2

First-Order Logic and Model Theory

2.1 Quantifiers

There are various kinds of logical inference that cannot be justified on the basis of the propositional calculus; for example:

1. Any friend of Martin is a friend of John.
   Peter is not John's friend.
   Hence, Peter is not Martin's friend.
2. All human beings are rational.
   Some animals are human beings.
   Hence, some animals are rational.
3. The successor of an even integer is odd.
   2 is an even integer.
   Hence, the successor of 2 is odd.

The correctness of these inferences rests not only upon the meanings of the truth-functional connectives, but also upon the meaning of such expressions as "any," "all," and "some," and other linguistic constructions.

In order to make the structure of complex sentences more transparent, it is convenient to introduce special notation to represent frequently occurring expressions. If $P(x)$ asserts that x has the property P, then $(\forall x)P(x)$ means that property P holds for all x or, in other words, that everything has the property P. On the other hand, $(\exists x)P(x)$ means that some x has the property P—that is, that there is at least one object having the property P. In $(\forall x)P(x)$, "$(\forall x)$" is called a *universal quantifier*; in $(\exists x)P(x)$, "$(\exists x)$" is called an *existential quantifier*. The study of quantifiers and related concepts is the principal subject of this chapter.

Examples

1′. Inference 1 above can be represented symbolically:

$$\begin{array}{l} (\forall x)(F(x,m) \Rightarrow F(x,j)) \\ \underline{\neg F(p,j) \qquad\qquad\qquad\quad} \\ \neg F(p,m) \end{array}$$

Here, $F(x, y)$ means that x is a friend of y, while m, j, and p denote Martin, John, and Peter, respectively. The horizontal line above "$\neg F(p, m)$" stands for "hence" or "therefore."

2′. Inference 2 becomes:

$$\begin{array}{l} (\forall x)(H(x) \Rightarrow R(x)) \\ \underline{(\exists x)(A(x) \wedge H(x))} \\ (\exists x)(A(x) \wedge R(x)) \end{array}$$

Here, H, R, and A designate the properties of being human, rational, and an animal, respectively.

3′. Inference 3 can be symbolized as follows:

$$\begin{array}{l} (\forall x)(I(x) \wedge E(x) \Rightarrow D(s(x))) \\ \underline{I(b) \wedge E(b) \qquad\qquad\qquad\quad} \\ D(s(b)) \end{array}$$

Here, I, E, and D designate respectively the properties of being an integer, even, and odd; $s(x)$ denotes the successor of x; and b denotes the integer 2.

Notice that the validity of these inferences does not depend upon the particular meanings of $F, m, j, p, H, R, A, I, E, D, s$, and b.

Just as statement forms were used to indicate logical structure dependent upon the logical connectives, so also the form of inferences involving quantifiers, such as inferences 1–3, can be represented abstractly, as in 1′–3′. For this purpose, we shall use commas, parentheses, the symbols ¬ and ⇒ of the propositional calculus, the universal quantifier symbol ∀, and the following groups of symbols:

Individual variables: $x_1, x_2, \ldots, x_n, \ldots$
Individual constants: $a_1, a_2, \ldots, a_n, \ldots$
Predicate letters: A_k^n (n and k are any positive integers)
Function letters: f_k^n (n and k are any positive integers)

The positive integer n that is a superscript of a predicate letter A_k^n or of a function letter f_k^n indicates the number of arguments, whereas the subscript k is just an indexing number to distinguish different predicate or function letters with the same number of arguments.*

In the preceding examples, x plays the role of an individual variable; m, j, p, and b play the role of individual constants; F is a binary predicate letter (i.e., a predicate letter with two arguments); H, R, A, I, E, and D are monadic predicate letters (i.e., predicate letters with one argument); and s is a function letter with one argument.

The function letters applied to the variables and individual constants generate the *terms*:

1. Variables and individual constants are terms.
2. If f_k^n is a function letter and $t_1, t_2, \ldots, t_n$ are terms, then $f_k^n(t_1, t_2, \ldots, t_n)$ is a term.
3. An expression is a term only if it can be shown to be a term on the basis of conditions 1 and 2.

Terms correspond to what in ordinary languages are nouns and noun phrases—for example, "two," "two plus three," and "two plus x."

The predicate letters applied to terms yield the *atomic formulas*; that is, if A_k^n is a predicate letter and $t_1, t_2, \ldots, t_n$ are terms, then $A_k^n(t_1, t_2, \ldots, t_n)$ is an atomic formula.

The *well-formed formulas* (wfs) of quantification theory are defined as follows:

1. Every atomic formula is a wf.
2. If $\mathscr{B}$ and $\mathscr{C}$ are wfs and y is a variable, then $(\neg\mathscr{B})$, $(\mathscr{B} \Rightarrow \mathscr{C})$, and $((\forall y)\mathscr{B})$ are wfs.
3. An expression is a wf only if it can be shown to be a wf on the basis of conditions 1 and 2.

* For example, in arithmetic both addition and multiplication take two arguments. So, we would use one function letter, say f_1^2, for addition, and a different function letter, say f_2^2 for multiplication.

In $((\forall y)\mathscr{B})$, "$\mathscr{B}$" is called the *scope* of the quantifier "$(\forall y)$." Notice that $\mathscr{B}$ need not contain the variable y. In that case, we understand $((\forall y)\mathscr{B})$ to mean the same thing as $\mathscr{B}$.

The expressions $(\mathscr{B} \wedge \mathscr{C})$, $(\mathscr{B} \vee \mathscr{C})$, and $(\mathscr{B} \Leftrightarrow \mathscr{C})$ are defined as in system L (see page 29). It was unnecessary for us to use the symbol $\exists$ as a primitive symbol because we can define existential quantification as follows:

$$((\exists x)\mathscr{B}) \quad \text{stands for} \quad (\neg((\forall x)(\neg\mathscr{B})))$$

This definition is faithful to the meaning of the quantifiers: $\mathscr{B}(x)$ is true for some x if and only if it is not the case that $\mathscr{B}(x)$ is false for all x.*

2.1.1 Parentheses

The same conventions as made in Chapter 1 (page 11) about the omission of parentheses are made here, with the additional convention that quantifiers $(\forall y)$ and $(\exists y)$ rank in strength between $\neg$, $\wedge$, $\vee$ and $\Rightarrow$, $\Leftrightarrow$. In other words, when we restore parentheses, negations, conjunctions, and disjunctions are handled first, then we take care of universal and existential quantifications, and then we deal with conditionals and biconditionals. As before, for connectives of the same kind, we proceed from left to right. For consecutive negations and quantifications, we proceed from right to left.

Examples

Parentheses are restored in the following steps.

1. $(\forall x_1)A_1^1(x_1) \Rightarrow A_1^2(x_2, x_1)$

   $((\forall x_1)A_1^1(x_1)) \Rightarrow A_1^2(x_2, x_1)$

   $(((\forall x_1)A_1^1(x_1)) \Rightarrow A_1^2(x_2, x_1))$

2. $(\forall x_1)A_1^1(x_1) \vee A_1^2(x_2, x_1)$

   $(\forall x_1)(A_1^1(x_1) \vee A_1^2(x_2, x_1))$

   $((\forall x_1)(A_1^1(x_1)) \vee A_1^2(x_2, x_1)))$

* We could have taken $\exists$ as primitive and then defined $((\forall x)\mathscr{B})$ as an abbreviation for $(\neg((\exists x)(\neg\mathscr{B})))$, since $\mathscr{B}(x)$ is true for all x if and only if it is not the case that $\mathscr{B}(x)$ is false for some x.

3. $(\forall x_1)\neg(\exists x_2)A_1^2(x_1,x_2)$

$(\forall x_1)\neg((\exists x_2)A_1^2(x_1,x_2))$

$(\forall x_1)(\neg((\exists x_2)A_1^2(x_1,x_2)))$

$((\forall x_1)(\neg((\exists x_2)A_1^2(x_1,x_2))))$

Exercises

2.1 Restore parentheses to the following.

a. $(\forall x_1)A_1^1(x_1)\wedge\neg A_1^1(x_2)$
b. $(\forall x_2)A_1^1(x_2)\Leftrightarrow A_1^1(x_2)$
c. $(\forall x_2)(\exists x_1)A_1^2(x_1,x_2)$
d. $(\forall x_1)(\forall x_3)(\forall x_4)A_1^1(x_1)\Rightarrow A_1^1(x_2)\wedge\neg A_1^1(x_1)$
e. $(\exists x_1)(\forall x_2)(\exists x_3)A_1^1(x_1)\vee(\exists x_2)\neg(\forall x_3)A_1^2(x_3,x_2)$
f. $(\forall x_2)\neg A_1^1(x_1)\Rightarrow A_1^3(x_1,x_1,x_2)\vee(\forall x_1)A_1^1(x_1)$
g. $\neg(\forall x_1)A_1^1(x_1)\Rightarrow(\exists x_2)A_1^1(x_2)\Rightarrow A_1^2(x_1,x_2)\wedge A_1^1(x_2)$

2.2 Eliminate parentheses from the following wfs as far as is possible.

a. $(((\forall x_1)(A_1^1(x_1)\Rightarrow A_1^1(x_1)))\vee((\exists x_1)A_1^1(x_1)))$
b. $((\neg((\exists x_2)(A_1^1(x_2)\vee A_1^1(a_1))))\Leftrightarrow A_1^1(x_2))$
c. $(((\forall x_1)(\neg(\neg A_1^1(a_3))))\Rightarrow(A_1^1(x_1)\Rightarrow A_1^1(x_2)))$

An occurrence of a variable x is said to be *bound* in a wf $\mathscr{B}$ if either it is the occurrence of x in a quantifier "$(\forall x)$" in $\mathscr{B}$ or it lies within the scope of a quantifier "$(\forall x)$" in $\mathscr{B}$. Otherwise, the occurrence is said to be *free* in $\mathscr{B}$.

Examples

1. $A_1^2(x_1,x_2)$
2. $A_1^2(x_1,x_2)\Rightarrow(\forall x_1)A_1^1(x_1)$
3. $(\forall x_1)(A_1^2(x_1,x_2)\Rightarrow(\forall x_1)A_1^1(x_1))$
4. $(\exists x_1)A_1^2(x_1,x_2)$

In Example 1, the single occurrence of x_1 is free. In Example 2, the occurrence of x_1 in $A_1^2(x_1,x_2)$ is free, but the second and third occurrences are bound. In Example 3, all occurrences of x_1 are bound, and in Example 4 both occurrences of x_1 are bound. (Remember that $(\exists x_1)A_1^2(x_1,x_2)$ is an abbreviation of $\neg(\forall x_1)\neg A_1^2(x_1,x_2)$.) In all four wfs, every occurrence of x_2 is free. Notice that, as in Example 2, a variable may have both free and bound occurrences in the same wf. Also observe that an occurrence of a variable may be bound in some wf $\mathscr{B}$ but free in a subformula of $\mathscr{B}$. For example, the first occurrence of x_1 is free in the wf of Example 2 but bound in the larger wf of Example 3.

A variable is said to be *free* (*bound*) in a wf $\mathscr{B}$ if it has a free (bound) occurrence in $\mathscr{B}$. Thus, a variable may be both free and bound in the same wf; for example, x_1 is free and bound in the wf of Example 2.

Exercises

2.3 Pick out the free and bound occurrences of variables in the following wfs.

a. $(\forall x_3)(((\forall x_1)A_1^2(x_1,x_2)) \Rightarrow A_1^2(x_3,a_1))$

b. $(\forall x_2)A_1^2(x_3,x_2) \Rightarrow (\forall x_3)A_1^2(x_3,x_2)$

c. $((\forall x_2)(\exists x_1)A_1^3(x_1,x_2,f_1^2(x_1,x_2))) \vee \neg(\forall x_1)A_1^2(x_2,f_1^1(x_1))$

2.4 Indicate the free and bound occurrences of all variables in the wfs of Exercises 2.1 and 2.2.

2.5 Indicate the free and bound variables in the wfs of Exercises 2.1–2.3.

We shall often indicate that some of the variables $x_{i_1},\ldots,x_{i_k}$ are free variables in a wf $\mathscr{B}$ by writing $\mathscr{B}$ as $\mathscr{B}(x_{i_1},\ldots,x_{i_k})$. This does not mean that $\mathscr{B}$ contains these variables as free variables, nor does it mean that $\mathscr{B}$ does not contain other free variables. This notation is convenient because we can then agree to write as $\mathscr{B}(t_1,\ldots,t_k)$ the result of substituting in $\mathscr{B}$ the terms $t_1,\ldots,t_k$ for all free occurrences (if any) of $x_{i_1},\ldots,x_{i_k}$, respectively.

If $\mathscr{B}$ is a wf and t is a term, then t is said to be *free for x_i in $\mathscr{B}$* if no free occurrence of x_i in $\mathscr{B}$ lies within the scope of any quantifier $(\forall x_j)$, where x_j is a variable in t. This concept of t being free for x_i in a wf $\mathscr{B}(x_i)$ will have certain technical applications later on. It means that, if t is substituted for all free occurrences (if any) of x_i in $\mathscr{B}(x_i)$, no occurrence of a variable in t becomes a bound occurrence in $\mathscr{B}(t)$.

Examples

1. The term x_2 is free for x_1 in $A_1^1(x_1)$, but x_2 is not free for x_1 in $(\forall x_2)A_1^1(x_1)$.
2. The term $f_1^2(x_1,x_3)$ is free for x_1 in $(\forall x_2)A_1^2(x_1,x_2) \Rightarrow A_1^1(x_1)$ but is not free for x_1 in $(\exists x_3)(\forall x_2)A_1^2(x_1,x_2) \Rightarrow A_1^1(x_1)$.

The following facts are obvious.

1. A term that contains no variables is free for any variable in any wf.
2. A term t is free for any variable in $\mathscr{B}$ if none of the variables of t is bound in $\mathscr{B}$.
3. x_i is free for x_i in any wf.
4. Any term is free for x_i in $\mathscr{B}$ if $\mathscr{B}$ contains no free occurrences of x_i.

Exercises

2.6 Is the term $f_1^2(x_1, x_2)$ free for x_1 in the following wfs?

a. $A_1^2(x_1, x_2) \Rightarrow (\forall x_2)A_1^1(x_2)$

b. $((\forall x_2)A_1^2(x_2, a_1)) \vee (\exists x_2)A_1^2(x_1, x_2)$

c. $(\forall x_1)A_1^2(x_1, x_2)$

d. $(\forall x_2)A_1^2(x_1, x_2)$

e. $(\forall x_2)A_1^1(x_2) \Rightarrow A_1^2(x_1, x_2)$

2.7 Justify facts 1–4 above.

When English sentences are translated into formulas, certain general guidelines will be useful:

1. A sentence of the form "All *A*s are *B*s" becomes $(\forall x)(A(x) \Rightarrow B(x))$. For example, *Every mathematician loves music* is translated as $(\forall x)(M(x) \Rightarrow L(x))$, where $M(x)$ means *x is a mathematician* and $L(x)$ means *x loves music*.
2. A sentence of the form "Some *A*s are *B*s" becomes $(\exists x)(A(x) \wedge B(x))$. For example, *Some New Yorkers are friendly* becomes $(\exists x)(N(x) \wedge F(x))$, where $N(x)$ means *x is a New Yorker* and $F(x)$ means *x is friendly*.
3. A sentence of the form "No *A*s are *B*s" becomes $(\forall x)(A(x) \Rightarrow \neg B(x))$.* For example, *No philosopher understands politics* becomes $(\forall x)(P(x) \Rightarrow \neg U(x))$, where $P(x)$ means *x is a philosopher* and $U(x)$ means *x understands politics*.

Let us consider a more complicated example: *Some people respect everyone*. This can be translated as $(\exists x)(P(x) \wedge (\forall y)(P(y) \Rightarrow R(x, y)))$, where $P(x)$ means *x is a person* and $R(x, y)$ means *x respects y*.

Notice that, in informal discussions, to make formulas easier to read we may use lower-case letters u, v, x, y, z instead of our official notation x_i for individual variables, capital letters $A, B, C,\ldots$ instead of our official notation A_k^n for predicate letters, lower-case letters $f, g, h,\ldots$ instead of our official notation f_k^n for function letters, and lower-case letters $a, b, c,\ldots$ instead of our official notation a_i for individual constants.

Exercises

2.8 Translate the following sentences into wfs.

a. Anyone who is persistent can learn logic.

b. No politician is honest.

* As we shall see later, this is equivalent to $\neg(\exists x)(A(x) \wedge B(x))$.

c. Not all birds can fly.
d. All birds cannot fly.
e. x is transcendental only if it is irrational.
f. Seniors date only juniors.
g. If anyone can solve the problem, Hilary can.
h. Nobody loves a loser.
i. Nobody in the statistics class is smarter than everyone in the logic class.
j. John hates all people who do not hate themselves.
k. Everyone loves somebody and no one loves everybody, or somebody loves everybody and someone loves nobody.
l. You can fool some of the people all of the time, and you can fool all the people some of the time, but you can't fool all the people all the time.
m. Any sets that have the same members are equal.
n. Anyone who knows Julia loves her.
o. There is no set belonging to precisely those sets that do not belong to themselves.
p. There is no barber who shaves precisely those men who do not shave themselves.

2.9 Translate the following into everyday English. Note that everyday English does not use variables.

a. $(\forall x)(M(x) \wedge (\forall y)\neg W(x, y) \Rightarrow U(x))$, where $M(x)$ means *x is a man*, $W(x, y)$ means *x is married to y*, and $U(x)$ means *x is unhappy*.
b. $(\forall x)(V(x) \wedge P(x) \Rightarrow A(x, b))$, where $V(x)$ means *x is an even integer*, $P(x)$ means *x is a prime integer*, $A(x, y)$ means $x = y$, and b denotes 2.
c. $\neg(\exists y)(I(y) \wedge (\forall x)(I(x) \Rightarrow L(x, y)))$, where $I(y)$ means *y is an integer* and $L(x, y)$ means $x \leq y$.
d. In the following wfs, $A_1^1(x)$ means *x is a person* and $A_1^2(x,y)$ means *x hates y*.
   i. $(\exists x)(A_1^1(x) \wedge (\forall y)(A_1^1(y) \Rightarrow A_1^2(x,y)))$
   ii. $(\forall x)(A_1^1(x) \Rightarrow (\forall y)(A_1^1(y) \Rightarrow A_1^2(x,y)))$
   iii. $(\exists x)(A_1^1(x) \wedge (\forall y)(A_1^1(y) \Rightarrow (A_1^2(x,y) \Leftrightarrow A_1^2(y,y))))$
e. $(\forall x)(H(x) \Rightarrow (\exists y)(\exists z)(\neg A(y, z) \wedge (\forall u)(P(u, x) \Leftrightarrow (A(u, y) \vee A(u, z)))))$, where $H(x)$ means *x is a person*, $A(u, v)$ means "$u = v$," and $P(u, x)$ means *u is a parent of x*.

2.2 First-Order Languages and Their Interpretations: Satisfiability and Truth: Models

Well-formed formulas have meaning only when an interpretation is given for the symbols. We usually are interested in interpreting wfs whose symbols come from a specific language. For that reason, we shall define the notion of a *first-order language*.*

Definition

A first-order language $\mathscr{L}$ contains the following symbols.

a. The propositional connectives ¬ and ⇒, and the universal quantifier symbol ∀.
b. Punctuation marks: the left parenthesis "(", the right parenthesis ")", and the comma ",".†
c. Denumerably many individual variables $x_1, x_2, \ldots$.
d. A finite or denumerable, possibly empty, set of function letters.
e. A finite or denumerable, possibly empty, set of individual constants.
f. A nonempty set of predicate letters.

By a *term of* $\mathscr{L}$ we mean a term whose symbols are symbols of $\mathscr{L}$.

By a *wf of* $\mathscr{L}$ we mean a wf whose symbols are symbols of $\mathscr{L}$.

Thus, in a language $\mathscr{L}$, some or all of the function letters and individual constants may be absent, and some (but not all) of the predicate letters may be absent.‡ The individual constants, function letters, and predicate letters of a language $\mathscr{L}$ are called the *nonlogical constants* of $\mathscr{L}$. Languages are designed in accordance with the subject matter we wish to study. A language for arithmetic might contain function letters for addition and multiplication and a

* The adjective "first-order" is used to distinguish the languages we shall study here from those in which there are predicates having other predicates or functions as arguments or in which predicate quantifiers or function quantifiers are permitted, or both. Most mathematical theories can be formalized within first-order languages, although there may be a loss of some of the intuitive content of those theories. Second-order languages are discussed in the appendix on second-order logic. Examples of higher-order languages are studied also in Gödel (1931), Tarski (1933), Church (1940), Leivant (1994), and van Bentham and Doets (1983). Differences between first-order and higher-order theories are examined in Corcoran (1980) and Shapiro (1991).

† The punctuation marks are not strictly necessary; they can be avoided by redefining the notions of term and wf. However, their use makes it easier to read and comprehend formulas.

‡ If there were no predicate letters, there would be no wfs.

predicate letter for equality, whereas a language for geometry is likely to have predicate letters for equality and the notions of *point* and *line*, but no function letters at all.

Definition

Let $\mathscr{L}$ be a first-order language. An *interpretation* M of $\mathscr{L}$ consists of the following ingredients.

a. A nonempty set D, called the *domain* of the interpretation.
b. For each predicate letter A_j^n of $\mathscr{L}$, an assignment of an n-place relation $(A_j^n)^M$ in D.
c. For each function letter f_j^n of $\mathscr{L}$, an assignment of an n-place operation $(f_j^n)^M$ in D (that is, a function from D^n into D).
d. For each individual constant a_i of $\mathscr{L}$, an assignment of some fixed element $(a_i)^M$ of D.

Given such an interpretation, variables are thought of as ranging over the set D, and $\neg$, $\Rightarrow$ and quantifiers are given their usual meaning. Remember that an n-place relation in D can be thought of as a subset of D^n, the set of all n-tuples of elements of D. For example, if D is the set of human beings, then the relation "father of" can be identified with the set of all ordered pairs $\langle x, y\rangle$ such that x is the father of y.

For a given interpretation of a language $\mathscr{L}$, a wf of $\mathscr{L}$ without free variables (called a *closed wf* or a *sentence*) represents a proposition that is true or false, whereas a wf with free variables may be satisfied (i.e., true) for some values in the domain and not satisfied (i.e., false) for the others.

Examples

Consider the following wfs:

1. $A_1^2(x_1, x_2)$
2. $(\forall x_2)A_1^2(x_1, x_2)$
3. $(\exists x_1)(\forall x_2)A_1^2(x_1, x_2)$

Let us take as domain the set of all positive integers and interpret $A_1^2(y, z)$ as $y \leq z$. Then wf 1 represents the expression "$x_1 \leq x_2$", which is satisfied by all the ordered pairs $\langle a, b\rangle$ of positive integers such that $a \leq b$. Wf 2 represents the expression "For all positive integers x_2, $x_1 \leq x_2$",* which is satisfied only by the integer 1. Wf 3 is a true sentence asserting that there is a smallest positive integer. If we were to take as domain the set of all integers, then wf 3 would be false.

* In ordinary English, one would say "x_1 is less than or equal to all positive integers."

Exercises

2.10 For the following wfs and for the given interpretations, indicate for what values the wfs are satisfied (if they contain free variables) or whether they are true or false (if they are closed wfs).

i. $A_1^2(f_1^2(x_1, x_2), a_1)$

ii. $A_1^2(x_1, x_2) \Rightarrow A_1^2(x_2, x_1)$

iii. $(\forall x_1)(\forall x_2)(\forall x_3)(A_1^2(x_1, x_2) \wedge A_1^2(x_2, x_3) \Rightarrow A_1^2(x_1, x_3))$

a. The domain is the set of positive integers, $A_1^2(y, z)$ is $y \geq z$, $f_1^2(y, z)$ is $y \cdot z$, and a_1 is 2.

b. The domain is the set of integers, $A_1^2(y, z)$ is $y = z$, $f_1^2(y, z)$ is $y + z$, and a_1 is 0.

c. The domain is the set of all sets of integers, $A_1^2(y, z)$ is $y \subseteq z$, $f_1^2(y, z)$ is $y \cap z$, and a_1 is the empty set $\varnothing$.

2.11 Describe in everyday English the assertions determined by the following wfs and interpretations.

a. $(\forall x)(\forall y)(A_1^2(x, y) \Rightarrow (\exists z)(A_1^1(z) \wedge A_1^2(x, z) \wedge A_1^2(z, y)))$, where the domain D is the set of real numbers, $A_1^2(x, y)$ means $x < y$, and $A_1^1(z)$ means z is a *rational number*.

b. $(\forall x)(A_1^1(x) \Rightarrow (\exists y)(A_2^1(y) \wedge A_1^2(y, x)))$, where D is the set of all days and people, $A_1^1(x)$ means *x is a day*, $A_2^1(y)$ means *y is a sucker*, and $A_1^2(y, x)$ means *y is born on day x*.

c. $(\forall x)(\forall y)(A_1^1(x) \wedge A_1^1(y) \Rightarrow A_2^1(f_1^2(x, y)))$, where D is the set of integers, $A_1^1(x)$ means *x is odd*, $A_2^1(x)$ means *x is even*, and $f_1^2(x, y)$ denotes $x + y$.

d. For the following wfs, D is the set of all people and $A_1^2(u,v)$ means *u loves v*.

i. $(\exists x)(\forall y)(A_1^2(x,y)$

ii. $(\forall y)(\exists x)A_1^2(x,y)$

iii. $(\exists x)(\forall y)((\forall z)(A_1^2(y,z)) \Rightarrow A_1^2(x,y))$

iv. $(\exists x)(\forall y)\neg A_1^2(x,y)$

e. $(\forall x)\ (\forall u)\ (\forall v)\ (\forall w)(\mathrm{E}(f(u, u), x) \wedge \mathrm{E}(f(v, v), x) \wedge \mathrm{E}(f(w, w), x) \Rightarrow \mathrm{E}(u, v) \vee \mathrm{E}(u, w) \vee \mathrm{E}(v, w))$, where D is the set of real numbers, $\mathrm{E}(x, y)$ means $x = y$, and f denotes the multiplication operation.

f. $A_1^1(x_1) \wedge (\exists x_3)(A_2^2(x_1, x_3) \wedge A_2^2(x_3, x_2))$ where D is the set of people, $A_1^1(u)$ means *u is a woman* and $A_2^2(u,v)$ means *u is a parent of v*.

g. $(\forall x_1)(\forall x_2)(A_1^1(x_1) \wedge A_1^1(x_2) \Rightarrow A_2^1(f_2^1(x_1, x_2)))$ where D is the set of real numbers, $A_1^1(u)$ means *u is negative*, $A_2^1(u)$ means *u is positive*, and $f_1^2(u,v)$ is the product of u and v.

The concepts of satisfiability and truth are intuitively clear, but, following Tarski (1936), we also can provide a rigorous definition. Such a definition is necessary for carrying out precise proofs of many metamathematical results.

Satisfiability will be the fundamental notion, on the basis of which the notion of truth will be defined. Moreover, instead of talking about the n-tuples of objects that satisfy a wf that has n free variables, it is much more convenient from a technical standpoint to deal uniformly with denumerable sequences. What we have in mind is that a denumerable sequence $s = (s_1, s_2, s_3, \ldots)$ is to be thought of as satisfying a wf $\mathscr{B}$ that has $x_{j_1}, x_{j_2}, \ldots, x_{j_n}$ as free variables (where $j_1 < j_2 < \cdots < j_n$) if the n-tuple $\langle s_{j_1}, s_{j_2}, \ldots, s_{j_n} \rangle$ satisfies $\mathscr{B}$ in the usual sense. For example, a denumerable sequence $(s_1, s_2, s_3, \ldots)$ of objects in the domain of an interpretation M will turn out to satisfy the wf $A_1^2(x_2, x_5)$ if and only if the ordered pair, $\langle s_2, s_5 \rangle$ is in the relation $(A_1^2)^{\mathrm{M}}$ assigned to the predicate letter A_1^2 by the interpretation M.

Let M be an interpretation of a language $\mathscr{L}$ and let D be the domain of M. Let Σ be the set of all denumerable sequences of elements of D. For a wf $\mathscr{B}$ of $\mathscr{L}$, we shall define what it means for a sequence $s = (s_1, s_2, \ldots)$ in Σ to satisfy $\mathscr{B}$ in M. As a preliminary step, for a given s in Σ we shall define a function s^* that assigns to each term t of $\mathscr{L}$ an element $s^*(t)$ in D.

1. If t is a variable x_j, let $s^*(t)$ be s_j.
2. If t is an individual constant a_j, then $s^*(t)$ is the interpretation $(a_j)^{\mathrm{M}}$ of this constant.
3. If f_k^n is a function letter, $(f_k^n)^{\mathrm{M}}$ is the corresponding operation in D, and $t_1, \ldots, t_n$ are terms, then

$$s^*(f_k^n(t_1, \ldots, t_n)) = (f_k^n)^{\mathrm{M}}(s^*(t_1), \ldots, s^*(t_n))$$

Intuitively, $s^*(t)$ is the element of D obtained by substituting, for each j, a name of s_j for all occurrences of x_j in t and then performing the operations of the interpretation corresponding to the function letters of t. For instance, if t is $f_2^2(x_3, f_1^2(x_1, a_1))$ and if the interpretation has the set of integers as its domain, f_2^2 and f_1^2 are interpreted as ordinary multiplication and addition, respectively, and a_1 is interpreted as 2, then, for any sequence $s = (s_1, s_2, \ldots)$ of integers, $s^*(t)$ is the integer $s_3 \cdot (s_1 + 2)$. This is really nothing more than the ordinary way of reading mathematical expressions.

Now we proceed to the definition of satisfaction, which will be an inductive definition.

1. If $\mathscr{B}$ is an atomic wf $A_k^n(t_1, \ldots, t_n)$ and $(A_k^n)^{\mathrm{M}}$ is the corresponding n-place relation of the interpretation, then a sequence $s = (s_1, s_2, \ldots)$ satisfies $\mathscr{B}$ if and only if $(A_k^n)^{\mathrm{M}}(s^*(t_1), \ldots, s^*(t_n))$—that is, if the n-tuple $\langle s^*(t_1), \ldots, s^*(t_n) \rangle$ is in the relation $(A_k^n)^{\mathrm{M}}$.*

* For example, if the domain of the interpretation is the set of real numbers, the interpretation of A_1^2 is the relation $\leq$, and the interpretation of f_1^1 is the function e^x, then a sequence $s = (s_1, s_2, \ldots)$ of real numbers satisfies $A_1^2(f_1^1(x_2), x_5)$ if and only if $e^{s_2} \leq s_5$. If the domain is the set of integers, the interpretation of $A_1^4(x, y, u, v)$ is $x \cdot v = u \cdot y$, and the interpretation of a_1 is 3, then a sequence $s = (s_1, s_2, \ldots)$ of integers satisfies $A_1^4(x_3, a_1, x_1, x_3)$ if and only if $(s_3)^2 = 3s_1$.

2. s satisfies $\neg\mathscr{B}$ if and only if s does not satisfy $\mathscr{B}$.
3. s satisfies $\mathscr{B} \Rightarrow \mathscr{C}$ if and only if s does not satisfy $\mathscr{B}$ or s satisfies $\mathscr{C}$.
4. s satisfies $(\forall x_i)\mathscr{B}$ if and only if every sequence that differs from s in at most the ith component satisfies $\mathscr{B}$.*

Intuitively, a sequence $s = (s_1, s_2, \ldots)$ satisfies a wf $\mathscr{B}$ if and only if, when, for each i, we replace all free occurrences of x_i (if any) in $\mathscr{B}$ by a symbol representing s_i, the resulting proposition is true under the given interpretation.

Now we can define the notions of truth and falsity of wfs for a given interpretation.

Definitions

1. A wf $\mathscr{B}$ is *true for the interpretation* M (written $\vDash_M \mathscr{B}$) if and only if every sequence in Σ satisfies $\mathscr{B}$.
2. $\mathscr{B}$ is said to be *false for* M if and only if no sequence in Σ satisfies $\mathscr{B}$.
3. An interpretation M is said to be a *model* for a set Γ of wfs if and only if every wf in Γ is true for M.

The plausibility of our definition of truth will be strengthened by the fact that we can derive all of the following expected properties I–XI of the notions of truth, falsity, and satisfaction. Proofs that are not explicitly given are left to the reader (or may be found in the answer to Exercise 2.12). ***Most of the results are also obvious if one wishes to use only the ordinary intuitive understanding of the notions of truth, falsity, and satisfaction.***

I. a. $\mathscr{B}$ is false for an interpretation M if and only if $\neg\mathscr{B}$ is true for M.
   b. $\mathscr{B}$ is true for M if and only if $\neg\mathscr{B}$ is false for M.

II. It is not the case that both $\vDash_M \mathscr{B}$ and $\vDash_M \neg\mathscr{B}$; that is, no wf can be both true and false for M.

III. If $\vDash_M \mathscr{B}$ and $\vDash_M \mathscr{B} \Rightarrow \mathscr{C}$, then $\vDash_M \mathscr{C}$.

IV. $\mathscr{B} \Rightarrow \mathscr{C}$ is false for M if and only if $\vDash_M \mathscr{B}$ and $\vDash_M \neg\mathscr{C}$.

V. †Consider an interpretation M with domain D.
   a. A sequence s satisfies $\mathscr{B} \wedge \mathscr{C}$ if and only if s satisfies $\mathscr{B}$ and s satisfies $\mathscr{C}$.
   b. s satisfies $\mathscr{B} \vee \mathscr{C}$ if and only if s satisfies $\mathscr{B}$ or s satisfies $\mathscr{C}$.
   c. s satisfies $\mathscr{B} \Leftrightarrow \mathscr{C}$ if and only if s satisfies both $\mathscr{B}$ and $\mathscr{C}$ or s satisfies neither $\mathscr{B}$ nor $\mathscr{C}$.

* In other words, a sequence $s = (s_1, s_2, \ldots, s_i, \ldots)$ satisfies $(\forall x_i)\mathscr{B}$ if and only if, for every element c of the domain, the sequence $(s_1, s_2, \ldots, c, \ldots)$ satisfies $\mathscr{B}$. Here, $(s_1, s_2, \ldots, c, \ldots)$ denotes the sequence obtained from $(s_1, s_2, \ldots, s_i, \ldots)$ by replacing the ith component s_i by c. Note also that, if s satisfies $(\forall x_i)\mathscr{B}$, then, as a special case, s satisfies $\mathscr{B}$.

† Remember that $\mathscr{B} \wedge \mathscr{C}$, $\mathscr{B} \vee \mathscr{C}$, $\mathscr{B} \Leftrightarrow \mathscr{C}$ and $(\exists x_i)\mathscr{B}$ are abbreviations for $\neg(\mathscr{B} \Rightarrow \neg\mathscr{C})$, $\neg\mathscr{B} \Rightarrow \mathscr{C}$, $(\mathscr{B} \Rightarrow \mathscr{C}) \wedge (\mathscr{C} \Rightarrow \mathscr{B})$ and $\neg(\forall x_i)\neg\mathscr{B}$, respectively.

d. s satisfies $(\exists x_i)\mathscr{B}$ if and only if there is a sequence s' that differs from s in at most the ith component such that s' satisfies $\mathscr{B}$. (In other words $s = (s_1, s_2, \ldots, s_i, \ldots)$ satisfies $(\exists x_i)\mathscr{B}$ if and only if there is an element c in the domain D such that the sequence $(s_1, s_2, \ldots, c, \ldots)$ satisfies $\mathscr{B}$.)

VI. $\vDash_M \mathscr{B}$ if and only if $\vDash_M (\forall x_i)\mathscr{B}$.

We can extend this result in the following way. By the *closure** of $\mathscr{B}$ we mean the closed wf obtained from $\mathscr{B}$ by prefixing in universal quantifiers those variables, in order of descending subscripts, that are free in $\mathscr{B}$. If $\mathscr{B}$ has no free variables, the closure of $\mathscr{B}$ is defined to be $\mathscr{B}$ itself. For example, if $\mathscr{B}$ is $A_1^2(x_2, x_5) \Rightarrow \neg(\exists x_2)A_1^3(x_1, x_2, x_3)$, its closure is $(\forall x_5)(\forall x_3)(\forall x_2)(\forall x_1)\mathscr{B}$. It follows from (VI) that a wf $\mathscr{B}$ is true if and only if its closure is true.

VII. Every instance of a tautology is true for any interpretation. (An *instance* of a statement form is a wf obtained from the statement form by substituting wfs for all statement letters, with all occurrences of the same statement letter being replaced by the same wf. Thus, an instance of $A_1 \Rightarrow \neg A_2 \vee A_1$ is $A_1^1(x_2) \Rightarrow (\neg(\forall x_1)A_1^1(x_1)) \vee A_1^1(x_2)$.)

To prove (VII), show that all instances of the axioms of the system L are true and then use (III) and Proposition 1.14.

VIII. If the free variables (if any) of a wf $\mathscr{B}$ occur in the list $x_{i_1}, \ldots, x_{i_k}$ and if the sequences s and s' have the same components in the i_1th, …, i_kth places, then s satisfies $\mathscr{B}$ if and only if s' satisfies $\mathscr{B}$ [*Hint*: Use induction on the number of connectives and quantifiers in $\mathscr{B}$. First prove this lemma: If the variables in a term t occur in the list $x_{i_1}, \ldots, x_{i_k}$, and if s and s' have the same components in the i_1th, …, i_kth places, then $s^*(t) = (s')^*(t)$. In particular, if t contains no variables at all, $s^*(t) = (s')^*(t)$ for any sequences s and s'.]

Although, by (VIII), a particular wf $\mathscr{B}$ with k free variables is essentially satisfied or not only by k-tuples, rather than by denumerable sequences, it is more convenient for a general treatment of satisfaction to deal with infinite rather than finite sequences. If we were to define satisfaction using finite sequences, conditions 3 and 4 of the definition of satisfaction would become much more complicated.

Let $x_{i_1}, \ldots, x_{i_k}$ be k distinct variables in order of increasing subscripts. Let $\mathscr{B}(x_{i_1}, \ldots, x_{i_k})$ be a wf that has $x_{i_1}, \ldots, x_{i_k}$ as its only free variables. The set of k-tuples $\langle b_1, \ldots, b_k\rangle$ of elements of the domain D such that any sequence with $b_1, \ldots, b_k$ in its i_1th, …, i_kth places, respectively, satisfies $\mathscr{B}(x_{i_1}, \ldots, x_{i_k})$ is called the *relation* (or *property*†) *of the interpretation defined by* $\mathscr{B}$. Extending our terminology, we shall say that every k-tuple $\langle b_1, \ldots, b_k\rangle$ in this relation *satisfies* $\mathscr{B}(x_{i_1}, \ldots, x_{i_k})$ in the interpretation M; this will be written $\vDash_M \mathscr{B}[b_1, \ldots, b_k]$. This extended notion of satisfaction corresponds to the original intuitive notion.

* A better term for *closure* would be *universal closure*.

† When $k = 1$, the relation is called a *property*.

Examples

1. If the domain D of M is the set of human beings, $A_1^2(x, y)$ is interpreted as *x is a brother of y*, and $A_2^2(x, y)$ is interpreted as *x is a parent of y*, then the binary relation on D corresponding to the wf $\mathscr{B}(x_1, x_2): (\exists x_3)(A_1^2(x_1, x_3) \wedge A_2^2(x_3, x_2))$ is the relation of unclehood. $\vDash_M \mathscr{B}[b, c]$ when and only when b is an uncle of c.
2. If the domain is the set of positive integers, A_1^2 is interpreted as =, f_1^2 is interpreted as multiplication, and a_1 is interpreted as 1, then the wf $\mathscr{B}(x_1)$:

$$\neg A_1^2(x_1, a_1) \wedge (\forall x_2)((\exists x_3)A_1^2(x_1, f_1^2(x_2, x_3)) \Rightarrow A_1^2(x_2, x_1) \vee A_1^2(x_2, a_1))$$

determines the property of being a prime number. Thus $\vDash_M \mathscr{B}[k]$ if and only if k is a prime number.

IX. If $\mathscr{B}$ is a closed wf of a language $\mathscr{L}$, then, for any interpretation M, either $\vDash_M \mathscr{B}$ or $\vDash_M \neg\mathscr{B}$—that is, either $\mathscr{B}$ is true for M or $\mathscr{B}$ is false for M. [*Hint*: Use (VIII).] Of course, $\mathscr{B}$ may be true for some interpretations and false for others. (As an example, consider $A_1^1(a_1)$. If M is an interpretation whose domain is the set of positive integers, A_1^1 is interpreted as the property of being a prime, and the interpretation of a_1 is 2, then $A_1^1(a_1)$ is true. If we change the interpretation by interpreting a_1 as 4, then $A_1^1(a_1)$ becomes false.)

If $\mathscr{B}$ is not closed—that is, if $\mathscr{B}$ contains free variables—$\mathscr{B}$ may be neither true nor false for some interpretation. For example, if $\mathscr{B}$ is $A_1^2(x_1, x_2)$ and we consider an interpretation in which the domain is the set of integers and $A_1^2(y, z)$ is interpreted as $y < z$, then $\mathscr{B}$ is satisfied by only those sequences $s = (s_1, s_2, \ldots)$ of integers in which $s_1 < s_2$. Hence, $\mathscr{B}$ is neither true nor false for this interpretation. On the other hand, there are wfs that are not closed but that nevertheless are true or false for every interpretation. A simple example is the wf $A_1^1(x_1) \vee \neg A_1^1(x_1)$, which is true for every interpretation.

X. Assume t is free for x_i in $\mathscr{B}(x_i)$. Then $(\forall x_i)\mathscr{B}(x_i) \Rightarrow \mathscr{B}(t)$ is true for all interpretations.

The proof of (X) is based upon the following lemmas.

Lemma 1

If t and u are terms, s is a sequence in Σ, t' results from t by replacing all occurrences of x_i by u, and s' results from s by replacing the ith component of s by $s^*(u)$, then $s^*(t') = (s')^*(t)$. [*Hint*: Use induction on the length of t.*]

* The *length* of an expression is the number of occurrences of symbols in the expression.

Lemma 2

Let t be free for x_i in $\mathscr{B}(x_i)$. Then:

a. A sequences $s = (s_1, s_2, \ldots)$ satisfies $\mathscr{B}(t)$ if and only if the sequence s', obtained from s by substituting $s^*(t)$ for s_i in the ith place, satisfies $\mathscr{B}(x_i)$. [*Hint*: Use induction on the number of occurrences of connectives and quantifiers in $\mathscr{B}(x_i)$, applying Lemma 1.]
b. If $(\forall x_i)\mathscr{B}(x_i)$ is satisfied by the sequence s, then $\mathscr{B}(t)$ also is satisfied by s.

XI. If $\mathscr{B}$ does not contain x_i free, then $(\forall x_i)(\mathscr{B} \Rightarrow \mathscr{C}) \Rightarrow (\mathscr{B} \Rightarrow (\forall x_i)\mathscr{C})$ is true for all interpretations.

Proof

Assume (XI) is not correct. Then $(\forall x_i)(\mathscr{B} \Rightarrow \mathscr{C}) \Rightarrow (\mathscr{B} \Rightarrow (\forall x_i)\mathscr{C})$ is not true for some interpretation. By condition 3 of the definition of satisfaction, there is a sequence s such that s satisfies $(\forall x_i)(\mathscr{B} \Rightarrow \mathscr{C})$ and s does not satisfy $\mathscr{B} \Rightarrow (\forall x_i)\mathscr{C}$. From the latter and condition 3, s satisfies $\mathscr{B}$ and s does not satisfy $(\forall x_i)\mathscr{C}$. Hence, by condition 4, there is a sequence s', differing from s in at most the ith place, such that s' does not satisfy $\mathscr{C}$. Since x_i is free in neither $(\forall x_i)(\mathscr{B} \Rightarrow \mathscr{C})$ nor $\mathscr{B}$, and since s satisfies both of these wfs, it follows by (VIII) that s' also satisfies both $(\forall x_i)(\mathscr{B} \Rightarrow \mathscr{C})$ and $\mathscr{B}$. Since s' satisfies $(\forall x_i)(\mathscr{B} \Rightarrow \mathscr{C})$, it follows by condition 4 that s' satisfies $\mathscr{B} \Rightarrow \mathscr{C}$. Since s' satisfies $\mathscr{B} \Rightarrow \mathscr{C}$ and $\mathscr{B}$, condition 3 implies that s' satisfies $\mathscr{C}$, which contradicts the fact that s' does not satisfy $\mathscr{C}$. Hence, (XI) is established.

Exercises

2.12 Verify (I)–(X).

2.13 Prove that a closed wf $\mathscr{B}$ is true for M if and only if $\mathscr{B}$ is satisfied by *some* sequence s in Σ. (Remember that Σ is the set of denumerable sequences of elements in the domain of M.)

2.14 Find the properties or relations determined by the following wfs and interpretations.

a. $[(\exists u)A_1^2(f_1^2(x,u),y)] \wedge [(\exists v)A_1^2(f_1^2(x,v),z)]$, where the domain D is the set of integers, A_1^2 is =, and f_1^2 is multiplication.
b. Here, D is the set of nonnegative integers, A_1^2 is =, a_1 denotes 0, f_1^2 is addition, and f_2^2 is multiplication.
   i. $[(\exists z)(\neg A_1^2(z,a_1) \wedge A_1^2(f_1^2(x,z),y))]$
   ii. $(\exists y)A_1^2(x, f_2^2(y,y))$
c. $(\exists x_3)A_1^2(f_1^2(x_1,x_3),x_2)$, where D is the set of positive integers, A_1^2 is =, and f_1^2 is multiplication,

d. $A_1^1(x_1) \wedge (\forall x_2)\neg A_1^2(x_1, x_2)$, where D is the set of all living people, $A_1^1(x)$ means *x is a man* and $A_1^2(x,y)$ means *x is married to y*.

e. i. $(\exists x_1)(\exists x_2)(A_1^2(x_1, x_3) \wedge A_1^2(x_2, x_4) \wedge A_2^2(x_1, x_2))$

ii. $(\exists x_3)(A_1^2(x_1, x_3) \wedge A_1^2(x_3, x_2))$

where D is the set of all people, $A_1^2(x, y)$ means *x is a parent of y*, and $A_2^2(x, y)$ means *x and y are siblings*.

f. $(\forall x_3)((\exists x_4)(A_1^2(f_1^2(x_4, x_3), x_1) \wedge (\exists x_4)(A_1^2(f_1^2(x_4, x_3), x_2)) \Rightarrow A_1^2(x_3, a_1))$, where D is the set of positive integers, A_1^2 is =, f_1^2 is multiplication, and a_1 denotes 1.

g. $\neg A_1^2(x_2, x_1) \wedge (\exists y)(A_1^2(y, x_1) \wedge A_2^2(x_2, y))$, where D is the set of all people, $A_1^2(u, v)$ means *u is a parent of v*, and $A_2^2(u, v)$ means *u is a wife of v*.

2.15 For each of the following sentences and interpretations, write a translation into ordinary English and determine its truth or falsity.

a. The domain D is the set of nonnegative integers, A_1^2 is =, f_1^2 is addition, f_2^2 is multiplication, a_1 denotes 0, and a_2 denotes 1.

i. $(\forall x)(\exists y)(A_1^2(x, f_1^2(y, y)) \vee A_1^2(x, f_1^2(f_1^2(y, y), a_2)))$

ii. $(\forall x)(\forall y)(A_1^2(f_2^2(x, y), a_1) \Rightarrow A_1^2(x, a_1) \vee A_1^2(y, a_1))$

iii. $(\exists y)A_1^2(f_1^2(y, y), a_2)$

b. Here, D is the set of integers, A_1^2 is =, and f_1^2 is addition.

i. $(\forall x_1)(\forall x_2)A_1^2(f_1^2(x_1, x_2), f_1^2(x_2, x_1))$

ii. $(\forall x_1)(\forall x_2)(\forall x_3)A_1^2(f_1^2(x_1, f_1^2(x_2, x_3)), f_1^2(f_1^2(x_1, x_2), x_3))$

iii. $(\forall x_1)(\forall x_2)(\exists x_3)A_1^2(f_1^2(x_1, x_3), x_2)$

c. The wfs are the same as in part (b), but the domain is the set of positive integers, A_1^2 is =, and $f_1^2(x, y)$ is x^y.

d. The domain is the set of rational numbers, A_1^2 is =, A_2^2 is <, f_1^2 is multiplication, $f_1^1(x)$ is $x + 1$, and a_1 denotes 0.

i. $(\exists x)A_1^2(f_1^2(x, x), f_1^1(f_1^1(a_1)))$

ii. $(\forall x)(\forall y)(A_2^2(x, y) \Rightarrow (\exists z)(A_2^2(x, z) \wedge A_2^2(z, y)))$

iii. $(\forall x)(\neg A_1^2(x, a_1) \Rightarrow (\exists y)A_1^2(f_1^2(x, y), f_1^1(a_1)))$

e. The domain is the set of nonnegative integers, $A_1^2(u, v)$ means $u \leq v$, and $A_1^3(u, v, w)$ means $u + v = w$.

i. $(\forall x)(\forall y)(\forall z)(A_1^3(x, y, z) \Rightarrow A_1^3(y, x, z))$

ii. $(\forall x)(\forall y)(A_1^3(x, x, y) \Rightarrow A_1^2(x, y))$

iii. $(\forall x)(\forall y)(A_1^2(x, y) \Rightarrow A_1^3(x, x, y))$

iv. $(\exists x)(\forall y)A_1^3(x, y, y)$

v. $(\exists y)(\forall x)A_1^2(x,y)$

vi. $(\forall x)(\forall y)(A_1^2(x,y) \Leftrightarrow (\exists z)A_1^3(x,z,y))$

f. The domain is the set of nonnegative integers, $A_1^2(u,v)$ means $u=v$, $f_1^2(u,v)=u+v$, and $f_2^2(u,v)=u\cdot v$

$$(\forall x)(\exists y)(\exists z)A_1^2(x, f_1^2(f_2^2(y,y), f_2^2(z,z)))$$

Definitions

A wf $\mathscr{B}$ is said to be *logically valid* if and only if $\mathscr{B}$ is true for every interpretation.*

$\mathscr{B}$ is said to be *satisfiable* if and only if there is an interpretation for which $\mathscr{B}$ is satisfied by at least one sequence.

It is obvious that $\mathscr{B}$ is logically valid if and only if $\neg\mathscr{B}$ is not satisfiable, and $\mathscr{B}$ is satisfiable if and only if $\neg\mathscr{B}$ is not logically valid.

If $\mathscr{B}$ is a closed wf, then we know that $\mathscr{B}$ is either true or false for any given interpretation; that is, $\mathscr{B}$ is satisfied by all sequences or by none. Therefore, if $\mathscr{B}$ is closed, then $\mathscr{B}$ is satisfiable if and only if $\mathscr{B}$ is true for some interpretation.

A set Γ of wfs is said to be *satisfiable* if and only if there is an interpretation in which there is a sequence that satisfies every wf of Γ.

It is impossible for both a wf $\mathscr{B}$ and its negation $\neg\mathscr{B}$ to be logically valid. For if $\mathscr{B}$ is true for an interpretation, then $\neg\mathscr{B}$ is false for that interpretation.

We say that $\mathscr{B}$ is *contradictory* if and only if $\mathscr{B}$ is false for every interpretation, or, equivalently, if and only if $\neg\mathscr{B}$ is logically valid.

$\mathscr{B}$ is said to *logically imply* $\mathscr{C}$ if and only if, in every interpretation, every sequence that satisfies $\mathscr{B}$ also satisfies $\mathscr{C}$. More generally, $\mathscr{C}$ is said to be a *logical consequence* of a set Γ of wfs if and only if, in every interpretation, every sequence that satisfies every wf in Γ also satisfies $\mathscr{C}$.

$\mathscr{B}$ and $\mathscr{C}$ are said to be *logically equivalent* if and only if they logically imply each other.

The following assertions are easy consequences of these definitions.

1. $\mathscr{B}$ logically implies $\mathscr{C}$ if and only if $\mathscr{B} \Rightarrow \mathscr{C}$ is logically valid.
2. $\mathscr{B}$ and $\mathscr{C}$ are logically equivalent if and only if $\mathscr{B} \Leftrightarrow \mathscr{C}$ is logically valid.
3. If $\mathscr{B}$ logically implies $\mathscr{C}$ and $\mathscr{B}$ is true in a given interpretation, then so is $\mathscr{C}$.
4. If $\mathscr{C}$ is a logical consequence of a set Γ of wfs and all wfs in Γ are true in a given interpretation, then so is $\mathscr{C}$.

* The mathematician and philosopher G.W. Leibniz (1646–1716) gave a similar definition: $\mathscr{B}$ is logically valid if and only if $\mathscr{B}$ is true in all "possible worlds."

Exercise

2.16 Prove assertions 1–4.

Examples

1. Every instance of a tautology is logically valid (VII).
2. If t is free for x in $\mathscr{B}(x)$, then $(\forall x)\mathscr{B}(x) \Rightarrow \mathscr{B}(t)$ is logically valid (X).
3. If $\mathscr{B}$ does not contain x free, then $(\forall x)(\mathscr{B} \Rightarrow \mathscr{C}) \Rightarrow (\mathscr{B} \Rightarrow (\forall x)\mathscr{C})$ is logically valid (XI).
4. $\mathscr{B}$ is logically valid if and only if $(\forall y_1) \ldots (\forall y_n)\mathscr{B}$ is logically valid (VI).
5. The wf $(\forall x_2)(\exists x_1)A_1^2(x_1, x_2) \Rightarrow (\exists x_1)(\forall x_2)A_1^2(x_1, x_2)$ is not logically valid. As a counterexample, let the domain D be the set of integers and let $A_1^2(y, z)$ mean $y < z$. Then $(\forall x_2)(\exists x_1)A_1^2(x_1, x_2)$ is true but $(\exists x_1)(\forall x_2)A_1^2(x_1, x_2)$ is false.

Exercises

2.17 Show that the following wfs are not logically valid.

a. $[(\forall x_1)A_1^1(x_1) \Rightarrow (\forall x_1)A_2^1(x_1)] \Rightarrow [(\forall x_1)(A_1^1(x_1) \Rightarrow A_2^1(x_1))]$

b. $[(\forall x_1)(A_1^1(x_1) \vee A_2^1(x_1))] \Rightarrow [((\forall x_1))A_1^1(x_1)) \vee (\forall x_1)A_2^1(x_1)]$

2.18 Show that the following wfs are logically valid.*

a. $\mathscr{B}(t) \Rightarrow (\exists x_i)\mathscr{B}(x_i)$ if t is free for x_i in $\mathscr{B}(x_i)$

b. $(\forall x_i)\mathscr{B} \Rightarrow (\exists x_i)\mathscr{B}$

c. $(\forall x_i)(\forall x_j)\mathscr{B} \Rightarrow (\forall x_j)(\forall x_i)\mathscr{B}$

d. $(\forall x_i)\mathscr{B} \Leftrightarrow \neg(\exists x_i)\neg\mathscr{B}$

e. $(\forall x_i)(\mathscr{B} \Rightarrow \mathscr{C}) \Rightarrow ((\forall x_i)\mathscr{B} \Rightarrow (\forall x_i)\mathscr{C})$

f. $((\forall x_i)\mathscr{B}) \wedge (\forall x_i)\mathscr{C} \Leftrightarrow (\forall x_i)(\mathscr{B} \wedge \mathscr{C})$

g. $((\forall x_i)\mathscr{B}) \vee (\forall x_i)\mathscr{C} \Rightarrow (\forall x_i)(\mathscr{B} \vee \mathscr{C})$

h. $(\exists x_i)(\exists x_j)\mathscr{B} \Leftrightarrow (\exists x_j)(\exists x_i)\mathscr{B}$

i. $(\exists x_i)(\forall x_j)\mathscr{B} \Rightarrow (\forall x_j)(\exists x_i)\mathscr{B}$

2.19 a. If $\mathscr{B}$ is a closed wf, show that $\mathscr{B}$ logically implies $\mathscr{C}$ if and only if $\mathscr{C}$ is true for every interpretation for which $\mathscr{B}$ is true.

b. Although, by (VI), $(\forall x_1)A_1^1(x_1)$ is true whenever $A_1^1(x_1)$ is true, find an interpretation for which $A_1^1(x_1) \Rightarrow (\forall x_1)A_1^1(x_1)$ is not true. (Hence, the hypothesis that $\mathscr{B}$ is a closed wf is essential in (a).)

* At this point, one can use intuitive arguments or one can use the rigorous definitions of satisfaction and truth, as in the argument above for (XI). Later on, we shall discover another method for showing logical validity.

2.20 Prove that, if the free variables of $\mathscr{B}$ are $y_1, \ldots, y_n$, then $\mathscr{B}$ is satisfiable if and only if $(\exists y_1), \ldots, (\exists y_n)\mathscr{B}$ is satisfiable.

2.21 Produce counterexamples to show that the following wfs are not logically valid (that is, in each case, find an interpretation for which the wf is not true).

a. $$[(\forall x)(\forall y)(\forall z)(A_1^2(x,y) \wedge A_1^2(y,z) \Rightarrow A_1^2(x,z)) \wedge (\forall x)\neg A_1^2(x,x)] \Rightarrow (\exists x)(\forall y)\neg A_1^2(x,y)$$

b. $(\forall x)(\exists y)A_1^2(x,y) \Rightarrow (\exists y)A_1^2(y,y)$

c. $(\exists x)(\exists y)A_1^2(x,y) \Rightarrow (\exists y)A_1^2(y,y)$

d. $[(\exists x)A_1^1(x) \Leftrightarrow (\exists x)A_2^1(x)] \Rightarrow (\forall x)(A_1^1(x) \Leftrightarrow A_2^1(x))$

e. $(\exists x)(A_1^1(x) \Rightarrow A_2^1(x)) \Rightarrow ((\exists x)A_1^1(x) \Rightarrow (\exists x)A_2^1(x))$

f. $$[(\forall x)(\forall y)(A_1^2(x,y) \Rightarrow A_1^2(y,x)) \wedge (\forall x)(\forall y)(\forall z)(A_1^2(x,y) \wedge A_1^2(y,z) \Rightarrow A_1^2(x,z))] \Rightarrow (\forall x)A_1^2(x,x)$$

g.[D] $(\exists x)(\forall y)(A_1^2(x,y) \wedge \neg A_1^2(y,x) \Rightarrow [A_1^2(x,x) \Leftrightarrow A_1^2(y,y)])$

h. $$(\forall x)(\forall y)(\forall z)(A_1^2(x,x) \wedge (A_1^2(x,z) \Rightarrow A_1^2(x,y) \vee A_1^2(y,z))) \Rightarrow (\exists y)(\forall z)A_1^2(y,z)$$

i. $(\exists x)(\forall y)(\exists z)((A_1^2(y,z) \Rightarrow A_1^2(x,z)) \Rightarrow (A_1^2(x,x) \Rightarrow A_1^2(y,x)))$

2.22 By introducing appropriate notation, write the sentences of each of the following arguments as wfs and determine whether the argument is correct, that is, determine whether the conclusion is logically implied by the conjunction of the premisses

a. All scientists are neurotic. No vegetarians are neurotic. Therefore, no vegetarians are scientists.

b. All men are animals. Some animals are carnivorous. Therefore, some men are carnivorous.

c. Some geniuses are celibate. Some students are not celibate. Therefore, some students are not geniuses.

d. Any barber in Jonesville shaves exactly those men in Jonesville who do not shave themselves. Hence, there is no barber in Jonesville.

e. For any numbers x, y, z, if $x > y$ and $y > z$, then $x > z$. $x > x$ is false for all numbers x. Therefore, for any numbers x and y, if $x > y$, then it is not the case that $y > x$.

f. No student in the statistics class is smarter than every student in the logic class. Hence, some student in the logic class is smarter than every student in the statistics class.

g. Everyone who is sane can understand mathematics. None of Hegel's sons can understand mathematics. No madmen are fit to vote. Hence, none of Hegel's sons is fit to vote.

h. For every set x, there is a set y such that the cardinality of y is greater than the cardinality of x. If x is included in y, the cardinality of x is not greater than the cardinality of y. Every set is included in V. Hence, V is not a set.

i. For all positive integers x, $x \leq x$. For all positive integers x, y, z, if $x \leq y$ and $y \leq z$, then $x \leq z$. For all positive integers x and y, $x \leq y$ or $y \leq x$. Therefore, there is a positive integer y such that, for all positive integers x, $y \leq x$.

j. For any integers x, y, z, if $x > y$ and $y > z$, then $x > z$. $x > x$ is false for all integers x. Therefore, for any integers x and y, if $x > y$, then it is not the case that $y > x$.

2.23 Determine whether the following sets of wfs are compatible—that is, whether their conjunction is satisfiable.

a. $(\exists x)(\exists y)A_1^2(x,y)$

$(\forall x)(\forall y)(\exists z)(A_1^2(x,z) \wedge A_1^2(z,y))$

b. $(\forall x)(\exists y)A_1^2(y,x)$

$(\forall x)(\forall y)(A_1^2(x,y) \Rightarrow \neg A_1^2(y,x))$

$(\forall x)(\forall y)(\forall z)(A_1^2(x,y) \wedge A_1^2(y,z) \Rightarrow A_1^2(x,z))$

c. All unicorns are animals.

No unicorns are animals.

2.24 Determine whether the following wfs are logically valid.

a. $\neg(\exists y)(\forall x)(A_1^2(x,y) \Leftrightarrow \neg A_1^2(x,x))$

b. $[(\exists x)A_1^1(x) \Rightarrow (\exists x)A_2^1(x)] \Rightarrow (\exists x)(A_1^1(x) \Rightarrow A_2^1(x))$

c. $(\exists x)(A_1^1(x) \Rightarrow (\forall y)A_1^1(y))$

d. $(\forall x)(A_1^1(x) \vee A_2^1(x)) \Rightarrow (((\forall x)A_1^1(x)) \vee (\exists x)A_2^1(x))$

e. $(\exists x)(\exists y)(A_1^2(x,y) \Rightarrow (\forall z)A_1^2(z,y))$

f. $(\exists x)(\exists y)(A_1^1(x) \Rightarrow A_2^1(y)) \Rightarrow (\exists x)(A_1^1(x) \Rightarrow A_2^1(x))$

g. $(\forall x)(A_1^1(x) \Rightarrow A_2^1(x)) \Rightarrow \neg(\forall x)(A_1^1(x) \Rightarrow \neg A_2^1(x))$

h. $(\exists x)A_1^2(x,x) \Rightarrow (\exists x)(\exists y)A_1^2(x,y)$

i. $((\exists x)A_1^1(x)) \wedge (\exists x)A_2^1(x) \Rightarrow (\exists x)(A_1^1(x) \wedge A_2^1(x))$

j. $((\forall x)A_1^1(x)) \vee (\forall x)A_2^1(x) \Rightarrow (\forall x)(A_1^1(x) \vee A_2^1(x))$

2.25 Exhibit a logically valid wf that is not an instance of a tautology. However, show that any logically valid *open* wf (that is, a wf without quantifiers) must be an instance of a tautology.

2.26 a. Find a satisfiable closed wf that is not true in any interpretation whose domain has only one member.

b. Find a satisfiable closed wf that is not true in any interpretation whose domain has fewer than three members.

2.3 First-Order Theories

In the case of the propositional calculus, the method of truth tables provides an effective test as to whether any given statement form is a tautology. However, there does not seem to be any effective process for determining whether a given wf is logically valid, since, in general, one has to check the truth of a wf for interpretations with arbitrarily large finite or infinite domains. In fact, we shall see later that, according to a plausible definition of "effective," it may actually be proved that there is no effective way to test for logical validity. The axiomatic method, which was a luxury in the study of the propositional calculus, thus appears to be a necessity in the study of wfs involving quantifiers,* and we therefore turn now to the consideration of first-order theories.

Let $\mathscr{L}$ be a first-order language. A *first-order theory* in the language $\mathscr{L}$ will be a formal theory K whose symbols and wfs are the symbols and wfs of $\mathscr{L}$ and whose axioms and rules of inference are specified in the following way.†

The axioms of K are divided into two classes: the logical axioms and the proper (or nonlogical) axioms.

2.3.1 Logical Axioms

If $\mathscr{B}$, $\mathscr{C}$, and $\mathscr{D}$ are wfs of $\mathscr{L}$, then the following are logical axioms of K:

(A1) $\mathscr{B} \Rightarrow (\mathscr{C} \Rightarrow \mathscr{B})$

(A2) $(\mathscr{B} \Rightarrow (\mathscr{C} \Rightarrow \mathscr{D})) \Rightarrow ((\mathscr{B} \Rightarrow \mathscr{C}) \Rightarrow (\mathscr{B} \Rightarrow \mathscr{D}))$

(A3) $(\neg\mathscr{C} \Rightarrow \neg\mathscr{B}) \Rightarrow ((\neg\mathscr{C} \Rightarrow \mathscr{B}) \Rightarrow \mathscr{C})$

* There is still another reason for a formal axiomatic approach. Concepts and propositions that involve the notion of interpretation and related ideas such as truth and model are often called *semantical* to distinguish them from *syntactical* concepts, which refer to simple relations among symbols and expressions of precise formal languages. Since semantical notions are set-theoretic in character, and since set theory, because of the paradoxes, is considered a rather shaky foundation for the study of mathematical logic, many logicians consider a syntactical approach, consisting of a study of formal axiomatic theories using only rather weak number-theoretic methods, to be much safer. For further discussions, see the pioneering study on semantics by Tarski (1936), as well as Kleene (1952), Church (1956), and Hilbert and Bernays (1934).

† The reader might wish to review the definition of *formal theory* in Section 1.4. We shall use the terminology (proof, theorem, consequence, axiomatic, $\vdash \mathscr{B}$, etc.) and notation ($\Gamma \vdash \mathscr{B}$, $\vdash \mathscr{B}$) introduced there.

(A4) $(\forall x_i)\mathscr{B}(x_i) \Rightarrow \mathscr{B}(t)$ if $\mathscr{B}(x_i)$ is a wf of $\mathscr{L}$ and t is a term of $\mathscr{L}$ that is free for x_i in $\mathscr{B}(x_i)$. Note here that t may be identical with x_i so that all wfs $(\forall x_i)\mathscr{B} \Rightarrow \mathscr{B}$ are axioms by virtue of axiom (A4).

(A5) $(\forall x_i)(\mathscr{B} \Rightarrow \mathscr{C}) \Rightarrow (\mathscr{B} \Rightarrow (\forall x_i)\mathscr{C})$ if $\mathscr{B}$ contains no free occurrences of x_i.

2.3.2 Proper Axioms

These cannot be specified, since they vary from theory to theory. A first-order theory in which there are no proper axioms is called a first-order *predicate calculus.*

2.3.3 Rules of Inference

The rules of inference of any first-order theory are:

1. Modus ponens: $\mathscr{C}$ follows from $\mathscr{B}$ and $\mathscr{B} \Rightarrow \mathscr{C}$.
2. Generalization: $(\forall x_i)\mathscr{B}$ follows from $\mathscr{B}$.

We shall use the abbreviations MP and Gen, respectively, to indicate applications of these rules.

Definition

Let K be a first-order theory in the language $\mathscr{L}$. By a *model* of K we mean an interpretation of $\mathscr{L}$ for which all the axioms of K are true.

By (III) and (VI) on page 57, if the rules of modus ponens and generalization are applied to wfs that are true for a given interpretation, then the results of these applications are also true. Hence *every theorem of* K *is true in every model of* K.

As we shall see, the logical axioms are so designed that the logical consequences (in the sense defined on pages 63–64) of the closures of the axioms of K are precisely the theorems of K. In particular, if K is a first-order predicate calculus, it turns out that the theorems of K are just those wfs of K that are logically valid.

Some explanation is needed for the restrictions in axiom schemas (A4) and (A5). In the case of (A4), if t were not free for x_i in $\mathscr{B}(x_i)$, the following unpleasant result would arise: let $\mathscr{B}(x_1)$ be $\neg(\forall x_2)A_1^2(x_1, x_2)$ and let t be x_2. Notice that t is not free for x_1 in $\mathscr{B}(x_1)$. Consider the following pseudo-instance of axiom (A4):

$$(\nabla) \quad (\forall x_1)\Big(\neg(\forall x_2)A_1^2(x_1, x_2)\Big) \Rightarrow \neg(\forall x_2)A_1^2(x_2, x_2)$$

Now take as interpretation any domain with at least two members and let A_1^2 stand for the identity relation. Then the antecedent of (∇) is true and the consequent false. Thus, (∇) is false for this interpretation.

In the case of axiom (A5), relaxation of the restriction that x_i not be free in $\mathscr{B}$ would lead to the following disaster. Let $\mathscr{B}$ and $\mathscr{C}$ both be $A_1^1(x_1)$. Thus, x_1 is free in $\mathscr{B}$. Consider the following pseudo-instance of axiom (A5):

$$(\nabla\nabla) \quad (\forall x_1)\left(A_1^1(x_1) \Rightarrow A_1^1(x_1)\right) \Rightarrow \left(A_1^1(x_1) \Rightarrow (\forall x_1)A_1^1(x_1)\right)$$

The antecedent of (∇∇) is logically valid. Now take as domain the set of integers and let $A_1^1(x)$ mean that *x is even*. Then $(\forall x_1)A_1^1(x_1)$ is false. So, any sequence $s = (s_1, s_2, \ldots)$ for which s_1 is even does not satisfy the consequent of (∇∇).* Hence, (∇∇) is not true for this interpretation.

Examples of first-order theories

1. *Partial order*. Let the language $\mathscr{L}$ have a single predicate letter A_2^2 and no function letters and individual constants. We shall write $x_i < x_j$ instead of $A_2^2(x_i, x_j)$. The theory K has two proper axioms.
   a. $(\forall x_1)(\neg\, x_1 < x_1)$ (irreflexivity)
   b. $(\forall x_1)(\forall x_2)(\forall x_3)(x_1 < x_2 \wedge x_2 < x_3 \Rightarrow x_1 < x_3)$ (transitivity)

   A model of the theory is called a *partially ordered structure*.
2. *Group theory*. Let the language $\mathscr{L}$ have one predicate letter A_1^2, one function letter f_1^2, and one individual constant a_1. To conform with ordinary notation, we shall write $t = s$ instead of $A_1^2(t, s)$, $t + s$ instead of $f_1^2(t, s)$, and 0 instead of a_1. The proper axioms of K are:
   a. $(\forall x_1)(\forall x_2)(\forall x_3)(x_1 + (x_2 + x_3)$
   $= (x_1 + x_2) + x_3)$ (associativity)
   b. $(\forall x_1)(0 + x_1 = x_1)$ (identity)
   c. $(\forall x_1)(\exists x_2)(x_2 + x_1 = 0)$ (inverse)
   d. $(\forall x_1)(x_1 = x_1)$ (reflexivity of =)
   e. $(\forall x_1)(\forall x_2)(x_1 = x_2 \Rightarrow x_2 = x_1)$ (symmetry of =)
   f. $(\forall x_1)(\forall x_2)(\forall x_3)(x_1 = x_2 \wedge x_2 = x_3 \Rightarrow x_1 = x_3)$ (transitivity of =)
   g. $(\forall x_1)(\forall x_2)(\forall x_3)(x_2 = x_3 \Rightarrow x_1 + x_2$
   $= x_1 + x_3 \wedge x_2 + x_1 = x_3 + x_1)$ (substitutivity of =)

A model for this theory, in which the interpretation of = is the identity relation, is called a *group*. A group is said to be *abelian* if, in addition, the wf $(\forall x_1)(\forall x_2)(x_1 + x_2 = x_2 + x_1)$ is true.

* Such a sequence would satisfy $A_1^1(x_1)$, since s_1 is even, but would not satisfy $(\forall x_1)A_1^1(x_1)$, since no sequence satisfies $(\forall x_1)A_1^1(x_1)$.

The theories of partial order and of groups are both axiomatic. In general, any theory with a finite number of proper axioms is axiomatic, since it is obvious that one can effectively decide whether any given wf is a logical axiom.

2.4 Properties of First-Order Theories

All the results in this section refer to an arbitrary first-order theory K. Instead of writing $\vdash_K \mathscr{B}$, we shall sometimes simply write $\vdash \mathscr{B}$. Moreover, we shall refer to first-order theories simply as *theories*, unless something is said to the contrary.

Proposition 2.1

Every wf $\mathscr{B}$ of K that is an instance of a tautology is a theorem of K, and it may be proved using only axioms (A1)–(A3) and MP.

Proof

$\mathscr{B}$ arises from a tautology $\mathscr{S}$ by substitution. By Proposition 1.14, there is a proof of $\mathscr{S}$ in L. In such a proof, make the same substitution of wfs of K for statement letters as were used in obtaining $\mathscr{B}$ from $\mathscr{S}$, and, for all statement letters in the proof that do not occur in $\mathscr{S}$, substitute an arbitrary wf of K. Then the resulting sequence of wfs is a proof of $\mathscr{B}$, and this proof uses only axiom schemes (A1)–(A3) and MP.

The application of Proposition 2.1 in a proof will be indicated by writing "Tautology."

Proposition 2.2

Every theorem of a first-order predicate calculus is logically valid.

Proof

Axioms (A1)–(A3) are logically valid by property (VII) of the notion of truth (see page 58), and axioms (A4) and (A5) are logically valid by properties (X) and (XI). By properties (III) and (VI), the rules of inference MP and Gen preserve logical validity. Hence, every theorem of a predicate calculus is logically valid.

Example

The wf $(\forall x_2)(\exists x_1)A_1^2(x_1, x_2) \Rightarrow (\exists x_1)(\forall x_2)A_1^2(x_1, x_2)$ is not a theorem of any first-order predicate calculus, since it is not logically valid (by Example 5, page 63).

Definition

A theory K is *consistent* if no wf $\mathscr{B}$ and its negation $\neg\mathscr{B}$ are both provable in K. A theory is *inconsistent* if it is not consistent.

Corollary 2.3

Any first-order predicate calculus is consistent.

Proof

If a wf $\mathscr{B}$ and its negation $\neg\mathscr{B}$ were both theorems of a first-order predicate calculus, then, by Proposition 2.2, both $\mathscr{B}$ and $\neg\mathscr{B}$ would be logically valid, which is impossible.

Notice that, in an inconsistent theory K, every wf $\mathscr{C}$ of K is provable in K. In fact, assume that $\mathscr{B}$ and $\neg\mathscr{B}$ are both provable in K. Since the wf $\mathscr{B} \Rightarrow (\neg\mathscr{B} \Rightarrow \mathscr{C})$ is an instance of a tautology, that wf is, by Proposition 2.1, provable in K. Then two applications of MP would yield $\vdash \mathscr{C}$.

It follows from this remark that, if some wf of a theory K is not a theorem of K, then K is consistent.

The deduction theorem (Proposition 1.9) for the propositional calculus cannot be carried over without modification to first-order theories. For example, for any wf $\mathscr{B}$, $\mathscr{B} \vdash_K (\forall x_i)\mathscr{B}$, but it is not always the case that $\vdash_K \mathscr{B} \Rightarrow (\forall x_i)\mathscr{B}$. Consider a domain containing at least two elements c and d. Let K be a predicate calculus and let $\mathscr{B}$ be $A_1^1(x_1)$. Interpret A_1^1 as a property that holds only for c. Then $A_1^1(x_1)$ is satisfied by any sequence $s = (s_1, s_2, \ldots)$ in which $s_1 = c$, but $(\forall x_1)A_1^1(x_1)$ is satisfied by no sequence at all. Hence, $A_1^1(x_1) \Rightarrow (\forall x_1)A_1^1(x_1)$ is not true in this interpretation, and so it is not logically valid. Therefore, by Proposition 2.2, $A_1^1(x_1) \Rightarrow (\forall x_1)A_1^1(x_1)$ is not a theorem of K.

A modified, but still useful, form of the deduction theorem may be derived, however. Let $\mathscr{B}$ be a wf in a set Γ of wfs and assume that we are given a deduction $\mathscr{D}_1, \ldots, \mathscr{D}_n$ from Γ , together with justification for each step in the deduction. We shall say that $\mathscr{D}_i$ *depends upon* $\mathscr{B}$ in this proof if and only if:

1. $\mathscr{D}_i$ is $\mathscr{B}$ and the justification for $\mathscr{D}_i$ is that it belongs to Γ, or
2. $\mathscr{D}_i$ is justified as a direct consequence by MP or Gen of some preceding wfs of the sequence, where at least one of these preceding wfs depends upon $\mathscr{B}$.

Example

$\mathscr{B}, (\forall x_1)\mathscr{B} \Rightarrow \mathscr{C} \vdash (\forall x_1)\mathscr{C}$

($\mathscr{D}_1$)	$\mathscr{B}$	Hyp
($\mathscr{D}_2$)	$(\forall x_1)\mathscr{B}$	($\mathscr{D}_1$), Gen
($\mathscr{D}_3$)	$(\forall x_1)\mathscr{B} \Rightarrow \mathscr{C}$	Hyp
($\mathscr{D}_4$)	$\mathscr{C}$	($\mathscr{D}_2$), ($\mathscr{D}_3$), MP
($\mathscr{D}_5$)	$(\forall x_1)\mathscr{C}$	($\mathscr{D}_4$), Gen

Here, ($\mathscr{D}_1$) depends upon $\mathscr{B}$, ($\mathscr{D}_2$) depends upon $\mathscr{B}$, ($\mathscr{D}_3$) depends upon $(\forall x_1)\mathscr{B} \Rightarrow \mathscr{C}$, ($\mathscr{D}_4$) depends upon $\mathscr{B}$ and $(\forall x_1)\mathscr{B} \Rightarrow \mathscr{C}$, and ($\mathscr{D}_5$) depends upon $\mathscr{B}$ and $(\forall x_1)\mathscr{B} \Rightarrow \mathscr{C}$.

Proposition 2.4

If $\mathscr{C}$ does not depend upon $\mathscr{B}$ in a deduction showing that $\Gamma, \mathscr{B} \vdash \mathscr{C}$, then $\Gamma \vdash \mathscr{C}$.

Proof

Let $\mathscr{D}_1 \ldots, \mathscr{D}_n$ be a deduction of $\mathscr{C}$ from Γ and $\mathscr{B}$, in which $\mathscr{C}$ does not depend upon $\mathscr{B}$. (In this deduction, $\mathscr{D}_n$ is $\mathscr{C}$.) As an inductive hypothesis, let us assume that the proposition is true for all deductions of length less than n. If $\mathscr{C}$ belongs to Γ or is an axiom, then $\Gamma \vdash \mathscr{C}$. If $\mathscr{C}$ is a direct consequence of one or two preceding wfs by Gen or MP, then, since $\mathscr{C}$ does not depend upon $\mathscr{B}$, neither do these preceding wfs. By the inductive hypothesis, these preceding wfs are deducible from Γ alone. Consequently, so is $\mathscr{C}$.

Proposition 2.5 (Deduction Theorem)

Assume that, in some deduction showing that $\Gamma, \mathscr{B} \vdash \mathscr{C}$, no application of Gen to a wf that depends upon $\mathscr{B}$ has as its quantified variable a free variable of $\mathscr{B}$. Then $\Gamma \vdash \mathscr{B} \Rightarrow \mathscr{C}$.

Proof

Let $\mathscr{D}_1, \ldots, \mathscr{D}_n$ be a deduction of $\mathscr{C}$ from Γ and $\mathscr{B}$, satisfying the assumption of our proposition. (In this deduction, $\mathscr{D}_n$ is $\mathscr{C}$.) Let us show by induction that $\Gamma \vdash \mathscr{B} \Rightarrow \mathscr{D}_i$ for each $i \leq n$. If $\mathscr{D}_i$ is an axiom or belongs to Γ, then $\Gamma \vdash \mathscr{B} \Rightarrow \mathscr{D}_i$, since $\mathscr{D}_i \Rightarrow (\mathscr{B} \Rightarrow \mathscr{D}_i)$ is an axiom. If $\mathscr{D}_i$ is $\mathscr{B}$, then $\Gamma \vdash \mathscr{B} \Rightarrow \mathscr{D}_i$, since, by Proposition 2.1, $\vdash \mathscr{B} \Rightarrow \mathscr{B}$. If there exist j and k less than i such that $\mathscr{D}_k$ is $\mathscr{D}_j \Rightarrow \mathscr{D}_i$, then, by inductive hypothesis, $\Gamma \vdash \mathscr{B} \Rightarrow \mathscr{D}_j$ and $\Gamma \vdash \mathscr{B} \Rightarrow (\mathscr{D}_j \Rightarrow \mathscr{D}_i)$. Now, by axiom (A2), $\vdash (\mathscr{B} \Rightarrow (\mathscr{D}_j \Rightarrow \mathscr{D}_i)) \Rightarrow ((\mathscr{B} \Rightarrow \mathscr{D}_j) \Rightarrow (\mathscr{B} \Rightarrow \mathscr{D}_i))$. Hence, by MP twice, $\Gamma \vdash \mathscr{B} \Rightarrow \mathscr{D}_i$. Finally, suppose that there is some $j < i$ such that $\mathscr{D}_i$ is $(\forall x_k)\mathscr{D}_j$. By the inductive hypothesis, $\Gamma \vdash \mathscr{B} \Rightarrow \mathscr{D}_j$, and, by the hypothesis of the theorem, either $\mathscr{D}_j$ does not depend upon $\mathscr{B}$ or x_k is not a free variable of $\mathscr{B}$. If $\mathscr{D}_j$ does not depend upon $\mathscr{B}$, then, by Proposition 2.4, $\Gamma \vdash \mathscr{D}_j$ and, consequently, by Gen, $\Gamma \vdash (\forall x_k)\mathscr{D}_j$. Thus, $\Gamma \vdash \mathscr{D}_i$. Now, by axiom (A1), $\vdash \mathscr{D}_i \Rightarrow (\mathscr{B} \Rightarrow \mathscr{D}_i)$. So, $\Gamma \vdash \mathscr{B} \Rightarrow \mathscr{D}_i$ by MP. If, on the other hand, x_k is not a free variable of $\mathscr{B}$, then, by axiom (A5), $\vdash (\forall x_k)(\mathscr{B} \Rightarrow \mathscr{D}_j) \Rightarrow (\mathscr{B} \Rightarrow (\forall x_k)\mathscr{D}_j)$. Since $\Gamma \vdash \mathscr{B} \Rightarrow \mathscr{D}_j$, we have, by Gen, $\Gamma \vdash (\forall x_k)(\mathscr{B} \Rightarrow \mathscr{D}_j)$, and so, by MP, $\Gamma \vdash \mathscr{B} \Rightarrow (\forall x_k)\mathscr{D}_j$; that is, $\Gamma \vdash \mathscr{B} \Rightarrow \mathscr{D}_i$. This completes the induction, and our proposition is just the special case $i = n$.

The hypothesis of Proposition 2.5 is rather cumbersome; the following weaker corollaries often prove to be more useful.

Corollary 2.6

If a deduction showing that $\Gamma, \mathscr{B} \vdash \mathscr{C}$ involves no application of Gen of which the quantified variables is free in $\mathscr{B}$, then $\Gamma \vdash \mathscr{B} \Rightarrow \mathscr{C}$.

Corollary 2.7

If $\mathscr{B}$ is a closed wf and $\Gamma, \mathscr{B} \vdash \mathscr{C}$, then $\Gamma \vdash \mathscr{B} \Rightarrow \mathscr{C}$.

Extension of Propositions 2.4–2.7

In Propositions 2.4–2.7, the following additional conclusion can be drawn from the proofs. The new proof of $\Gamma \vdash \mathscr{B} \Rightarrow \mathscr{C}$ (in Proposition 2.4, of $\Gamma \vdash \mathscr{C}$) involves an application of Gen to a wf depending upon a wf $\mathscr{E}$ of Γ only if there is an application of Gen in the given proof of $\Gamma, \mathscr{B} \vdash \mathscr{C}$ that involves the same quantified variable and is applied to a wf that depends upon $\mathscr{E}$. (In the proof of Proposition 2.5, one should observe that $\mathscr{D}_j$ depends upon a premiss $\mathscr{E}$ of Γ in the original proof if and only if $\mathscr{B} \Rightarrow \mathscr{D}_j$ depends upon $\mathscr{E}$ in the new proof.)

This supplementary conclusion will be useful when we wish to apply the deduction theorem several times in a row to a given deduction—for example, to obtain $\Gamma \vdash \mathscr{D} \Rightarrow (\mathscr{B} \Rightarrow \mathscr{C})$ from $\Gamma, \mathscr{D}, \mathscr{B} \vdash \mathscr{C}$; from now on, it is to be considered an integral part of the statements of Propositions 2.4–2.7.

Example

$$\vdash (\forall x_1)(\forall x_2)\mathscr{B} \Rightarrow (\forall x_2)(\forall x_1)\mathscr{B}$$

Proof

1. $(\forall x_1)(\forall x_2)\mathscr{B}$ Hyp
2. $(\forall x_1)(\forall x_2)\mathscr{B} \Rightarrow (\forall x_2)\mathscr{B}$ (A4)

3. $(\forall x_2)\mathscr{B}$ — 1, 2, MP
4. $(\forall x_2)\mathscr{B} \Rightarrow \mathscr{B}$ — (A4)
5. $\mathscr{B}$ — 3, 4, MP
6. $(\forall x_1)\mathscr{B}$ — 5, Gen
7. $(\forall x_2)(\forall x_1)\mathscr{B}$ — 6, Gen

Thus, by 1–7, we have $(\forall x_1)(\forall x_2)\mathscr{B} \vdash (\forall x_2)(\forall x_1)\mathscr{B}$, where, in the deduction, no application of Gen has as a quantified variable a free variable of $(\forall x_1)(\forall x_2)\mathscr{B}$. Hence, by Corollary 2.6, $\vdash (\forall x_1)(\forall x_2)\mathscr{B} \Rightarrow (\forall x_2)(\forall x_1)\mathscr{B}$.

Exercises

2.27 Derive the following theorems.

a. $\vdash (\forall x)(\mathscr{B} \Rightarrow \mathscr{C}) \Rightarrow ((\forall x)\mathscr{B} \Rightarrow (\forall x)\mathscr{C})$

b. $\vdash (\forall x)(\mathscr{B} \Rightarrow \mathscr{C}) \Rightarrow ((\exists x)\mathscr{B} \Rightarrow (\exists x)\mathscr{C})$

c. $\vdash (\forall x)(\mathscr{B} \wedge \mathscr{C}) \Leftrightarrow (\forall x)\mathscr{B}) \wedge (\forall x)\mathscr{C}$

d. $\vdash (\forall y_1) \ldots (\forall y_n)\mathscr{B} \Rightarrow \mathscr{B}$

e. $\vdash \neg(\forall x)\mathscr{B} \Rightarrow (\exists x)\neg\mathscr{B}$

2.28[D] Let K be a first-order theory and let $K^{\#}$ be an axiomatic theory having the following axioms:

a. $(\forall y_1) \ldots (\forall y_n)\mathscr{B}$, where $\mathscr{B}$ is any axiom of K and $y_1, \ldots, y_n (n \geq 0)$ are any variables (none at all when $n = 0$);

b. $(\forall y_1) \ldots (\forall y_n)(\mathscr{B} \Rightarrow \mathscr{C}) \Rightarrow [(\forall y_1) \ldots (\forall y_n)\mathscr{B} \Rightarrow (\forall y_1) \ldots (\forall y_n)\mathscr{C}]$ where $\mathscr{B}$ and $\mathscr{C}$ are any wfs and $y_1 \ldots, y_n$ are any variables.

Moreover, $K^{\#}$ has modus ponens as its only rule of inference. Show that $K^{\#}$ has the same theorems as K. Thus, at the expense of adding more axioms, the generalization rule can be dispensed with.

2.29 Carry out the proof of the Extension of Propositions 2.4–2.7 above.

2.5 Additional Metatheorems and Derived Rules

For the sake of smoothness in working with particular theories later, we shall introduce various techniques for constructing proofs. In this section it is assumed that we are dealing with an arbitrary theory K.

Often one wants to obtain $\mathscr{B}(t)$ from $(\forall x)\mathscr{B}(x)$, where t is a term free for x in $\mathscr{B}(x)$. This is allowed by the following *derived rule*.

2.5.1 Particularization Rule A4

If t is free for x in $\mathscr{B}(x)$, then $(\forall x)\mathscr{B}(x) \vdash \mathscr{B}(t)$.*

Proof

From $(\forall x)\mathscr{B}(x)$ and the instance $(\forall x)\mathscr{B}(x) \Rightarrow \mathscr{B}(t)$ of axiom (A4), we obtain $\mathscr{B}(t)$ by modus ponens.

Since x is free for x in $\mathscr{B}(x)$, a special case of rule A4 is: $(\forall x)\mathscr{B} \vdash \mathscr{B}$.

There is another very useful derived rule, which is essentially the contrapositive of rule A4.

2.5.2 Existential Rule E4

Let t be a term that is free for x in a wf $\mathscr{B}(x, t)$, and let $\mathscr{B}(t, t)$ arise from $\mathscr{B}(x, t)$ by replacing all free occurrences of x by t. ($\mathscr{B}(x, t)$ may or may not contain occurrences of t.) Then, $\mathscr{B}(t, t) \vdash (\exists x)\mathscr{B}(x, t)$

Proof

It suffices to show that $\vdash \mathscr{B}(t, t) \Rightarrow (\exists x)\mathscr{B}(x, t)$. But, by axiom (A4), $\vdash (\forall x)\neg\mathscr{B}(x, t) \Rightarrow \neg\mathscr{B}(t, t)$. Hence, by the tautology $(A \Rightarrow \neg B) \Rightarrow (B \Rightarrow \neg A)$ and MP, $\vdash \mathscr{B}(t, t) \Rightarrow \neg(\forall x)\neg\mathscr{B}(x, t)$, which, in abbreviated form, is $\vdash \mathscr{B}(t, t) \Rightarrow (\exists x)\mathscr{B}(x, t)$.

A special case of rule E4 is $\mathscr{B}(t) \vdash (\exists x)\mathscr{B}(x)$, whenever t is free for x in $\mathscr{B}(x)$. In particular, when t is x itself, $\mathscr{B}(x) \vdash (\exists x)\mathscr{B}(x)$.

Example

$\vdash (\forall x)\mathscr{B} \Rightarrow (\exists x)\mathscr{B}$

1. $(\forall x)\mathscr{B}$	Hyp
2. $\mathscr{B}$	1, rule A4
3. $(\exists x)\mathscr{B}$	2, rule E4
4. $(\forall x)\mathscr{B} \vdash (\exists x)\mathscr{B}$	1–3
5. $\vdash (\forall x)\mathscr{B} \Rightarrow (\exists x)\mathscr{B}$	1–4, Corollary 2.6

The following derived rules are extremely useful.

Negation elimination:	$\neg\neg\mathscr{B} \vdash \mathscr{B}$
Negation introduction:	$\mathscr{B} \vdash \neg\neg\mathscr{B}$
Conjunction elimination:	$\mathscr{B} \wedge \mathscr{C} \vdash \mathscr{B}$
	$\mathscr{B} \wedge \mathscr{C} \vdash \mathscr{C}$
	$\neg(\mathscr{B} \wedge \mathscr{C}) \vdash \neg\mathscr{B} \vee \neg\mathscr{C}$

* From a strict point of view, $(\forall x)\mathscr{B}(x) \vdash \mathscr{B}(t)$ states a fact about derivability. Rule A4 should be taken to mean that, if $(\forall x)\mathscr{B}(x)$ occurs as a step in a proof, we may write $\mathscr{B}(t)$ as a later step (if t is free for x in $\mathscr{B}(x)$). As in this case, we shall often state a derived rule in the form of the corresponding derivability result that justifies the rule.

Conjunction introduction:	$\mathscr{B}, \mathscr{C} \vdash \mathscr{B} \wedge \mathscr{C}$
Disjunction elimination:	$\mathscr{B} \vee \mathscr{C}, \neg\mathscr{B} \vdash \mathscr{C}$
	$\mathscr{B} \vee \mathscr{C}, \neg\mathscr{C} \vdash \mathscr{B}$
	$\neg(\mathscr{B} \vee \mathscr{C}) \vdash \neg\mathscr{B} \wedge \neg\mathscr{C}$
	$\mathscr{B} \Rightarrow \mathscr{D}, \mathscr{C} \Rightarrow \mathscr{D}, \mathscr{B} \vee \mathscr{C} \vdash \mathscr{D}$
Disjunction introduction:	$\mathscr{B} \vdash \mathscr{B} \vee \mathscr{C}$
	$\mathscr{C} \vdash \mathscr{B} \vee \mathscr{C}$
Conditional elimination:	$\mathscr{B} \Rightarrow \mathscr{C}, \neg\mathscr{C} \vdash \neg\mathscr{B}$
	$\mathscr{B} \Rightarrow \neg\mathscr{C}, \mathscr{C} \vdash \neg\mathscr{B}$
	$\neg\mathscr{B} \Rightarrow \mathscr{C}, \neg\mathscr{C} \vdash \mathscr{B}$
	$\neg\mathscr{B} \Rightarrow \neg\mathscr{C}, \mathscr{C} \vdash \mathscr{B}$
	$\neg(\mathscr{B} \Rightarrow \mathscr{C}) \vdash \mathscr{B}$
	$\neg(\mathscr{B} \Rightarrow \mathscr{C}) \vdash \neg\mathscr{C}$
Conditional introduction:	$\mathscr{B}, \neg\mathscr{C} \vdash \neg(\mathscr{B} \Rightarrow \mathscr{C})$
Conditional contrapositive:	$\mathscr{B} \Rightarrow \mathscr{C} \vdash \neg\mathscr{C} \Rightarrow \neg\mathscr{B}$
	$\neg\mathscr{C} \Rightarrow \neg\mathscr{B} \vdash \mathscr{B} \Rightarrow \mathscr{C}$
Biconditional elimination:	$\mathscr{B} \Leftrightarrow \mathscr{C}, \mathscr{B} \vdash \mathscr{C}\ \mathscr{B} \Leftrightarrow \mathscr{C}, \neg\mathscr{B} \vdash \neg\mathscr{C}$
	$\mathscr{B} \Leftrightarrow \mathscr{C}, \mathscr{C} \vdash \mathscr{B}\ \mathscr{B} \Leftrightarrow \mathscr{C}, \neg\mathscr{C} \vdash \neg\mathscr{B}$
	$\mathscr{B} \Leftrightarrow \mathscr{C} \vdash \mathscr{B} \Rightarrow \mathscr{C}\ \mathscr{B} \Leftrightarrow \mathscr{C} \vdash \mathscr{C} \Rightarrow \mathscr{B}$
Biconditional introduction:	$\mathscr{B} \Rightarrow \mathscr{C}, \mathscr{C} \Rightarrow \mathscr{B} \vdash \mathscr{B} \Leftrightarrow \mathscr{C}$
Biconditional negation:	$\mathscr{B} \Leftrightarrow \mathscr{C} \vdash \neg\mathscr{B} \Leftrightarrow \neg\mathscr{C}$
	$\neg\mathscr{B} \Leftrightarrow \neg\mathscr{C} \vdash \mathscr{B} \Leftrightarrow \mathscr{C}$

Proof by contradiction: If a proof of $\Gamma, \neg\mathscr{B} \vdash \mathscr{C} \wedge \neg\mathscr{C}$ involves no application of Gen using a variable free in $\mathscr{B}$, then $\Gamma \vdash \mathscr{B}$. (Similarly, one obtains $\Gamma \vdash \neg\mathscr{B}$ from $\Gamma, \mathscr{B} \vdash \mathscr{C} \wedge \neg\mathscr{C}$.)

Exercises

2.30 Justify the derived rules listed above.

2.31 Prove the following.

a. $\vdash (\forall x)(\forall y)A_1^2(x, y) \Rightarrow (\forall x)A_1^2(x, x)$

b. $\vdash [(\forall x)\mathscr{B}] \vee [(\forall x)\mathscr{C}] \Rightarrow (\forall x)(\mathscr{B} \vee \mathscr{C})$

c. $\vdash \neg(\exists x)\mathscr{B} \Rightarrow (\forall x)\neg\mathscr{B}$

d. $\vdash (\forall x)\mathscr{B} \Rightarrow (\forall x)(\mathscr{B} \vee \mathscr{C})$

e. $\vdash (\forall x)(\forall y)(A_1^2(x, y) \Rightarrow \neg A_1^2(y, x)) \Rightarrow (\forall x)\neg A_1^2(x, x)$

f. $\vdash [(\exists x)\mathscr{B} \Rightarrow (\forall x)\mathscr{C}] \Rightarrow (\forall x)(\mathscr{B} \Rightarrow \mathscr{C})$

g. $\vdash (\forall x)(\mathscr{B} \vee \mathscr{C}) \Rightarrow [(\forall x)\mathscr{B}] \vee (\exists x)\mathscr{C}$

h. $\vdash (\forall x)(A_1^2(x, x) \Rightarrow (\exists y)A_1^2(x, y))$

i. $\vdash (\forall x)(\mathscr{B} \Rightarrow \mathscr{C}) \Rightarrow [(\forall x) \neg\mathscr{C} \Rightarrow (\forall x) \neg\mathscr{B}]$
j. $\vdash (\exists y)[A_1^1(y) \Rightarrow (\forall y)A_1^1(y)]$
k. $\vdash [(\forall x)(\forall y)(\mathscr{B}(x, y) \Rightarrow \mathscr{B}(y, x)) \wedge (\forall x)(\forall y)(\forall z)(\mathscr{B}(x, y) \wedge \mathscr{B}(y, z) \Rightarrow \mathscr{B}(x, z))] \Rightarrow (\forall x)(\forall y)(\mathscr{B}(x, y) \Rightarrow \mathscr{B}(x, x))$
l. $\vdash (\exists x)A_1^2(x,x) \Rightarrow (\exists x)(\exists y)A_1^2(x,y)$

2.32 Assume that $\mathscr{B}$ and $\mathscr{C}$ are wfs and that x is not free in $\mathscr{B}$. Prove the following.

a. $\vdash \mathscr{B} \Rightarrow (\forall x)\mathscr{B}$
b. $\vdash (\exists x)\mathscr{B} \Rightarrow \mathscr{B}$
c. $\vdash (\mathscr{B} \Rightarrow (\forall x)\mathscr{C}) \Leftrightarrow (\forall x)(\mathscr{B} \Rightarrow \mathscr{C})$
d. $\vdash ((\exists x)\mathscr{C} \Rightarrow \mathscr{B}) \Leftrightarrow (\forall x)(\mathscr{C} \Rightarrow \mathscr{B})$

We need a derived rule that will allow us to replace a part $\mathscr{C}$ of a wf $\mathscr{B}$ by a wf that is provably equivalent to $\mathscr{C}$. For this purpose, we first must prove the following auxiliary result.

Lemma 2.8

For any wfs $\mathscr{B}$ and $\mathscr{C}$, $\vdash (\forall x)(\mathscr{B} \Leftrightarrow \mathscr{C}) \Rightarrow ((\forall x)\mathscr{B} \Leftrightarrow (\forall x)\mathscr{C})$.

Proof

1. $(\forall x)(\mathscr{B} \Leftrightarrow \mathscr{C})$	Hyp
2. $(\forall x)\mathscr{B}$	Hyp
3. $\mathscr{B} \Leftrightarrow \mathscr{C}$	1, rule A4
4. $\mathscr{B}$	2, rule A4
5. $\mathscr{C}$	3, 4, biconditional elimination
6. $(\forall x)\mathscr{C}$	5, Gen
7. $(\forall x)(\mathscr{B} \Leftrightarrow \mathscr{C}), (\forall x)\mathscr{B} \vdash (\forall x)\mathscr{C}$	1–6
8. $(\forall x)(\mathscr{B} \Leftrightarrow \mathscr{C}) \vdash (\forall x)\mathscr{B} \Rightarrow (\forall x)\mathscr{C}$	1–7, Corollary 2.6
9. $(\forall x)(\mathscr{B} \Leftrightarrow \mathscr{C}) \vdash (\forall x)\mathscr{C} \Rightarrow (\forall x)\mathscr{B}$	Proof like that of 8
10. $(\forall x)(\mathscr{B} \Leftrightarrow \mathscr{C}) \vdash (\forall x)\mathscr{B} \Leftrightarrow (\forall x)\mathscr{C}$	8, 9, Biconditional introduction
11. $\vdash (\forall x)(\mathscr{B} \Leftrightarrow \mathscr{C}) \Rightarrow ((\forall x)\mathscr{B} \Leftrightarrow (\forall x)\mathscr{C})$	1–10, Corollary 2.6

Proposition 2.9

If $\mathscr{C}$ is a subformula of $\mathscr{B}$, $\mathscr{B}'$ is the result of replacing zero or more occurrences of $\mathscr{C}$ in $\mathscr{B}$ by a wf $\mathscr{D}$, and every free variable of $\mathscr{C}$ or $\mathscr{D}$ that is also a bound variable of $\mathscr{B}$ occurs in the list $y_1, \ldots, y_k$, then:

a. $\vdash [(\forall y_1) \ldots (\forall y_k)(\mathscr{C} \Leftrightarrow \mathscr{D})] \Rightarrow (\mathscr{B} \Leftrightarrow \mathscr{B}')$ (Equivalence theorem)
b. If $\vdash \mathscr{C} \Leftrightarrow \mathscr{D}$, then $\vdash \mathscr{B} \Leftrightarrow \mathscr{B}'$ (Replacement theorem)
c. If $\vdash \mathscr{C} \Leftrightarrow \mathscr{D}$ and $\vdash \mathscr{B}$, then $\vdash \mathscr{B}'$

Example

a. $\vdash (\forall x)(A_1^1(x) \Leftrightarrow A_2^1(x)) \Rightarrow [(\exists x)A_1^1(x) \Leftrightarrow (\exists x)A_2^1(x)]$

Proof

a. We use induction on the number of connectives and quantifiers in $\mathscr{B}$. Note that, if zero occurrences are replaced, $\mathscr{B}'$ is $\mathscr{B}$ and the wf to be proved is an instance of the tautology $A \Rightarrow (B \Leftrightarrow B)$. Note also that, if $\mathscr{C}$ is identical with $\mathscr{B}$ and this occurrence of $\mathscr{C}$ is replaced by $\mathscr{D}$, the wf to be proved, $[(\forall y_1) \dots (\forall y_k)(\mathscr{C} \Leftrightarrow \mathscr{D})] \Rightarrow (\mathscr{B} \Leftrightarrow \mathscr{B}')$, is derivable by Exercise 2.27(d). Thus, we may assume that $\mathscr{C}$ is a proper part of $\mathscr{B}$ and that at least one occurrence of $\mathscr{C}$ is replaced. Our inductive hypothesis is that the result holds for all wfs with fewer connectives and quantifiers than $\mathscr{B}$.

   Case 1. $\mathscr{B}$ is an atomic wf. Then $\mathscr{C}$ cannot be a proper part of $\mathscr{B}$.

   Case 2. $\mathscr{B}$ is $\neg\mathscr{E}$. Let $\mathscr{B}'$ be $\neg\mathscr{E}'$. By inductive hypothesis, $\vdash [(\forall y_1) \dots (\forall y_k)(\mathscr{C} \Leftrightarrow \mathscr{D})] \Rightarrow (\mathscr{E} \Leftrightarrow \mathscr{E}')$. Hence, by a suitable instance of the tautology $(C \Rightarrow (A \Leftrightarrow B)) \Rightarrow (C \Rightarrow (\neg A \Leftrightarrow \neg B))$ and MP, we obtain $\vdash [(\forall y_1) \dots (\forall y_k)(\mathscr{C} \Leftrightarrow \mathscr{D})] \Rightarrow (\mathscr{B} \Leftrightarrow \mathscr{B}')$.

   Case 3. $\mathscr{B}$ is $\mathscr{E} \Rightarrow \mathscr{F}$. Let $\mathscr{B}'$ be $\mathscr{E}' \Rightarrow \mathscr{F}'$. By inductive hypothesis, $\vdash [(\forall y_1) \dots (\forall y_k)(\mathscr{C} \Leftrightarrow \mathscr{D})] \Rightarrow (\mathscr{E} \Leftrightarrow \mathscr{E}')$ and $\vdash [(\forall y_1) \dots (\forall y_k)(\mathscr{C} \Leftrightarrow \mathscr{D})] \Rightarrow (\mathscr{F} \Leftrightarrow \mathscr{F}')$. Using a suitable instance of the tautology

$$(A \Rightarrow (B \Leftrightarrow C)) \wedge (A \Rightarrow (D \Leftrightarrow E)) \Rightarrow (A \Rightarrow [(B \Rightarrow D) \Leftrightarrow (C \Rightarrow E)])$$

   we obtain $\vdash [(\forall y_1) \dots (\forall y_k)(\mathscr{C} \Leftrightarrow \mathscr{D})] \Rightarrow (\mathscr{B} \Leftrightarrow \mathscr{B}')$.

   Case 4. $\mathscr{B}$ is $(\forall x)\mathscr{E}$. Let $\mathscr{B}'$ be $(\forall x)\mathscr{E}'$. By inductive hypothesis, $\vdash [(\forall y_1) \dots (\forall y_k)(\mathscr{C} \Leftrightarrow \mathscr{D})] \Rightarrow (\mathscr{E} \Leftrightarrow \mathscr{E}')$. Now, x does not occur free in $(\forall y_1) \dots (\forall y_k)(\mathscr{C} \Leftrightarrow \mathscr{D})$ because, if it did, it would be free in $\mathscr{C}$ or $\mathscr{D}$ and, since it is bound in $\mathscr{B}$, it would be one of $y_1, \dots, y_k$ and it would not be free in $(\forall y_1) \dots (\forall y_k)(\mathscr{C} \Leftrightarrow \mathscr{D})$. Hence, using axiom (A5), we obtain $\vdash (\forall y_1) \dots (\forall y_k)(\mathscr{C} \Leftrightarrow \mathscr{D}) \Rightarrow (\forall x)(\mathscr{E} \Leftrightarrow \mathscr{E}')$. However, by Lemma 2.8, $\vdash (\forall x)(\mathscr{E} \Leftrightarrow \mathscr{E}) \Rightarrow ((\forall x)\mathscr{E} \Leftrightarrow (\forall x)\mathscr{E}')$. Then, by a suitable tautology and MP, $\vdash [(\forall y_1) \dots (\forall y_k)(\mathscr{C} \Leftrightarrow \mathscr{D})] \Rightarrow (\mathscr{B} \Leftrightarrow \mathscr{B}')$.

b. From $\vdash \mathscr{C} \Leftrightarrow \mathscr{D}$, by several applications of Gen, we obtain $\vdash (\forall y_1) \dots (\forall y_k)(\mathscr{C} \Leftrightarrow \mathscr{D})$. Then, by (a) and MP, $\vdash \mathscr{B} \Leftrightarrow \mathscr{B}'$.

c. Use part (b) and biconditional elimination.

Exercises

2.33 Prove the following:

   a. $\vdash (\exists x) \neg\mathscr{B} \Leftrightarrow \neg(\forall x)\mathscr{B}$

   b. $\vdash (\forall x)\mathscr{B} \Leftrightarrow \neg(\exists x) \neg\mathscr{B}$

   c. $\vdash (\exists x)(\mathscr{B} \Rightarrow \neg(\mathscr{C} \vee \mathscr{D})) \Rightarrow (\exists x)(\mathscr{B} \Rightarrow \neg\mathscr{C} \wedge \neg\mathscr{D})$

d. $\vdash (\forall x)(\exists y)(\mathscr{B} \Rightarrow \mathscr{C}) \Leftrightarrow (\forall x)(\exists y)(\neg\mathscr{B} \vee \mathscr{C})$
e. $\vdash (\forall x)(\mathscr{B} \Rightarrow \neg\mathscr{C}) \Leftrightarrow \neg(\exists x)(\mathscr{B} \wedge \mathscr{C})$

2.34 Show by a counterexample that we cannot omit the quantifiers $(\forall y_1) \ldots (\forall y_k)$ in Proposition 2.9(a).

2.35 If $\mathscr{C}$ is obtained from $\mathscr{B}$ by erasing all quantifiers $(\forall x)$ or $(\exists x)$ whose scope does not contain x free, prove that $\vdash \mathscr{B} \Leftrightarrow \mathscr{C}$.

2.36 For each wf $\mathscr{B}$ below, find a wf $\mathscr{C}$ such that $\vdash \mathscr{C} \Leftrightarrow \neg\mathscr{B}$ and negation signs in $\mathscr{C}$ apply only to atomic wfs.

a. $(\forall x)(\forall y)(\exists z)A_1^3(x, y, z)$
b. $(\forall \varepsilon)(\varepsilon > 0 \Rightarrow (\exists \delta)(\delta > 0 \wedge (\forall x)(|x - c| < \delta \Rightarrow |f(x) - f(c)| < \varepsilon))$
c. $(\forall \varepsilon)(\varepsilon > 0 \Rightarrow (\exists n)(\forall m)(m > n \Rightarrow |a_m - b| < \varepsilon))$

2.37 Let $\mathscr{B}$ be a wf that does not contain $\Rightarrow$ and $\Leftrightarrow$. Exchange universal and existential quantifiers and exchange $\wedge$ and $\vee$. The result $\mathscr{B}^*$ is called the *dual* of $\mathscr{B}$.

a. In any predicate calculus, prove the following.
  i. $\vdash \mathscr{B}$ if and only if $\vdash \neg\mathscr{B}^*$
  ii. $\vdash \mathscr{B} \Rightarrow \mathscr{C}$ if and only if $\vdash \mathscr{C}^* \Rightarrow \mathscr{B}^*$
  iii. $\vdash \mathscr{B} \Leftrightarrow \mathscr{C}$ if and only if $\vdash \mathscr{B}^* \Leftrightarrow \mathscr{C}^*$
  iv. $\vdash (\exists x)(\mathscr{B} \vee \mathscr{C}) \Leftrightarrow [((\exists x)\mathscr{B}) \vee (\exists x)\mathscr{C}]$. [*Hint*: Use Exercise 2.27(c).]

b. Show that the duality results of part (a), (i)–(iii), do not hold for arbitrary theories.

2.6 Rule C

It is very common in mathematics to reason in the following way. Assume that we have proved a wf of the form $(\exists x)\mathscr{B}(x)$. Then we say, let b be an object such that $\mathscr{B}(b)$. We continue the proof, finally arriving at a formula that does not involve the arbitrarily chosen element b.

For example, let us say that we wish to show that $(\exists x)(\mathscr{B}(x) \Rightarrow \mathscr{C}(x))$, $(\forall x)\mathscr{B}(x) \vdash (\exists x)\mathscr{C}(x)$.

1. $(\exists x)(\mathscr{B}(x) \Rightarrow \mathscr{C}(x))$	Hyp
2. $(\forall x)\mathscr{B}(x)$	Hyp
3. $\mathscr{B}(b) \Rightarrow \mathscr{C}(b)$ for some b	1
4. $\mathscr{B}(b)$	2, rule A4
5. $\mathscr{C}(b)$	3, 4, MP
6. $(\exists x)\mathscr{C}(x)$	5, rule E4

Such a proof seems to be perfectly legitimate on an intuitive basis. In fact, we can achieve the same result without making an arbitrary choice of an element b as in step 3. This can be done as follows:

1. $(\forall x)\mathscr{B}(x)$	Hyp
2. $(\forall x)\neg\mathscr{C}(x)$	Hyp
3. $\mathscr{B}(x)$	1, rule A4
4. $\neg\mathscr{C}(x)$	2, rule A4
5. $\neg(\mathscr{B}(x) \Rightarrow \mathscr{C}(x))$	3, 4, conditional introduction
6. $(\forall x)\neg(\mathscr{B}(x) \Rightarrow \mathscr{C}(x))$	5, Gen
7. $(\forall x)\mathscr{B}(x), (\forall x)\neg\mathscr{C}(x) \vdash (\forall x)\neg(\mathscr{B}(x) \Rightarrow \mathscr{C}(x))$	1–6
8. $(\forall x)\mathscr{B}(x) \vdash (\forall x)\neg\mathscr{C}(x) \Rightarrow (\forall x)\neg(\mathscr{B}(x) \Rightarrow \mathscr{C}(x))$	1–7, corollary 2.6
9. $(\forall x)\mathscr{B}(x) \vdash \neg(\forall x)\neg(\mathscr{B}(x) \Rightarrow \mathscr{C}(x)) \Rightarrow \neg(\forall x)\neg\mathscr{C}(x)$	8, contrapositive
10. $(\forall x)\mathscr{B}(x) \vdash (\exists x)(\mathscr{B}(x) \Rightarrow \mathscr{C}(x)) \Rightarrow (\exists x)\mathscr{C}(x)$	Abbreviation of 9
11. $(\exists x)(\mathscr{B}(x) \Rightarrow \mathscr{C}(x)), (\forall x)\mathscr{B}(x) \vdash (\exists x)\mathscr{C}(x)$	10, MP

In general, any wf that can be proved using a finite number of arbitrary choices can also be proved without such acts of choice. We shall call the rule that permits us to go from $(\exists x)\mathscr{B}(x)$ to $\mathscr{B}(b)$, *rule C* ("C" for "choice"). More precisely, a rule C deduction in a first-order theory K is defined in the following manner: $\Gamma \vdash_C \mathscr{B}$ if and only if there is a sequence of wfs $\mathscr{D}_1, \ldots, \mathscr{D}_n$ such that $\mathscr{D}_n$ is $\mathscr{B}$ and the following four conditions hold:

1. For each $i < n$, either
   a. $\mathscr{D}_i$ is an axiom of K, or
   b. $\mathscr{D}_i$ is in Γ, or
   c. $\mathscr{D}_i$ follows by MP or Gen from preceding wfs in the sequence, or
   d. there is a preceding wf $(\exists x)\mathscr{C}(x)$ such that $\mathscr{D}_i$ is $\mathscr{C}(d)$, where d is a new individual constant (rule C).
2. As axioms in condition 1(a), we also can use all logical axioms that involve the new individual constants already introduced in the sequence by applications of rule C.

3. No application of Gen is made using a variable that is free in some $(\exists x)\mathscr{C}(x)$ to which rule C has been previously applied.
4. $\mathscr{B}$ contains none of the new individual constants introduced in the sequence in any application of rule C.

A word should be said about the reason for including condition 3. If an application of rule C to a wf $(\exists x)\mathscr{C}(x)$ yields $\mathscr{C}(d)$, then the object referred to by d may depend on the values of the free variables in $(\exists x)\mathscr{C}(x)$. So that one object may not satisfy $\mathscr{C}(x)$ for *all* values of the free variables in $(\exists x)\mathscr{C}(x)$. For example, without clause 3, we could proceed as follows:

1. $(\forall x)(\exists y)A_1^2(x,y)$	Hyp
2. $(\exists y)A_1^2(x,y)$	1, rule A4
3. $A_1^2(x,d)$	2, rule C
4. $(\forall x)A_1^2(x,d)$	3, Gen
5. $(\exists y)(\forall x)A_1^2(x,y)$	4, rule E4

However, there is an interpretation for which $(\forall x)(\exists y)A_1^2(x,y)$ is true but $(\exists y)(\forall x)A_1^2(x,y)$ is false. Take the domain to be the set of integers and let $A_1^2(x,y)$ mean that $x < y$.

Proposition 2.10

If $\Gamma \vdash_C \mathscr{B}$, then $\Gamma \vdash \mathscr{B}$. Moreover, from the following proof it is easy to verify that, if there is an application of Gen in the new proof of $\mathscr{B}$ from Γ using a certain variable and applied to a wf depending upon a certain wf of Γ, then there was such an application of Gen in the original proof.*

Proof

Let $(\exists y_1)\mathscr{C}_1(y_1)$, …, $(\exists y_k)\mathscr{C}_k(y_k)$ be the wfs in order of occurrence to which rule C is applied in the proof of $\Gamma \vdash_C \mathscr{B}$, and let d_1, …, d_k be the corresponding new individual constants. Then $\Gamma, \mathscr{C}_1(d_1), \ldots, \mathscr{C}_k(d_k) \vdash \mathscr{B}$. Now, by condition 3 of the definition above, Corollary 2.6 is applicable, yielding $\Gamma, \mathscr{C}_1(d_1), \ldots, \mathscr{C}_{k-1}(d_{k-1}) \vdash \mathscr{C}_k(d_k) \Rightarrow \mathscr{B}$. We replace d_k everywhere by a variable z that does not occur in the proof.

Then

$$\Gamma, \mathscr{C}_1(d_1), \ldots, \mathscr{C}_{k-1}(d_{k-1}) \vdash \mathscr{C}_k(z) \Rightarrow \mathscr{B}$$

* The first formulation of a version of rule C similar to that given here seems to be due to Rosser (1953).

and, by Gen,

$$\Gamma, \mathscr{C}_1(d_1), \dots, \mathscr{C}_{k-1}(d_{k-1}) \vdash (\forall z)(\mathscr{C}_k(z) \Rightarrow \mathscr{B})$$

Hence, by Exercise 2.32(d),

$$\Gamma, \mathscr{C}_1(d_1), \dots, \mathscr{C}_{k-1}(d_{k-1}) \vdash (\exists y_k)\mathscr{C}_k(y_k) \Rightarrow \mathscr{B}$$

But,

$$\Gamma, \mathscr{C}_1(d_1), \dots, \mathscr{C}_{k-1}(d_{k-1}) \vdash (\exists y_k)\mathscr{C}_k(y_k)$$

Hence, by MP,

$$\Gamma, \mathscr{C}_1(d_1), \dots, \mathscr{C}_{k-1}(d_{k-1}) \vdash \mathscr{B}$$

Repeating this argument, we can eliminate $\mathscr{C}_{k-1}(d_{k-1}), \dots, \mathscr{C}_1(d_1)$ one after the other, finally obtaining $\Gamma \vdash \mathscr{B}$.

Example

$$\vdash (\forall x)(\mathscr{B}(x) \Rightarrow \mathscr{C}(x)) \Rightarrow ((\exists x)\mathscr{B}(x) \Rightarrow (\exists x)\mathscr{C}(x))$$

1. $(\forall x)((\mathscr{B}x) \Rightarrow \mathscr{C}(x))$	Hyp
2. $(\exists x)\mathscr{B}(x)$	Hyp
3. $\mathscr{B}(d)$	2, rule C
4. $\mathscr{B}(d) \Rightarrow \mathscr{C}(d)$	1, rule A4
5. $\mathscr{C}(d)$	3, 4, MP
6. $(\exists x)\mathscr{C}(x)$	5, rule E4
7. $(\forall x)(\mathscr{B}(x) \Rightarrow \mathscr{C}(x)), (\exists x)\mathscr{B}(x) \vdash_C (\exists x)\mathscr{C}(x)$	1–6
8. $(\forall x)(\mathscr{B}(x) \Rightarrow \mathscr{C}(x)), (\exists x)\mathscr{B}(x) \vdash (\exists x)\mathscr{C}(x)$	7, Proposition 2.10
9. $(\forall x)(\mathscr{B}(x) \Rightarrow \mathscr{C}(x)) \vdash (\exists x)\mathscr{B}(x) \Rightarrow (\exists x)\mathscr{C}(x)$	1–8, corollary 2.6
10. $\vdash (\forall x)(\mathscr{B}(x) \Rightarrow \mathscr{C}(x)) \Rightarrow ((\exists x)\mathscr{B}(x) \Rightarrow (\exists x)\mathscr{C}(x))$	1–9, corollary 2.6

Exercises

Use rule C and Proposition 2.10 to prove Exercises 2.38–2.45.

2.38 $\vdash (\exists x)(\mathscr{B}(x) \Rightarrow \mathscr{C}(x)) \Rightarrow ((\forall x)\mathscr{B}(x) \Rightarrow (\exists x)\mathscr{C}(x))$

2.39 $\vdash \neg(\exists y)(\forall x)(A_1^2(x,y) \Leftrightarrow \neg A_1^2(x,x))$

2.40 $\vdash [(\forall x)(A_1^1(x) \Rightarrow A_2^1(x) \vee A_3^1(x)) \wedge \neg(\forall x)(A_1^1(x) \Rightarrow A_2^1(x))] \Rightarrow (\exists x)(A_1^1(x) \wedge A_3^1(x))$

2.41 $\vdash [(\exists x)\mathscr{B}(x)] \wedge [(\forall x)\mathscr{C}(x)] \Rightarrow (\exists x)(\mathscr{B}(x) \wedge \mathscr{C}(x))$

2.42 $\vdash (\exists x)\mathscr{C}(x) \Rightarrow (\exists x)(\mathscr{B}(x) \vee \mathscr{C}(x))$

2.43 $\vdash (\exists x)(\exists y)\mathscr{B}(x, y) \Leftrightarrow (\exists y)(\exists x)\mathscr{B}(x, y)$

2.44 $\vdash (\exists x)(\forall y)\mathscr{B}(x, y) \Rightarrow (\forall y)(\exists x)\mathscr{B}(x, y)$

2.45 $\vdash (\exists x)(\mathscr{B}(x) \wedge \mathscr{C}(x)) \Rightarrow ((\exists x)\mathscr{B}(x)) \wedge (\exists x)\mathscr{C}(x)$

2.46 What is wrong with the following alleged derivations?

a.

1.	$(\exists x)\mathscr{B}(x)$	Hyp
2.	$\mathscr{B}(d)$	1, rule C
3.	$(\exists x)\mathscr{C}(x)$	Hyp
4.	$\mathscr{C}(d)$	3, rule C
5.	$\mathscr{B}(d) \wedge \mathscr{C}(d)$	2, 4, conjunction introduction
6.	$(\exists x)(\mathscr{B}(x) \wedge \mathscr{C}(x))$	5, rule E4
7.	$(\exists x)\mathscr{B}(x), (\exists x)\mathscr{C}(x) \vdash (\exists x)(\mathscr{B}(x) \wedge \mathscr{C}(x))$	1–6, Proposition 2.10

b.

1.	$(\exists x)(\mathscr{B}(x) \Rightarrow \mathscr{C}(x))$	Hyp
2.	$(\exists x)\mathscr{B}(x)$	Hyp
3.	$\mathscr{B}(d) \Rightarrow \mathscr{C}(d)$	1, rule C
4.	$\mathscr{B}(d)$	2, rule C
5.	$\mathscr{C}(d)$	3, 4, MP
6.	$(\exists x)\mathscr{C}(x)$	5, rule E4
7.	$(\exists x)(\mathscr{B}(x) \Rightarrow \mathscr{C}(x)), (\exists x)\mathscr{B}(x) \vdash (\exists x)\mathscr{C}(x)$	1–6, Proposition 2.10

2.7 Completeness Theorems

We intend to show that the theorems of a first-order predicate calculus K are precisely the same as the logically valid wfs of K. Half of this result was proved in Proposition 2.2. The other half will follow from a much more general proposition established later. First we must prove a few preliminary lemmas.

If x_i and x_j are distinct, then $\mathscr{B}(x_i)$ and $\mathscr{B}(x_j)$ are said to be *similar* if and only if x_j is free for x_i in $\mathscr{B}(x_i)$ and $\mathscr{B}(x_i)$ has no free occurrences of x_j. It is assumed here that $\mathscr{B}(x_j)$ arises from $\mathscr{B}(x_i)$ by substituting x_j for all free occurrences of x_i. It is easy to see that, if $\mathscr{B}(x_i)$ and $\mathscr{B}(x_j)$ are similar, then x_i is free for x_j in

$\mathscr{B}(x_j)$ and $\mathscr{B}(x_j)$ has no free occurrences of x_i. Thus, if $\mathscr{B}(x_i)$ and $\mathscr{B}(x_j)$ are similar, then $\mathscr{B}(x_j)$ and $\mathscr{B}(x_i)$ are similar. Intuitively, $\mathscr{B}(x_i)$ and $\mathscr{B}(x_j)$ are similar if and only if $\mathscr{B}(x_i)$ and $\mathscr{B}(x_j)$ are the same except that $\mathscr{B}(x_i)$ has free occurrences of x_i in exactly those places where $\mathscr{B}(x_j)$ has free occurrences of x_j.

Example

$$(\forall x_3)\left[A_1^2(x_1, x_3) \vee A_1^1(x_1)\right] \quad \text{and} \quad (\forall x_3\left[[A_1^2(x_2, x_3) \vee A_1^1(x_2)\right] \text{are similar.}$$

Lemma 2.11

If $\mathscr{B}(x_i)$ and $\mathscr{B}(x_j)$ are similar, then $\vdash (\forall x_i)\mathscr{B}(x_i) \Leftrightarrow (\forall x_j)\mathscr{B}(x_j)$.

Proof

$\vdash (\forall x_i)\mathscr{B}(x_i) \Rightarrow \mathscr{B}(x_j)$ by axiom (A4). Then, by Gen, $\vdash (\forall x_j)((\forall x_i)\mathscr{B}(x_i) \Rightarrow \mathscr{B}(x_j))$, and so, by axiom (A5) and MP, $\vdash (\forall x_i)\mathscr{B}(x_i) \Rightarrow (\forall x_j)\mathscr{B}(x_j)$. Similarly, $\vdash (\forall x_j)\ \mathscr{B}(x_j) \Rightarrow (\forall x_i)\mathscr{B}(x_i)$. Hence, by biconditional introduction, $\vdash (\forall x_i)\mathscr{B}(x_i) \Leftrightarrow (\forall x_j)\mathscr{B}(x_j)$.

Exercises

2.47 If $\mathscr{B}(x_i)$ and $\mathscr{B}(x_j)$ are similar, prove that $\vdash (\exists x_i)\mathscr{B}(x_i) \Leftrightarrow (\exists x_j)\mathscr{B}(x_j)$.

2.48 *Change of bound variables*. If $\mathscr{B}(x)$ is similar to $\mathscr{B}(y)$, $(\forall x)\mathscr{B}(x)$ is a subformula of $\mathscr{C}$, and $\mathscr{C}'$ is the result of replacing one or more occurrences of $(\forall x)\mathscr{B}(x)$ in $\mathscr{C}$ by $(\forall y)\mathscr{B}(y)$, prove that $\vdash \mathscr{C} \Leftrightarrow \mathscr{C}'$.

Lemma 2.12

If a closed wf $\neg\mathscr{B}$ of a theory K is not provable in K, and if K′ is the theory obtained from K by adding $\mathscr{B}$ as a new axiom, then K′ is consistent.

Proof

Assume K′ inconsistent. Then, for some wf $\mathscr{C}$, $\vdash_{K'} \mathscr{C}$ and $\vdash_{K'} \neg\mathscr{C}$. Now, $\vdash_{K'} \mathscr{C} \Rightarrow (\neg\mathscr{C} \Rightarrow \neg\mathscr{B})$ by Proposition 2.1. So, by two applications of MP, $\vdash_{K'} \neg\mathscr{B}$. Now, any use of $\mathscr{B}$ as an axiom in a proof in K′ can be regarded as a hypothesis in a proof in K. Hence, $\mathscr{B} \vdash_K \neg\mathscr{B}$. Since $\mathscr{B}$ is closed, we have $\vdash_K \mathscr{B} \Rightarrow \neg\mathscr{B}$ by Corollary 2.7. However, by Proposition 2.1, $\vdash_K(\mathscr{B} \Rightarrow \neg\mathscr{B}) \Rightarrow \neg\mathscr{B}$. Therefore, by MP, $\vdash_K \neg\mathscr{B}$, contradicting our hypothesis.

Corollary

If a closed wf $\mathscr{B}$ of a theory K is not provable in K, and if K′ is the theory obtained from K by adding $\neg\mathscr{B}$ as a new axiom, then K′ is consistent.

Lemma 2.13

The set of expressions of a language $\mathscr{L}$ is denumerable. Hence, the same is true of the set of terms, the set of wfs and the set of closed wfs.

Proof

First assign a distinct positive integer $g(u)$ to each symbol u as follows: $g(() = 3$, $g()) = 5$, $g(,) = 7$, $g(\neg) = 9$, $g(\Rightarrow) = 11$, $g(\forall) = 13$, $g(x_k) = 13 + 8k$, $g(a_k) = 7 + 8k$, $g(f_k^n) = 1 + 8(2^n 3^k)$, and $g(A_k^n) = 3 + 8(2^n 3^k)$. Then, to an expression $u_0 u_1 \dots u_r$ associate the number $2^{g(u_0)} 3^{g(u_1)} \dots p_r^{g(u_r)}$, where p_j is the jth prime number, starting with $p_0 = 2$. (Example: the number of $A_1^1(x_2)$ is $2^{51}3^3 5^{29} 7^5$.) We can enumerate all expressions in the order of their associated numbers; so, the set of expressions is denumerable.

If we can effectively tell whether any given symbol is a symbol of $\mathscr{L}$, then this enumeration can be effectively carried out, and, in addition, we can effectively decide whether any given number is the number of an expression of $\mathscr{L}$. The same holds true for terms, wfs and closed wfs. If a theory K in the language $\mathscr{L}$ is axiomatic, that is, if we can effectively decide whether any given wf is an axiom of K, then we can effectively enumerate the theorems of K in the following manner. Starting with a list consisting of the first axiom of K in the enumeration just specified, add to the list all the direct consequences of this axiom by MP and by Gen used only once and with x_1 as quantified variable. Add the second axiom to this new list and write all new direct consequences by MP and Gen of the wfs in this augmented list, with Gen used only once and with x_1 and x_2 as quantified variables. If at the kth step we add the kth axiom and apply MP and Gen to the wfs in the new list (with Gen applied only once for each of the variables $x_1, \dots, x_k$), we eventually obtain in this manner all theorems of K. However, in contradistinction to the case of expressions, terms, wfs and closed wfs, it turns out that there are axiomatic theories K for which we cannot tell in advance whether any given wf of K will eventually appear in the list of theorems.

Definitions

i. A theory K is said to be *complete* if, for every closed wf $\mathscr{B}$ of K, either $\vdash_K \mathscr{B}$ or $\vdash_K \neg\mathscr{B}$.
ii. A theory K′ is said to be an *extension* of a theory K if every theorem of K is a theorem of K′. (We also say in such a case that K is a *subtheory* of K′.)

Proposition 2.14 (Lindenbaum's Lemma)

If K is a consistent theory, then there is a consistent, complete extension of K.

Proof

Let $\mathscr{B}_1, \mathscr{B}_2, \ldots$ be an enumeration of all closed wfs of the language of K, by Lemma 2.13. Define a sequence $J_0, J_1, J_2, \ldots$ of theories in the following way. J_0 is K. Assume J_n is defined, with $n \geq 0$. If it is not the case that $\vdash_{J_n} \neg\mathscr{B}_{n+1}$, then let J_{n+1} be obtained from J_n by adding $\mathscr{B}_{n+1}$ as an additional axiom. On the other hand, if $\vdash_{J_n} \neg\mathscr{B}_{n+1}$, let $J_{n+1} = J_n$. Let J be the theory obtained by taking as axioms all the axioms of all the J_is. Clearly, J_{i+1} is an extension of J_i, and J is an extension of all the J_is, including $J_0 = K$. To show that J is consistent, it suffices to prove that every J_i is consistent because a proof of a contradiction in J, involving as it does only a finite number of axioms, is also a proof of a contradiction in some J_i. We prove the consistency of the J_is, by induction. By hypothesis, $J_0 = K$ is consistent. Assume that J_i is consistent. If $J_{i+1} = J_i$, then J_{i+1} is consistent. If $J_i \neq J_{i+1}$, and therefore, by the definition of J_{i+1}, $\neg\mathscr{B}_{i+1}$ is not provable in J_i, then, by Lemma 2.12, J_{i+1} is also consistent. So, we have proved that all the J_is are consistent and, therefore, that J is consistent. To prove the completeness of J, let $\mathscr{C}$ be any closed wf of K. Then $\mathscr{C} = \mathscr{B}_{j+1}$ for some $j \geq 0$. Now, either $\vdash_{J_j} \neg\mathscr{B}_{j+1}$ or $\vdash_{J_{j+1}} \mathscr{B}_{j+1}$, since, if it is not the case that $\vdash_{J_j} \neg\mathscr{B}_{j+1}$, then $\mathscr{B}_{j+1}$ is added as an axiom in J_{j+1}. Therefore, either $\vdash_J \neg\mathscr{B}_{j+1}$ or $\vdash_J \mathscr{B}_{j+1}$. Thus, J is complete.

Note that even if one can effectively determine whether any wf is an axiom of K, it may not be possible to do the same with (or even to enumerate effectively) the axioms of J; that is, J may not be axiomatic even if K is. This is due to the possibility of not being able to determine, at each step, whether or not $\neg\mathscr{B}_{n+1}$ is provable in J_n.

Exercises

2.49 Show that a theory K is complete if and only if, for any closed wfs $\mathscr{B}$ and $\mathscr{C}$ of K, if $\vdash_K \mathscr{B} \vee \mathscr{C}$, then $\vdash_K \mathscr{B}$ or $\vdash_K \mathscr{C}$.

2.50[D] Prove that every consistent decidable theory has a consistent, decidable, complete extension.

Definitions

1. A *closed term* is a term without variables.
2. A theory K is a *scapegoat theory** if, for any wf $\mathscr{B}(x)$ that has x as its only free variable, there is a closed term t such that

$$\vdash_K (\exists x)\neg\mathscr{B}(x) \Rightarrow \neg\mathscr{B}(t)$$

* If a scapegoat theory assumes that a given property B fails for at least one object, then there must be a name (that is, a suitable closed term t) of a specific object for which B provably fails. So, t would play the role of a scapegoat, in the usual meaning of that idea. Many theories lack the linguistic resources (individual constants and function letters) to be scapegoat theories, but the notion of *scapegoat theory* will be very useful in proving some deep properties of first-order theories.

Lemma 2.15

Every consistent theory K has a consistent extension K′ such that K′ is a scapegoat theory and K′ contains denumerably many closed terms.

Proof

Add to the symbols of K a denumerable set $\{b_1, b_2, \ldots\}$ of new individual constants. Call this new theory K_0. Its axioms are those of K plus those logical axioms that involve the symbols of K and the new constants. K_0 is consistent. For, if not, there is a proof in K_0 of a wf $\mathscr{B} \wedge \neg\mathscr{B}$. Replace each b_i appearing in this proof by a variable that does not appear in the proof. This transforms axioms into axioms and preserves the correctness of the applications of the rules of inference. The final wf in the proof is still a contradiction, but now the proof does not involve any of the b_is and therefore is a proof in K. This contradicts the consistency of K. Hence, K_0 is consistent.

By Lemma 2.13, let $F_1(x_{i_1}), F_2(x_{i_2}), \ldots, F_k(x_{i_k}), \ldots$ be an enumeration of all wfs of K_0 that have one free variable. Choose a sequence $b_{j_1}, b_{j_2}, \ldots$ of some of the new individual constants such that each b_{j_k} is not contained in any of the wfs $F_1(x_{i_1}), \ldots, F_k(x_{i_k})$ and such that b_{j_k} is different from each of $b_{j_1}, \ldots, b_{j_{k-1}}$. Consider the wf

$$(S_k) \quad (\exists x_{i_k})\neg F_k(x_{i_k}) \Rightarrow \neg F_k(b_{j_k})$$

Let K_n be the theory obtained by adding $(S_1), \ldots, (S_n)$ to the axioms of K_0, and let K_∞ be the theory obtained by adding all the (S_i)s as axioms to K_0. Any proof in K_∞ contains only a finite number of the (S_i)s and, therefore, will also be a proof in some K_n. Hence, if all the K_ns are consistent, so is K_∞. To demonstrate that all the K_ns are consistent, proceed by induction. We know that K_0 is consistent. Assume that K_{n-1} is consistent but that K_n is inconsistent ($n \geq 1$). Then, as we know, any wf is provable in K_n (by the tautology $\neg A \Rightarrow (A \Rightarrow B)$, Proposition 2.1 and MP). In particular, $\vdash_{K_n} \neg(S_n)$. Hence, $(S_n) \vdash_{K_{n-1}} \neg(S_n)$. Since (S_n) is closed, we have, by Corollary 2.7, $\vdash_{K_{n-1}} (S_n) \Rightarrow \neg(S_n)$. But, by the tautology $(A \Rightarrow \neg A) \Rightarrow \neg A$, Proposition 2.1 and MP, we then have $\vdash_{K_{n-1}} \neg(S_n)$; that is, $\vdash_{K_{n-1}} \neg[(\exists x_{i_n})\neg F_n(x_{i_n}) \Rightarrow \neg F_n(b_{j_n})]$. Now, by conditional elimination, we obtain $\vdash_{K_{n-1}} (\exists x_{i_n})\neg F_n(x_{i_n})$ and $\vdash_{K_{n-1}} \neg\neg F_n(b_{j_n})$, and then, by negation elimination, $\vdash_{K_{n-1}} F_n(b_{j_n})$. From the latter and the fact that b_{j_n} does not occur in $(S_0), \ldots, (S_{n-1})$, we conclude $\vdash_{K_{n-1}} F_n(x_r)$, where x_r is a variable that does not occur in the proof of $F_n(b_{j_n})$. (Simply replace in the proof all occurrences of b_{j_n} by x_r.) By Gen, $\vdash_{K_{n-1}} (\forall x_r)F_n(x_r)$, and then, by Lemma 2.11 and biconditional elimination, $\vdash_{K_{n-1}} (\forall x_{i_n})F_n(x_{i_n})$. (We use the fact that $F_n(x_r)$ and $F_n(x_{i_n})$ are similar.) But we already have $\vdash_{K_{n-1}} (\exists x_{i_n})\neg F_n(x_{i_n})$, which is an abbreviation of $\vdash_{K_{n-1}} \neg(\forall x_{i_n})\neg\neg F_n(x_{i_n})$, whence, by the replacement theorem, $\vdash_{K_{n-1}} \neg(\forall x_{i_n})F_n(x_{i_n})$, contradicting the hypothesis that K_{n-1} is consistent. Hence, K_n must also be consistent. Thus K_∞ is consistent, it is an extension of K, and it is clearly a scapegoat theory.

Lemma 2.16

Let J be a consistent, complete scapegoat theory. Then J has a model M whose domain is the set D of closed terms of J.

Proof

For any individual constant a_i of J, let $(a_i)^M = a_i$. For any function letter f_k^n of J and for any closed terms $t_1, \ldots, t_n$ of J, let $(f_k^n)^M(t_1, \ldots, t_n) = f_k^n(t_1, \ldots, t_n)$. (Notice that $f_k^n(t_1, \ldots, t_n)$ is a closed term. Hence, $(f_k^n)^M$ is an n-ary operation on D.) For any predicate letter A_k^n of J, let $(A_k^n)^M$ consist of all n-tuples $\langle t_1, \ldots, t_n \rangle$ of closed terms $t_1, \ldots, t_n$ of J such that $\vdash_J A_k^n(t_1, \ldots, t_n)$. It now suffices to show that, for any closed wf $\mathscr{C}$ of J:

$$(\square) \quad \models_M \mathscr{C} \quad \text{if and only if} \quad \vdash_J \mathscr{C}$$

(If this is established and $\mathscr{B}$ is any axiom of J, let $\mathscr{C}$ be the closure of $\mathscr{B}$. By Gen, $\vdash_J \mathscr{C}$. By ($\square$), $\models_M \mathscr{C}$. By (VI) on page 58, $\models_M \mathscr{B}$. Hence, M would be a model of J.) The proof of ($\square$) is by induction on the number r of connectives and quantifiers in $\mathscr{C}$. Assume that ($\square$) holds for all closed wfs with fewer than r connectives and quantifiers.

Case 1. $\mathscr{C}$ is a closed atomic wf $A_k^n(t_1, \ldots, t_n)$. Then ($\square$) is a direct consequence of the definition of $(A_k^n)^M$.

Case 2. $\mathscr{C}$ is $\neg\mathscr{D}$. If $\mathscr{C}$ is true for M, then $\mathscr{D}$ is false for M and so, by inductive hypothesis, not-$\vdash_J \mathscr{D}$. Since J is complete and $\mathscr{D}$ is closed, $\vdash_J \neg\mathscr{D}$—that is, $\vdash_J \mathscr{C}$. Conversely, if $\mathscr{C}$ is not true for M, then $\mathscr{D}$ is true for M. Hence, $\vdash_J \mathscr{D}$. Since J is consistent, not-$\vdash_J \neg\mathscr{D}$, that is, not-$\vdash_J \mathscr{C}$.

Case 3. $\mathscr{C}$ is $\mathscr{D} \Rightarrow \mathscr{E}$. Since $\mathscr{C}$ is closed, so are $\mathscr{D}$ and $\mathscr{E}$. If $\mathscr{C}$ is false for M, then $\mathscr{D}$ is true and $\mathscr{E}$ is false. Hence, by inductive hypothesis, $\vdash_J \mathscr{D}$ and not-$\vdash_J \mathscr{E}$. By the completeness of J, $\vdash_J \neg\mathscr{E}$. Therefore, by an instance of the tautology $D \Rightarrow (\neg E \Rightarrow \neg(D \Rightarrow E))$ and two applications of MP, $\vdash_J \neg(\mathscr{D} \Rightarrow \mathscr{E})$, that is, $\vdash_J \neg\mathscr{C}$, and so, by the consistency of J, not-$\vdash_J \mathscr{C}$. Conversely, if not-$\vdash_J \mathscr{C}$, then, by the completeness of J, $\vdash_J \neg\mathscr{C}$, that is, $\vdash_J \neg(\mathscr{D} \Rightarrow \mathscr{E})$. By conditional elimination, $\vdash_J \mathscr{D}$ and $\vdash_J \neg\mathscr{E}$. Hence, by ($\square$) for $\mathscr{D}$, $\mathscr{D}$ is true for M. By the consistency of J, not-$\vdash_J \mathscr{E}$ and, therefore, by ($\square$) for $\mathscr{E}$, $\mathscr{E}$ is false for M. Thus, since $\mathscr{D}$ is true for M and $\mathscr{E}$ is false for M, $\mathscr{C}$ is false for M.

Case 4. $\mathscr{C}$ is $(\forall x_m)\mathscr{D}$.

Case 4a. $\mathscr{D}$ is a closed wf. By inductive hypothesis, $\models_M \mathscr{D}$ if and only if $\vdash_J \mathscr{D}$. By Exercise 2.32(a), $\vdash_J \mathscr{D} \Leftrightarrow (\forall x_m)\mathscr{D}$. So, $\vdash_J \mathscr{D}$ if and only if $\vdash_J (\forall x_m)\mathscr{D}$, by biconditional elimination. Moreover, $\vDash_M \mathscr{D}$ if and only if $\vDash_M (\forall x_m)\mathscr{D}$ by property (VI) on page 58. Hence, $\vDash_M \mathscr{C}$ if and only if $\vdash_J \mathscr{C}$.

Case 4b. $\mathscr{D}$ is not a closed wf. Since $\mathscr{C}$ is closed, $\mathscr{D}$ has x_m as its only free variable, say $\mathscr{D}$ is $F(x_m)$. Then $\mathscr{C}$ is $(\forall x_m)F(x_m)$.

i. Assume $\vDash_M \mathscr{C}$ and not-$\vdash_J \mathscr{C}$. By the completeness of J, $\vdash_J \neg\mathscr{C}$, that is, $\vdash_J \neg(\forall x_m)F(x_m)$. Then, by Exercise 2.33(a) and biconditional elimination, $\vdash_J (\exists x_m) \neg F(x_m)$. Since J is a scapegoat theory, $\vdash_J \neg F(t)$ for some closed term t of J. But $\vDash_M \mathscr{C}$, that is, $\vDash_M (\forall x_m)F(x_m)$. Since $(\forall x_m)F(x_m) \Rightarrow F(t)$ is true for M by property (X) on page 59, $\vDash_M F(t)$. Hence, by ($\square$) for $F(t)$, $\vdash_J F(t)$. This contradicts the consistency of J. Thus, if $\vDash_M \mathscr{C}$, then, $\vdash_J \mathscr{C}$.

ii. Assume $\vdash_J \mathscr{C}$ and not-$\vDash_M \mathscr{C}$. Thus,

$$(\#) \quad \vdash_J (\forall x_m)F(x_m) \quad \text{and} \quad (\#\#) \quad \text{not}-\vDash_M (\forall x_m)F(x_m).$$

By (##), some sequence of elements of the domain D does not satisfy $(\forall x_m)$ $F(x_m)$. Hence, some sequence s does not satisfy $F(x_m)$. Let t be the ith component of s. Notice that $s^*(u) = u$ for all closed terms u of J (by the definition of $(a_i)^M$ and $(f_k^n)^M$). Observe also that $F(t)$ has fewer connectives and quantifiers than $\mathscr{C}$ and, therefore, the inductive hypothesis applies to $F(t)$, that is, ($\square$) holds for $F(t)$. Hence, by Lemma 2(a) on page 60, s does not satisfy $F(t)$. So, $F(t)$ is false for M. But, by (#) and rule A4, $\vdash_J F(t)$, and so, by ($\square$) for $F(t)$, $\vDash_M F(t)$. This contradiction shows that, if $\vdash_J \mathscr{C}$, then $\vDash_M \mathscr{C}$.

Now we can prove the fundamental theorem of quantification theory. By a *denumerable model* we mean a model in which the domain is denumerable.

Proposition 2.17*

Every consistent theory K has a denumerable model.

Proof

By Lemma 2.15, K has a consistent extension K′ such that K′ is a scapegoat theory and has denumerably many closed terms. By Lindenbaum's lemma, K′ has a consistent, complete extension J that has the same symbols as K′. Hence, J is also a scapegoat theory. By Lemma 2.16, J has a model M whose domain is the denumerable set of closed terms of J. Since J is an extension of K, M is a denumerable model of K.

* The proof given here is essentially due to Henkin (1949), as simplified by Hasenjaeger (1953). The result was originally proved by Gödel (1930). Other proofs have been published by Rasiowa and Sikorski (1951, 1952) and Beth (1951), using (Boolean) algebraic and topological methods, respectively. Still other proofs may be found in Hintikka (1955a,b) and in Beth (1959).

Corollary 2.18

Any logically valid wf $\mathscr{B}$ of a theory K is a theorem of K.

Proof

We need consider only closed wfs $\mathscr{B}$, since a wf $\mathscr{D}$ is logically valid if and only if its closure is logically valid, and $\mathscr{D}$ is provable in K if and only if its closure is provable in K. So, let $\mathscr{B}$ be a logically valid closed wf of K. Assume that not-$\vdash_K \mathscr{B}$. By Lemma 2.12, if we add $\neg\mathscr{B}$ as a new axiom to K, the new theory K′ is consistent. Hence, by Proposition 2.17, K′ has a model M. Since $\neg\mathscr{B}$ is an axiom of K′, $\neg\mathscr{B}$ is true for M. But, since $\mathscr{B}$ is logically valid, $\mathscr{B}$ is true for M. Hence, $\mathscr{B}$ is both true and false for M, which is impossible (by (II) on page 57). Thus, $\mathscr{B}$ must be a theorem of K.

Corollary 2.19 (Gödel's Completeness Theorem, 1930)

In any predicate calculus, the theorems are precisely the logically valid wfs.

Proof

This follows from Proposition 2.2 and Corollary 2.18. (Gödel's original proof runs along quite different lines. For other proofs, see Beth (1951), Dreben (1952), Hintikka (1955a,b) and Rasiowa and Sikorski (1951, 1952).)

Corollary 2.20

Let K be any theory.

a. A wf $\mathscr{B}$ is true in every denumerable model of K if and only if $\vdash_K \mathscr{B}$.
b. If, in every model of K, every sequence that satisfies all wfs in a set Γ of wfs also satisfies a wf $\mathscr{B}$, then $\Gamma \vdash_K \mathscr{B}$.
c. If a wf $\mathscr{B}$ of K is a logical consequence of a set Γ of wfs of K, then $\Gamma \vdash_K \mathscr{B}$.
d. If a wf $\mathscr{B}$ of K is a logical consequence of a wf $\mathscr{C}$ of K, then $\mathscr{C} \vdash_K \mathscr{B}$.

Proof

a. We may assume $\mathscr{B}$ is closed (Why?). If not-$\vdash_K \mathscr{B}$, then the theory K′ = K + $\{\neg\mathscr{B}\}$ is consistent, by Lemma 2.12.* Hence, by Proposition 2.17, K′ has a denumerable model M. However, $\neg\mathscr{B}$, being an axiom of K′, is true for M. By hypothesis, since M is a denumerable model of K, $\mathscr{B}$ is true for M. Therefore, $\mathscr{B}$ is true and false for M, which is impossible.

* If K is a theory and Δ is a set of wfs of K, then K + Δ denotes the theory obtained from K by adding the wfs of Δ as axioms.

b. Consider the theory K + Γ. By the hypothesis, $\mathscr{B}$ is true for every model of this theory. Hence, by (a), $\vdash_{K+\Gamma} \mathscr{B}$. So, $\Gamma \vdash_K \mathscr{B}$.

Part (c) is a consequence of (b), and part (d) is a special case of (c).

Corollaries 2.18–2.20 show that the "syntactical" approach to quantification theory by means of first-order theories is equivalent to the "semantical" approach through the notions of interpretations, models, logical validity, and so on. For the propositional calculus, Corollary 1.15 demonstrated the analogous equivalence between the semantical notion (tautology) and the syntactical notion (theorem of L). Notice also that, in the propositional calculus, the completeness of the system L (see Proposition 1.14) led to a solution of the decision problem. However, for first-order theories, we cannot obtain a decision procedure for logical validity or, equivalently, for provability in first-order predicate calculi. We shall prove this and related results in Section 3.6.

Corollary 2.21 (Skolem–Löwenheim Theorem, 1920, 1915)

Any theory that has a model has a denumerable model.

Proof

If K has a model, then K is consistent, since no wf can be both true and false for the same model M. Hence, by Proposition 2.17, K has a denumerable model.

The following stronger consequence of Proposition 2.17 is derivable.

Corollary 2.22[A]

For any cardinal number $\mathfrak{m} \geq \aleph_0$, any consistent theory K has a model of cardinality $\mathfrak{m}$.

Proof

By Proposition 2.17, we know that K has a denumerable model. Therefore, it suffices to prove the following lemma.

Lemma

If $\mathfrak{m}$ and $\mathfrak{n}$ are two cardinal numbers such that $\mathfrak{m} \leqslant \mathfrak{n}$ and if K has a model of cardinality $\mathfrak{m}$, then K has a model of cardinality $\mathfrak{n}$.

Proof

Let M be a model of K with domain D of cardinality $\mathfrak{m}$. Let D' be a set of cardinality n that includes D. Extend the model M to an interpretation M′ that has D' as domain in the following way. Let c be a fixed element of D. We stipulate that the elements of $D' - D$ behave like c. For example, if B_j^n is the interpretation in M of the predicate letter A_j^n and $(B_j^n)'$ is the new interpretation in M′, then for any $d_1, \ldots, d_n$ in D′, $(B_j^n)'$ holds for $(d_1, \ldots, d_n)$ if and only if B_j^n holds for $(u_1, \ldots, u_n)$, where $u_i = d_i$ if $d_i \in$ D and $u_i = c$ if $d_i \in D' - D$. The interpretation of the function letters is extended in an analogous way, and the individual constants have the same interpretations as in M. It is an easy exercise to show, by induction on the number of connectives and quantifiers in a wf $\mathscr{B}$, that $\mathscr{B}$ is true for M′ if and only if it is true for M. Hence, M′ is a model of K of cardinality $\mathfrak{n}$.

Exercises

2.51 For any theory K, if $\Gamma \vdash_K \mathscr{B}$ and each wf in Γ is true for a model M of K, show that $\mathscr{B}$ is true for M.

2.52 If a wf $\mathscr{B}$ without quantifiers is provable in a predicate calculus, prove that $\mathscr{B}$ is an instance of a tautology and, hence, by Proposition 2.1, has a proof without quantifiers using only axioms (A1)–(A3) and MP. [*Hint*: if $\mathscr{B}$ were not a tautology, one could construct an interpretation, having the set of terms that occur in $\mathscr{B}$ as its domain, for which $\mathscr{B}$ is not true, contradicting Proposition 2.2.]

Note that this implies the consistency of the predicate calculus and also provides a decision procedure for the provability of wfs without quantifiers.

2.53 Show that $\vdash_K \mathscr{B}$ if and only if there is a wf $\mathscr{C}$ that is the closure of the conjunction of some axioms of K such that $\mathscr{C} \Rightarrow \mathscr{B}$ is logically valid.

2.54 *Compactness*. If all finite subsets of the set of axioms of a theory K have models, prove that K has a model.

2.55 a. For any wf $\mathscr{B}$, prove that there is only a finite number of interpretations of $\mathscr{B}$ on a given domain of finite cardinality k.

b. For any wf $\mathscr{B}$, prove that there is an effective way of determining whether $\mathscr{B}$ is true for all interpretations with domain of some fixed cardinality k.

c. Let a wf $\mathscr{B}$ be called *k-valid* if it is true for all interpretations that have a domain of k elements. Call $\mathscr{B}$ *precisely k-valid* if it is k-valid but not $(k + 1)$-valid. Show that $(k + 1)$-validity implies k-validity and give an example of a wf that is precisely k-valid. (See Hilbert and Bernays (1934, § 4–5) and Wajsberg (1933).)

2.56 Show that the following wf is true for all finite domains but is false for some infinite domain.

$$(\forall x)(\forall y)(\forall z)\Big[A_1^2(x,x) \wedge \Big(A_1^2(x,y) \wedge A_1^2(y,z) \Rightarrow A_1^2(x,z)\Big) \wedge \Big(A_1^2(x,y) \vee A_1^2(y,x)\Big)\Big]$$
$$\Rightarrow (\exists y)(\forall x)A_1^2(y,x)$$

2.57 Prove that there is no theory K whose models are exactly the interpretations with finite domains.

2.58 Let $\mathscr{B}$ be any wf that contains no quantifiers, function letters, or individual constants.

a. Show that a closed *prenex* wf $(\forall x_1) \ldots (\forall x_n)(\exists y_1) \ldots (\exists y_m)\mathscr{B}$, with $m \geq 0$ and $n \geq 1$, is logically valid if and only if it is true for every interpretation with a domain of n objects.

b. Prove that a closed prenex wf $(\exists y_1) \ldots (\exists y_m)\mathscr{B}$ is logically valid if and only if it is true for all interpretations with a domain of one element.

c. Show that there is an effective procedure to determine the logical validity of all wfs of the forms given in (a) and (b).

2.59 Let K_1 and K_2 be theories in the same language $\mathscr{L}$. Assume that any interpretation M of $\mathscr{L}$ is a model of K_1 if and only if M is not a model of K_2. Prove that K_1 and K_2 are finitely axiomatizable, that is, there are finite sets of sentences Γ and Δ such that, for any sentence $\mathscr{B}$, $\vdash_{K_1} \mathscr{B}$ if and only if $\Gamma \vdash \mathscr{B}$, and $\vdash_{K_2} \mathscr{B}$ if and only if $\Delta \vdash \mathscr{B}$.*

2.60 A set Γ of sentences is called an *independent axiomatization* of a theory K if (a) all sentences in Γ are theorems of K, (b) $\Gamma \vdash \mathscr{B}$ for every theorem $\mathscr{B}$ of K, and (c) for every sentence $\mathscr{C}$ of Γ, it is not the case that $\Gamma - \{\mathscr{C}\} \vdash \mathscr{C}$.* Prove that every theory K has an independent axiomatization.

2.61[A] If, for some cardinal $\mathfrak{m} \geq \aleph_0$, a wf $\mathscr{B}$ is true for every interpretation of cardinality $\mathfrak{m}$, prove that $\mathscr{B}$ is logically valid.

2.62[A] If a wf $\mathscr{B}$ is true for all interpretations of cardinality $\mathfrak{m}$ prove that $\mathscr{B}$ is true for all interpretations of cardinality less than or equal to $\mathfrak{m}$.

2.63 a. Prove that a theory K is a scapegoat theory if and only if, for any wf $\mathscr{B}(x)$ with x as its only free variable, there is a closed term t such that $\vdash_K (\exists x)\mathscr{B}(x) \Rightarrow \mathscr{B}(t)$.

b. Prove that a theory K is a scapegoat theory if and only if, for any wf $\mathscr{B}(x)$ with x as its only free variable such that $\vdash_K (\exists x)\mathscr{B}(x)$, there is a closed term t such that $\vdash_K \mathscr{B}(t)$.

c. Prove that no predicate calculus is a scapegoat theory.

* Here, an expression $\Gamma \vdash \mathscr{B}$, without any subscript attached to ⊢, means that $\mathscr{B}$ is derivable from Γ using only logical axioms, that is, within the predicate calculus.

2.8 First-Order Theories with Equality

Let K be a theory that has as one of its predicate letters A_1^2. Let us write $t = s$ as an abbreviation for $A_1^2(t,s)$, and $t \neq s$ as an abbreviation for $\neg A_1^2(t,s)$. Then K is called a *first-order theory with equality* (or simply a *theory with equality*) if the following are theorems of K:

$$(A6)\ (\forall x_1)x_1 = x_1 \qquad \text{(reflexivity of equality)}$$
$$(A7)\ x = y \Rightarrow (\mathscr{B}(x,x) \Rightarrow \mathscr{B}(x,y)) \qquad \text{(substitutivity of equality)}$$

where x and y are any variables, $\mathscr{B}(x, x)$ is any wf, and $\mathscr{B}(x, y)$ arises from $\mathscr{B}(x, x)$ by replacing some, but not necessarily all, free occurrences of x by y, with the proviso that y is free for x in $\mathscr{B}(x, x)$. Thus, $\mathscr{B}(x, y)$ may or may not contain free occurrences of x.

The numbering (A6) and (A7) is a continuation of the numbering of the logical axioms.

Proposition 2.23

In any theory with equality,

a. $\vdash t = t$ for any term t;
b. $\vdash t = s \Rightarrow s = t$ for any terms t and s;
c. $\vdash t = s \Rightarrow (s = r \Rightarrow t = r)$ for any terms t, s, and r.

Proof

a. By (A6), $\vdash (\forall x_1)x_1 = x_1$. Hence, by rule A4, $\vdash t = t$.
b. Let x and y be variables not occurring in t or s. Letting $\mathscr{B}(x, x)$ be $x = x$ and $\mathscr{B}(x, y)$ be $y = x$ in schema (A7), $\vdash x = y \Rightarrow (x = x \Rightarrow y = x)$. But, by (a), $\vdash x = x$. So, by an instance of the tautology $(A \Rightarrow (B \Rightarrow C)) \Rightarrow (B \Rightarrow (A \Rightarrow C))$ and two applications of MP, we have $\vdash x = y \Rightarrow y = x$. Two applications of Gen yield $\vdash (\forall x)(\forall y)(x = y \Rightarrow y = x)$, and then two applications of rule A4 give $\vdash t = s \Rightarrow s = t$.
c. Let x, y, and z be three variables not occurring in t, s, or r. Letting $\mathscr{B}(y, y)$ be $y = z$ and $\mathscr{B}(y, x)$ be $x = z$ in (A7), with x and y interchanged, we obtain $\vdash y = x \Rightarrow (y = z \Rightarrow x = z)$. But, by (b), $\vdash x = y \Rightarrow y = x$. Hence, using an instance of the tautology $(A \Rightarrow B) \Rightarrow ((B \Rightarrow C) \Rightarrow (A \Rightarrow C))$ and two applications of MP, we obtain $\vdash x = y \Rightarrow (y = z \Rightarrow x = z)$. By three applications of Gen, $\vdash (\forall x)(\forall y)(\forall z)(x = y \Rightarrow (y = z \Rightarrow x = z))$, and then, by three uses of rule A4, $\vdash t = s \Rightarrow (s = r \Rightarrow t = r)$.

Exercises

2.64 Show that (A6) and (A7) are true for any interpretation M in which $(A_1^2)^M$ is the identity relation on the domain of the interpretation.

2.65 Prove the following in any theory with equality.

a. $\vdash (\forall x)(\mathscr{B}(x) \Leftrightarrow (\exists y)(x = y \wedge \mathscr{B}(y)))$ if y does not occur in $\mathscr{B}(x)$
b. $\vdash (\forall x)(\mathscr{B}(x) \Leftrightarrow (\forall y)(x = y \Rightarrow \mathscr{B}(y)))$ if y does not occur in $\mathscr{B}(x)$
c. $\vdash (\forall x)(\exists y)x = y$
d. $\vdash x = y \Rightarrow f(x) = f(y)$, where f is any function letter of one argument
e. $\vdash \mathscr{B}(x) \wedge x = y \Rightarrow \mathscr{B}(y)$, if y is free for x in $\mathscr{B}(x)$
f. $\vdash \mathscr{B}(x) \wedge \neg\mathscr{B}(y) \Rightarrow x \neq y$, if y is free for x in $\mathscr{B}(x)$

We can reduce schema (A7) to a few simpler cases.

Proposition 2.24

Let K be a theory for which (A6) holds and (A7) holds for all atomic wfs $\mathscr{B}(x, x)$ in which there are no individual constants. Then K is a theory with equality, that is, (A7) holds for all wfs $\mathscr{B}(x, x)$.

Proof

We must prove (A7) for all wfs $\mathscr{B}(x, x)$. It holds for atomic wfs by assumption. Note that we have the results of Proposition 2.23, since its proof used (A7) only with atomic wfs without individual constants. Note also that we have (A7) for all atomic wfs $\mathscr{B}(x, x)$. For if $\mathscr{B}(x, x)$ contains individual constants, we can replace those individual constants by new variables, obtaining a wf $\mathscr{B}^*(x, x)$ without individual constants. By hypothesis, the corresponding instance of (A7) with $\mathscr{B}^*(x, x)$ is a theorem; we can then apply Gen with respect to the new variables, and finally apply rule A4 one or more times to obtain (A7) with respect to $\mathscr{B}(x, x)$.

Proceeding by induction on the number n of connectives and quantifiers in $\mathscr{B}(x, x)$, we assume that (A7) holds for all $k < n$.

Case 1. $\mathscr{B}(x, x)$ is $\neg\mathscr{C}(x, x)$. By inductive hypothesis, we have $\vdash y = x \Rightarrow (\mathscr{C}(x, y) \Rightarrow \mathscr{C}(x, x))$, since $\mathscr{C}(x, x)$ arises from $\mathscr{C}(x, y)$ by replacing some occurrences of y by x. Hence, by Proposition 2.23(b), instances of the tautologies $(A \Rightarrow B) \Rightarrow (\neg\mathscr{B} \Rightarrow \neg A)$ and $(A \Rightarrow B) \Rightarrow ((B \Rightarrow C) \Rightarrow (A \Rightarrow C))$ and MP, we obtain $\vdash x = y \Rightarrow (\mathscr{B}(x, x) \Rightarrow \mathscr{B}(x, y))$.

Case 2. $\mathscr{B}(x, x)$ is $\mathscr{C}(x, x) \Rightarrow \mathscr{D}(x, x)$. By inductive hypothesis and Proposition 2.23(b), $\vdash x = y \Rightarrow (\mathscr{C}(x, y) \Rightarrow \mathscr{C}(x, x))$ and $\vdash x = y \Rightarrow (\mathscr{D}(x, x) \Rightarrow \mathscr{D}(x, y))$. Hence, by the tautology $(A \Rightarrow (C_1 \Rightarrow C)) \Rightarrow [(A \Rightarrow (D \Rightarrow D_1)) \Rightarrow (A \Rightarrow ((C \Rightarrow D) \Rightarrow (C_1 \Rightarrow D_1)))]$, we have $\vdash x = y \Rightarrow (\mathscr{B}(x, x) \Rightarrow \mathscr{B}(x, y))$.

Case 3. $\mathscr{B}(x, x)$ is $(\forall z)\mathscr{C}(x, x, z)$. By inductive hypothesis, $\vdash x = y \Rightarrow (\mathscr{C}(x, x, z) \Rightarrow \mathscr{C}(x, y, z))$. Now, by Gen and axiom (A5), $\vdash x = y \Rightarrow (\forall z)\,(\mathscr{C}(x, x, z) \Rightarrow \mathscr{C}(x, y, z))$. By Exercise 2.27(a), $\vdash (\forall z)(\mathscr{C}(x, x, z) \Rightarrow \mathscr{C}(x, y, z)) \Rightarrow [(\forall z)\mathscr{C}(x, x, z) \Rightarrow (\forall z)\mathscr{C}(x, y, z)]$, and so, by the tautology $(A \Rightarrow B) \Rightarrow ((B \Rightarrow C) \Rightarrow (A \Rightarrow C))$, $\vdash x = y \Rightarrow (\mathscr{B}(x, x) \Rightarrow \mathscr{B}(x, y))$.

The instances of (A7) can be still further reduced.

Proposition 2.25

Let K be a theory in which (A6) holds and the following are true.

a. Schema (A7) holds for all atomic wfs $\mathscr{B}(x, x)$ such that no function letters or individual constants occur in $\mathscr{B}(x, x)$ and $\mathscr{B}(x, y)$ comes from $\mathscr{B}(x, x)$ by replacing exactly one occurrence of x by y.
b. $\vdash x = y \Rightarrow f_j^n(z_1, \ldots, z_n) = f_j^n(w_1, \ldots, w_n)$, where f_j^n is any function letter of K, $z_1, \ldots, z_n$ are variables, and $f_j^n(w_1, \ldots, w_n)$ arises from $f_j^n(z_1, \ldots, z_n)$ by replacing exactly one occurrence of x by y.

Then K is a theory with equality.

Proof

By repeated application, our assumptions can be extended to replacements of more than one occurrence of x by y. Also, Proposition 2.23 is still derivable. By Proposition 2.24, it suffices to prove (A7) for only atomic wfs without individual constants. But, hypothesis (a) enables us easily to prove

$$\vdash (y_1 = z_1 \wedge \ldots \wedge y_n = z_n) \Rightarrow (\mathscr{B}(y_1, \ldots, y_n) \Rightarrow \mathscr{B}(z_1, \ldots, z_n))$$

for all variables $y_1, \ldots, y_n, z_1, \ldots, z_n$ and any atomic wf $\mathscr{B}(y_1, \ldots, y_n)$ without function letters or individual constants. Hence, it suffices to show:

> (*) If $t(x, x)$ is a term without individual constants and $t(x, y)$ comes from $t(x, x)$ by replacing some occurrences of x by y, then $\vdash x = y \Rightarrow t(x, x) = t(x, y)$.*

But (*) can be proved, using hypothesis (b), by induction on the number of function letters in $t(x,x)$, and we leave this as an exercise.

It is easy to see from Proposition 2.25 that, when the language of K has only finitely many predicate and function letters, it is only necessary to verify (A7) for a finite list of special cases (in fact, n wfs for each A_j^n and n wfs for each f_j^n).

* The reader can clarify how (*) is applied by using it to prove the following instance of (A7): $\vdash x = y \Rightarrow (A_1^1(f_1^1(x)) \Rightarrow A_1^1(f_1^1(y)))$. Let $t(x, x)$ be $f_1^1(x)$ and let $t(x, y)$ be $f_1^1(y)$.

Exercises

2.66 Let K_1 be a theory whose language has only = as a predicate letter and no function letters or individual constants. Let its proper axioms be $(\forall x_1)x_1 = x_1$, $(\forall x_1)(\forall x_2)(x_1 = x_2 \Rightarrow x_2 = x_1)$, and $(\forall x_1)(\forall x_2)(\forall x_3)(x_1 = x_2 \Rightarrow (x_2 = x_3 \Rightarrow x_1 = x_3))$. Show that K_1 is a theory with equality. [*Hint*: It suffices to prove that $\vdash x_1 = x_3 \Rightarrow (x_1 = x_2 \Rightarrow x_3 = x_2)$ and $\vdash x_2 = x_3 \Rightarrow (x_1 = x_2 \Rightarrow x_1 = x_3)$.] K_1 is called the *pure first-order theory of equality*.

2.67 Let K_2 be a theory whose language has only = and < as predicate letters and no function letters or individual constants. Let K_2 have the following proper axioms.

a. $(\forall x_1)x_1 = x_1$
b. $(\forall x_1)(\forall x_2)(x_1 = x_2 \Rightarrow x_2 = x_1)$
c. $(\forall x_1)(\forall x_2)(\forall x_3)(x_1 = x_2 \Rightarrow (x_2 = x_3 \Rightarrow x_1 = x_3))$
d. $(\forall x_1)(\exists x_2)(\exists x_3)(x_1 < x_2 \wedge x_3 < x_1)$
e. $(\forall x_1)(\forall x_2)(\forall x_3)(x_1 < x_2 \wedge x_2 < x_3 \Rightarrow x_1 < x_3)$
f. $(\forall x_1)(\forall x_2)(x_1 = x_2 \Rightarrow \neg\, x_1 < x_2)$
g. $(\forall x_1)(\forall x_2)(x_1 < x_2 \vee x_1 = x_2 \vee x_2 < x_1)$
h. $(\forall x_1)(\forall x_2)(x_1 < x_2 \Rightarrow (\exists x_3)(x_1 < x_3 \wedge x_3 < x_2))$

Using Proposition 2.25, show that K_2 is a theory with equality. K_2 is called the *theory of densely ordered sets with neither first nor last element*.

2.68 Let K be any theory with equality. Prove the following.

a. $\vdash x_1 = y_1 \wedge \ldots \wedge x_n = y_n \Rightarrow t(x_1, \ldots, x_n) = t(y_1, \ldots, y_n)$, where $t(y_1, \ldots, y_n)$ arises from the term $t(x_1, \ldots, x_n)$ by substitution of $y_1, \ldots, y_n$ for $x_1, \ldots, x_n$, respectively.
b. $\vdash x_1 = y_1 \wedge \ldots \wedge x_n = y_n \Rightarrow (\mathscr{B}(x_1, \ldots, x_n) \Leftrightarrow \mathscr{B}(y_1, \ldots, y_n))$, where $\mathscr{B}(y_1, \ldots, y_n)$ is obtained by substituting $y_1, \ldots, y_n$ for one or more occurrences of $x_1, \ldots, x_n$, respectively, in the wf $\mathscr{B}(x_1, \ldots, x_n)$, and $y_1, \ldots, y_n$ are free for $x_1, \ldots, x_n$, respectively, in the wf $\mathscr{B}(x_1, \ldots, x_n)$.

Examples

(In the literature, "elementary" is sometimes used instead of "first-order.")

1. *Elementary theory* G *of groups*: predicate letter =, function letter f_1^2, and individual constant a_1. We abbreviate $f_1^2(t, s)$ by $t + s$ and a_1 by 0. The proper axioms are the following.

a. $x_1 + (x_2 + x_3) = (x_1 + x_2) + x_3$
b. $x_1 + 0 = x_1$
c. $(\forall x_1)(\exists x_2)x_1 + x_2 = 0$

d. $x_1 = x_1$

e. $x_1 = x_2 \Rightarrow x_2 = x_1$

f. $x_1 = x_2 \Rightarrow (x_2 = x_3 \Rightarrow x_1 = x_3)$

g. $x_1 = x_2 \Rightarrow (x_1 + x_3 = x_2 + x_3 \wedge x_3 + x_1 = x_3 + x_2)$

That G is a theory with equality follows easily from Proposition 2.25. If one adds to the axioms the following wf:

h. $x_1 + x_2 = x_2 + x_1$

the new theory is called the *elementary theory of abelian groups*.

2. *Elementary theory* F *of fields*: predicate letter =, function letters f_1^2 and f_2^2, and individual constants a_1 and a_2. Abbreviate $f_1^2(t, s)$ by $t + s$, $f_2^2(t, s)$ by $t \cdot s$, and a_1 and a_2 by 0 and 1. As proper axioms, take (a)–(h) of Example 1 plus the following.

i. $x_1 = x_2 \Rightarrow (x_1 \cdot x_3 = x_2 \cdot x_3 \wedge x_3 \cdot x_1 = x_3 \cdot x_2)$

j. $x_1 \cdot (x_2 \cdot x_3) = (x_1 \cdot x_2) \cdot x_3$

k. $x_1 \cdot (x_2 + x_3) = (x_1 \cdot x_2) + (x_1 \cdot x_3)$

l. $x_1 \cdot x_2 = x_2 \cdot x_1$

m. $x_1 \cdot 1 = x_1$

n. $x_1 \neq 0 \Rightarrow (\exists x_2) x_1 \cdot x_2 = 1$

o. $0 \neq 1$

F is a theory with equality. Axioms (a)–(m) define the elementary theory R_C of commutative rings with unit. If we add to F the predicate letter A_2^2, abbreviate $A_2^2(t,s)$ by $t < s$, and add axioms (e), (f), and (g) of Exercise 2.67, as well as $x_1 < x_2 \Rightarrow x_1 + x_3 < x_2 + x_3$ and $x_1 < x_2 \wedge 0 < x_3 \Rightarrow x_1 \cdot x_3 < x_2 \cdot x_3$, then the new theory $F_<$ is called the *elementary theory of ordered fields*.

Exercise

2.69 a. What formulas must be derived in order to use Proposition 2.25 to conclude that the theory G of Example 1 is a theory with equality?

b. Show that the axioms (d)–(f) of equality mentioned in Example 1 can be replaced by (d) and

$$(f') : x_1 = x_2 \Rightarrow (x_3 = x_2 \Rightarrow x_1 = x_3).$$

One often encounters theories K in which=may be defined; that is, there is a wf $\mathscr{E}(x, y)$ with two free variables x and y, such that, if we abbreviate $\mathscr{E}(t, s)$ by $t = s$, then axioms (A6) and (A7) are provable in K. We make the convention that, if t and s are terms that are not free for x and y, respectively, in $\mathscr{E}(x, y)$, then, by suitable changes of bound variables (see Exercise 2.48), we replace $\mathscr{E}(x, y)$ by a logically equivalent wf $\mathscr{E}^*(x, y)$ such that t and s are free for x and y, respectively, in $\mathscr{E}^*(x, y)$; then $t = s$ is to be the abbreviation of $\mathscr{E}^*(t, s)$. Proposition 2.23 and analogues of Propositions 2.24 and 2.25 hold for such theories. There is no harm in extending the term *theory with equality* to cover such theories.

In theories with equality it is possible to define in the following way phrases that use the expression "There exists one and only one x such that…."

Definition

$$(\exists_1 x)\mathscr{B}(x) \quad \text{for} \quad (\exists x)\mathscr{B}(x) \wedge (\forall x)(\forall y)(\mathscr{B}(x) \wedge \mathscr{B}(y) \Rightarrow x = y)$$

In this definition, the new variable y is assumed to be the first variable that does not occur in $\mathscr{B}(x)$. A similar convention is to be made in all other definitions where new variables are introduced.

Exercise

2.70 In any theory with equality, prove the following.

a. $\vdash (\forall x)(\exists_1 y)x = y$
b. $\vdash (\exists_1 x)\mathscr{B}(x) \Leftrightarrow (\exists x)(\forall y)(x = y \Leftrightarrow \mathscr{B}(y))$
c. $\vdash (\forall x)(\mathscr{B}(x) \Leftrightarrow \mathscr{C}(x)) \Rightarrow [(\exists_1 x)\mathscr{B}(x) \Leftrightarrow (\exists_1 x)\mathscr{C}(x)]$
d. $\vdash (\exists_1 x)(\mathscr{B} \vee \mathscr{C}) \Rightarrow ((\exists_1 x)\mathscr{B}) \vee (\exists_1 x)\mathscr{C}$
e. $\vdash (\exists_1 x)\mathscr{B}(x) \Leftrightarrow (\exists x)(\mathscr{B}(x) \wedge (\forall y)(\mathscr{B}(y) \Rightarrow y = x))$

In any model for a theory K with equality, the relation E in the model corresponding to the predicate letter = is an equivalence relation (by Proposition 2.23). If this relation E is the identity relation in the domain of the model, then the model is said to be *normal*.

Any model M for K can be *contracted* to a normal model M* for K by taking the domain D^* of M* to be the set of equivalence classes determined by the relation E in the domain D of M. For a predicate letter A_j^n and for any equivalence classes $[b_1], \dots, [b_n]$ in D^* determined by elements $b_1, \dots, b_n$ in D, we let $(A_j^n)^{M^*}$ hold for $([b_1], \dots, [b_n])$ if and only if $(A_j^n)^M$ holds for $(b_1, \dots, b_n)$. Notice that it makes no difference which representatives $b_1, \dots, b_n$ we select in the given equivalence classes because, from (A7), $\vdash x_1 = y_1 \wedge \ldots \wedge x_n = y_n \Rightarrow (A_j^n(x_1, \ldots, x_n) \Leftrightarrow A_j^n(y_1, \ldots, y_n))$. Likewise, for any function letter f_j^n and any equivalence classes $[b_1], \dots, [b_n]$ in D^*, let $(f_j^n)^{M^*}([b_1], \ldots, [b_n]) = [(f_j^n)^M(b_1, \ldots, b_n)]$. Again note that this is independent of the choice of the representatives $b_1, \dots, b_n$, since, from (A7), we can prove $\vdash x_1 = y_1 \wedge \ldots \wedge x_n = y_n \Rightarrow f_j^n(x_1, \ldots, x_n) = f_j^n(y_1, \ldots, y_n)$. For any individual constant a_i let $(a_i)^{M^*} = [(a_i)^M]$. The relation E* corresponding to = in the model M* is the identity relation in D^*: $E^*([b_1], [b_2])$ if and only if $E(b_1, b_2)$, that is, if and only if $[b_1] = [b_2]$. Now one can easily prove by induction the following lemma: If $s = (b_1, b_2, \ldots)$ is a denumerable sequence of elements of D, and $s' = ([b_1], [b_2], \ldots)$ is the corresponding sequence of equivalence classes, then a wf $\mathscr{B}$ is satisfied by s in M if and only if $\mathscr{B}$ is satisfied by s' in M*. It follows that, for any wf $\mathscr{B}$, $\mathscr{B}$ is true for M if and only if $\mathscr{B}$ is true for M*. Hence, because M is a model of K, M* is a normal model of K.

Proposition 2.26 (Extension of Proposition 2.17)

(Gödel, 1930) Any consistent theory with equality K has a finite or denumerable normal model.

Proof

By Proposition 2.17, K has a denumerable model M. Hence, the contraction of M to a normal model yields a finite or denumerable normal model M* because the set of equivalence classes in a denumerable set D is either finite or denumerable.

Corollary 2.27 (Extension of the Skolem–Löwenheim Theorem)

Any theory with equality K that has an infinite normal model M has a denumerable normal model.

Proof

Add to K the denumerably many new individual constants $b_1, b_2, \ldots$ together with the axioms $b_i \neq b_j$ for $i \neq j$. Then the new theory K′ is consistent. If K′ were inconsistent, there would be a proof in K′ of a contradiction $\mathscr{C} \wedge \neg\mathscr{C}$, where we may assume that $\mathscr{C}$ is a wf of K. But this proof uses only a finite number of the new axioms: $b_{i1} \neq b_{j1}, \ldots, b_{in} \neq b_{jn}$. Now, M can be extended to a model M# of K plus the axioms $b_{i1} \neq b_{j1}, \ldots, b_{in} \neq b_{jn}$; in fact, since M is an infinite normal model, we can choose interpretations of $b_{i1}, b_{j1}, \ldots, b_{in}, b_{jn}$, so that the wfs $b_{i1} \neq b_{j1}, \ldots, b_{in} \neq b_{jn}$ are true. But, since $\mathscr{C} \wedge \neg\mathscr{C}$ is derivable from these wfs and the axioms of K, it would follow that $\mathscr{C} \wedge \neg\mathscr{C}$ is true for M#, which is impossible. Hence, K′ must be consistent. Now, by Proposition 2.26, K′ has a finite or denumerable normal model N. But, since, for $i \neq j$, the wfs $b_i \neq b_j$ are axioms of K′, they are true for N. Thus, the elements in the domain of N that are the interpretations of $b_1, b_2, \ldots$ must be distinct, which implies that the domain of N is infinite and, therefore, denumerable.

Exercises

2.71 We define $(\exists_n x)\mathscr{B}(x)$ by induction on $n \geq 1$. The case $n = 1$ has already been taken care of. Let $(\exists_{n+1} x)\mathscr{B}(x)$ stand for $(\exists y)(\mathscr{B}(y) \wedge (\exists_n x)\ (x \neq y \wedge \mathscr{B}(x)))$.

a. Show that $(\exists_n x)\mathscr{B}(x)$ asserts that there are exactly n objects for which $\mathscr{B}$ holds, in the sense that in any normal model for $(\exists_n x)\mathscr{B}(x)$ there are exactly n objects for which the property corresponding to $\mathscr{B}(x)$ holds.

b. i. For each positive integer n, write a closed wf $\mathscr{B}_n$ such that $\mathscr{B}_n$ is true in a normal model when and only when that model contains at least n elements.

ii. Prove that the theory K, whose axioms are those of the pure theory of equality K_1 (see Exercise 2.66), plus the axioms $\mathscr{B}_1$, $\mathscr{B}_2$, …, is not finitely axiomatizable, that is, there is no theory K′ with a finite number of axioms such that K and K′ have the same theorems.

iii. For a normal model, state in ordinary English the meaning of $\neg\mathscr{B}_{n+1}$.

c. Let n be a positive integer and consider the wf $(\mathscr{E}_n)$ $(\exists_n x)x = x$. Let L_n be the theory $K_1 + \{\mathscr{E}_n\}$, where K_1 is the pure theory of equality.

i. Show that a normal model M is a model of L_n if and only if there are exactly n elements in the domain of M.

ii. Define a procedure for determining whether any given sentence is a theorem of L_n and show that L_n is a complete theory.

2.72 a. Prove that, if a theory with equality K has arbitrarily large finite normal models, then it has a denumerable normal model.

b. Prove that there is no theory with equality whose normal models are precisely all finite normal interpretations.

2.73 Prove that any predicate calculus with equality is consistent. (A predicate calculus with equality is assumed to have (A1)–(A7) as its only axioms.)

2.74[D] Prove the independence of axioms (A1)–(A7) in any predicate calculus with equality.

2.75 If $\mathscr{B}$ is a wf that does not contain the = symbol and $\mathscr{B}$ is provable in a predicate calculus with equality K, show that $\mathscr{B}$ is provable in K without using (A6) or (A7).

2.76[D] Show that = can be defined in any theory whose language has only a finite number of predicate letters and no function letters.

2.77 a.[A] Find a nonnormal model of the elementary theory of groups G.

b. Show that any model M of a theory with equality K can be extended to a nonnormal model of K. [*Hint*: Use the argument in the proof of the lemma within the proof of Corollary 2.22.]

2.78 Let $\mathscr{B}$ be a wf of a theory with equality. Show that $\mathscr{B}$ is true in every normal model of K if and only if $\vdash_K \mathscr{B}$.

2.79 Write the following as wfs of a theory with equality.

a. There are at least three moons of Jupiter.

b. At most two people know everyone in the class.

c. Everyone in the logic class knows at least two members of the geometry class.

d. Every person loves at most one other person.

2.80 If $P(x)$ means *x is a person*, $A(x, y)$ means *x is a parent of y*, $G(x, y)$ means *x is a grandparent of y*, and $x = y$ means *x and y are identical*, translate the following wfs into ordinary English.

$$\text{i. } (\forall x)(P(x) \Rightarrow [(\forall y)(G(y,x) \Leftrightarrow (\exists w)(A(y,w) \wedge A(w,x)))])$$

$$\begin{aligned}\text{ii. } (\forall x)(P(x) \Rightarrow (\exists x_1)(\exists x_2)(\exists x_3)(\exists x_4)(&x_1 \neq x_2 \wedge x_1 \neq x_3 \wedge x_1 \neq x_4 \wedge \\ &x_2 \neq x_3 \wedge x_2 \neq x_4 \wedge x_3 \neq x_4 \wedge G(x_1,x) \wedge G(x_2,x) \wedge G(x_3,x) \wedge \\ &G(x_4,x) \wedge (\forall y)(G(y,x) \Rightarrow y = x_1 \vee y = x_2 \vee y = x_3 \vee y = x_4)))\end{aligned}$$

2.81 Consider the wf

$$(*) \quad (\forall x)(\forall y)(\exists z)(z \neq x \wedge z \neq y \wedge A(z)).$$

Show that (*) is true in a normal model M of a theory with equality if and only if there exist in the domain of M at least three things having property $A(z)$.

2.82 Let the language $\mathscr{L}$ have the four predicate letters =, P, S, and L. Read $u = v$ as *u and v are identical*, $P(u)$ as *u is a point*, $S(u)$ as *u is a line*, and $L(u, v)$ as *u lies on v*. Let the theory of equality $\mathbb{G}$ of *planar incidence geometry* have, in addition to axioms (A1)–(A7), the following nonlogical axioms.

1. $P(x) \Rightarrow \neg S(x)$
2. $L(x, y) \Rightarrow P(x) \wedge S(y)$
3. $S(x) \Rightarrow (\exists y)(\exists z)(y \neq z \wedge L(y, x) \wedge L(z, x))$
4. $P(x) \wedge P(y) \wedge x \neq y \Rightarrow (\exists_1 z)(S(z) \wedge L(x, z) \wedge L(y, z))$
5. $(\exists x)(\exists y)(\exists z)(P(x) \wedge P(y) \wedge P(z) \wedge \neg \mathscr{C}(x, y, z))$

where $\mathscr{C}(x, y, z)$ is the wf $(\exists u)(S(u) \wedge L(x, u) \wedge L(y, u) \wedge L(z, u))$, which is read as *x, y, z are collinear.*

a. Translate (1)–(5) into ordinary geometric language.
b. Prove $\vdash_{\mathbb{G}} (\forall u)(\forall v)(S(u) \wedge S(v) \wedge u \neq v \Rightarrow (\forall x)(\forall y)(L(x, u) \wedge L(x, v) \wedge L(y, u) \wedge L(y, v) \Rightarrow x = y))$, and translate this theorem into ordinary geometric language.
c. Let $R(u, v)$ stand for $S(u) \wedge S(v) \wedge \neg(\exists w)(L(w, u) \wedge L(w, v))$. Read $R(u, v)$ as *u and v are distinct parallel lines.*
   i. Prove: $\vdash_{\mathbb{G}} R(u, v) \Rightarrow u \neq v$
   ii. Show that there exists a normal model of $\mathbb{G}$ with a finite domain in which the following sentence is true:

$$(\forall x)(\forall y)(S(x) \wedge P(y) \wedge \neg L(y, x) \Rightarrow (\exists_1 z)(L(y, z) \wedge R(z, x)))$$

d. Show that there exists a model of $\mathbb{G}$ in which the following sentence is true:

$$(\forall x)(\forall y)(S(x) \wedge S(y) \wedge x \neq y \Rightarrow \neg R(x, y))$$

2.9 Definitions of New Function Letters and Individual Constants

In mathematics, once we have proved, for any $y_1, \ldots, y_n$, the existence of a unique object u that has a property $\mathscr{B}(u, y_1, \ldots, y_n)$, we often introduce a new function letter $f(y_1, \ldots, y_n)$ such that $\mathscr{B}(f(y_1, \ldots, y_n), y_1, \ldots, y_n)$ holds for all $y_1, \ldots, y_n$. In cases where we have proved the existence of a unique object u that satisfies a wf $\mathscr{B}(u)$ and $\mathscr{B}(u)$ contains u as its only free variable, then we introduce a new individual constant b such that $\mathscr{B}(b)$ holds. It is generally acknowledged that such definitions, though convenient, add nothing really new to the theory. This can be made precise in the following manner.

Proposition 2.28

Let K be a theory with equality. Assume that $\vdash_K (\exists_1 u)\mathscr{B}(u, y_1, \ldots, y_n)$. Let $K^{\#}$ be the theory with equality obtained by adding to K a new function letter f of n arguments and the proper axiom $\mathscr{B}(f(y_1, \ldots, y_n), y_1, \ldots, y_n)$,* as well as all

* It is better to take this axiom in the form $(\forall u)(u = f(y_1, \ldots, y_n) \Rightarrow \mathscr{B}(u, y_1, \ldots, y_n))$, since $f(y_1, \ldots, y_n)$ might not be free for u in $\mathscr{B}(u, y_1, \ldots, y_n)$.

instances of axioms (A1)–(A7) that involve f. Then there is an effective transformation mapping each wf $\mathscr{C}$ of K$^{\#}$ into a wf $\mathscr{C}^{\#}$ of K such that:

a. If f does not occur in $\mathscr{C}$, then $\mathscr{C}^{\#}$ is $\mathscr{C}$.
b. $(\neg\,\mathscr{C})^{\#}$ is $\neg\,(\mathscr{C}^{\#})$.
c. $(\mathscr{C} \Rightarrow \mathscr{D})^{\#}$ is $\mathscr{E}^{\#} \Rightarrow \mathscr{D}^{\#}$.
d. $((\forall x)\mathscr{C})^{\#}$ is $(\forall x)(\mathscr{C}^{\#})$.
e. $\vdash_{K^{\#}}(\mathscr{C} \Leftarrow \mathscr{C}^{\#})$.
f. If $\vdash_{K^{\#}}\mathscr{C}$, then $\vdash_{K}\mathscr{C}^{\#}$.

Hence, if $\mathscr{C}$ does not contain f and $\vdash_{K}{}^{\#}\,\mathscr{C}$, then $\vdash_{K}\mathscr{C}$.

Proof

By a *simple f-term* we mean an expression $f(t_1, \ldots, t_n)$ in which $t_1, \ldots, t_n$ are terms that do not contain f. Given an atomic wf $\mathscr{C}$ of K$^{\#}$, let $\mathscr{C}^*$ be the result of replacing the leftmost occurrence of a simple term $f(t_1, \ldots, t_n)$ in $\mathscr{C}$ by the first variable v not in $\mathscr{C}$ or $\mathscr{B}$. Call the wf $(\exists v)(\mathscr{B}(v, t_1, \ldots, t_n) \wedge \mathscr{C}^*)$ the *f-transform* of $\mathscr{C}$. If $\mathscr{C}$ does not contain f, then let $\mathscr{C}$ be its own f-transform. Clearly, $\vdash_{K^{\#}}(\exists v)(\mathscr{B}(v, t_1, \ldots, t_n) \wedge \mathscr{C}^*) \Leftarrow \mathscr{C}$. (Here, we use $\vdash_K (\exists_1 u)\mathscr{B}(u, y_1, \ldots, y_n)$ and the axiom $\mathscr{B}(f(y_1, \ldots, y_n), y_1, \ldots, y_n)$ of K$^{\#}$.) Since the f-transform $\mathscr{C}'$ of $\mathscr{C}$ contains one less f than $\mathscr{C}$ and $\vdash_{K^{\#}}\mathscr{C}' \Leftrightarrow \mathscr{C}$, if we take successive f-transforms, eventually we obtain a wf C$^{\#}$ that does not contain f and such that $\vdash_{K^{\#}}\mathscr{C}^{\#} \Leftrightarrow \mathscr{C}$. Call $\mathscr{C}^{\#}$ the *f-less transform* of $\mathscr{C}$. Extend the definition to all wfs of K$^{\#}$ by letting $(\neg\,\mathscr{D})^{\#}$ be $\neg\,(\mathscr{D}^{\#})$, $(\mathscr{D} \Rightarrow \mathscr{E})^{\#}$ be $\mathscr{D}^{\#} \Rightarrow \mathscr{E}^{\#}$, and $((\forall\, x)\mathscr{D})^{\#}$ be $(\forall\, x)\mathscr{D}^{\#}$. Properties (a)–(e) of Proposition 2.28 are then obvious. To prove property (f), it suffices, by property (e), to show that, if $\mathscr{C}$ does not contain f and $\vdash_{K^{\#}}\mathscr{C}$, then $\vdash_K \mathscr{C}$. We may assume that $\mathscr{C}$ is a closed wf, since a wf and its closure are deducible from each other.

Assume that M is a model of K. Let M$_1$ be the normal model obtained by contracting M. We know that a wf is true for M if and only if it is true for M$_1$. Since $\vdash_K (\exists_1 u)\mathscr{B}(u, y_1, \ldots, y_n)$, then, for any $b_1, \ldots, b_n$ in the domain of M$_1$, there is a unique c in the domain of M$_1$ such that $\vDash_{M_1} \mathscr{B}[c, b_1, \ldots, b_n]$. If we define $f_1(b_1, \ldots, b_n)$ to be c, then, taking f_1 to be the interpretation of the function letter f, we obtain from M$_1$ a model M$^{\#}$ of K$^{\#}$. For the logical axioms of K$^{\#}$ (including the equality axioms of K$^{\#}$) are true in any normal interpretation, and the axiom $\mathscr{B}(f(y_1, \ldots, y_n), y_1, \ldots, y_n)$ also holds in M$^{\#}$ by virtue of the definition of f_1. Since the other proper axioms of K$^{\#}$ do not contain f and since they are true for M$_1$, they are also true for M$^{\#}$. But $\vdash_{K^{\#}}\mathscr{C}$. Therefore, $\mathscr{C}$ is true for M$^{\#}$, but since $\mathscr{C}$ does not contain f, $\mathscr{C}$ is true for M$_1$ and hence also for M. Thus, $\mathscr{C}$ is true for every model of K. Therefore, by Corollary 2.20(a), $\vdash_K \mathscr{C}$. (In the case where $\vdash_K (\exists_1 u)\mathscr{B}(u)$ and $\mathscr{B}(u)$ contains only u as a free variable, we form K$^{\#}$ by adding a new individual constant b and the axiom $\mathscr{B}(b)$. Then the analogue of Proposition 2.28 follows from practically the same proof as the one just given.)

Exercise

2.83 Find the f-less transforms of the following wfs.

a. $(\forall x)(\exists y)(A_1^3(x,y,f(x,y_1,\dots,y_n)) \Rightarrow f(y,x,\dots,x)=x)$

b. $A_1^1(f(y_1,\dots,y_{n-1},f(y_1,\dots,y_n))) \wedge (\exists x)A_1^2(x,f(y_1,\dots,y_n))$

Note that Proposition 2.28 also applies when we have introduced several new symbols $f_1, \dots, f_m$ because we can assume that we have added each f_i to the theory already obtained by the addition of $f_1, \dots, f_{i-1}$; then m successive applications of Proposition 2.28 are necessary. The resulting wf $\mathscr{C}^{\#}$ of K can be considered an $(f_1, \dots, f_m)$-free transform of $\mathscr{C}$ into the language of K.

Examples

1. In the elementary theory G of groups, one can prove $(\exists_1 y)x + y = 0$. Then introduce a new function f of one argument, abbreviate $f(t)$ by $(-t)$, and add the new axiom $x + (-x) = 0$. By Proposition 2.28, we now are not able to prove any wf of G that we could not prove before. Thus, the definition of $(-t)$ adds no really new power to the original theory.
2. In the elementary theory F of fields, one can prove that $(\exists_1 y)((x \neq 0 \wedge x \cdot y = 1) \vee (x = 0 \wedge y = 0))$. We then introduce a new function letter g of one argument, abbreviate $g(t)$ by t^{-1}, and add the axiom $(x \neq 0 \wedge x \cdot x^{-1} = 1) \vee (x = 0 \wedge x^{-1} = 0)$, from which one can prove $x \neq 0 \Rightarrow x \cdot x^{-1} = 1$.

From the statement and proof of Proposition 2.28 we can see that, in theories with equality, only predicate letters are needed; function letters and individual constants are dispensable. If f_j^n is a function letter, we can replace it by a new predicate letter A_k^{n+1} if we add the axiom $(\exists_1 u)A_k^{n+1}(u, y_1, \dots, y_n)$. An individual constant is to be replaced by a new predicate letter A_k^1 if we add the axiom $(\exists_1 u)A_k^1(u)$.

Example

In the elementary theory G of groups, we can replace + and 0 by predicate letters A_1^3 and A_1^1 if we add the axioms $(\forall x_1)(\forall x_2)(\exists_1 x_3)A_1^3(x_1,x_2,x_3)$ and $(\exists_1 x_1)A_1^1(x_1)$, and if we replace axioms (a), (b), (c), and (g) by the following:

a′. $A_1^3(x_2,x_3,u) \wedge A_1^3(x_1,u,v) \wedge A_1^3(x_1,x_2,w) \wedge A_1^3(w,x_3,y) \Rightarrow v = y$

b′. $A_1^1(y) \wedge A_1^3(x,y,z) \Rightarrow z = x$

c′. $(\exists y)(\forall u)(\forall v)(A_1^1(u) \wedge A_1^3(x,y,v) \Rightarrow v = u)$

g′. $[x_1 = x_2 \wedge A_1^3(x_1,y,z) \wedge A_1^3(x_2,y,u) \wedge A_1^3(y,x_1,v) \wedge A_1^3(y,x_2,w)] \Rightarrow z = u \wedge v = w$

Notice that the proof of Proposition 2.28 is highly nonconstructive, since it uses semantical notions (model, truth) and is based upon Corollary 2.20(a), which was proved in a nonconstructive way. Constructive syntactical proofs have been given for Proposition 2.28 (see Kleene, 1952, § 74), but, in general, they are quite complex.

Descriptive phrases of the kind "the u such that $\mathscr{B}(u, y_1, \ldots, y_n)$" are very common in ordinary language and in mathematics. Such phrases are called *definite descriptions*. We let $\iota u(\mathscr{B}(u, y_1, \ldots, y_n))$ denote the unique object u such that $\mathscr{B}(u, y_1, \ldots, y_n)$ if there is such a unique object. If there is no such unique object, either we may let $\iota u(\mathscr{B}(u, y_1, \ldots, y_n))$ stand for some fixed object, or we may consider it meaningless. (For example, we may say that the phrases "the present king of France" and "the smallest integer" are meaningless or we may arbitrarily make the convention that they denote 0.) There are various ways of incorporating these ι-terms in formalized theories, but since in most cases the same results are obtained by using new function letters or individual constants as above, and since they all lead to theorems similar to Proposition 2.28, we shall not discuss them any further here. For details, see Hilbert and Bernays (1934) and Rosser (1939, 1953).

2.10 Prenex Normal Forms

A wf $(Q_1y_1) \ldots (Q_ny_n)\mathscr{B}$, where each (Q_iy_i) is either $(\forall y_i)$ or $(\exists y_i)$, y_i is different from y_j for $i \neq j$, and $\mathscr{B}$ contains no quantifiers, is said to be in *prenex normal form*. (We include the case $n = 0$, when there are no quantifiers at all.) We shall prove that, for every wf, we can construct an equivalent prenex normal form.

Lemma 2.29

In any theory, if y is not free in $\mathscr{D}$, and $\mathscr{C}(x)$ and $\mathscr{C}(y)$ are similar, then the following hold.

a. $\vdash ((\forall x)\mathscr{C}(x) \Rightarrow \mathscr{D}) \Leftrightarrow (\exists y)(\mathscr{C}(y) \Rightarrow \mathscr{D})$
b. $\vdash ((\exists x)\mathscr{C}(x) \Rightarrow \mathscr{D}) \Leftrightarrow (\forall y)(\mathscr{C}(y) \Rightarrow \mathscr{D})$
c. $\vdash (\mathscr{D} \Rightarrow (\forall x)\mathscr{C}(x)) \Leftrightarrow (\forall y)(\mathscr{D} \Rightarrow \mathscr{C}(y))$
d. $\vdash \neg(\mathscr{D} \Rightarrow (\exists x)\mathscr{C}(x)) \Leftrightarrow (\exists y)(\mathscr{D} \Rightarrow \mathscr{C}(y))$
e. $\vdash \neg(\forall x)\mathscr{C} \Leftrightarrow (\exists x) \neg\mathscr{C}$
f. $\vdash \neg(\exists x)\mathscr{C} \Leftrightarrow (\forall x) \neg\mathscr{C}$

Proof

For part (a):

1. $(\forall x)\mathscr{C}(x) \Rightarrow \mathscr{D}$	Hyp
2. $\neg(\exists y)(\mathscr{C}(y) \Rightarrow \mathscr{D})$	Hyp
3. $\neg\neg(\forall y)\,\neg(\mathscr{C}(y) \Rightarrow \mathscr{D})$	2, abbreviation
4. $(\forall y)\,\neg(\mathscr{C}(y) \Rightarrow \mathscr{D})$	3, negation elimination
5. $(\forall y)(\mathscr{C}(y) \wedge \neg\mathscr{D})$	4, tautology, Proposition 2.9(c)
6. $\mathscr{C}(y) \wedge \neg\mathscr{D}$	5, rule A4
7. $\mathscr{C}(y)$	6, conjunction elimination
8. $(\forall y)\mathscr{C}(y)$	7, Gen
9. $(\forall x)\mathscr{C}(x)$	8, Lemma 2.11, Biconditional elimination
10. $\mathscr{D}$	1, 9, MP
11. $\neg\mathscr{D}$	6, conjunction elimination
12. $\mathscr{D} \wedge \neg\mathscr{D}$	10, 11, conjunction introduction
13. $(\forall x)\mathscr{C}(x) \Rightarrow \mathscr{D}$, $\neg(\exists y)(\mathscr{C}(y) \Rightarrow \mathscr{D}) \vdash \mathscr{D} \wedge \neg\mathscr{D}$	1–12
14. $(\forall x)\mathscr{C}(x) \Rightarrow \mathscr{D} \vdash (\exists y)(\mathscr{C}(y) \Rightarrow \mathscr{D})$	1–13, proof by contradiction
15. $\vdash (\forall x)\mathscr{C}(x) \Rightarrow \mathscr{D} \Rightarrow (\exists y)(\mathscr{C}(y) \Rightarrow \mathscr{D})$	1–14, Corollary 2.6

The converse is proven in the following manner.

1. $(\exists y)(\mathscr{C}(y) \Rightarrow \mathscr{D})$	Hyp
2. $(\forall x)\mathscr{C}(x)$	Hyp
3. $\mathscr{C}(b) \Rightarrow \mathscr{D}$	1, rule C
4. $\mathscr{C}(b)$	2, rule A4
5. $\mathscr{D}$	3, 4, MP
6. $(\exists y)(\mathscr{C}(y) \Rightarrow \mathscr{D}), (\forall x)\mathscr{C}(x) \vdash_C \mathscr{D}$	1–5
7. $(\exists y)(\mathscr{C}(y) \Rightarrow \mathscr{D}), (\forall x)\mathscr{C}(x) \vdash \mathscr{D}$	6, Proposition 2.10
8. $\vdash (\exists y)(\mathscr{C}(y) \Rightarrow \mathscr{D}) \Rightarrow ((\forall x)\mathscr{C}(x) \Rightarrow \mathscr{D})$	1–7, Corollary 2.6 twice

Part (a) follows from the two proofs above by biconditional introduction. Parts (b)–(f) are proved easily and left as an exercise. (Part (f) is trivial, and (e) follows from Exercise 2.33(a); (c) and (d) follow easily from (b) and (a), respectively.)

Lemma 2.29 allows us to move interior quantifiers to the front of a wf. This is the essential process in the proof of the following proposition.

Proposition 2.30

There is an effective procedure for transforming any wf $\mathscr{B}$ into a wf $\mathscr{C}$ in prenex normal form such that $\vdash \mathscr{B} \Leftrightarrow \mathscr{C}$.

Proof

We describe the procedure by induction on the number k of occurrences of connectives and quantifiers in $\mathscr{B}$. (By Exercise 2.32(a,b), we may assume that the quantified variables in the prefix that we shall obtain are distinct.) If $k = 0$, then let $\mathscr{C}$ be $\mathscr{B}$ itself. Assume that we can find a corresponding $\mathscr{C}$ for all wfs with $k < n$, and assume that $\mathscr{B}$ has n occurrences of connectives and quantifiers.

Case 1. If $\mathscr{B}$ is $\neg\mathscr{D}$, then, by inductive hypothesis, we can construct a wf $\mathscr{E}$ in prenex normal form such that $\vdash \mathscr{D} \Leftrightarrow \mathscr{E}$. Hence, $\vdash \neg\mathscr{D} \Leftrightarrow \neg\mathscr{E}$ by biconditional negation. Thus, $\vdash \mathscr{B} \Leftrightarrow \neg\mathscr{E}$, and, by applying parts (e) and (f) of Lemma 2.29 and the replacement theorem (Proposition 2.9(b)), we can find a wf $\mathscr{C}$ in prenex normal form such that $\vdash \neg\mathscr{E} \Leftrightarrow \mathscr{C}$. Hence, $\vdash \mathscr{B} \Leftrightarrow \mathscr{C}$.

Case 2. If $\mathscr{B}$ is $\mathscr{D} \Rightarrow \mathscr{E}$, then, by inductive hypothesis, we can find wfs $\mathscr{D}_1$ and $\mathscr{E}_1$ in prenex normal form such that $\vdash \mathscr{D} \Leftrightarrow \mathscr{D}_1$ and $\vdash \mathscr{E} \Leftrightarrow \mathscr{E}_1$. Hence, by a suitable tautology and MP, $\vdash (\mathscr{D} \Rightarrow \mathscr{E}) \Leftrightarrow (\mathscr{D}_1 \Rightarrow \mathscr{E}_1)$, that is, $\vdash \mathscr{B} \Leftrightarrow (\mathscr{D}_1 \Rightarrow \mathscr{E}_1)$. Now, applying parts (a)–(d) of Lemma 2.29 and the replacement theorem, we can move the quantifiers in the prefixes of $\mathscr{D}_1$ and $\mathscr{E}_1$ to the front, obtaining a wf $\mathscr{C}$ in prenex normal form such that $\vdash \mathscr{B} \Leftrightarrow \mathscr{C}$.

Case 3. If $\mathscr{B}$ is $(\forall x)\mathscr{D}$, then, by inductive hypothesis, there is a wf $\mathscr{D}_1$ in prenex normal form such that $\vdash \mathscr{D} \Leftrightarrow \mathscr{D}_1$; hence, $\vdash \mathscr{B} \Leftrightarrow (\forall x)\mathscr{D}_1$ by Gen, Lemma 2.8, and MP. But $(\forall x)\mathscr{D}_1$ is in prenex normal form.

Examples

1. Let $\mathscr{B}$ be $(\forall x)(A_1^1(x) \Rightarrow (\forall y)(A_2^2(x,y) \Rightarrow \neg(\forall z)(A_3^2(y,z))))$. By part (e) of Lemma 2.29: $(\forall x)(A_1^1(x) \Rightarrow (\forall y)[A_2^2(x,y) \Rightarrow (\exists z)\neg A_3^2(y,z)])$.

   By part (d): $(\forall x)(A_1^1(x) \Rightarrow (\forall y)(\exists u)[A_2^2(x,y) \Rightarrow \neg A_3^2(y,u)])$.

   By part (c): $(\forall x)(\forall v)(A_1^1(x) \Rightarrow (\exists u)[A_2^2(x,v) \Rightarrow \neg A_3^2(v,u)])$.

   By part (d): $(\forall x)(\forall v)(\exists w)(A_1^1(x) \Rightarrow (A_2^2(x,v) \Rightarrow \neg A_3^2(v,w)))$.

   Changing bound variables: $(\forall x)(\forall y)(\exists z)(A_1^1(x) \Rightarrow (A_2^2(x,y) \Rightarrow \neg A_3^2(y,z)))$.

2. Let $\mathscr{B}$ be $A_1^2(x,y) \Rightarrow (\exists y)[A_1^1(y) \Rightarrow ([(\exists x)A_1^1(x)] \Rightarrow A_2^1(y))]$.

   By part (b): $A_1^2(x,y) \Rightarrow (\exists y)(A_1^1(y) \Rightarrow (\forall u)[A_1^1(u) \Rightarrow A_2^1(y)])$.

   By part (c): $A_1^2(x,y) \Rightarrow (\exists y)(\forall v)(A_1^1(y) \Rightarrow [A_1^1(v) \Rightarrow A_2^1(y)])$.

   By part (d): $(\exists w)(A_1^2(x,y) \Rightarrow (\forall v)[A_1^1(w) \Rightarrow (A_1^1(v) \Rightarrow A_2^1(w))])$.

   By part (c): $(\exists w)(\forall z)(A_1^2(x,y) \Rightarrow [A_1^1(w) \Rightarrow (A_1^1(z) \Rightarrow A_2^1(w))])$.

Exercise

2.84 Find prenex normal forms equivalent to the following wfs.

a. $[(\forall x)(A_1^1(x) \Rightarrow A_1^2(x,y))] \Rightarrow ([(\exists y)A_1^1(y)] \Rightarrow (\exists z)A_1^2(y,z))$
b. $(\exists x)A_1^2(x,y) \Rightarrow (A_1^1(x) \Rightarrow \neg(\exists u)A_1^2(x,u))$

A predicate calculus in which there are no function letters or individual constants and in which, for any positive integer n, there are infinitely many predicate letters with n arguments, will be called a *pure predicate calculus*. For pure predicate calculi we can find a very simple prenex normal form theorem. A wf in prenex normal form such that all existential quantifiers (if any) precede all universal quantifiers (if any) is said to be in *Skolem normal form*.

Proposition 2.31

In a pure predicate calculus, there is an effective procedure assigning to each wf $\mathscr{B}$ another wf $\mathscr{S}$ in Skolem normal form such that $\vdash \mathscr{B}$ if and only if $\vdash \mathscr{S}$ (or, equivalently, by Gödel's completeness theorem, such that $\mathscr{B}$ is logically valid if and only if $\mathscr{S}$ is logically valid).

Proof

First we may assume that $\mathscr{B}$ is a closed wf, since a wf is provable if and only if its closure is provable. By Proposition 2.30 we may also assume that $\mathscr{B}$ is in prenex normal form. Let the *rank* r of $\mathscr{B}$ be the number of universal quantifiers in $\mathscr{B}$ that precede existential quantifiers. By induction on the rank, we shall describe the process for finding Skolem normal forms. Clearly, when the rank is 0, we already have the Skolem normal form. Let us assume that we can construct Skolem normal forms when the rank is less than r, and let r be the rank of $\mathscr{B}$. $\mathscr{B}$ can be written as follows: $(\exists y_1) \dots (\exists y_n) (\forall u)\mathscr{C}(y_1, \dots, y_n, u)$, where $\mathscr{C}(y_1, \dots, y_n, u)$ has only $y_1, \dots, y_n, u$ as its free variables. Let A_j^{n+1} be the first predicate letter of $n + 1$ arguments that does not occur in $\mathscr{B}$. Construct the wf

$$(\mathscr{B}_1) \quad (\exists y_1)\dots(\exists y_n)([(\forall u)(\mathscr{C}(y_1,\dots,y_n,u) \Rightarrow A_j^{n+1}(y_1,\dots,y_n,u))] \\ \Rightarrow (\forall u)A_j^{n+1}(y_1,\dots,y_n,u))$$

Let us show that $\vdash \mathscr{B}$ if and only if $\vdash \mathscr{B}_1$. Assume $\vdash \mathscr{B}_1$. In the proof of $\mathscr{B}_1$, replace all occurrences of $A_j^{n+1}(z_1,\dots,z_n,w)$ by $\mathscr{C}^*(z_1, \dots, z_n, w))$, where $\mathscr{C}^*$ is obtained from $\mathscr{C}$ by replacing all bound variables having free occurrences in the proof by new variables not occurring in the proof. The result is a proof of

$$(\exists y_1)\dots(\exists y_n)(((\forall u)(\mathscr{C}(y_1,\dots,y_n,u) \Rightarrow \mathscr{C}^*(y_1,\dots,y_n,u))) \\ \Rightarrow (\forall u)\mathscr{C}^*(y_1,\dots,y_n,u))$$

($\mathscr{C}^*$ was used instead of $\mathscr{C}$ so that applications of axiom (A4) would remain applications of the same axiom.) Now, by changing the bound variables back again, we see that

$$\vdash (\exists y_1)\dots(\exists y_n)[(\forall u)(\mathscr{C}(y_1,\dots,y_n,u) \Rightarrow \mathscr{C}(y_1,\dots,y_n,u)) \\ \Rightarrow (\forall u)\mathscr{C}(y_1,\dots,y_n,u)]$$

Since $\vdash (\forall u)(\mathscr{C}(y_1, \dots, y_n, u) \Rightarrow \mathscr{C}(y_1, \dots, y_n, u))$, we obtain, by the replacement theorem, $\vdash (\exists y_1) \dots (\exists y_n)(\forall u)\mathscr{C}(y_1, \dots, y_n, u)$, that is, $\vdash \mathscr{B}$. Conversely, assume that $\vdash \mathscr{B}$. By rule C, we obtain $(\forall u)\mathscr{C}(b_1, \dots, b_n, u)$. But, $\vdash (\forall u)\mathscr{D} \Rightarrow ((\forall u)(\mathscr{D} \Rightarrow \mathscr{E}) \Rightarrow (\forall u)\mathscr{E})$ (see Exercise 2.27 (a)) for any wfs $\mathscr{D}$ and $\mathscr{E}$. Hence, $\vdash_C (\forall u)(\mathscr{C}(b_1,\dots,b_n,u) \Rightarrow A_j^{n+1}(b_1,\dots,b_n,u)) \Rightarrow (\forall u)A_j^{n+1}\ (b_1, \dots, b_n, u)$. So, by rule E4, $\vdash_C (\exists y_1)\dots(\exists y_n)([(\forall u)(\mathscr{C}\left(b_1,\dots,b_n,u\right) \Rightarrow A_j^{n+1}\left(y_1,\dots,y_n,u\right))] \Rightarrow (\forall u)A_j^{n+1}(y_1,\dots,y_n,u))$, that is, $\vdash_C \mathscr{B}_1$. By Proposition 2.10, $\vdash \mathscr{B}_1$. A prenex normal form of $\mathscr{B}_1$ has the form $\mathscr{B}_2$: $(\exists y_1) \dots (\exists y_n)\ (\exists u)(Q_1z_1) \dots (Q_sz_s)(\forall v)\mathscr{G}$, where $\mathscr{G}$ has no quantifiers and $(Q_1z_1) \dots (Q_sz_s)$ is the prefix of $\mathscr{C}$. [In deriving the prenex normal form, first, by Lemma 2.29(a), we pull out the first $(\forall u)$, which changes to $(\exists u)$; then we pull out of the first conditional the quantifiers in the prefix of $\mathscr{C}$. By Lemma 2.29(a,b), this exchanges existential and universal quantifiers, but then we again pull these out of the second conditional of $\mathscr{B}_1$, which brings the prefix back to its original form. Finally, by Lemma 2.29(c), we bring the second $(\forall u)$ out to the prefix, changing it to a new quantifier $(\forall v)$.] Clearly, $\mathscr{B}_2$ has rank one less than the rank of $\mathscr{B}$ and, by Proposition 2.30, $\vdash \mathscr{B}_1 \Leftrightarrow \mathscr{B}_2$. But, $\vdash \mathscr{B}$ if and only if $\vdash \mathscr{B}_1$. Hence, $\vdash \mathscr{B}$ if and only if $\vdash \mathscr{B}_2$. By inductive hypothesis, we can find a Skolem normal form for $\mathscr{B}_2$, which is also a Skolem normal form for $\mathscr{B}$.

Example

$\mathscr{B}$: $(\forall x)(\forall y)(\exists z)\mathscr{C}(x, y, z)$, where $\mathscr{C}$ contains no quantifiers

$$\mathscr{B}_1 : (\forall x)((\forall y)(\exists z)\mathscr{C}(x,y,z) \Rightarrow A_j^1(x)) \Rightarrow (\forall x)A_j^1(x)\text{, where } A_j^1 \text{ is not in } \mathscr{C}.$$

We obtain the prenex normal form of $\mathscr{B}_1$:

$$(\exists x)\Big(\Big[(\forall y)(\exists z)\mathscr{C}(x,y,z) \Rightarrow A_j^1(x)\Big] \Rightarrow (\forall x)A_j^1(x)\Big) \qquad 2.29(a)$$

$$(\exists x)\Big((\exists y)\Big[[(\exists z)\mathscr{C}(x,y,z) \Rightarrow A_j^1(x)\Big] \Rightarrow (\forall x)A_j^1(x)\Big] \qquad 2.29(a)$$

$$(\exists x)\Big((\exists y)(\forall z)\Big[\mathscr{C}(x,y,z) \Rightarrow A_j^1(x)\Big] \Rightarrow (\forall x)A_j^1(x)\Big) \qquad 2.29(b)$$

$$(\exists x)(\forall y)\Big[(\forall z)\Big(\mathscr{C}(x,y,z) \Rightarrow A_j^1(x)\Big) \Rightarrow (\forall x)A_j^1(x)\Big] \qquad 2.29(b)$$

$$(\exists x)(\forall y)(\exists z)\Big(\Big[\mathscr{C}(x,y,z) \Rightarrow A_j^1(x)\Big] \Rightarrow (\forall x)A_j^1(x)\Big] \qquad 2.29(a)$$

$$(\exists x)(\forall y)(\exists z)(\forall v)\Big[\Big(\mathscr{C}(x,y,z) \Rightarrow A_j^1(x)\Big) \Rightarrow A_j^1(v)\Big] \qquad 2.29(c)$$

We repeat this process again: Let $\mathscr{D}(x, y, z, v)$ be $(\mathscr{C}(x,y,z) \Rightarrow A_j^1(x)) \Rightarrow A_j^1(v)$. Let A_k^2 not occur in $\mathscr{D}$. Form:

$$(\exists x)\Big(\Big[(\forall y)\Big[(\exists z)(\forall v)(\mathscr{D}(x,y,z,v)) \Rightarrow A_k^2(x,y)\Big]\Big] \Rightarrow (\forall y)A_k^2(x,y)\Big)$$

$$(\exists x)(\exists y)[[(\exists z)(\forall v)(\mathscr{D}(x,y,z,v)) \Rightarrow A_k^2(x,y)] \Rightarrow (\forall y)A_k^2(x,y)] \qquad 2.29(a)$$

$$(\exists x)(\exists y)(\exists z)(\exists v)([(\mathscr{D}(x,y,z,v) \Rightarrow A_k^2(x,y)] \Rightarrow (\forall y)A_k^2(x,y)) \qquad 2.29(a,b)$$

$$(\exists x)(\exists y)(\exists z)(\forall v)(\forall w)([\mathscr{D}(x,y,z,v) \Rightarrow A_k^2(x,y)] \Rightarrow A_k^2(x,w)] \qquad 2.29(c)$$

Thus, a Skolem normal form of $\mathscr{B}$ is:

$$(\exists x)(\exists y)(\exists z)(\forall v)(\forall w)([((\mathscr{C}(x,y,z) \Rightarrow A_j^1(x)) \Rightarrow A_j^1(v)) \Rightarrow A_k^2(x,y)] \Rightarrow A_k^2(x,w))$$

Exercises

2.85 Find Skolem normal forms for the following wfs.

a. $\neg(\exists x)A_1^1(x) \Rightarrow (\forall u)(\exists y)(\forall x)A_1^3(u,x,y)$

b. $(\forall x)(\exists y)(\forall u)(\exists v)A_1^4(x,y,u,v)$

2.86 Show that there is an effective procedure that gives, for each wf $\mathscr{B}$ of a pure predicate calculus, another wf $\mathscr{D}$ of this calculus of the form $(\forall y_1) \ldots (\forall y_n)(\exists z_1) \ldots (\exists z_m)\mathscr{C}$, such that $\mathscr{C}$ is quantifier-free, $n, m \geq 0$, and $\mathscr{B}$ is satisfiable if and only if $\mathscr{D}$ is satisfiable. [*Hint*: Apply Proposition 2.31 to $\neg\mathscr{B}$.]

2.87 Find a Skolem normal form $\mathscr{S}$ for $(\forall x)(\exists y)A_1^2(x,y)$ and show that it is not the case that $\vdash \mathscr{S} \Leftrightarrow (\forall x)(\exists y)A_1^2(x,y)$. Hence, a Skolem normal form for a wf $\mathscr{B}$ is not necessarily logically equivalent to $\mathscr{B}$, in contradistinction to the prenex normal form given by Proposition 2.30.

2.11 Isomorphism of Interpretations: Categoricity of Theories

We shall say that an interpretation M of some language $\mathscr{L}$ is *isomorphic* with an interpretation M* of $\mathscr{L}$ if and only if there is a one–one correspondence g (called an *isomorphism*) of the domain D of M with the domain D^* of M* such that:

1. For any predicate letter A_j^n of $\mathscr{L}$ and for any $b_1, \ldots, b_n$ in D, $\vDash_M A_j^n[b_1, \ldots, b_n]$ if and only if $\vDash_{M^*} A_j^n[g(b_1), \ldots, g(b_n)]$.
2. For any function letter f_j^n of $\mathscr{L}$ and for any $b_1, \ldots, b_n$ in D, $g((f_j^n)^M(b_1, \ldots, b_n)) = (f_j^n)^{M^*}(g(b_1), \ldots, g(b_n))$.
3. For any individual constant a_j of $\mathscr{L}$, $g((a_j)^M) = (a_j)^{M^*}$.

The notation M ≈ M* will be used to indicate that M is isomorphic with M*. Notice that, if M ≈ M*, then the domains of M and M* must be of the same cardinality.

Proposition 2.32

If g is an isomorphism of M with M*, then:

a. for any wf $\mathscr{B}$ of $\mathscr{L}$, any sequence $s = (b_1, b_2, \ldots)$ of elements of the domain D of M, and the corresponding sequence $g(s) = (g(b_1), g(b_2), \ldots)$, s satisfies $\mathscr{B}$ in M if and only if $g(s)$ satisfies $\mathscr{B}$ in M*;
b. hence, $\vDash_M \mathscr{B}$ if and only if $\vDash_{M^*} \mathscr{B}$.

Proof

Part (b) follows directly from part (a). The proof of part (a) is by induction on the number of connectives and quantifiers in $\mathscr{B}$ and is left as an exercise.

From the definition of isomorphic interpretations and Proposition 2.32 we see that isomorphic interpretations have the same "structure" and, thus, differ in no essential way.

Exercises

2.88 Prove that, if M is an interpretation with domain D and D^* is a set that has the same cardinality as D, then one can define an interpretation M* with domain D^* such that M is isomorphic with M*.

2.89 Prove the following: (a) M is isomorphic with M. (b) If M_1 is isomorphic with M_2, then M_2 is isomorphic with M_1. (c) If M_1 is isomorphic with M_2 and M_2 is isomorphic with M_3, then M_1 is isomorphic with M_3.

A theory with equality K is said to be $\mathfrak{m}$—*categorical*, where $\mathfrak{m}$ is a cardinal number, if and only if: any two normal models of K of cardinality m are isomorphic, and K has at least one normal model of cardinality $\mathfrak{m}$ (see Loś, 1954c).

Examples

1. Let K^2 be the pure theory of equality K_1 (see page 96) to which has been added axiom (E2): $(\exists x_1)(\exists x_2)(x_1 \neq x_2 \wedge (\forall x_3)(x_3 = x_1 \vee x_3 = x_2))$. Then K^2 is 2-categorical. Every normal model of K^2 has exactly two elements. More generally, define (E_n) to be:

$$(\exists x_1)\dots(\exists x_n)\left(\bigwedge_{1\le i<j\le n} x_i \neq x_j \wedge (\forall y)(y = x_1 \vee \dots \vee y = x_n)\right)$$

   where $\wedge_{1\le i<j\le n}\, x_i \neq x_j$ is the conjunction of all wfs $x_i \neq x_j$ with $1 \le i < j \le n$. Then, if K^n is obtained from K_1 by adding (E_n) as an axiom, K^n is n-categorical, and every normal model of K^n has exactly n elements.

2. The theory K_2 (see page 96) of densely ordered sets with neither first nor last element is $\aleph_0$–categorical (see Kamke, 1950, p. 71: every denumerable normal model of K_2 is isomorphic with the model consisting of the set of rational numbers under their natural ordering). But one can prove that K_2 is not $\mathfrak{m}$–categorical for any $\mathfrak{m}$ different from $\aleph_0$.

Exercises

2.90[A] Find a theory with equality that is not $\aleph_0$–categorical but is $\mathfrak{m}$–categorical for all $\mathfrak{m} > \aleph_0$. [*Hint*: Consider the theory G_C of abelian groups (see page 96). For each integer n, let ny stand for the term $(y + y) + \cdots + y$ consisting of the sum of n ys. Add to G_C the axioms $(\mathscr{B}_n)$:$(\forall x)(\exists_1 y)(ny = x)$ for all $n \geq 2$. The new theory is the theory of uniquely divisible abelian groups. Its normal models are essentially vector spaces over the field of rational numbers. However, any two vector spaces over the rational numbers of the same nondenumerable cardinality are isomorphic, and there are denumerable vector spaces over the rational numbers that are not isomorphic (see Bourbaki, 1947).]

2.91[A] Find a theory with equality that is $\mathfrak{m}$–categorical for all infinite cardinals $\mathfrak{m}$. [*Hint*: Add to the theory G_C of abelian groups the axiom $(\forall x_1)(2x_1 = 0)$. The normal models of this theory are just the vector spaces over the field of integers modulo 2. Any two such vector spaces of the same cardinality are isomorphic (see Bourbaki, 1947).]

2.92 Show that the theorems of the theory K^n in Example 1 above are precisely the set of all wfs of K^n that are true in all normal models of cardinality n.

2.93[A] Find two nonisomorphic densely ordered sets of cardinality $2^{\aleph_0}$ with neither first nor last element. (This shows that the theory K_2 of Example 2 is not $2^{\aleph_0}$–categorical.)

Is there a theory with equality that is $\mathfrak{m}$–categorical for some noncountable cardinal $\mathfrak{m}$ but not $\mathfrak{m}$–categorical for some other noncountable cardinal $\mathfrak{n}$? In Example 2 we found a theory that is only $\aleph_0$-categorical; in Exercise 2.90 we found a theory that is $\mathfrak{m}$–categorical for all infinite $\mathfrak{m} > \aleph_0$ but not $\aleph_0$–categorical, and in Exercise 2.91, a theory that is $\mathfrak{m}$–categorical for any infinite $\mathfrak{m}$. The elementary theory G of groups is not $\mathfrak{m}$–categorical for any infinite $\mathfrak{m}$. The problem is whether these four cases exhaust all the possibilities. That this is so was proved by Morley (1965).

2.12 Generalized First-Order Theories: Completeness and Decidability*

If, in the definition of the notion of first-order language, we allow a noncountable number of predicate letters, function letters, and individual constants, we arrive at the notion of a *generalized first-order language*. The notions of *interpretation* and *model* extend in an obvious way to a generalized first-order language. A *generalized first-order theory* in such a language is obtained by taking as proper axioms any set of wfs of the language. Ordinary first-order theories are special cases of generalized first-order theories. The reader may easily check that all the results for first-order theories, through Lemma 2.12, hold also for generalized first-order theories without any changes in the proofs. Lemma 2.13 becomes Lemma 2.13′: if the set of symbols of a generalized theory K has cardinality $\aleph\alpha$, then the set of expressions of K also can be well-ordered and has cardinality $\aleph\alpha$. (First, fix a well-ordering of the symbols of K. Second, order the expressions by their length, which is some positive integer, and then stipulate that if e_1 and e_2 are two distinct expressions of the same length k, and j is the first place in which they differ, then e_1 *precedes* e_2 if the jth symbol of e_1 precedes the jth symbol of e_2 according to the given well-ordering of the symbols of K.) Now, under the same assumption as for Lemma 2.13′, Lindenbaum's Lemma 2.14′ can be proved for generalized theories much as before, except that all the enumerations (of the wfs $\mathscr{B}_i$ and of the theories J_i) are transfinite, and the proof that J is consistent and complete uses transfinite induction. The analogue of Henkin's Proposition 2.17 runs as follows.

* Presupposed in parts of this section is a slender acquaintance with ordinal and cardinal numbers (see Chapter 4; or Kamke, 1950; or Sierpinski, 1958).

Proposition 2.33

If the set of symbols of a consistent generalized theory K has cardinality $\aleph_\alpha$, then K has a model of cardinality $\aleph_\alpha$.

Proof

The original proof of Lemma 2.15 is modified in the following way. Add $\aleph_\alpha$ new individual constants $b_1, b_2, \ldots, b_\lambda, \ldots$. As before, the new theory K_0 is consistent. Let $F_1(x_{i_1}), \ldots, F_\lambda(x_{i_\lambda}), \ldots (\lambda < \omega_\alpha)$ be a sequence consisting of all wfs of K_0 with exactly one free variable. Let (S_λ) be the sentence $(\exists x_{i_\lambda})\neg F_\lambda(x_{i_\lambda}) \Rightarrow \neg F_\lambda(b_{j_\lambda})$, where the sequence $b_{j1}, b_{j2}, \ldots b_{j_\lambda}, \ldots$ of distinct individual constants is chosen so that b_{j_λ} does not occur in $F_\beta(x_{i_\beta})$ for $\beta \leq \lambda$. The new theory K_∞, obtained by adding all the wfs (S_λ) as axioms, is proved to be consistent by a transfinite induction analogous to the inductive proof in Lemma 2.15. K_∞ is a scapegoat theory that is an extension of K and contains $\aleph_\alpha$ closed terms. By the extended Lindenbaum Lemma 2.14′, K_∞ can be extended to a consistent, complete scapegoat theory J with $\aleph_\alpha$ closed terms. The same proof as in Lemma 2.16 provides a model M of J of cardinality $\aleph_\alpha$.

Corollary 2.34

a. If the set of symbols of a consistent generalized theory with equality K has cardinality $\aleph_\alpha$, then K has a normal model of cardinality less than or equal to $\aleph_\alpha$.
b. If, in addition, K has an infinite normal model (or if K has arbitrarily large finite normal models), then K has a normal model of any cardinality $\aleph_\beta \geq \aleph_\alpha$.
c. In particular, if K is an ordinary theory with equality (i.e., $\aleph_\alpha = \aleph_0$) and K has an infinite normal model (or if K has arbitrarily large finite normal models), then K has a normal model of any cardinality $\aleph_\beta (\beta \geq 0)$.

Proof

a. The model guaranteed by Proposition 2.33 can be contracted to a normal model consisting of equivalence classes in a set of cardinality $\aleph_\alpha$. Such a set of equivalence classes has cardinality less than or equal to $\aleph_\alpha$.
b. Assume $\aleph_\beta \geq \aleph_\alpha$. Let $b_1, b_2, \ldots$ be a set of new individual constants of cardinality $\aleph_\beta$, and add the axioms $b_\lambda \neq b_\mu$ for $\lambda \neq \mu$. As in the proof of Corollary 2.27, this new theory is consistent and so, by (a), has a

normal model of cardinality less than or equal to $\aleph_\beta$ (since the new theory has $\aleph_\beta$ new symbols). But, because of the axioms $b_\lambda \neq b_\mu$, the normal model has exactly $\aleph_\beta$ elements.

c. This is a special case of (b).

Exercise

2.94 If the set of symbols of a predicate calculus with equality K has cardinality $\aleph_\alpha$, prove that there is an extension K′ of K (with the same symbols as K) such that K′ has a normal model of cardinality $\aleph_\alpha$, but K′ has no normal model of cardinality less than $\aleph_\alpha$.

From Lemma 2.12 and Corollary 2.34(a,b), it follows easily that, if a generalized theory with equality K has $\aleph_\alpha$ symbols, is $\aleph_\beta$-categorical for some $\beta \geq \alpha$, and has no finite models, then K is complete, in the sense that, for any closed wf $\mathscr{B}$, either $\vdash_K \mathscr{B}$ or $\vdash_K \neg\mathscr{B}$ (Vaught, 1954). If not-$\vdash_K \mathscr{B}$ and not-$\vdash_K \neg\mathscr{B}$, then the theories $K' = K + \{\neg\mathscr{B}\}$ and $K'' = K + \{\mathscr{B}\}$ are consistent by Lemma 2.12, and so, by Corollary 2.34(a), there are normal models M′ and M″ of K′ and K″, respectively, of cardinality less than or equal to $\aleph_\alpha$. Since K has no finite models, M′ and M″ are infinite. Hence, by Corollary 2.34(b), there are normal models N′ and N″ of K′ and K″, respectively, of cardinality $\aleph_\beta$. By the $\aleph_\beta$-categoricity of K, N′ and N″ must be isomorphic. But, since $\neg\mathscr{B}$ is true in N′ and $\mathscr{B}$ is true in N″, this is impossible by Proposition 2.32(b). Therefore, either $\vdash_K \mathscr{B}$ or $\vdash_K \neg\mathscr{B}$.

In particular, if K is an ordinary theory with equality that has no finite models and is $\aleph_\beta$-categorical for some $\beta \geq 0$, then K is complete. As an example, consider the theory K_2 of densely ordered sets with neither first nor last element (see page 96). K_2 has no finite models and is $\aleph_0$-categorical.

If an ordinary theory K is axiomatic (i.e., one can effectively decide whether any wf is an axiom) and complete, then K is decidable, that is, there is an effective procedure to determine whether any given wf is a theorem. To see this, remember (see page 84) that if a theory is axiomatic, one can effectively enumerate the theorems. Any wf $\mathscr{B}$ is provable if and only if its closure is provable. Hence, we may confine our attention to closed wfs $\mathscr{B}$. Since K is complete, either $\mathscr{B}$ is a theorem or $\neg\mathscr{B}$ is a theorem, and, therefore, one or the other will eventually turn up in our enumeration of theorems. This provides an effective test for theoremhood. Notice that, if K is inconsistent, then every wf is a theorem and there is an obvious decision procedure; if K is consistent, then not both $\mathscr{B}$ and $\neg\mathscr{B}$ can show up as theorems and we need only wait until one or the other appears.

If an ordinary axiomatic theory with equality K has no finite models and is $\aleph_\beta$-categorical for some $\beta \geq 0$, then, by what we have proved, K is decidable. In particular, the theory K_2 discussed above is decidable.

In certain cases, there is a more direct method of proving completeness or decidability. Let us take as an example the theory K_2 of densely ordered sets with neither first nor last element. Langford (1927) has given the following procedure for K_2. Consider any closed wf $\mathscr{B}$. By Proposition 2.30, we can assume that $\mathscr{B}$ is in prenex normal form $(Q_1y_1) \ldots (Q_ny_n)\mathscr{C}$, where $\mathscr{C}$ contains no quantifiers. If (Q_ny_n) is $(\forall y_n)$, replace $(\forall y_n)\mathscr{C}$ by $\neg(\exists y_n)\neg\mathscr{C}$. In all cases, then, we have, at the right side of the wf, $(\exists y_n)\mathscr{D}$, where $\mathscr{D}$ has no quantifiers. Any negation $x \neq y$ can be replaced by $x < y \vee y < x$, and $\neg(x < y)$ can be replaced by $x = y \vee y < x$. Hence, all negation signs can be eliminated from $\mathscr{D}$. We can now put $\mathscr{D}$ into disjunctive normal form, that is, a disjunction of conjunctions of atomic wfs (see Exercise 1.42). Now $(\exists y_n)(\mathscr{D}_1 \vee \mathscr{D}_2 \vee \ldots \vee \mathscr{D}_k)$ is equivalent to $(\exists y_n)\mathscr{D}_1 \vee (\exists y_n)\mathscr{D}_2 \vee \ldots \vee (\exists y_n)\mathscr{D}_k$. Consider each $(\exists y_n)\mathscr{D}_i$ separately. $\mathscr{D}_i$ is a conjunction of atomic wfs of the form $t < s$ and $t = s$. If $\mathscr{D}_i$ does not contain y_n, just erase $(\exists y_n)$. Note that, if a wf $\mathscr{E}$ does not contain y_n, then $(\exists y_n)(\mathscr{E} \wedge \mathscr{F})$ may be replaced by $\mathscr{E} \wedge (\exists y_n)\mathscr{F}$. Hence, we are reduced to the consideration of $(\exists y_n)\mathscr{F}$, where $\mathscr{F}$ is a conjunction of atomic wfs of the form $t < s$ or $t = s$, each of which contains y_n. Now, if one of the conjuncts is $y_n = z$ for some z different from y_n, then replace in $\mathscr{F}$ all occurrences of y_n by z and erase $(\exists y_n)$. If we have $y_n = y_n$ alone, then just erase $(\exists y_n)$. If we have $y_n = y_n$ as one conjunct among others, then erase $y_n = y_n$. If $\mathscr{F}$ has a conjunct $y_n < y_n$, then replace all of $(\exists y_n)\mathscr{F}$ by $y_n < y_n$. If $\mathscr{F}$ consists of $y_n < z_1 \wedge \ldots \wedge y_n < z_j \wedge u_1 < y_n \wedge \ldots \wedge u_m < y_n$, then replace $(\exists y_n)\mathscr{F}$ by the conjunction of all the wfs $u_i < z_p$ for $1 \leq i \leq m$ and $1 \leq p \leq j$. If all the u_is or all the z_ps are missing, replace $(\exists y_n)\mathscr{F}$ by $y_n = y_n$. This exhausts all possibilities and, in every case, we have replaced $(\exists y_n)\mathscr{F}$ by a wf containing no quantifiers, that is, we have eliminated the quantifier $(\exists y_n)$. We are left with $(Q_1y_1) \ldots (Q_{n-1}y_{n-1})\mathscr{G}$, where $\mathscr{G}$ contains no quantifiers. Now we apply the same procedure successively to $(Q_{n-1}y_{n-1}), \ldots, (Q_1y_1)$. Finally we are left with a wf without quantifiers, built up of wfs of the form $x = x$ and $x < x$. If we replace $x = x$ by $x = x \Rightarrow x = x$ and $x < x$ by $\neg(x = x \Rightarrow x = x)$, the result is either an instance of a tautology or the negation of such an instance. Hence, by Proposition 2.1, either the result or its negation is provable. Now, one can easily check that all the replacements we have made in this whole reduction procedure applied to $\mathscr{B}$ have been replacements of wfs $\mathscr{H}$ by other wfs $\mathscr{U}$ such that $\vdash_K \mathscr{H} \Leftrightarrow \mathscr{U}$. Hence, by the replacement theorem, if our final result $\mathscr{B}^*$ is provable, then so is the original wf $\mathscr{B}$, and, if $\neg\mathscr{B}^*$ is provable, then so is $\neg\mathscr{B}$. Thus, K_2 is complete and decidable.

The method used in this proof, the successive elimination of existential quantifiers, has been applied to other theories. It yields a decision procedure (see Hilbert and Bernays, 1934, §5) for the pure theory of equality K_1 (see page 96). It has been applied by Tarski (1951) to prove the completeness and decidability of elementary algebra (i.e., of the theory of real-closed fields; see van der Waerden, 1949) and by Szmielew (1955) to prove the decidability of the theory G_C of abelian groups.

Exercises

2.95 (Henkin, 1955) If an ordinary theory with equality K is finitely axiomatizable and $\aleph_\alpha$-categorical for some α, prove that K is decidable.

2.96 a. Prove the decidability of the pure theory K_1 of equality.

b. Give an example of a theory with equality that is $\aleph_\alpha$-categorical for some α, but is incomplete.

2.12.1 Mathematical Applications

1. Let F be the elementary theory of fields (see page 96). We let n stand for the term $1 + 1 + \cdots + 1$, consisting of the sum of n 1s. Then the assertion that a field has characteristic p can be expressed by the wf $\mathscr{C}_p$: $p = 0$. A field has characteristic 0 if and only if it does not have characteristic p for any prime p. Then for any closed wf $\mathscr{B}$ of F that is true for all fields of characteristic 0, there is a prime number q such that $\mathscr{B}$ is true for all fields of characteristic greater than or equal to q. To see this, notice that, if F_0 is obtained from F by adding as axioms $\neg\mathscr{C}_2, \neg\mathscr{C}_3, \ldots, \neg\mathscr{C}_p, \ldots$ (for all primes p), the normal models of F_0 are the fields of characteristic 0. Hence, by Exercise 2.77, $\vdash_{F_0} \mathscr{B}$. But then, for some finite set of new axioms $\neg\mathscr{C}_{q_1}, \neg\mathscr{C}_{q_2}, \ldots, \neg\mathscr{C}_{q_n}$, we have $\neg\mathscr{C}_{q_1}, \neg\mathscr{C}_{q_2}, \ldots, \neg\mathscr{C}_{q_n} \vdash_F \mathscr{B}$. Let q be a prime greater than all $q_1, \ldots, q_n$, In every field of characteristic greater than or equal to q, the wfs $\neg\mathscr{C}_{q_1}, \neg\mathscr{C}_{q_2}, \ldots, \neg\mathscr{C}_{q_n}$ are true; hence, $\mathscr{B}$ is also true. (Other applications in algebra may be found in A. Robinson (1951) and Cherlin (1976).)
2. A *graph* may be considered as a set with a symmetric binary relation R (i.e., the relation that holds between two vertices if and only if they are connected by an edge). Call a graph k-colorable if and only if the graph can be divided into k disjoint (possibly empty) sets such that no two elements in the same set are in the relation R. (Intuitively, these sets correspond to k colors, each color being painted on the points in the corresponding set, with the proviso that two points connected by an edge are painted different colors.) Notice that any subgraph of a k-colorable graph is k-colorable. Now we can show that, if every finite subgraph of a graph $\mathscr{G}$ is k-colorable, and if $\mathscr{G}$ can be well-ordered, then the whole graph $\mathscr{G}$ is k-colorable. To prove this, construct the following generalized theory with equality K (Beth, 1953). There are two binary predicate letters, $A_1^2(=)$ and A_2^2 (corresponding to the relation R on $\mathscr{G}$); there are k monadic predicate letters $A_1^1, \ldots, A_k^1$ (corresponding to the k subsets into which we hope to divide the graph); and there are individual constants a_c, one for each element c of the graph $\mathscr{G}$.

As proper axioms, in addition to the usual assumptions (A6) and (A7), we have the following wfs:

I. $\neg A_2^2(x, x)$	(irreflexivity of R)
II. $A_2^2(x, y) \Rightarrow A_2^2(y, x)$	(symmetry of R)
III. $(\forall x)(A_1^1(x) \vee A_2^1(x) \vee \ldots \vee A_k^1(x))$	(division into k classes)
IV. $(\forall x)\neg(A_i^1(x) \wedge A_j^1(x))$	(disjointness of the k classes) for $1 \le i < j \le k$
V. $(\forall x)(\forall y)(A_i^1(x) \wedge A_i^1(y) \Rightarrow \neg A_2^2(x, y))$ for $1 \le i \le k$	(two elements of the same class are not in the relation R)
VI. $a_b \neq a_c$,	for any two distinct elements b and c of $\mathscr{G}$
VII. $A_2^2(a_b, a_c)$, ,	if $R(b, c)$ holds in $\mathscr{G}$

Now, any finite set of these axioms involves only a finite number of the individual constants $a_{c_1}, \ldots, a_{c_n}$, and since the corresponding subgraph $\{c_1, \ldots, c_n\}$ is, by assumption, k-colorable, the given finite set of axioms has a model and is, therefore, consistent. Since any finite set of axioms is consistent, K is consistent. By Corollary 2.34(a), K has a normal model of cardinality less than or equal to the cardinality of $\mathscr{G}$. This model is a k-colorable graph and, by (VI)–(VII), has $\mathscr{G}$ as a subgraph. Hence $\mathscr{G}$ is also k-colorable. (Compare this proof with a standard mathematical proof of the same result by de Bruijn and Erdös (1951). Generally, use of the method above replaces complicated applications of Tychonoff's theorem or König's Unendlichkeits lemma.)

Exercises

2.97[A] (Loś, 1954b) A group B is said to be *orderable* if there exists a binary relation R on B that totally orders B such that, if xRy, then $(x + z) R(y + z)$ and $(z + x)R(z + y)$. Show, by a method similar to that used in Example 2 above, that a group B is orderable if and only if every finitely generated subgroup is orderable (if we assume that the set B can be well-ordered).

2.98[A] Set up a theory for algebraically closed fields of characteristic $p(\geq 0)$ by adding to the theory F of fields the new axioms P_n, where P_n states that every nonconstant polynomial of degree n has a root, as well as axioms that determine the characteristic. Show that every wf of F that holds for one algebraically closed field of characteristic 0 holds for all of them. [*Hint*: This theory is $\aleph_\beta$-categorical for $\beta > 0$, is axiomatizable, and has no finite models. See A. Robinson (1952).]

2.99 By ordinary mathematical reasoning, solve the *finite marriage problem*. Given a finite set M of m men and a set N of women such that each man knows only a finite number of women and, for $1 \le k \le m$, any subset of M having k elements knows at least k women of N (i.e., there are at least k women in N who know at least one of the k given men), then it is possible to marry (monogamously) all the men of M to women in N so that every man is married to a women whom he knows. [*Hint* (Halmos and Vaughn, 1950): $m = 1$ is trivial. For $m > 1$, use induction, considering the cases: (I) for all k with $1 \leq k < m$, every set of k men knows at least $k + 1$ women; and (II) for some k with $1 \leq k < m$, there is a set of k men knowing exactly k women.] Extend this result to the infinite case, that is, when M is infinite and well-orderable and the assumptions above hold for all finite k. [*Hint*: Construct an appropriate generalized theory with equality, analogous to that in Example 2 above, and use Corollary 2.34(a).]

2.100 Prove that there is no generalized theory with equality K, having one predicate letter $<$ in addition to $=$, such that the normal models of K are exactly those normal interpretations in which the interpretation of $<$ is a well-ordering of the domain of the interpretation.

Let $\mathscr{B}$ be a wf in prenex normal form. If $\mathscr{B}$ is not closed, form its closure instead. Suppose, for example, $\mathscr{B}$ is $(\exists y_1)(\forall y_2)(\forall y_3)(\exists y_4)(\exists y_5)(\forall y_6)\mathscr{C}(y_1, y_2, y_3, y_4, y_5, y_6)$, where $\mathscr{C}$ contains no quantifiers. Erase $(\exists y_1)$ and replace y_1 in $\mathscr{C}$ by a new individual constant b_1: $(\forall y_2)(\forall y_3)(\exists y_4)(\exists y_5)(\forall y_6)\ \mathscr{C}(b_1, y_2, y_3, y_4, y_5, y_6)$. Erase $(\forall y_2)$ and $(\forall y_3)$, obtaining $(\exists y_4)(\exists y_5)(\forall y_6)\mathscr{C}(b_1, y_2, y_3, y_4, y_5, y_6)$. Now erase $(\exists y_4)$ and replace y_4 in $\mathscr{C}$ by $g(y_2, y_3)$, where g is a new function letter: $(\exists y_5)(\forall y_6)\mathscr{C}(b_1, y_2, y_3, g(y_2, y_3), y_5, y_6)$. Erase $(\exists y_5)$ and replace y_5 by $h(y_2, y_3)$, where h is another new function letter: $(\forall y_6)\mathscr{C}(b_1, y_2, y_3, g(y_2, y_3), h(y_2, y_3), y_6)$. Finally, erase $(\forall y_6)$. The resulting wf $\mathscr{C}(b_1, y_2, y_3, g(y_2, y_3), h(y_2, y_3), y_6)$ contains no quantifiers and will be denoted by $\mathscr{B}^*$. Thus, by introducing new function letters and individual constants, we can eliminate the quantifiers from a wf.

Examples

1. If $\mathscr{B}$ is $(\forall y_1)(\exists y_2)(\forall y_3)(\forall y_4)(\exists y_5)\mathscr{C}(y_1, y_2, y_3, y_4, y_5)$, where $\mathscr{C}$ is quantifier-free, then $\mathscr{B}^*$ is of the form $\mathscr{C}(y_1, g(y_1), y_3, y_4, h(y_1, y_3, y_4))$.
2. If $\mathscr{B}$ is $(\exists y_1)(\exists y_2)(\forall y_3)(\forall y_4)(\exists y_5)\mathscr{C}(y_1, y_2, y_3, y_4, y_5)$, where $\mathscr{C}$ is quantifier-free, then $\mathscr{B}^*$ is of the form $\mathscr{C}(b, c, y_3, y_4, g(y_3, y_4))$.

Notice that $\mathscr{B}^* \vdash \mathscr{B}$, since we can put the quantifiers back by applications of Gen and rule E4. (To be more precise, in the process of obtaining $\mathscr{B}^*$, we drop all quantifiers and, for each existentially quantified variable y_i, we substitute a term $g(z_1, \ldots, z_k)$, where g is a new function letter and $z_1, \ldots, z_k$ are the variables that were universally quantified in the prefix preceding $(\exists y_i)$. If there are no such variables $z_1, \ldots, z_k$, we replace y_i by a new individual constant.)

Proposition 2.35 (Second ε-Theorem)

(Rasiowa, 1956; Hilbert and Bernays, 1939) Let K be a generalized theory. Replace each axiom $\mathscr{B}$ of K by $\mathscr{B}$*. (The new function letters and individual constants introduced for one axiom are to be different from those introduced for another axiom.) Let K* be the generalized theory with the proper axioms $\mathscr{B}$*. Then:

a. If $\mathscr{D}$ is a wf of K and $\vdash_{K^*} \mathscr{D}$, then $\vdash_K \mathscr{D}$.
b. K is consistent if and only if K* is consistent.

Proof

a. Let $\mathscr{D}$ be a wf of K such that $\vdash_{K^*} \mathscr{D}$. Consider the ordinary theory K° whose axioms $\mathscr{B}_1, \ldots, \mathscr{B}_n$ are such that $\mathscr{B}_1^*, \ldots, \mathscr{B}_n^*$ are the axioms used in the proof of $\mathscr{D}$. Let K○* be the theory whose axioms are $\mathscr{B}_1^*, \ldots, \mathscr{B}_n^*$. Hence $\vdash_{K^*} \mathscr{D}$. Assume that M is a denumerable model of K°. We may assume that the domain of M is the set *P* of positive integers (see Exercise 2.88). Let $\mathscr{B}$ be any axiom of K°. For example, suppose that $\mathscr{B}$ has the form $(\exists y_1)(\forall y_2)(\forall y_3)(\exists y_4)\mathscr{C}(y_1, y_2, y_3, y_4)$, where $\mathscr{C}$ is quantifier-free. $\mathscr{B}$* has the form $\mathscr{C}(b, y_2, y_3, g(y_2, y_3))$. Extend the model M step by step in the following way (noting that the domain always remains *P*); since $\mathscr{B}$ is true for M, $(\exists y_1)(\forall y_2)(\forall y_3)(\exists y_4)\mathscr{C}(y_1, y_2, y_3, y_4)$ is true for M. Let the interpretation b^* of b be the least positive integer y_1 such that $(\forall y_2)(\forall y_3)(\exists y_4)\mathscr{C}(y_1, y_2, y_3, y_4)$ is true for M. Hence, $(\exists y_4)\mathscr{C}(b, y_2, y_3, y_4)$ is true in this extended model. For any positive integers y_2 and y_3, let the interpretation of $g(y_2, y_3)$ be the least positive integer y_4 such that $\mathscr{C}(b, y_2, y_3, y_4)$ is true in the extended model. Hence, $\mathscr{C}(b, y_2, y_3, g(y_2, y_3))$ is true in the extended model. If we do this for all the axioms $\mathscr{B}$ of K°, we obtain a model M* of K°*. Since $\vdash_{K^{\circ *}} \mathscr{D}$, $\mathscr{D}$ is true for M*. Since M* differs from M only in having interpretations of the new individual constants and function letters, and since $\mathscr{D}$ does not contain any of those symbols, $\mathscr{D}$ is true for M. Thus, $\mathscr{D}$ is true in every denumerable model of K°. Hence, $\vdash_{K^\circ} \mathscr{D}$, by Corollary 2.20(a). Since the axioms of K° are axioms of K, we have $\vdash_K \mathscr{D}$. (For a constructive proof of an equivalent result, see Hilbert and Bernays (1939).)
b. Clearly, K* is an extension of K, since $\mathscr{B}^* \vdash \mathscr{B}$. Hence, if K* is consistent, so is K. Conversely, assume K is consistent. Let $\mathscr{D}$ be any wf of K. If K* is inconsistent, $\vdash_{K^*} \mathscr{D} \wedge \neg \mathscr{D}$. By (a), $\vdash_K \mathscr{D} \wedge \neg \mathscr{D}$, contradicting the consistency of K.

Let us use the term *generalized completeness theorem* for the proposition that every consistent generalized theory has a model. If we assume that every set can be well-ordered (or, equivalently, the axiom of choice), then the generalized completeness theorem is a consequence of Proposition 2.33.

By the *maximal ideal theorem* (MI) we mean the proposition that every proper ideal of a Boolean algebra can be extended to a maximal ideal.* This is equivalent to the Boolean representation theorem, which states that every Boolean algebra is isomorphic to a Boolean algebra of sets (Compare Stone 1936). For the theory of Boolean algebras, see Sikorski (1960) or Mendelson (1970). The usual proofs of the MI theorem use the axiom of choice, but it is a remarkable fact that the MI theorem is equivalent to the generalized completeness theorem, and this equivalence can be proved without using the axiom of choice.

Proposition 2.36

(Loś, 1954a; Rasiowa and Sikorski, 1951, 1952) The generalized completeness theorem is equivalent to the maximal ideal theorem.

Proof

a. Assume the generalized completeness theorem. Let B be a Boolean algebra. Construct a generalized theory with equality K having the binary function letters ∪ and ∩, the singulary function letter f_1^1 [we denote $f_1^1(t)$ by $\overline{t}$], predicate letters=and A_1^1, and, for each element b in B, an individual constant a_b. By the complete description of B, we mean the following sentences: (i) $a_b \neq a_c$ if b and c are distinct elements of B; (ii) $a_b \cup a_c = a_d$ if b, c, d are elements of B such that $b \cup c = d$ in B; (iii) $a_b \cap a_c = a_e$ if b, c, e are elements of b such that $b \cap c = e$ in B; and (iv) $\overline{a}_b = a_c$ if b and c are elements of B such that $\overline{b} = c$ in B, where $\overline{b}$ denotes the complement of b. As axioms of K we take a set of axioms for a Boolean algebra, axioms (A6) and (A7) for equality, the complete description of B, and axioms asserting that A_1^1 determines a maximal ideal (i.e., $A_1^1(x \cap \overline{x})$, $A_1^1(x) \wedge A_1^1(y) \Rightarrow A_1^1(x \cup y)$, $A_1^1(x) \Rightarrow A_1^1(x \cap y)$, $A_1^1(x) \vee A_1^1(\overline{x})$, and $\neg A_1^1(x \cup \overline{x})$). Now K is consistent, for, if there were a proof in K of a contradiction, this proof would contain only a finite number of the symbols $a_b, a_c, \ldots$—say, $a_{b_1}, \ldots, a_{b_n}$. The elements $b_1, \ldots, b_n$ generate a finite subalgebra B′ of B. Every finite Boolean algebra clearly has a maximal ideal. Hence, B′ is a model for the wfs that occur in the proof of the contradiction, and therefore the contradiction is true in B′, which is impossible. Thus, K is consistent and, by the generalized completeness theorem, K has a model. That model can be contracted to a normal model of K, which is a Boolean algebra A with a maximal ideal I. Since the complete description of B is included in the axioms of K, B is a subalgebra of A, and then I ∩ B is a maximal ideal in B.

* Since {0} is a proper ideal of a Boolean algebra, this implies (and is implied by) the proposition that every Boolean algebra has a maximal ideal.

b. Assume the maximal ideal theorem. Let K be a consistent generalized theory. For each axiom $\mathscr{B}$ of K, form the wf $\mathscr{B}^*$ obtained by constructing a prenex normal form for $\mathscr{B}$ and then eliminating the quantifiers through the addition of new individual constants and function letters (see the example preceding the proof of Proposition 2.35). Let $K^{\#}$ be a new theory having the wfs $\mathscr{B}^*$, plus all instances of tautologies, as its axioms, such that its wfs contain no quantifiers and its rules of inference are modus ponens and a rule of substitution for variables (namely, substitution of terms for variables). Now, $K^{\#}$ is consistent, since the theorems of $K^{\#}$ are also theorems of the consistent K^* of Proposition 2.35. Let B be the Lindenbaum algebra determined by $K^{\#}$ (i.e., for any wfs $\mathscr{C}$ and $\mathscr{D}$, let $\mathscr{C}$ Eq $\mathscr{D}$ mean that $\vdash_{K\#} \mathscr{C} \Leftrightarrow \mathscr{D}$; Eq is an equivalence relation; let $[\mathscr{C}]$ be the equivalence class of $\mathscr{C}$; define $[\mathscr{C}] \cup [\mathscr{D}] = [\mathscr{C} \vee \mathscr{D}]$, $[\mathscr{C}] \cap [\mathscr{D}] = [\mathscr{C} \wedge \mathscr{D}]$, $\overline{[\mathscr{C}]} = [\neg \mathscr{C}]$; under these operations, the set of equivalence classes is a Boolean algebra, called the Lindenbaum algebra of $K^{\#}$). By the maximal ideal theorem, let I be a maximal ideal in B. Define a model M of $K^{\#}$ having the set of terms of $K^{\#}$ as its domain; the individual constants and function letters are their own interpretations, and, for any predicate letter A_j^n, we say that $A_j^n(t_1, \ldots, t_n)$ is true in M if and only if $[A_j^n(t_1, \ldots, t_n)]$ is not in I. One can show easily that a wf $\mathscr{C}$ of $K^{\#}$ is true in M if and only if $[\mathscr{C}]$ is not in I. But, for any theorem $\mathscr{D}$ of $K^{\#}$, $[\mathscr{D}] = 1$, which is not in I. Hence, M is a model for $K^{\#}$. For any axiom $\mathscr{B}$ of K, every substitution instance of $\mathscr{B}^*(y_1, \ldots, y_n)$ is a theorem in $K^{\#}$; therefore, $\mathscr{B}^*(y_1, \ldots, y_n)$ is true for all $y_1, \ldots, y_n$ in the model. It follows easily, by reversing the process through which $\mathscr{B}^*$ arose from $\mathscr{B}$, that $\mathscr{B}$ is true in the model. Hence, M is a model for K.

The maximal ideal theorem (and, therefore, also the generalized completeness theorem) turns out to be strictly weaker than the axiom of choice (see Halpern, 1964).

Exercise

2.101 Show that the generalized completeness theorem implies that every set can be totally ordered (and, therefore, that the axiom of choice holds for any set of nonempty disjoint finite sets).

The natural algebraic structures corresponding to the propositional calculus are Boolean algebras (see Exercise 1.60, and Rosenbloom, 1950, Chapters 1 and 2). For first-order theories, the presence of quantifiers introduces more algebraic structure. For example, if K is a first-order theory, then, in the corresponding Lindenbaum algebra B, $[(\exists x)\mathscr{B}(x)] = \Sigma_t[\mathscr{B}(t)]$, where Σ_t indicates the least upper bound in B, and t ranges over all terms of K that are free for x in $\mathscr{B}(x)$. Two types of algebraic structure have been proposed to serve as algebraic counterparts of quantification theory. The first, cylindrical algebras, have been studied extensively by Tarski, Thompson, Henkin, Monk, and others (see Henkin et al., 1971). The other approach is the theory of polyadic algebras, invented and developed by Halmos (1962).

2.13 Elementary Equivalence: Elementary Extensions

Two interpretations M_1 and M_2 of a generalized first-order language $\mathscr{L}$ are said to be *elementarily equivalent* (written $M_1 \equiv M_2$) if the sentences of $\mathscr{L}$ true for M_1 are the same as the sentences true for M_2. Intuitively, $M_1 \equiv M_2$ if and only if M_1 and M_2 cannot be distinguished by means of the language $\mathscr{L}$. Of course, since $\mathscr{L}$ is a generalized first-order language, $\mathscr{L}$ may have nondenumerably many symbols.

Clearly, (1) $M \equiv M$; (2) if $M_1 \equiv M_2$, then $M_2 \equiv M_1$; (3) if $M_1 \equiv M_2$ and $M_2 \equiv M_3$, then $M_1 \equiv M_3$.

Two models of a complete theory K must be elementarily equivalent, since the sentences true in these models are precisely the sentences provable in K. This applies, for example, to any two densely ordered sets without first or last elements (see page 115).

We already know, by Proposition 2.32(b), that isomorphic models are elementarily equivalent. The converse, however, is not true. Consider, for example, any complete theory K that has an infinite normal model. By Corollary 2.34(b), K has normal models of any infinite cardinality $\aleph_\alpha$. If we take two normal models of K of different cardinality, they are elementarily equivalent but not isomorphic. A concrete example is the complete theory K_2 of densely ordered sets that have neither first nor last element. The rational numbers and the real numbers, under their natural orderings, are elementarily equivalent nonisomorphic models of K_2.

Exercises

2.102 Let K_∞, the theory of infinite sets, consist of the pure theory K_1 of equality plus the axioms $\mathscr{B}_n$, where $\mathscr{B}_n$ asserts that there are at least n elements. Show that any two models of K∞ are elementarily equivalent (see Exercises 2.66 and 2.96(a)).

2.103^D If M_1 and M_2 are elementarily equivalent normal models and M_1 is finite, prove that M_1 and M_2 are isomorphic.

2.104 Let K be a theory with equality having $\aleph_\alpha$ symbols.

a. Prove that there are at most $2^{\aleph_\alpha}$ models of K, no two of which are elementarily equivalent.

b. Prove that there are at most $2^{\aleph_\lambda}$ mutually nonisomorphic models of K of cardinality $\aleph_\beta$, where γ is the maximum of α and β.

2.105 Let M be any infinite normal model of a theory with equality K having $\aleph_\alpha$ symbols. Prove that, for any cardinal $\aleph_\gamma \geq \aleph_\alpha$, there is a normal model M* of K of cardinality $\aleph_\alpha$ such that $M \equiv M^*$.

A model M_2 of a language $\mathscr{L}$ is said to be an *extension* of a model M_1 of $\mathscr{L}$ (written $M_1 \subseteq M_2$)* if the following conditions hold:

1. The domain D_1 of M_1 is a subset of the domain D_2 of M_2.
2. For any individual constant c of $\mathscr{L}$, $c^{M_2} = c^{M_1}$, where c^{M_2} and c^{M_1} are the interpretations of c in M_2 and M_1.
3. For any function letter f_j^n of $\mathscr{L}$ and any $b_1,\dots, b_n$ in D_1, $(f_j^n)^{M_2}(b_1,\dots,b_n) = (f_j^n)^{M_1}(b_1,\dots,b_n)$.
4. For any predicate letter A_j^n of $\mathscr{L}$ and any $b_1,\dots,b_n$ in D_1, $\models_{M_1} A_j^n[b_1,\dots,b_n]$ if and only if $\models_{M_2} A_j^n[b_1,\dots,b_n]$.

When $M_1 \subseteq M_2$, one also says that M_1 is a *substructure* (or *submodel*) of M_2.

Examples

1. If $\mathscr{L}$ contains only the predicate letters = and <, then the set of rational numbers under its natural ordering is an extension of the set of integers under its natural ordering.
2. If $\mathscr{L}$ is the language of field theory (with the predicate letter =, function letters + and ×, and individual constants 0 and 1), then the field of real numbers is an extension of the field of rational numbers, the field of rational numbers is an extension of the ring of integers, and the ring of integers is an extension of the "semiring" of nonnegative integers. For any fields F_1 and F_2, $F_1 \subseteq F_2$ if and only if F_1 is a subfield of F_2 in the usual algebraic sense.

Exercises

2.106 Prove:

a. $M \subseteq M$;

b. if $M_1 \subseteq M_2$ and $M_2 \subseteq M_3$, then $M_1 \subseteq M_3$;

c. If $M_1 \subseteq M_2$ and $M_2 \subseteq M_1$, then $M_1 = M_2$.

2.107 Assume $M_1 \subseteq M_2$.

a. Let $\mathscr{B}(x_1, \dots, x_n)$ be a wf of the form $(\forall y_1) \dots (\forall y_m)\mathscr{C}(x_1, \dots, x_n, y_1, \dots, y_m)$, where $\mathscr{C}$ is quantifier-free. Show that, for any $b_1, \dots, b_n$ in the domain of M_1, if $\models_{M_2} \mathscr{B}[b_1,\dots,b_n]$, then $\models_{M_1} \mathscr{B}[b_1,\dots,b_n]$. In particular, any sentence $(\forall y_1) \dots (\forall y_m)\ \mathscr{C}(y_1, \dots, y_m)$, where $\mathscr{C}$ is quantifier-free, is true in M_1 if it is true in M_2.

* The reader will have no occasion to confuse this use of ⊆ with that for the inclusion relation.

b. Let $\mathscr{B}(x_1, \ldots, x_n)$ be a wf of the form $(\exists y_1) \ldots (\exists y_m)\mathscr{C}(x_1, \ldots, x_n, y_1, \ldots, y_m)$, where $\mathscr{C}$ is quantifier-free. Show that, for any $b_1, \ldots, b_n$ in the domain of M_1, if $\vDash_{M_1} \mathscr{B}[b_1, \ldots, b_n]$, then $\vDash_{M_2} \mathscr{B}[b_1, \ldots, b_n]$. In particular, any sentence $(\exists y_1) \ldots (\exists y_m) \mathscr{C}(y_1, \ldots, y_m)$, where $\mathscr{C}$ is quantifier-free, is true in M_2 if it is true in M_1.

2.108 a. Let K be the predicate calculus of the language of field theory. Find a model M of K and a nonempty subset X of the domain D of M such that there is no substructure of M having domain X.

b. If K is a predicate calculus with no individual constants or function letters, show that, if M is a model of K and X is a subset of the domain D of M, then there is one and only one substructure of M having domain X.

c. Let K be any predicate calculus. Let M be any model of K and let X be any subset of the domain D of M. Let Y be the intersection of the domains of all submodels M* of M such that X is a subset of the domain D_{M^*} of M*. Show that there is one and only one submodel of M having domain Y. (This submodel is called the *submodel generated* by X.)

A somewhat stronger relation between interpretations than "extension" is useful in model theory. Let M_1 and M_2 be models of some language $\mathscr{L}$. We say that M_2 is an *elementary extension* of M_1 (written $M_1 \leq_e M_2$) if (1) $M_1 \subseteq M_2$ and (2) for any wf $\mathscr{B}(y_1, \ldots, y_n)$ of $\mathscr{L}$ and for any $b_1, \ldots, b_n$ in the domain D_1 of M_1, $\vDash_{M_1} \mathscr{B}[b_1, \ldots, b_n]$ if and only if $\vDash_{M_2} \mathscr{B}[b_1, \ldots, b_n]$. (In particular, for any sentence $\mathscr{B}$ of $\mathscr{L}$, $\mathscr{B}$ is true for M_1 if and only if $\mathscr{B}$ is true for M_2.) When $M_1 \leq_e M_2$, we shall also say that M_1 is an *elementary substructure* (or *elementary submodel*) of M_2.

It is obvious that, if $M_1 \leq_e M_2$, then $M_1 \subseteq M_2$ and $M_1 \equiv M_2$. The converse is not true, as the following example shows. Let G be the elementary theory of groups (see page 96). G has the predicate letter =, function letter +, and individual constant 0. Let I be the group of integers and E the group of even integers. Then $E \subseteq I$ and $I \cong E$. (The function g such that $g(x) = 2x$ for all x in I is an isomorphism of I with E.) Consider the wf $\mathscr{B}(y)$: $(\exists x)(x + x = y)$. Then $\vDash_I \mathscr{B}[2]$, but not-$\vDash_E \mathscr{B}[2]$. Thus, I is not an elementary extension of E. (This example shows the stronger result that even assuming $M_1 \subseteq M_2$ and $M_1 \cong M_2$ does not imply $M_1 \leq_e M_2$.)

The following theorem provides an easy method for showing that $M_1 \leq_e M_2$.

Proposition 2.37 (Tarski and Vaught, 1957)

Let $M_1 \subseteq M_2$. Assume the following condition:

($) For every wf $\mathscr{B}(x_1, \ldots, x_k)$ of the form $(\exists y)\mathscr{C}(x_1, \ldots, x_k, y)$ and for all $b_1, \ldots, b_k$ in the domain D_1 of M_1, if $\vDash_{M_2} \mathscr{B}[b_1, \ldots, b_k]$, then there is some d in D_1 such that $\vDash_{M_2} \mathscr{C}[b_1, \ldots, b_k, d]$.

Then $M_1 \leq_e M_2$.