}$ for which (exactly one) $x_n = F(x_1,\ldots,x_{n-1})$ exists.*
2. *The* range *of F is the set of all $x_n \in S_n$ such that $x_n = F(x_1,\ldots,x_{n-1})$ for at least one $(x_1,\ldots,x_{n-1})$.*
3. *F is* total *if the domain of F is all of $S_1 \times \cdots \times S_{n-1}$; otherwise, F is* partial.
4. *F is* injective *or* one-to-one *iff $(x_1,\ldots,x_{n-1}) \neq (y_1,\ldots,y_{n-1})$ implies that*

$$F(x_1,\ldots,x_{n-1}) \neq F(y_1,\ldots,y_{n-1}).$$

5. *F is* surjective *or* onto *iff its range is all of S_n.*
6. *F is* bijective *(*one-to-one and onto*) iff it is injective and surjective.*

Notation

The notation $F : S_1 \times \cdots S_{n-1} \mapsto S_n$ is used instead of $F(x_1,\ldots,x_n) \in S_n$, where $(x_1,\ldots,x_{n-1}) \in S_1 \times \cdots \times S_{n-1}$.

Similarly, the value of the function for an element of the domain can use functional notation or an expression can be given. For the example above, $n_2 = SQ(n_1) = n_1^2$ for $n_1, n_2 \in \mathbb{Z}$.

Example Figure 1 shows the function:

$$M2 : \{0,1,2,3,4,5,6\} \mapsto \{0,1\}\,, \text{ where } M2(n) = n \bmod 2\,,$$

which is surjective (onto) but not injective (one-to-one).

Figure 2 shows the function:

$$M2' : \{0,1,2,3,4,5,6\} \mapsto \{0,1,2,3,4,5,6\}\,, \text{ where } M2'(n) = n \bmod 2\,,$$

which is neither surjective nor injective:

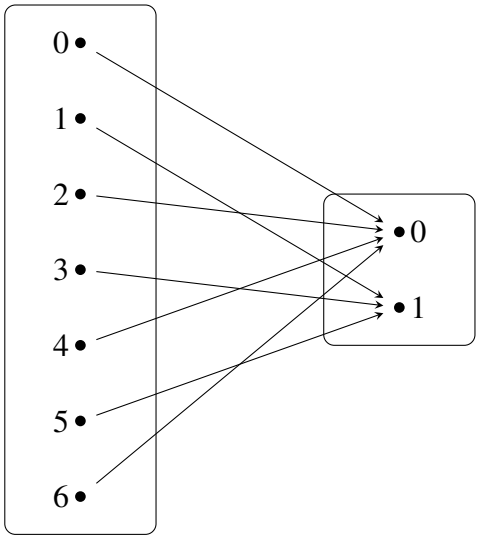


Figure 1: Surjective but not injective

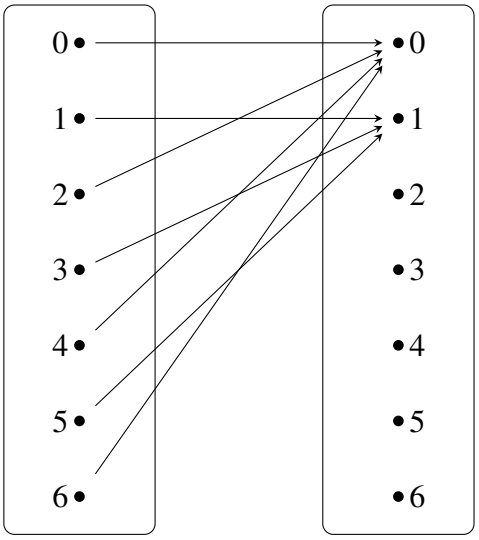


Figure 2: Neither surjective nor injective

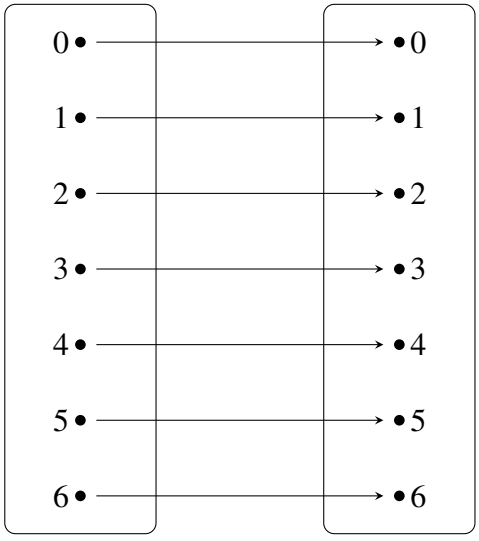


Figure 3: Surjective and injective

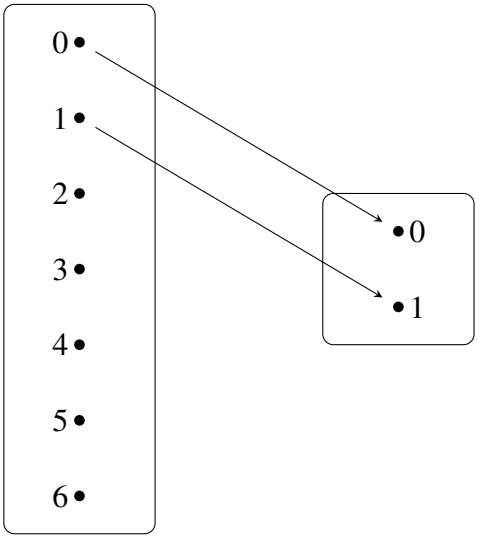


Figure 4: Surjective and injective but not total

Figure 3 shows the function:

$$M7 : \{0,1,2,3,4,5,6\} \mapsto \{0,1,2,3,4,5,6\}, \text{ where } M7(n) = n \bmod 7,$$

which is bijective, because it is both surjective (onto) and injective (one-to-one).

Figure 4 shows the function:

$$M2'' : \{0,1,2,3,4,5,6\} \mapsto \{0,1\}, \text{ where } M2''(n) = n \bmod 2,$$

which is bijective, but it is a partial function because the domain is not all of $\{0,1,2,3,4,5,6\}$.

Example $SQ = \{(n_1,n_2), n_1,n_2 \in \mathbb{Z} \mid n_2 = n_1^2\}$ is a total function. Its domain is all of $\mathbb{Z}$, but its range is only the subset of $\mathbb{Z}$ consisting of all squares. Therefore SQ is not surjective and thus not bijective. It is also not injective because it is possible that $n_1 \neq n_2$, but $n_1^2 = n_2^2$, for example, $7 \neq -7$ but $49 = 49$.

Example Let $\{a_0,\ldots,a_{n-1}\} \in \mathbb{N}$ and define the set:

$$CY_{\{a_0,\ldots,a_{n-1}\}} = \{(a_k, a_{k+1}, \ldots, a_{(k+n-1) \bmod n}) \mid k \in \{0,\ldots,n-1\}\}.^3$$

Each CY is a set of sequences whose elements are cyclic. For example:

$$\begin{aligned} CY_{\{2,16,11,14,82\}} &= \{(2,16,11,14,82),\ (16,11,14,82,2), \\ &\quad (11,14,82,2,16),\ (14,82,2,16,11),\ (82,2,16,11,14)\}. \end{aligned}$$

We can define function for shifting left and right:

$$\begin{aligned} \text{left\_shift}(a_0,a_1,\ldots,a_{n-2},a_{n-1}) &= (a_1,\ldots,a_{n-2},a_{n-1},a_0) \\ \text{right\_shift}(a_0,a_1,\ldots,a_{n-2},a_{n-1}) &= (a_{n-1},a_0,a_1,\ldots,a_{n-2}). \end{aligned}$$

Clearly, the domain and the range of these functions is all of $CY_{\{a_0,\ldots,a_{n-1}\}}$. The functions are bijective since for every $c \in CY$ there is exactly one c_l and one c_r such that $c_l = \text{left\_shift}(c)$ and $c_r = \text{right\_shift}(c)$.

9 Cardinality

Definition 18 *The* cardinality *of a set S is the number of elements in the set. The cardinality of a set S is* finite *iff there is an integer n such that the number of elements in S is the same as the number of elements in the set* $\{1,2,\ldots,n\}$. *Otherwise the cardinality is* infinite.

A set S is countable *iff it is finite or its cardinality is the same as the cardinality of* $\mathbb{N}$. *Otherwise the set is* uncountable. ∎

[3] For $k = n-1$, we need to compute $(k+n-1) \bmod n = (2n-2) \bmod n$. But adding a multiple of n does not change the modulus, so $(2n-2) \bmod n = (2n-2+(-n)) \bmod n = (n-2) \bmod n = n-2$ as expected. When $n = 5$ as in the example, $(k+n-1) \bmod 5 = (5-1+5-1) \bmod 5 = (10-2+(-5)) \bmod 5 = 3 \bmod 5 = 3$, and the last element of CY is (a_4,a_0,a_1,a_2,a_3).

The cardinality of a finite set S is n if there is a bijective function $f : \{1,\ldots,n\} \mapsto S$. An infinite set is countable if there is a bijective function $f : \mathbb{N} \mapsto S$.

Example The cardinality of $S = \{a,b,c,\ldots,x,y,z\}$, the set of lower-case English letters, is 26. The following function is bijective:

$$f(1) = a,\ f(2) = b,\ f(3) = c,\ \ldots,\ f(24) = x,\ f(25) = y,\ f(26) = z.$$

Example E, the set of even natural numbers, is countable and infinite. Let f be:

$$f(0) = 0,\ f(1) = 2,\ f(2) = 4,\ \ldots.$$

Of course we can't list the value of the function for every natural number, but we can give the definition of a function f whose domain is $\mathbb{N}$ and whose range is E: $f(n) = 2n$. We leave it to the reader to show that f is bijective.

Infinite numbers are non-intuitive. The set of even natural numbers is a *proper* subset of the set of natural numbers (because, for example, $3 \in \mathbb{N}$, $3 \notin E$), but the cardinality of E (the number of elements in E) is the same as the cardinality of $\mathbb{N}$ (the number of elements in $\mathbb{N}$)!

Theorem 19 *$\mathbb{Z}$, the set of integers, is countable.*

Proof At first glance this seems impossible because $\mathbb{Z}$ has no "first element." However, the definition of cardinality does not require that order be preserved, only that there be a bijective function. If we arrange the integers as follows it is obvious that $\mathbb{Z}$ is countable:

$$0, 1, -1, 2, -2, 3, -3, 4, -4, \ldots,$$

or in general:

$$\begin{array}{ll} f(1) = 0 & \\ f(k) = k/2 & \text{if } k > 0 \text{ is even} \\ f(k) = -(k-1)/2 & \text{if } k \text{ is odd}. \end{array}$$

∎

It is worthwhile becoming familiar with this "trick" because it is used, for example, to prove that a Turing machine with a two-way tape, a tape that is infinite in two directions, can be simulated by a Turing machine with a one-way tape, a tape that is infinite in one direction. The two-way tape is "folded over" to become a one-way tape where each symbol on the tape encodes the symbol of the left-hand part of the tape and the right-hand part of the tape.

Theorem 20 *The set of rational numbers $\mathbb{Q}$ is countable.*

Proof This is more difficult than proving that $\mathbb{Z}$ is countable because rational numbers are infinite both in the numerator and the denominator. The trick is to order the rational numbers by the *sum* of the numerator and the denominator. Within each sum the rational numbers are

ordered by increasing demoninators:

Sum	Rational numbers
0	0
1	1
2	2
3	$3, \frac{1}{2}$
4	$4, \frac{1}{3}$
5	$5, \frac{3}{2}, \frac{2}{3}, \frac{1}{4}$
6	$6, \frac{1}{5}$
7	$7, \frac{5}{2}, \frac{4}{3}, \frac{3}{4}, \frac{2}{5}, \frac{1}{6}$

We have assumed that the rational numbers are reduced so that there is no common factor in both the numerator and the denominator. The negative rational numbers can be included as we did for the negative integers. ∎

Georg Cantor first proved the following theorem:

Theorem 21 *The set of real numbers $\mathbb{R}$ is uncountable.*

Proof Suppose to the contrary that there is a bijective function $f : \mathbb{N} \mapsto \mathbb{R}$, so that it makes sense to talk about r_i, the i-th real number. In fact, let us suppose that there is a bijective function $f : \mathbb{N} \mapsto \mathbb{R}, 0 \leq r < 1$. Each real number can be represented as an infinite decimal number:[4]

$$r_i = 0.d_i^1 d_i^2 d_i^3 d_i^4 d_i^5 \cdots.$$

Consider now the real number r defined by:

$$r = 0.e_1 e_2 e_3 e_4 e_5 \cdots,$$

where $e_i = (d_i^i + 1) \bmod 10$. That is, the first digit of r is different from the first digit of r_1, the second digit of r is different from the second digit of r_2, and so on. It follows that $r \neq r_i$ for all $i \in \mathbb{N}$, contradicting the assumption that f was surjective. ∎

This method of proof, called *diagonalization*, is frequently used in computer science, for example, to prove that no Turing machine can decide whether an arbitrary Turing machine halts for a arbitrary input.

[4]The full proof must take account of a technical problem: two real numbers can have different sequences of digits, for example, $1.0000\cdots = 0.9999\cdots$.

10 Powersets

Definition 22 *Let S be a set. The* powerset *of S, denoted 2^S, is the set of all subsets of S.*

Example Here is the powerset of the finite set $S = \{red, yellow, green\}$:

$$\begin{array}{l}\{\\ \{red, yellow, green\},\\ \{red, yellow\}, \{red, green\}, \{yellow, green\},\\ \{red\}, \{yellow\}, \{green\},\\ \emptyset\\ \}.\end{array}$$

The cardinality of S is 3, while the cardinality of the powerset is $8 = 2^3$.

Theorem 23 *Let S be a finite set of cardinality n; then the cardinality of its powerset is 2^n.*

Proof A subset S' is constructed by choosing for each $s \in S$ whether $s \in S'$ or $s \notin S'$. These choices are independent: for any s_i, once we have chosen $s_i \in S'$ or $s_i \notin S'$, for each $s_j \in S$, $j \neq i$ the choice $s_j \in S'$ or $s_j \notin S'$ does not depend on our previous choice for s_i. Therefore, the number of subsets is:

$$\overbrace{2 \cdot 2 \cdot 2 \cdots 2 \cdot 2}^{n} = 2^n.$$

∎

A diagonalization argument can be used to show that $2^{\mathbb{N}}$ is not countable so its cardinality is larger than the cardinality of $\mathbb{N}$ which is denoted $\aleph_0$ (read "aleph 0"). The cardinality of $2^{\mathbb{N}}$ is denoted $\aleph_1$ and $\aleph_1 > \aleph_0$. If we again take powersets, we find an infinite hierarchy of ever larger cardinalities.

11 Induction

To prove a property of a set, mathematical induction can be used. We give an overview of induction; for more detail and many examples see [2, 3].

Axiom 24 (Mathematical induction) *Let $P(n)$ be a property, $n \in \mathbb{N}, n > 0$. If you can:*

- Base case*: Prove $P(1)$.*
- Inductive step*: For arbitrary $m \in \mathbb{N}$, prove $P(m+1)$ under the assumption that $P(m)$ is true.*

Then you have proved $P(n)$ for all $n \geq 1$.
The assumption that $P(m)$ is true for arbitrary m is called the inductive hypothesis.

Here is a simple theorem that can be proved using mathematical induction:

Theorem 25 *For $n \geq 1$:*

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}.$$

Proof The base case is trivial:

$$\sum_{i=1}^{1} i = 1 = \frac{1(1+1)}{2}.$$

The inductive hypothesis is that the equation is true for m:

$$\sum_{i=1}^{m} i = \frac{m(m+1)}{2}.$$

The inductive step is to prove the equation for $m+1$:

$$\begin{aligned}
\sum_{i=1}^{m+1} i &= \sum_{i=1}^{m} i + (m+1) && (1)\\
&\stackrel{\bullet}{=} \frac{m(m+1)}{2} + (m+1) && (2)\\
&= \frac{m(m+1)+2(m+1)}{2} = \frac{(m+1)(m+2)}{2}. && (3)
\end{aligned}$$

By the axiom of mathematical induction:

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$$

is true for any $n \geq 1$. ∎

Let us justify the reasoning in the inductive step. In (1) the sum is rewritten as two terms: the first term is the sum of the numbers from 1 to m and the second term is the number $(m+1)$. In (2), the notation $\stackrel{\bullet}{=}$ denotes that the *inductive hypothesis* is used to substitute $\frac{m(m+1)}{2}$ for $\sum_{i=1}^{m} i$. Equations (3) use elementary algebra.

Axiom 24 is the simplest form of induction. It can be generalized:

- The base case need not be $n = 1$.
- The inductive step need not be $n+1$.
- There may be more than one base case.
- There may be more than one inductive step.
- The inductive hypothesis can assume $P(k)$ for *all* $k \leq m$ and not just for $k = m$.

In computer science *structural induction* is commonly used. Instead of an inductive hypothesis $P(n)$ being used in the inductive step to prove $P(n+1)$, in structural induction the inductive hypothesis is that a property is true for "simple" structures and the inductive step

proves that the property is true for "complex" structures that are built from the "simple" structures.

Here is an example of the use of structural induction in computer science. The proof is also interesting because there are several base cases and several inductive steps. Knowledge of nondeterministic finite automata (NFA) and regular expressions (RE) is assumed.

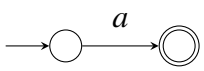
Theorem 26 *Let r be an RE. Then there is an NFA that accepts the language of r.*

Proof There are *three* base cases:

- r is the empty set $\emptyset$. The NFA consisting of one initial and one final state with no transitions. It accepts no strings.

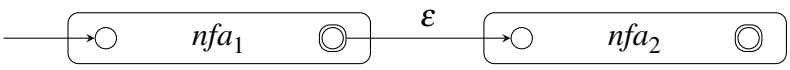
- r is the null string ε. The NFA consisting of one state which is both initial and final. It accepts the null string:

- r is a single symbol a. The following NFA accepts the language $\{a\}$:

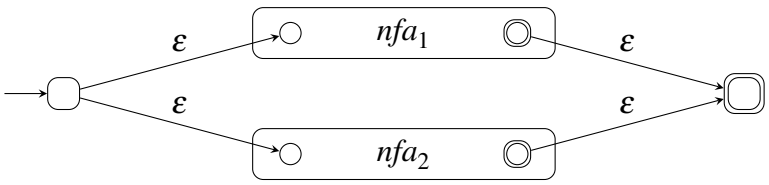


There are *three* inductive steps:

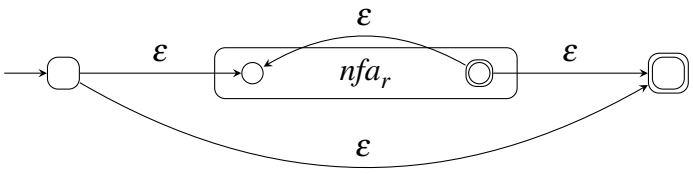
- Concatenation $r_1 r_2$: By the inductive hypothesis there are NFAs nfa_1 and nfa_2 that accept the languages of r_1 and r_2, respectively. Construct nfa_{12} by adding a null transition from the final state of nfa_1 to the initial state of nfa_2. The initial state of nfa_{12} is the initial state of nfa_1 and its final state is the final state of nfa_2.



- Union $r_1 + r_2$: By the inductive hypothesis there are NFAs nfa_1 and nfa_2 that accept the languages of r_1 and r_2, respectively. Construct nfa_{12} by adding new start and final states and null transitions.



- Closure r^*: By the inductive hypothesis there is an NFA nfa_r that accepts the language of r. Add new initial and final states and null transitions as shown in the diagram:



The null transition from the initial state to the final state is taken for strings generated by zero instances of r, and the internal transition from the final state of nfa_r to its initial state is for more than one repetition of r.

∎

12 The Well-ordering Principle

The well-ordering principle is an axiom that is more intuitive than mathematical induction, although induction is much easier to use in practice.

Definition 27 *Let S be a set with a binary relational operator $\leq$.*

1. *S is* totally ordered *iff for any $x, y \in S$, either $x \leq y$ or $y \leq x$ or $x = y$.*
2. *A totally ordered set S has a* lower bound *iff there is some b such that $b \leq n$ for all $n \in S$.*
3. *A totally ordered set S has a* least element *iff there is some $b \in S$ such that $b \leq n$ for all $n \in S$. A least element is also a lower bound, but in addition it is an element of the set.*

Example Subsets of $\mathbb{Z}$ are totally ordered:

$$S = \{8, 3, 19, 5, 6, 23\}, \quad E_Z = \{\ldots, -4, -2, 0, 2, 4, \ldots\},$$

where E_Z is the set of all even *integers*. Some lower bounds for S are $3, 0, -10$. The least element of S is 3. *E_Z does not* have a lower bound and therefore does not have a least element.

Example The set of *positive* rational numbers has an infinite number of lower bounds (zero and all negative rational numbers), but it has no least element because for any positive rational number x, $x/2$ is a smaller rational number.

Definition 28 *Let S be a totally ordered set. S is* well-ordered *iff* every *nonempty subset of S has a least element.*

$S = \{8, 3, 19, 5, 6, 23\}$ is well-ordered because every nonempty subset has a least element. The least element of S itself is 3, the least element of $\{8, 19, 5\}$ is 5.

$E_Z = \{\ldots, -4, -2, 0, 2, 4, \ldots\}$ is not well-ordered because $E_Z \subseteq E_Z$ but it has no least element. The set of even numbers in $\mathbb{N}$ is well-ordered.

Axiom 29 ***(The well-ordering principle)*** *Every nonempty subset of the integers that has a lower bound is well-ordered.*

The set of positive integers is non-empty and has many lower bounds—zero and all negative integers (and the set even has a least element, namely 1), therefore by Axiom 29 it is well-ordered. If we assume Axiom 24 we can prove Axiom 29 and conversely [2, 3].

13 References

[1] M. Ben-Ari. *Mathematical Logic for Computer Science (Third Edition)*, Springer, 2012, ISBN 978-1-4471-4128-0, `http://www.springer.com/978-1-4471-4128-0`.

[2] M. Ben-Ari. *The Many Guises of Induction*, 2019, `https://www.weizmann.ac.il/sci-tea/benari/mathematics#induction`.

[3] D.S. Gunderson. *Handbook of Mathematical Induction: Theory and Applications*, Mathematical Association of America, 2010.

[4] D.J. Velleman. *How to Prove It: A Structured Approach (Second Edition)*, Cambridge University Press, 2006.

Index

$\in$, 2
$\notin$, 2
$\emptyset$, 2
$\{\cdots\}$, 2
$\mathbb{Z}$, 2
$\mathbb{N}$, 2
$\{\cdots \mid \cdots\}$, 2
$\mathbb{R}$, 2
$\subseteq$, 3
$\subset$, 3
$\cup$, 4
$\cap$, 4
$-$, 4
$\overline{T}$, 4
$(\cdots)$, 7
$\times$, 7
R^*, 9
$\mapsto$, 10
2^S, 15

Bijective, 10

Cardinality, 12
Cartesian product, 7
Complement, 4
Countable, 12

Diagonalization, 14
Difference, 4
Disjoint, 4
Domain, 10

Element, 2
Empty set, 2

Function, 10

Injective, 10
Integers, 2
Intersection, 4

Least element, 18
Lower bound, 18

Mathematical induction, 15

Natural numbers, 2

Partial, 10
Powerset, 15
Proper, 3

Range, 10
Real numbers, 2
Reflexive transitive closure, 9
Relation, 8

Sequence, 6
Set, 2
Set comprehension, 2
Singleton, 2
Structural induction, 16
Subset, 3
Surjective, 10

Total, 10
Totally ordered, 18

Uncountable, 12
Union, 4

Venn diagram, 3

Well-ordering principle, 19