Problems and Comments on the Foundations: Logic and Proofs, Sets and Functions

Chapter 1 and Chapter 2

Section 1.1, Problems 1,3,17,25

Comments. You should pay special attention to implication $p \to q$, read *"If p then q"*. The formula $p \to q$ is considered as false only if p is true and q is false. The conjunction of $p \to q$ and $q \to p$ is called **equivalence**, $p \leftrightarrow q$ and read "p if and only if q". Notice that "*p if q" is actually* $q \to p$ and *"p only if q" is the same as* $p \to q$.

The expression $\neg q \to \neg p$ is called the **contrapositive** of $p \to q$ while $q \to p$ is called the **converse** of $p \to q$. The expression $\neg p \to \neg q$ is called the **inverse** of $p \to q$.

The implication $p \to q$ has the same meaning as its contrapositive $\neg q \to \neg p$.

Problem: (a) What is the converse of the inverse? (b) What is the inverse of the converse?

Answer: (a) $p \to q$ given. Inverse: $\neg p \to \neg q$, converse of inverse: $\neg q \to \neg p$ which is the contrapositive of $p \to q$. (b) $p \to q$ given. Converse: $q \to p$, inverse of converse: $\neg q \to \neg p$. Which is again the contrapositive.

Section 1.3, Problems: 16, 23, 24, 25

Comments. Let $P(p_1,\ldots,p_n)$ be a formula in propositional variables $p_1,\ldots,p_n$. Like $P(p_1,p_2) = p_1 \vee \neg p_2$. We say that $P(p_1,\ldots,p_n)$ and $Q(p_1,\ldots,p_n)$ are *equivalent* if they have for any truth valuation of the p_i the same truth tables. For this we write $P \equiv Q$. For example $p_1 \to p_2 \equiv \neg p_1 \vee p_2$.

A formula $P(p_1,\ldots,p_n)$ is called a **tautology** if $P(p_1,\ldots,p_n)$ has always truth value $\mathbf{T}$.

Proposition *$P \equiv Q$ if and only if $P \leftrightarrow Q$ is a tautology.*

Proof $P \leftrightarrow Q$ has truth value **F** iff $(P \to Q) \wedge (Q \to P)$ is false. That is, $(P \to Q)$ is false or $(Q \to P)$ is false. Now, $(P \to Q)$ is false iff (P is **T**) and (Q is **F**) or (Q is **T**) and (P is **F**). Thus $P \leftrightarrow Q$ has truth value **T** iff P and Q have he same truth values. Hence $P \leftrightarrow Q$ has always truth value **T** if P and Q have always the same truth values.

Example *$p \to q \equiv \neg q \to \neg p$. Hence an implication is equivalent to its contrapositive. Thus $q \to p$ is equivalent to $\neg p \to \neg q$, the converse of an implication is equivalent to the inverse.*

Section 1.4, Problems: 16, 43, 45, 46, 47

Comments. A formula like $x < y$ does not have a fixed truth value. If we interpret $<$ as the ordinary ordering in $\mathbb{R}$, then $2 < 3$ has truth value $\mathbf{T}$ while $3 < 2$ has truth value $\mathbf{F}$. We say that $x < y$ is a *predicate* or *propositional function* in variables x and y. A propositional function $P(x_{1,}\ldots x_n)$ takes on a certain truth value if the variables have been specified by elements $x_1 = a_1,\ldots,x_n = a_n$. The elements a_i are taken from a fixed domain D that is associated with $P(x_{1,}\ldots x_n)$.

First order formulas are defined inductively by the following three rules. We start with a certain set of **basic** predicates, like $P(a,y)$ etc. and say that

1. All basic predicates are formulas. They are called the **atomic** formulas.
2. If α and β are formulas then $(\alpha \wedge \beta), (\alpha \vee \beta), (\alpha \to \beta), (\alpha \leftrightarrow \beta), \neg a$ are formulas. That is, propositional combinations of formulas are formulas.
3. If α is a formula and $x$any variable then $\forall x\alpha$ and $\exists x\alpha$ are formulas. We assume that we have an infinite set of variables which designate elements from a fixed domain D.

Formulas are derived by finitely many applications of the rules 1-3. It is assumed that equality $E(x,y) \equiv x = y$ is amongst the basic predicates. In rule 3. we say that α is the scope of the quantifier $\forall x$ or $\exists x$, respectively. An occurrence of a variable x is called *free* in the formula α if it is not in the scope of any quantifier. Otherwise this occurrence is called *bound*. A formula without free variables is called a sentence.

Example *The first occurrence (read from left to right) of the variable x in the formula* $(\forall x \exists y\, P(x,y) \vee \neg Q(x,y,z))$ *is bound, the other occurrence is free. The only occurrence of z is free.*

Given a formula $\alpha(x_1, \dots x_n)$ and elements $a_1, \dots a_n$ from the domain D, then $\alpha(a_1, \dots, a_n)$ takes on a certain truth value. This can be proven proven inductively and is anyway quite obvious. For example, if we know already the truth value of $\alpha(a)$ for every $a \in D$, then the truth value of $\forall x\alpha(x)$ is $\mathbf{T}$ in case that the truth value of $\alpha(a)$ is $\mathbf{T}$ for every $a \in D$.

Example *Let* $D = \mathbb{N}$ *be the set of natural numbers. Find the truth value of* $\forall x \exists y\ (y > x)$. *This formula does not contain any free variable, it is a sentence. Thus it is either true or false. Let n be any natural number. then we need to know the truth value of* $\exists y\ (y > n)$. *But* $n + 1 > n$, *and for the chosen n we see that* $y = n + 1$ *makes* $\exists y\ (y > n)$ *true. Because this works for every n the sentence* $\forall x \exists y\ (y > x)$ *is true in* $\mathbb{N}$.

Section 1.5, Problems: 9, 19

Comments. It is important to observe the order of quantifiers. If$L(x,y)$ stands for "x *likes to buy* y" *where* x stands for persons and y stands for expensive cars then the sentence $\forall x \exists y\, L(x,y)$ says that every person likes to buy an expensive car. While $\exists y \forall x\, L(x,y)$ says something different. Namely, there is some expensive car everybody would like to buy. A more mathematical example is $\forall x \exists y\ (x < y)$ which is true for the set of real numbers but $\exists y \forall x\ x < y$ is false for real numbers.

Section 1.6, Problems: 3, 7

This section gives a short but complete description of first order logic. Table 1 and Table 2 list all the necessary rules. Many proofs are *indirect*: Instead of proving directly the implication $p \to q$, one shows $\neg q \to \neg p$. This is then called an *Indirect Proof*.

That $p \to q$ is true can also be shown by demonstrating that $p \wedge \neg q$ is false. This is called a *Proof by Contradiction*.

Proof $p \to q \equiv \mathbf{T}$ iff $p \wedge \neg q \equiv \mathbf{F}$ iff $\neg(p \wedge \neg q) \equiv \mathbf{T}$ iff $\neg p \vee q \equiv \mathbf{T}$ iff p$\to q \equiv \mathbf{T}$.

While reading this section you should keep in mind that you can learn how to prove mathematical facts essentially only by doing proofs and not so much by studying the concept of formal proofs and logic.

Section 2.1, Problems: 1, 2, 3, 21, 23

Definition 1 is somewhat circular. In order to understand "A *set* is an unordered *collection* of objects", you have to understand what a collection is. We will later treat sets as undefined objects $x, y, \ldots$ that are subject to certain axioms concerning the membership relation $"\in"$.

One of the axioms will be *Leibnitz's Extensionality Axiom*. It is treated in the book as a definition: *Two sets are equal if and only if they have the same elements.*

The practical aspect here is that if a set is given by listing its elements, there might be repetitions. The elements of the set $\mathbb{N}$ of natural numbers can be listed as: $1, 2, 3, \ldots$ or $1, 2, 2, 3, 3, 3, \ldots$ or $1, 2, 4, 3, 5, 6, 8, 7\ldots$. As long as any natural numbers appears in the list, we are given as underlying set of the list, the set of natural numbers. We will later learn in the axiomatic approach how to define the set of natural numbers precisely.

Given a set A, we can define its subsets, and its set of subsets, the powerset $\mathbf{P}(\mathbf{A})$. If A is finite and has n elements, then the powerset of A has 2^n elements.

If A and B are sets, then the set of all *ordered pairs* $(a, b), a \in A, b \in B$ is called the Cartesian product $A \times B$. If A and B are finite with n and m elements, respectively, then $A \times B$ has $n \cdot m$ many elements. We will later give a precise definition of what an ordered pair is.

We stipulate the existence of a set with no element. This is called the *empty* set. It must be unique and its notation is ∅.

Section 2.2, Problems: 20, 46, 47, 59

The complement of a set A can be defined only relative to a set that contains A. Such a set is often called the *universe* U. Then $\overline{A}$ is the set of elements that belong to U but not to A. Union and intersection can be defined for any collection of sets. For example, let $D_r = \{P \mid |P| < r\}$ be the set of all points in the plane whose distance to the origin O is less than r. Then

$$\bigcup\{D_r \mid r > 0\} = \mathbb{R}^2, \bigcap\{D_r \mid r > 0\} = O$$

Table 1 and Table 2 relate logical identities to set identities. The logical equivalence $p \vee q \equiv q \vee p$ corresponds to the set theoretical identity $A \cup B = B \cup A$. This becomes obvious if we let: $p = (x$ *belongs to* $A)$, $q = (x$ *belongs to* $B)$.

Section 2.3, Problems: 1, 10, 12, 14, 16, 40

It is important to note that for a function $f : A \to B$ one has that for **every** $a \in A$ one has **exactly** one $b \in B$ assigned as function value. In lower level courses, this is called the *vertical line* test for the graph of functions from real numbers to real numbers.

An incorrect but often used notation for a function is $y = f(x)$. A correct notation must specify the *domain* A and the co-domain B. A function f can be described by a formula, like $f(x) = x^2$, or by a recipe that allows us to calculate for any argument its value. Here is such a recipe: $f(0) = 0$, $f(1) = 1$ and $f(n) = f(n-1) + f(n-2)$. This gives $f(2) = f(1) + f(0) = 1 + 0 = 1$, $f(3) = f(2) + f(1) = 1 + 1 = 2, f(4) = f(3) + f(2) = 2 + 1 = 3.\ldots$

The point is that one can calculate for every natural number n the value $f(n)$.

A function $f : A \to B$ is *one-one* or *injective* if $x \neq y \to f(x) \neq f(y)$. This is logically equivalent to $f(x) = f(y) \to x = y$. Notice that $x = y \to f(x) = f(y)$ is always true, it doesn't say anything about f. The function $f : \mathbb{R} \to \mathbb{R}, x \mapsto x^2$ is not injective, because we have $f(-1) = f(1)$. However, $\mathbb{R}^+ \to \mathbb{R}^+, x \mapsto x^2$ is injective.

In lower level classes, a function is called injective if it passes the *horizontal line* test. A function from a set of real numbers to the set of real numbers is injective if every horizontal line intersects the graph of the function in at most one point. By the way, we can define the graph of any function as a subset of the Cartesian product of domain and co-domain:

$$f : A \to B, \text{has } graph(f) = \{(a, f(a)) \mid a \in A\}$$

One must distinguish between *image* and codomain. The image of a function is the set of all images:

$$f : A \to B \text{ has } im(f) = \{f(a) | a \in A\}$$

A function $f : A \to B$ is *onto* or *surjective* if $im(f) = B$. A function which is injective as well as surjective is called *bijective*. You should note the following:

$f : A \to B$*is* ***injective*** *if and only if for every* $b \in B$ *there is* ***at most one*** $a \in A$ *such that* $f(a) = b$

$f : A \to B$ *is* ***surjective*** *if there is for every* $b \in B$ **at least one** $a \in A$ *such that* $f(a) = b$

$f : A \to B$ *is* ***bijective*** *if for every* $b \in B$ *there is* **exactly one** $a \in A$ *such that*$f(a) = b$

Bijective functions are also called *one-to-one correspondences*. If $f : A \to B$ is bijecive, then one has a function which is called the *inverse* f^{-1} of f :

$$f^{-1} : B \to A, b \mapsto a \text{ where } f(a) = b$$

If $f : A \to B$ and $g : B \to C$ are functions, then the *composition* of g and f is the function $h : A \to C, a \mapsto g(f(a))$ and one writes $h = g \circ f$, read g *after* f.

For every set A one has the identity function $id_A : A \to A, a \mapsto a$. We now have:

$$f^{-1} \circ f = id_A,\ f \circ f^{-1} = id_B$$

Theorem *A function $f : A \to B$ is surjective if and only if there is a function $g : B \to A$ such that $f \circ g = id_B$. That is, f has a right inverse g. The function $f : A \to B$ is injective if and only if there is a function $h : B \to A$ such that $h \circ f = id_A$. 'That is, f has a left inverse. The function f is bijective if and only if it has a left as well a right inverse both of which have to be equal to the inverse f^{-1} of f.*

Proof *A right inverse $g : B \to A$ calculates the counter image a for the element $b : g(b) = a$ where $f(a) = b$ If f is surjective then there is for every $b \in B$ some $a \in A$ such that $f(a) = b$. Thus we may pick for every $b \in B$ any such a in order to define a right inverse g for f. This process requires the Axiom of Choice. We will discuss this axiom later.*

If f has a left inverse $h : B \to A$ then $f(a_1) = f(a_2)$ yields $h(f(a_1) = h(f(a_2))$, that is $a_1 = a_2$. If f is injective then in order to define a left inverse, we pick for every $b \in im(f)$ the unique $a \in A$ such that $f(a) = b$. For any $b \notin im(f)$ we may choose some $a \in A$ arbitrarily. This defines a left inverse $h : B \to A$ for f.

If f has a left inverse $h : B \to A$ as well as a right inverse then $g : B \to A$, then $g = h = f^{-1}$. Indeed,

If $f \circ g = id_B$, and $h \circ f = id_A$ then $h \circ (f \circ g) = h \circ id_B = h$ and $(h \circ f) \circ g = id_A \circ g = g$

However, composition of functions is associative, that is $h \circ (f \circ g) = (h \circ f) \circ g$.

Example *The function $f : \mathbb{N} \to \mathbb{N}, n \mapsto 2n$, is injective. For any left inverse $h : \mathbb{N} \to \mathbb{N}$, we must have $h(2n) = n$ but it doesn't matter how any such h is defined on the odd numbers in order to have that $h(f(n)) = h(2n) = n$.*

The function $f : \mathbb{N} \to \mathbb{N}, 0 \mapsto 0, 1 \mapsto 0, n \mapsto n-1$ for $n \geq 2$ is surjective. We can define a right inverse g by $g(n) = n+1$. But we could also have defined $g(0) = 0$ and $g(n) = n+1$ for $n \geq 1$, in order to have $f(g(n)) = n$.