Section 1.2 Subsets and the Boolean operations on sets

If every element of the set A is an element of the set B, we say that A is a *subset of* B, or that A *is contained in* B, or that B *contains* A, and we write $A \subseteq B$, or $B \supseteq A$.

Any set is a subset of itself: $A \subseteq A$. When we want to say that A is a subset of B and A is diffrent from B, we say that A is a *proper subset of* B, and we write $A \subset B$.

For instance, $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$; the five basic number-sets are increasingly more comprehensive sets. $\mathbb{Z} \subset \mathbb{Q}$ because every integer is a rational number ($n = \frac{n}{1}$) , but $not(\mathbb{Q} \subseteq \mathbb{Z})$, since $\frac{1}{2}$ is a rational number but not an integer.

Note that for any set A, $\emptyset \subseteq A$: any thing that belongs to $\emptyset$ (there is nothing like that) belongs to A as well.

The following general law, the

Antisymmetry law for $\subseteq$:

$$A \subseteq B \;\;\&\;\; B \subseteq A \implies A = B \tag{1}$$

(here A and B are sets) is equivalent to the principle of extensionality (see the first section), namely the principle that says that sets are determined by their elements. In fact, the left-hand side of the implication in (1) says that all elements of A are elements of B and vice versa, which is to say that A and B have the same elements.

The antisymmetry law serves as the formal setup for showing that two sets given by conditions are the same (if that is the case). One shows two things: **one**, that the first set is contained in the other, **two**, that the other is contained in the first. The reader should try this on the example of the sets given in (2) and (3) in Section 1.1. Later in this section, we will see another example.

An obvious law for $\subseteq$ is the law of

Transitivity for $\subseteq$:

$$A \subseteq B \text{ and } B \subseteq C \implies A \subseteq C .$$

Although this is completely obvious, it serves as the basis for an important generalization, in the notion of ordering, considered in the next chapter.

Note that the relations "being an element of" , denoted by $\in$, and "being a subset of", denoted by $\subseteq$, are very different. E.g., $2 \subseteq \{2, 7\}$ does not hold; it does not even make sense, since 2 is not a set. Of course, $2 \in \{2, 7\}$ does hold. Also, $\{2\} \subseteq \{2, 7\}$ holds, and this is the same as $2 \in \{2, 7\}$. On the other hand, $\{2\} \in \{2, 7\}$ *does not* hold; in fact, the set $\{2, 7\}$ has no element that is a set.

It is possible that both $A \in B$ and $A \subseteq B$ hold at the same time; e.g.,

$$\{1, 2\} \in \{\{1, 2\}, 1, 2\} , \qquad \{1, 2\} \subseteq \{\{1, 2\}, 1, 2\} ;$$

however, such situations are rather rare in practice.

The set of all subsets of a given set A is called the *power set* of A , and it is denoted by $\mathcal{P}(A)$. That is to say,

$$X \in \mathcal{P}(A) \iff X \subseteq A .$$

E.g.,

$$\mathcal{P}(\{1, 2\}) = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\} .$$

$\mathcal{P}(A)$ is never empty; the empty set is always in it. Also, $A \in \mathcal{P}(A)$, and thus, if A is non-empty, $\mathcal{P}(A)$ has at least two elements.

With A an *alphabet*, a set of symbols, A^* the set of strings over A, a subset of A^* is called a *language* over A. The idea is that the strings in the language are the *well-formed sentences* of the language (a sentence is considered a single string, by treating the blanks as occurrences of a special character called **blank**). Initially, we do not put any condition on how the sentences should be formed; hence the complete generality of the definition of language. The theory of formal languages, an important part of theoretical computer science, deals mainly with how one can generate a language by rules. The legal programs of PASCAL form a (formal) language; the relevance of formal language theory should be indicated by this remark alone. Later we will also see particular formal languages related to logic.

The bracket notation for sets is used in a modified form to denote a subset of a given set.

$$\{n\in\mathbb{N} \mid n \text{ is prime}\}$$

is the same as

$$\{n \mid n\in\mathbb{N} \text{ and } n \text{ is prime}\}\,,$$

the set of positive primes, and of course, it is a subset of $\mathbb{N}$.

In what follows, capital letters always denote sets.

The *intersection* of X and Y, in symbols $X \cap Y$, is the set whose elements are the things that are elements of both X and Y at the same time:

$$a \in X \cap Y \iff a \in X \text{ and } a \in Y\,.$$

E.g.,

$$\{1, 2, 4, 7, 10\} \cap \{3, 4, 10, 13\} = \{4, 10\}\,,$$

$$\{1, 2, 4\} \cap \{3, 7\} = \emptyset\,,$$

$$\{n\in\mathbb{N} \mid n \text{ is even}\} \cap \{n\in\mathbb{N} \mid n \text{ is prime}\} = \{2\}\,.$$

If A_i is a set for all values $i = 1, 2, ..., n$ of the subscript i, then $\bigcap_{i=1}^{n} A_i$, the *intersection* of the A_i, is the set of all those x which are elements of all A_i:

$$x \in \bigcap_{i=1}^{n} A_i \iff \text{for all } i \text{ such that } 1 \le i \le n, \text{ we have } x \in A_i .$$

E.g., if A_i is the set of natural numbers divisible by i, then

$$\bigcap_{i=1}^{10} A_i = A_{8 \cdot 9 \cdot 5 \cdot 7} = A_{2520}$$

(why?).

One can take the intersection of any family of sets *except the intersection of the empty family* (which would have to be the set of all things, a non-existent set). If I is any set, and A_i is a particular set for each $i \in I$ (in which case we talk about the *family* $\langle A_i \rangle_{i \varepsilon I}$ of sets), then $\bigcap_{i \in I} A_i$, the intersection of the A_i, is the set of all things that belong to every A_i, $i \in I$. The notation $\bigcap_{i=1}^{n} A_i$ means the same as $\bigcap_{i \in \{1, ..., n\}} A_i$.

E.g., $\bigcap_{n \in \mathbb{N} - \{0\}} \{k \cdot n \mid k \in \mathbb{N}\} = \{0\}$ (why?) ($\mathbb{N} - \{0\} = \{n \in \mathbb{N} \mid n \neq 0\}$).

The *union* of two sets X and Y, denoted $X \cup Y$, is the set of all things that are elements of either X, *or* Y, or both:

$$a \in X \cup Y \iff \text{either } a \in X, \text{ or } a \in Y \text{ (or both).}$$

E.g.,

$$\{1, 2, 4, 7, 10\} \cup \{3, 4, 10, 13\} = \{1, 2, 3, 4, 7, 10, 13\},$$

$$\{n \in \mathbb{N} \mid n \text{ is even}\} \cup \{n \in \mathbb{N} \mid n \text{ is odd}\} = \mathbb{N} .$$

We may take the union of more than two sets. $\bigcup_{i\in I} A_i$ denotes the set, called the *union* of the sets A_i , $i\in I$, whose elements are those things that belong to at least one of the sets A_i . E.g., if $B_i = \{5k+i \mid k \in \mathbb{N}\}$, then

$$\bigcup_{i\in\{0,\ 1,\ 2,\ 3,\ 4\}} B_i = \mathbb{N}$$

(why is that?). We write $\bigcup_{i=0}^{4}$ instead of $\bigcup_{i\varepsilon\{0,\ 1,\ 2,\ 3,\ 4\}}$, thus $\bigcup_{i=0}^{4} B_i = \mathbb{N}$. Or, to give another example: if A_i is the set

$$A_i = \{n\in\mathbb{N} \mid n \text{ is divisible by } i \text{, and } n \le 120\} ,$$

then

$$\bigcup_{i=2}^{7} A_i = \{n\in\mathbb{N} \mid n \le 120 \text{ and } n \text{ is not prime}\} \cup \{2,\ 3,\ 5,\ 7\}$$

(this is related to the equality of the sets (2) and (3) in Section 1.1).

Here is another way we can use the union ($\bigcup$) and intersection ($\bigcap$) symbols.

Assume that X is a set of sets: that is, all elements of X are themselves sets. Then $\bigcup X$ denotes the union of all the sets in X .

For instance, if $X = \{A_i \mid i\in\{2,\ 3,\ 4,\ 5,\ 6,\ 7,\ \}\}$, then $\bigcup X$ is the same as $\bigcup_{i=2}^{7} A_i$ considered earlier.

We may give the definition of the notation $\bigcup X$ thus: for any x ,

$$x\in \bigcup X \iff \text{there is } A\in X \text{ such that } x\in A \text{ .}$$

Note that $\bigcup \emptyset = \emptyset$.

The notation $\bigcap X$ is similar: it denotes the intersection of all the sets in the set X . In symbols:

$$x \in \bigcap X \iff \text{for all } A \in X \text{, we have } x \in A .$$

Here, there is an exclusion: $\bigcap \emptyset$ does not make sense. The reason is that the last display would give that *all* things x belong to $\bigcap \emptyset$; however, "the set of *all* things" is not a legitimate concept.

The third operation we consider here is the *difference* of two sets. $X - Y$ denotes the set of those things that are in X , but are not in Y :

$$a \in X - Y \iff a \in X \text{ and } a \notin Y .$$

E.g.,

$$\{1, 2, 4, 7, 10\} - \{3, 4, 10, 13\} = \{1, 2, 7\} .$$

Let A be a fixed set, and consider the operations of intersection, union and difference performed on subsets of A ; the result is always a subset of A again. This is clear since those operations never involve elements that are not in either of the sets in question. We say that the collection of all subsets of A is *closed* under the operation of intersection, union and difference; the latter operations are collectively called the *Boolean operations*.

If all sets under consideration are understood to be subsets of the fixed set A , then the difference $A - X$ is abbreviated as $-X$, and it is called the *complement* of X , or the *complement of* X *with respect to* A in more detail. E.g., if we are talking about subsets of $\mathbb{N}$, that is, $A=\mathbb{N}$, then

$$-\{n \in \mathbb{N} \mid n \text{ is even}\} = \{n \in \mathbb{N} \mid n \text{ is odd}\} .$$

The Boolean operations obey certain laws. Here is a list of them. In what follows, A is a fixed set, X , Y and Z denote arbitrary subsets of A ; $-X$ means $A-X$.

Commutative laws:

$$X \cap Y = Y \cap X\,, \qquad X \cup Y = Y \cup X\,.$$

Associative laws:

$$X \cap (Y \cap Z) = (X \cap Y) \cap Z\,, \qquad X \cup (Y \cup Z) = (X \cup Y) \cup Z\,.$$

Just like in the case of addition and multiplication of numbers, these laws have the consequence that in expressions using several intersection operations, or alternatively, several union operations, parentheses may be omitted or restored in any meaningful way, and the order of terms may be changed, without altering the value of the expression. E.g.,

$$(X \cap (U \cap V)) \cap Z = (U \cap Z) \cap (V \cap X) = U \cap V \cap X \cap Z\,.$$

Absorption laws:

$$X \cap (X \cup Y) = X\,, \qquad X \cup (X \cap Y) = X\,.$$

Idempotent laws:

$$X \cap X = X \qquad X \cup X = X$$

Distributive laws:

$$(X \cup Y) \cap Z = (X \cap Z) \cup (Y \cap Z)\,,$$
$$(X \cap Y) \cup Z = (X \cup Z) \cap (Y \cup Z)\,.$$

Laws for complements:

$$X \cap -X = 0\,, \qquad X \cup -X = A\,;$$

De Morgan's laws:

$$-(X \cap Y) = (-X) \cup (-Y)\,, \qquad -(X \cup Y) = (-X) \cap (-Y)\,.$$

The proofs of these identities can be done via antisymmetry for $\subseteq$. We consider the first distributive law.

To show $(X \cup Y) \cap Z \subseteq (X \cap Z) \cup (Y \cap Z)$, let x belong to the left-hand-side, to show that it belongs to the right-hand side. Then x belongs both to $X \cup Y$ and Z. Since it belongs to $X \cup Y$, it either belongs to X (**Case 1**), or to Y (**Case 2**), or possibly both. In **Case 1**, x belongs both to X and Z, hence, to $X \cap Z$, and thus to $(X \cap Z) \cup (Y \cap Z)$. In **Case 2**, we obtain the same conclusion similarly. We have shown that x belongs to the right-hand side in any case.

To show $(X \cup Y) \cap Z \supseteq (X \cap Z) \cup (Y \cap Z)$, let x belong to the right-hand side, to show that it belongs to the left-hand side. Since $x \in (X \cap Z) \cup (Y \cap Z)$, either $x \in X \cap Z$ (**Case 1**), or $x \in Y \cap Z$ (**Case 2**). In the first case, $x \in X$ and $x \in Z$; from $x \in X$ it follows that $x \in X \cup Y$; hence, x belongs to both $X \cup Y$ and Z, and thus to the left-hand side. In Case 2, the argument is similar.