Chapter 1. Sets and Mathematical Statements

1.1 Remark: A little over 100 years ago, it was found that some mathematical proofs contained paradoxes, and these paradoxes could be used to prove statements that were known to be false. One well-known paradox, outside of the realm of mathematics, is the statement

"This statement is false".

The above statement is true if and only if it is false. It is one form of a paradox known as the **liar's paradox**. After examining some lengthy and convoluted mathematical proofs which contained paradoxes, Bertrand Russell came up with the following mathematical paradox, which is somewhat similar to the liar's paradox:

Let X be the set of all sets, and let $S = \{A \in X \big| A \notin A\}$.
Note for example that $\mathbf{Z} \notin \mathbf{Z}$ so $\mathbf{Z} \in S$, and $X \in X$ so $X \notin S$.
Then we have $S \in S$ if and only if $S \notin S$.

This paradox is known as **Russell's paradox**. With Russell's paradox, it was possible to construct a proof by contradiction, which followed all the accepted rules of mathematical proof, of any statement whatsoever. Mathematicians realized that they would need to modify the accepted framework of mathematics in order to ensure that mathematical paradoxes could no longer arise. They were led to consider the following three questions.

1. Exactly what is an allowable mathematical object?
2. Exactly what is an allowable mathematical statement?
3. Exactly what is an allowable mathematical proof?

Eventually, after a great deal of work by many mathematicians, a consensus was reached as to the answers to these three questions. Roughly speaking, the answers are as follows. Essentially every mathematical object is a mathematical **set** (this includes objects that we would not normally consider to be sets, such as integers and functions), and a mathematical set can be constructed using certain specific rules, known as the **Zermelo-Fraenkel Axioms** along with the **Axiom of Choice** which, together, are referred to as the **ZFC axioms**. Mathematical statements are normally expressed using a combination of mathematical symbols and words from a natural language, such as English, but every mathematical statement can be expressed as a so-called **formula** in a certain specific formal symbolic language, called the language of **first-order set theory**, which uses symbols rather than words. Mathematical proofs are likewise normally expressed using a combination of symbols and words, but every mathematical proof can be translated into a very precise symbolic form of proof called a **derivation**. One kind of derivation consists of a finite list of ordered pairs $(\mathcal{S}_n, F_n)$ (which we think of as *proven theorems*), where each S_n is a finite set of formulas (called the *premises*) and each F_n is a single formula (called the *conclusion*), such that each pair $(\mathcal{S}_n, F_n)$ can be obtained from previous pairs $(\mathcal{S}_i, F_i)$ with $i < n$, using certain specific proof rules. In this chapter we shall provide more detailed answers to the the first two of the above three questions, and in the next chapter we shall consider the third question.

1.2 Definition: Every mathematical object is either a (mathematical) **set** or a (mathematical) **class**. Every set or class is a collection of sets. When S is a set or a class and x is a set, we write $x \in S$ to indicate that x is an **element** of S. When A and B are sets, we say that A is **equal** to B, and we write $A = B$, when A and B have the same elements, we say that A is a **subset** of B, and we write $A \subseteq B$ (some books write $A \subset B$), when every element of A is also an element of B, and we say that A is a **proper subset** of B, and we write $A \subset B$, or for emphasis $A \subsetneqq B$, when $A \subseteq B$ but $A \neq B$. When $F(x)$ is a mathematical statement about a set x, we write $\{x \,|\, F(x)\}$ to denote the collection of all sets x for which the statement $F(x)$ is true. When $F(x)$ is a statement about a set x and A is a set, we write $\{x \in A \,|\, F(x)\}$ to denote the collection $\big\{x \,\big|\, x \in A \text{ and } F(x)\big\}$.

A (mathematical) **class** is any collection of sets of the form $\{x \,|\, F(x)\}$ where $F(x)$ is a mathematical statement about x.

A (mathematical) **set** is a collection of sets which can be constructed using certain specific rules, which are known as the **ZFC axioms** (or the **Zermelo-Fraenkel Axioms** along with the **Axiom of Choice**). The ZFC axioms include (or imply) each of the following.

Equality Axiom: Two sets are equal if and only if they have the same elements.

Empty Set Axiom: There is set called the **empty set**, denoted by $\emptyset$, with no elements.

Pair Axiom: If A and B are sets then $\{A, B\} = \big\{x \,\big|\, x = A \text{ or } x = B\big\}$ is a set.

Union Axiom: If S is a set of sets then $\bigcup S = \bigcup\limits_{A \in S} A = \big\{x \,\big|\, x \in A \text{ for some } A \in S\big\}$ is a set.

Power Set Axiom: If A is a set then $\mathcal{P}(A) = \big\{X \big| X \subseteq A\big\}$ is a set, which we call the **power set** of A.

Axiom of Infinity: If we define the **natural numbers** to be the sets $0 = \emptyset$, $1 = \{0\}$, $2 = \{0, 1\}$, $3 = \{0, 1, 2\}$ and so on, then $\mathbf{N} = \{0, 1, 2, 3, \cdots\}$ is a set.

Separation Axiom: If A is a set and $F(x)$ is a statement about x, $\big\{x \in A \big| F(x)\big\}$ is a set.

Replacement Axiom: If A is a set and $F(x, y)$ is a statement about x and y with the property that for every set x there exists a unique set $y = f(x)$ for which $F(x, y)$ is true, then $\big\{f(x) \,\big|\, x \in A\big\} = \big\{y \,\big|\, \exists x \in A \ F(x, y)\big\}$ is a set.

Axiom of Choice: Given a nonempty set S of non-empty disjoint sets, there exists a set C which contains exacly one element from each of the sets in S.

1.3 Definition: For sets A and B, we use the following notation. We denote the **union** of A and B by $A \cup B$, the **intersection** of A and B by $A \cap B$, the set A **remove** B by $A \setminus B$ and the **product** of A and B by $A \times B$, that is

$$\begin{aligned}
A \cup B &= \bigcup \{A, B\} = \big\{x \big| x \in A \text{ or } x \in B\big\}, \\
A \cap B &= \big\{x \in A \cup B \big| x \in A \text{ and } x \in B\big\}, \\
A \setminus B &= \big\{x \in A \big| x \notin B\big\}, \text{ and} \\
A \times B &= \big\{(x, y) \in \mathcal{P}\big(\mathcal{P}(A \cup B)\big) \big| x \in A \text{ and } b \in B\big\}
\end{aligned}$$

where the **ordered pair** (x, y) is defined to be the set $(x, y) = \big\{\{x\}, \{x, y\}\big\}$. We say that A and B are **disjoint** when $A \cap B = \emptyset$. We also write $A^2 = A \times A$.

1.4 Theorem: *(Properties of Sets) Let $A, B, C \subseteq X$. Then*

(1) (Idempotence) $A \cup A = A$, $A \cap A = A$,

(2) (Identity) $A \cup \emptyset = A$, $A \cap \emptyset = \emptyset$, $A \cup X = X$, $A \cap X = A$,

(3) (Associativity) $(A \cup B) \cup C = A \cup (B \cup C)$ *and* $(A \cap B) \cap C = A \cap (B \cap C)$,

(4) (Commutativity) $A \cup B = B \cup A$ *and* $A \cap B = B \cap A$,

(5) (Distributivity) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ *and* $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$,

(6) (De Morgan's Laws) $X \setminus (A \cup B) = (X \setminus A) \cap (X \setminus B)$ *and* $X \setminus (A \cap B) = (X \setminus A) \cup (X \setminus B)$.

Proof: We shall provide some proofs later once we have listed some methods of proof.

1.5 Example: When A is a set, we have $\{A, A\} = \{A\}$ by the Equality Axiom and so $\{A\}$ is a set by the Pair Axiom. In particular, since $\emptyset$ is a set, so is $\{\emptyset\}$. Note that $\emptyset \neq \{\emptyset\}$, indeed the set $\emptyset$ has no elements but $\{\emptyset\}$ has one element. Since $\emptyset$ and $\{\emptyset\}$ are sets, so is the set $\big\{\emptyset, \{\emptyset\}\big\}$ by the Pair Axiom. Using the Pair Axiom and the Union Axiom we can then construct the set $\big\{\emptyset\,,\{\emptyset\}\big\} \cup \big\{\{\emptyset, \{\emptyset\}\}\big\} = \big\{\emptyset\,, \{\emptyset\}\,, \{\emptyset, \{\emptyset\}\}\big\}$. The first few natural numbers are given by $0 = \emptyset$, $1 = \{0\} = \{\emptyset\}$, $2 = \{0, 1\} = \big\{\emptyset, \{\emptyset\}\big\}$ and $3 = \{0, 1, 2\} = \big\{\emptyset\,, \{\emptyset\}\,, \{\emptyset, \{\emptyset\}\}\big\}$. Having constructed the natural number n as a set, the number $n+1$ is defined to be the set $n+1 = n \cup \{n\}$ (which is a set by the Pair and Union Axioms). We remark that although we only need to use the Pair Axion and the Union Axiom to construct any given natural number n, we need to use the Axiom of Infinity to conclude that the collection of all natural numbers is a set.

1.6 Example: Since the natural number $3 = \{0, 1, 2\}$ is a set, so is its power set

$$\mathcal{P}(3) = \mathcal{P}\big(\{0, 1, 2\}\big) = \Big\{\emptyset\,, \{0\}\,, \{1\}\,, \{2\}\,, \{1, 2\}\,, \{0, 2\}\,, \{0, 1\}\,, \{0, 1, 2\}\Big\}.$$

We remark that when a set A has n elements, its power set $\mathcal{P}(A)$ has 2^n elements. We also remark that when A is a finite set we do not need to use the Power Set Axiom to construct the set $\mathcal{P}(A)$ since we can construct the set $\mathcal{P}(A)$ using the Pair and Union Axioms.

1.7 Remark: It was mentioned earlier that essentially all mathematical objects are sets, including objects that we do not normally consider to be sets such as numbers and functions. We have seen that the natural numbers $0, 1, 2, \cdots$ are defined to be sets. In the following two definitions we indicate how functions and relations are defined to be sets.

1.8 Definition: When A and B are sets, a **function** from A to B is defined to be a set $F \subseteq A \times B$ with the property that for every $x \in A$ there exists a unique element $y \in B$ such that $(x, y) \in F$. We write $F : A \to B$ to indicate that F is a function from A to B, and we write $y = F(x)$ to indicate that $(x, y) \in F$. Thus a function is in fact defined to be equal to what we would normally consider to be its graph.

1.9 Definition: When A is a set, a **binary relation** on A is a subset $R \subseteq A^2$. When $x, y \in A$ we write xRy to indicate that $(x, y) \in R$.

1.10 Example: The operation $+$ on $\mathbf{N}$ is a function $+ : \mathbf{N}^2 \to \mathbf{N}$ and for $x, y \in \mathbf{N}$ we write $x + y$ to denote $+(x, y)$. The relation $<$ on $\mathbf{N}$ is a subset $< \subseteq \mathbf{N}^2$ and for $x, y \in \mathbf{N}$ we write $x < y$ to indicate that $(x, y) \in <$.

1.11 Remark: Using the axioms of set theory, it is possible to construct the set of integers $\mathbf{Z} = \{0, \pm 1, \pm 2, \cdots\}$, the sets of rational numbers $\mathbf{Q} = \left\{\frac{k}{n} \,\middle|\, k, n \in \mathbf{Z}, n > 0\right\}$ and the set of real numbers $\mathbf{R}$, along with their usual operations $+, -, \times, \div$ and their usual inequality relations $<, \leq, >, \geq$. This procedure is outlined in the Appendix 1.

1.12 Definition: Every mathematical statement can be expressed in a formal symbolic language called the language of **first-order set theory**, which we shall describe below. For the moment, we describe a very simple formal symbolic language that captures some of the features of mathematics. In the language of **propositional logic**, we use symbols from the **symbol set** $\{\neg\,,\,\wedge\,,\,\vee\,,\,\rightarrow\,,\,\leftrightarrow\,,\,(\,,\,)\}$ along some variable symbols which we denote by $P, Q, R, \cdots$. The symbols (and) are called parentheses or brackets and the other symbols in the symbol set represent English words as follows:

$\neg$	$\wedge$	$\vee$	$\rightarrow$	$\leftrightarrow$
not	and	or	implies	if and only if

The propositional variables $P, Q, R, \cdots$ are intended to represent certain unknown mathematical statements which are assumed to be either true or false, but never both. When X and Y are strings of symbols, we shall write $X \equiv Y$ when X and Y are **identical**.

A **formula**, in propositional logic, is a string of symbols which can be obtained using the following rules:

F1. Every propositional variable symbol is a formula.
F2. If F is a formula then so is the string $\neg F$.
F3. If F and G are formulas then so are the strings $(F \wedge G)$, $(F \vee G)$, $(F \rightarrow G)$ and $(F \leftrightarrow G)$.

A **derivation** for a formula F is a list of formulas $F_1, F_2, \cdots, F_n$ with $F \equiv F_m$ for some $m \leq n$ (usually $F \equiv F_n$) such that each formula F_k is obtained by applying one of the above 3 rules to previous formulas in the list.

We shall often omit the outermost pair of brackets from formulas, for example we might write the formula $(P \vee (Q \rightarrow R))$ as $P \vee (Q \rightarrow R)$.

1.13 Example: The string $F \equiv \neg\big(P \rightarrow \neg(Q \wedge P)\big)$ is a formula. One derivation for F is as follows.

$$P\,,\,Q\,,\,(Q \wedge P)\,,\,\neg(Q \wedge P)\,,\,\big(P \rightarrow \neg(Q \wedge P)\big)\,,\,F.$$

1.14 Definition: An **assignment** of truth-values is a function $\alpha : \{P, Q, R, \cdots\} \rightarrow \{0, 1\}$. When $\alpha(P) = 1$ we say that P is **true** under the assignment α and when $\alpha(P) = 0$ we say that P is **false** under α, and similarly for the variables $Q, R, \cdots$.

Given an assignment α, we define $\alpha(F)$, for any formula F, recursively as follows:

A1. $\alpha(P)$, $\alpha(Q)$ and $\alpha(R)$, and so on, are already known.
A2. If G is a formula then $\alpha(\neg G)$ is defined according to the table

G	$\neg G$
1	0
0	1

A3. If G and H are formulas then $\alpha(G \wedge H)$, $\alpha(G \vee H)$, $\alpha(G \rightarrow H)$ and $\alpha(G \leftrightarrow H)$ are defined according to the table

G	H	$(G \wedge H)$	$(G \vee H)$	$(G \rightarrow H)$	$(G \leftrightarrow H)$
1	1	1	1	1	1
1	0	0	1	0	0
0	1	0	1	1	0
0	0	0	0	1	1

When $\alpha(F) = 1$ we say that F is **true** under the assignment α, and when $\alpha(F) = 0$ we say that F is **false** under α.

1.15 Note: When $\alpha(G) = 0$ we have $\alpha(G \to H) = 1$. This agrees with the way in which we use the word "implies" in mathematics. For example, for every integer x, the statement "if $0 = 1$ then $x = 3$" is considered to be true.

1.16 Example: Let $F \equiv (P \wedge \neg(Q \to P)) \vee (R \leftrightarrow \neg Q)$, and let α be any assignment with $\alpha(P) = 1$, $\alpha(Q) = 0$ and $\alpha(R) = 1$. Determine whether F is true under α.

Solution: We make a derivation $F_1 \; F_2 \; \cdots \; F_n$ for F, and under each formula F_i we put the truth value $\alpha(F_i)$, which we find using the above definition.

P	Q	R	$Q \to P$	$\neg(Q \to P)$	$P \wedge \neg(Q \to P)$	$\neg Q$	$R \leftrightarrow \neg Q$	F
1	0	1	1	0	0	1	1	1

The table shows that $\alpha(F) = 1$.

1.17 Definition: An assignment on the propositional variables $P_1, \cdots, P_n$ is a function $\alpha : \{P_1, P_2, \cdots, P_n\} \to \{0, 1\}$ (there are 2^n such assignments). A **truth-table**, on the variables $P_1, P_2, \cdots P_n$, for the formula F, is a table in which

T1. The header row is a derivation $F_1 \; F_2 \; \cdots \; F_n \; \cdots \; F_l$ for F, where the formulas F_k use no propositional variables other than $P_1, \cdots, P_n$, with $F_k = P_k$ for $1 \le k \le n$.

T2. There are 2^n rows (not counting the header row): for each of the 2^n assignments α on $P_1, \cdots, P_n$, there is a row of the form $\alpha(F_1) \; \alpha(F_2) \; \cdots \; \alpha(F_l)$.

T3. The rows are ordered so that first n columns (headed by $P_1, \cdots, P_n$) list the binary numbers in decreasing order from $11 \cdots 1$ at the top down to $00 \cdots 0$ at the bottom.

1.18 Example: Make a truth-table on P, Q and R for the formula $F \equiv \neg((P \vee \neg Q) \to R)$.

Solution: We make a table, as in example 1.16, but with $2^3 = 8$ rows.

P	Q	R	$\neg Q$	$P \vee \neg Q$	$(P \vee \neg Q) \to R$	F
1	1	1	0	1	1	0
1	1	0	0	1	0	1
1	0	1	1	1	1	0
1	0	0	1	1	0	1
0	1	1	0	0	1	0
0	1	0	0	0	1	0
0	0	1	1	1	1	0
0	0	0	1	1	0	1

1.19 Definition: Let F and G be formulas and let $\mathcal{S}$ be a set of formulas.

(1) We say F is a **tautology**, and write $\models F$, when for all assignments α we have $\alpha(F) = 1$.
(2) We say that F is a **contradiction** when $\models \neg F$.
(3) We say F is **equivalent** to G, and write $F \cong G$, when for all assignments α, $\alpha(F) = \alpha(G)$.
(4) We say the argument "$\mathcal{S}$ therefore G" is **valid**, or we say that $\mathcal{S}$ **induces** G, or that G is a **consequence** of $\mathcal{S}$, and we write $\mathcal{S} \models G$, when for all assignments α, if $\alpha(F) = 1$ for every $F \in \mathcal{S}$ then $\alpha(G) = 1$. In the case that $\mathcal{S} = \{F_1, F_2, \cdots, F_n\}$ we often omit the set brackets and write $\mathcal{S} \models G$ as $F_1, F_2, \cdots, F_n \models G$. The formulas in $\mathcal{S}$ are called the **premises** of the argument and G is called the **conclusion**.

1.20 Theorem: *Let F, G and $F_1, \cdots, F_n$ be formulas. Then*
(1) $\models F \iff \emptyset \models F$,
(2) $F \models G \iff \models (F \to G)$,
(3) $F \cong G \iff \big(F \models G \text{ and } G \models F\big) \iff \models (F \leftrightarrow G)$, *and*
(4) $\{F_1, F_2, \cdots, F_n\} \models G \iff (\cdots((F_1 \wedge F_2) \wedge F_3) \wedge \cdots \wedge F_n) \models G$.

Proof: We shall provide some proofs once we have discussed proof methods.

1.21 Example: Let $F \equiv (P \leftrightarrow ((Q \wedge \neg R) \vee S)) \vee (P \to \neg S)$. Determine whether $\models F$.

Solution: We make a truth-table for F.

P	Q	R	S	$\neg R$	$Q \wedge \neg R$	$(Q \wedge \neg R) \vee S$	$P \leftrightarrow ((Q \wedge \neg R) \vee S)$	$\neg S$	$P \to \neg S$	F
1	1	1	1	0	0	1	1	0	0	1
1	1	1	0	0	0	0	0	1	1	1
1	1	0	1	1	1	1	1	0	0	1
1	1	0	0	1	1	1	1	1	1	1
1	0	1	1	0	0	1	1	0	0	1
1	0	1	0	0	0	0	0	1	1	1
1	0	0	1	1	0	1	1	0	0	1
1	0	0	0	1	0	0	0	1	1	1
0	1	1	1	0	0	1	0	0	1	1
0	1	1	0	0	0	0	1	1	1	1
0	1	0	1	1	1	1	0	0	1	1
0	1	0	0	1	1	1	0	1	1	1
0	0	1	1	0	0	1	0	0	1	1
0	0	1	0	0	0	0	1	1	1	1
0	0	0	1	1	0	1	0	0	1	1
0	0	0	0	1	0	0	1	1	1	1

Since all the entries in the F-column are equal to 1, we have $\models F$.

1.22 Example: Let $F \equiv (P \vee Q) \to R$ and $G \equiv (P \to R) \vee (Q \to R)$. Determine whether $F \cong G$.

Solution: We make a truth-table for F and G,

P	Q	R	$P \vee Q$	F	$P \to R$	$Q \to R$	G
1	1	1	1	1	1	1	1
1	1	0	1	0	0	0	0
1	0	1	1	1	1	1	1
1	0	0	1	0	0	1	1
0	1	1	1	1	1	1	1
0	1	0	1	0	1	0	1
0	0	1	0	1	1	1	1
0	0	0	0	1	1	1	1

The F-column is not the same as the G-column, for example on the 4^{th} row, F is false and G is true, and so $F \ncong G$.

1.23 Example: Let $F \equiv (P \vee \neg Q) \rightarrow R$, $G \equiv P \leftrightarrow (Q \wedge R)$, and $H \equiv (Q \rightarrow R)$, and let $K = \neg(\neg Q \wedge R)$. Determine whether $\{F, G, H\} \models K$.

Solution: In general, we have $\{F_1, \cdots, F_n\} \models K$ if and only if in a truth-table for the formulas $F_1, \cdots, F_n$ and K, for every row in which $F_1, \cdots, F_n$ are all true, we also have K true. We make a truth-table for F, G, H and K:

P	Q	R	$\neg Q$	$P \vee \neg Q$	F	$Q \wedge R$	G	H	$\neg Q \wedge R$	K
1	1	1	0	1	1	1	1	1	0	1
1	1	0	0	1	0	0	0	0	0	1
1	0	1	1	1	1	0	0	1	1	0
1	0	0	1	1	0	0	0	1	0	1
0	1	1	0	0	1	1	0	1	0	1
0	1	0	0	0	1	0	1	0	0	1
0	0	1	1	1	1	0	1	1	1	0
0	0	0	1	1	0	0	1	1	0	1

On row 7, F G and H are all true but K is false. This implies that $\{F, G, H\} \not\models K$.

1.24 Example: Determine whether $\{P \vee Q, \neg Q, P \rightarrow Q\} \models \neg P$.

Solution: We have

P	Q	$P \vee Q$	$\neg Q$	$P \rightarrow Q$	$\neg P$
1	1	1	0	1	0
1	0	1	1	0	0
0	1	1	0	1	1
0	0	0	1	1	1

Notice that there are no rows in which the premises are all true. In this situation, we can conclude that $\{P \vee Q, \neg Q, P \rightarrow Q\} \models \neg P$ (we don't even need to look at the last column of the table).

1.25 Example: Let F and G be formulas. Consider the following table.

F	G	$F \leftrightarrow G$	$\neg G$	$F \rightarrow \neg G$
1	1	1	0	0
1	0	0	1	1
0	1	0	0	1
0	0	1	1	1

Notice that on the first row (when F and G are both true) we have $F \leftrightarrow G$ true and $F \rightarrow \neg G$ false. This might seem to imply that $(F \leftrightarrow G) \not\models (F \rightarrow \neg G)$, but it does not! For example, if $F \equiv P$ and $G \equiv \neg P \wedge Q$, then we never have F and G both true, so the combination of truth-values shown in the first row of the above table never actually occurs. The above table is not actually a truth-table as defined in 1.17. Rather, it is a table of possible combinations of truth-values which may or may not actually occur.

1.26 Definition: Let A be a set. Recall that $A^2 = A \times A$. A **unary function** on A is a function $f : A \to A$. A **binary function** on A is a function $g : A^2 \to A$. Many binary functions g are used with **infix notation** which means that we write $g(x, y)$ as xgy. For example, $+$ is a binary function on $\mathbf{N}$ and we write $+(x, y)$ as $x+y$. A **unary relation** on A is a subset $P \subseteq A$, and we write $P(x)$ to indicate that $x \in P$. A **binary relation** on A is a subset $R \subseteq A^2$. When R is written with **prefix notation** we write $R(x, y)$ to indicate that $(x, y) \in R$, and when R is used with **infix notation**, we write xRy to indicate that $(x, y) \in R$. For example, $<$ is a binary relation on the set $\mathbf{N}$, which means that $< \subseteq \mathbf{N}^2$, and we write $x < y$ to indicate that $(x, y) \in <$.

1.27 Definition: We now describe a type of formal symbolic language, called a **first-order language**. Every first-order language uses symbols from the **common symbol set**

$$\{\neg, \wedge, \vee, \rightarrow, \leftrightarrow, =, \forall, \exists, (,), ,\}$$

together with some **variable** symbols (such as x, y, z, u, v and w), and possibly some **additional symbols** which might include some **constant** symbols (such as a, b, c, $\emptyset$, 0, 1, e, π), some **function** symbols (such as f, g, h, $\cap$, $\cup$, $+$, $-$, $\times$), and some **relation** symbols (such as P, Q, R, $\in$, $\subset$, $\subseteq$, $<$, $\leq$). The variable symbols are intended to represent elements in a certain set (or class) U, called the **universal set** (or the **universal class**), which is usually understood from the context. The symbol $\forall$ is read as "for all" or "for every" and the symbol $\exists$ is read as "for some" or "there exists".

A **term** in the first-order language is a string of symbols using only variable, constant and function symbols, along with parentheses and commas if necessary, which can be obtained using the following rules.

T1. Every variable symbol is a term and every constant symbol is a term.
T2. If t is a term and f is a unary function symbol then the string $f(t)$ is a term.
T3. If s and t are terms and g is a binary function symbol then the string $g(s, t)$ (or the string (sgt) in the case that g is used with infix notation) is a term. We sometimes omit the brackets from the string (sgt).

Here are some examples of terms (with some brackets omitted):

$$u\ ,\ u \cap v\ ,\ u \cap (v \cup \emptyset)\ ,\ x\ ,\ x+1\ ,\ x \times (y+1)\ ,\ f(x)\ ,\ g(x, y)\ ,\ g(x+1, f(y))$$

Each term represents an element in the universal set (or class) U.

A **formula** is a string of symbols which can be obtained using the following rules.

F1. If t is a term and P is a unary relation symbol, then the string $P(t)$ is a formula.
F2. If s and t are terms and R is a binary relation symbol then the string $R(s, t)$ (or the string sRt in the case that R is used with infix notation) is a formula.
F3. If F is a formula then so is the string $\neg F$.
F4. If F and G are formulas then so are the strings $(F \wedge G)$, $(F \vee G)$, $(F \rightarrow G)$ and $(F \leftrightarrow G)$.
F5. If F is a formula and x is a variable symbol, the strings $\forall x\ F$ and $\exists x\ G$ are formulas.

Here are some examples of formulas (with some brackets omitted):

$$u \subseteq v\ ,\ u \cap v = \emptyset\ ,\ (x=0 \vee x=1)\ ,\ \forall x\,(x \neq 0 \rightarrow \exists y\ x \times y = 1)$$

A formula is a precise way of expressing a mathematical statement about elements in U.

1.28 Definition: In the language of **first-order number theory**, in addition to symbols from the common symbol set $\{\neg, \wedge, \vee, \rightarrow, \leftrightarrow, =, \forall, \exists, (,), ,\}$ along with variable symbols such as x, y, z, we allow ourselves to use additional symbols from the **additional symbol set** $\{0, 1, +, \times, <\}$. (the symbols 0 and 1 are constant symbols, the symbols $+$ and $\times$ are binary function symbols used with infix notation, and the symbol $<$ is a binary relation symbol used with infix notation). Unless explicitly stated otherwise, we do not allow ourselves to use any other symbols (such as 2, $-$, $>$).

1.29 Example: Express each of the following statements about integers as formulas in the language of first-order number theory, taking the universal set to be $U = \mathbf{Z}$.
(a) x is a factor of y.
(b) $z = \min\{x, y\}$.
(c) x is prime.
(d) x is a power of 2.

Solution: The statement "x is a factor of y" can be expressed as $\exists z\ y = x \times z$.
The statement "$z = \min\{x, y\}$" can be expressed as $(x < y \rightarrow z = x) \wedge (\neg\, x < y \rightarrow z = y)$.
The statement "x is prime" can be expressed as

$$1 < x \wedge \forall x\, \forall y\, ((1 < y \wedge 1 < z) \rightarrow \neg\, x = y \times z).$$

The statement "x is a power of 2" is equivalent to the statement "x is positive and every factor of x which is greater than 1 is even" which can be expressed as

$$0 < x \wedge \forall y\, ((1 < y \wedge \exists z\ x = y \times z) \rightarrow \exists z\ y = z + z).$$

1.30 Example: Express each of the following statements about a function $f : \mathbf{R} \rightarrow \mathbf{R}$ as formulas in the language of first-order number theory, taking the universal set to be $U = \mathbf{R}$ and allowing the use of the additional unary function symbol f.
(a) f is nondecreasing.
(b) f is bijective.
(c) $\lim\limits_{x \to a} f(x) = b$.

Solution: The statement "f is nondecreasing" means "for all $x, y \in \mathbf{R}$, if $x \leq y$ then $f(x) \leq f(y)$" which can be expressed as $\forall x\ \forall y\ \big(\neg\, y < x \rightarrow \neg\, f(y) < f(x)\big)$.
The statement "f is bijective" means "for all $y \in \mathbf{R}$ there exists a unique $x \in \mathbf{R}$ such that $y = f(x)$" which can be expressed as $\forall y\ \exists x\big(y = f(x) \wedge \forall z(y = f(z) \rightarrow z = x)\big)$.
The statement "$|x - a| < \delta$" is equivalent to $-\delta < x - a$ and $x - a < \delta$" which can be expressed as the formula $(a < x + \delta \wedge x < a + \delta)$. The statement "$\lim\limits_{x \to a} f(x) = b$" means "for every $\epsilon > 0$ there exists $\delta > 0$ such that for all x, if $0 < |x - a| < \delta$ then $|f(x) - b| < \epsilon$", which can be expressed as the formula

$$\forall \epsilon\big(0 < \epsilon\ \rightarrow\ \exists \delta\ \big(0 < \delta \wedge \forall x\big((\neg\, x = a \wedge (a < x + \delta \wedge x < a + \delta))\ \rightarrow\ (b < f(x) + \epsilon \wedge f(x) < b + \epsilon)\big)\big)\big).$$

In the above formula, the symbols ϵ and δ are being used as variable symbols.

1.31 Definition: In the language of **first-order set theory**, in addition to the symbols from the common symbol set along with variable symbols, the only additional symbol that we allow ourselves to use (unless explicitly stated otherwise) is the symbol $\in$, which is a binary relation symbol used with infix notation. When we use the language of first-order set theory, unless indicated otherwise we shall take the universal class to be the class of all sets.

1.32 Remark: Every mathematical statement can, in principle, be expressed in the language of first-order set theory.

1.33 Example: Let u, v and w be sets. The mathematical statement $u \subseteq v$ can be expressed as the formula $\forall x(x \in u \rightarrow x \in v)$. The mathematical statement $w = \{u, v\}$ is equivalent (by the Equality Axiom) to the statement "for every set x, we have $x \in w$ if and only if $x \in \{u, v\}$" which can be expressed as the formula $\forall x\big(x \in w \leftrightarrow (x = u \vee x = v)\big)$. The mathematical statement $w = u \cup v$ can be expressed as $\forall x\,\big(x \in w \leftrightarrow (x \in u \vee x \in v)\big)$.

1.34 Example: Each of the ZFC axioms can be expressed as a formula in the language of first-order set theory. Here are a few of the axioms expressed as formulas.

Equality Axiom: $\forall u \,\forall v\, \big(u = v \leftrightarrow \forall x(x \in u \leftrightarrow x \in v)\big)$
Empty Set Axiom: $\exists u \,\forall x\, \neg x \in u$
Pair Axiom: $\forall u \,\forall v \,\exists w \,\forall x\big(x \in w \leftrightarrow (x = u \vee x = v)\big)$
Union Axiom: $\forall u \,\exists w \,\forall x\big(x \in w \leftrightarrow \exists v(v \in u \wedge x \in v)\big)$
Power Set Axiom: $\forall u \,\exists w \,\forall v\big(v \in w \leftrightarrow \forall x(x \in v \rightarrow x \in u)\big)$

1.35 Example: Express the mathematical statement $u = 2$ (that is, u is equal to the natural number 2) as a formula in first-order set theory.

Solution: The following statements are equivalent, and the final statement in the list is expressed in the form of a formula:

$$
\begin{aligned}
&u = 2\\
&u = \big\{\emptyset, \{\emptyset\}\big\}\\
&\forall x\,\big(x \in u \leftrightarrow x \in \big\{\emptyset, \{\emptyset\}\big\}\big)\\
&\forall x\,\big(x \in u \leftrightarrow \big(x = \emptyset \vee x = \{\emptyset\}\big)\big)\\
&\forall x\,\big(x \in u \leftrightarrow \big(\forall y\, \neg y \in x \;\vee\; \forall y\,(y \in x \leftrightarrow y = \emptyset)\big)\big)\\
&\forall x\,\big(x \in u \leftrightarrow \big(\forall y\, \neg y \in x \;\vee\; \forall y\,(y \in x \leftrightarrow \forall z\, \neg z \in y)\big)\big)
\end{aligned}
$$

1.36 Remark: As the above example illustrates, although every mathematical statement can, in principle, be expressed as a formula in the language of first-order set theory in practice even fairly simple mathematical statements (such as the statement $u = 2$) can become extremely long and complicated and difficult to read when expressed as formulas. For this reason, as we build mathematics from the foundations of set theory by introducing new concepts and proving new theorems, we continually add new symbols to the symbol set and allow additional notation to be used.

1.37 Example: When the universal set is U and $A \subseteq U$ (in other words A is a unary relation on U), the statement "$x \in A$" can be expressed as the formula $A(x)$. When $S(x)$ is a mathematical statement about the variable x which can be expressed as the formula F, the statement "for all $x \in A$, $S(x)$ is true" can be expressed as the formula $\forall x(A(x) \rightarrow F)$, and the statement "there exists $x \in A$ such that $S(x)$" can be expressed as $\exists x(A(x) \wedge F)$.

1.38 Definition: Every occurrence of each variable symbol, which does not immediately follow a quantifier symbol, in a formula H is either **free** or **bound**, as follows. When H is of the form $P(t)$ or $R(s, t)$, every occurrence of each variable symbol is free. When H is of one of the forms $\neg F$, $(F \wedge G)$, $(F \vee G)$, $(F \rightarrow G)$ or $(F \leftrightarrow G)$, every occurrence of each variable symbol in H is free or bound in accordance with whether it was free or bound in F or in G. When H is of one of the forms $\forall x\, F$ or $\exists x\, F$, each occurrence of any variable symbol y other than x is free or bound in H in accordance with whether it was free or bound in F, every bound occurrence of x in F remains bound in H (and it is bound in H by the same quantifier symbol which bound it in F), and every free occurrence of x in F becomes bound in H (and it is bound by the initial quantifier symbol).

1.39 Example: The mathematical statement "x is a factor of y", which is a statement about integers x and y, can be expressed as the formula $\exists z\; y = x \times z$ in first-order number theory. In this formula, the variables x and y are free and the variable z is bound by the quantifier. Note that the statement is a statement about x and y but not about z.

1.40 Definition: An **interpretation** for a first-order language is given by specifying a non-empty universal set U, and by specifying exactly which constants, functions and relations are represented by each of the constant, function and relation symbols.

1.41 Note: Until we have chosen an interpretation, a formula F in a first-order language is nothing more than a meaningless string of symbols. Once we have chosen an interpretation, the formula becomes a meaningful statement about its free variables (that is about the elements in the universal set which are represented by the variable symbols which occur freely in F). The truth or falsehood of F may still depend on the values in U assigned to the free variables in F.

1.42 Example: Let F be the formula $\forall y\; x\times y = y\times x$. If we use the interpretation $\mathbf{R}$ (that is is we specify that the universal set is $\mathbf{R}$ and that the function symbol $\times$ represents multiplication) then the the formula F becomes the meaningful statement "the real number x commutes with every real number y" (which is true, no matter what number is assigned to the variable x). If we use the interpretation $\mathbf{R}^3$ (in which the universal set is $\mathbf{R}^3$ and the function symbol $\times$ represents cross product) then the formula becomes the meaningful statement "the vector $x \in \mathbf{R}^3$ has the property that $x\times y = y\times x$ for every vector $y \in \mathbf{R}^3$" (which is true if and only if x is the zero vector).

1.43 Definition: Given an interpretation U, an **assignment** (of values to the variable symbols) in U is a function

$$\alpha : \{\text{variable symbols}\} \to U\,.$$

For a formula F, we write $\alpha(F) = 1$ when F is true under the assignment α in the interpretation U, and we write $\alpha(F) = 0$ when F is false under the assignment α in U.

1.44 Definition: Let F and G be formulas and let $\mathcal{S}$ be a set of formulas.

(1) We say that F is a **tautology**, and we write $\models F$, when for all interpretations U and for all assignments α in U we have $\alpha(F) = 1$.
(2) We say that F is a **contradiction** when $\models \neg F$.
(3) We say that F and G are **equivalent**, and we write $F \cong G$, when for all interpretations U and all assignments α in U we have $\alpha(F) = \alpha(G)$.
(4) We say that the argument "$\mathcal{S}$ therefore G" is **valid**, or we say that $\mathcal{S}$ **induces** G, or that G is a **consequence** of $\mathcal{S}$, and we write $\mathcal{S} \models G$, when for all interpretations U and all assignments α in U, if $\alpha(H) = 1$ for every formula $H \in \mathcal{S}$ then $\alpha(F) = 1$.

1.45 Example: For any term t, we have $\models\; t{=}t$. When x and y are variables and f is a unary function symbol, we have $\models\; \exists y\; y = f(x)$. When a and b are constant symbols, we have $\not\models\; \neg\, a{=}b$. For any terms s and t, we have $s{=}t \cong t{=}s$. Although it is true in the interpretation $\mathbf{Z}$ that the formula $x \leq y$ has the same meaning as the formula $\neg\, y < x$, we have $x \leq y \not\cong \neg\, y < x$. When x is a variable symbol, a is a constant symbol, t is a term and R is a binary relation symbol, we have $\big\{\forall x\; xRa\big\} \models tRa$.

1.46 Note: Unlike the situation in propositional logic, in first order logic there is no routine algorithmic procedure that one can apply to determine whether a given formula is a tautology, or whether two given formulas are equivalent, or whether a given argument is valid. Sometimes we can solve such problems by constructing a mathematical proof.