2.4 Definability in a Structure

When choosing a language and a structure, we must keep in mind what objects (functions, relations, constants) of the universe we want to be able to discuss in the language. Clearly we can discuss any function or relation for which the language has a symbol, but there are others: for example, we have seen that any term defines a function as well. Now we ask the question: which functions and relations in a structure can be defined and discussed in a given first-order language?

2.4.A First-Order Definability

First, we must say what it means to define something. Let L be a first-order language and $\mathcal{M} = \langle M, \dots \rangle$ a structure for L.

If $A(x_1, \dots, x_n)$ is a formula of L, the *graph* of A in $\mathcal{M}$ is the n-ary relation $GR_A^{\mathcal{M}} \subseteq M^n$ defined by

$$GR_A^{\mathcal{M}}(a_1, \dots, a_n) \text{ iff } \mathcal{M} \models A[a_1, \dots, a_n].$$

Example 2.4.1 Let $L = \{+, \cdot, <\}$ and $\mathcal{M} = \langle \mathbb{R}, +, \cdot, <\}$, and consider the following formulas, with their graphs.

$$A_0 : x = y$$
$$GR_{A_0}^{\mathcal{M}} = \{(a,b) : a = b\}$$
$$(GR_{A_0}^{\mathcal{M}}(a,b) \text{ iff } a = b)$$

$$A_1 : x < y$$
$$GR_{A_1}^{\mathcal{M}} = \{(a,b) : a < b\}$$
$$(GR_{A_1}^{\mathcal{M}}(a,b) \text{ iff } a < b)$$

$$A_2(x,y) : \exists z(z \cdot z = z \wedge \neg z + z = z \wedge x \cdot x + y \cdot y = z)$$
$$GR_{A_2}^{\mathcal{M}}(a,b) \text{ iff } a^2 + b^2 = 1$$

$$A_3(x,y) : \exists z(z \cdot z = z \wedge \neg z + z = z \wedge x \cdot x + y \cdot y < z)$$
$$GR_{A_3}^{\mathcal{M}}(a,b) \text{ iff } a^2 + b^2 < 1$$

$$A_4(x,y,z) : \exists w(w+w=w \wedge x+y+z=w)$$
$$GR^{\mathcal{M}}_{A_4}(a,b,c) \text{ iff } a+b+c=0$$

$$A_5(x,y) : A_3 \wedge (x\cdot x+y\cdot y<x+x)$$
$$GR^{\mathcal{M}}_{A_5}(a,b) \text{ iff } a^2+b^2<1 \text{ and } (a-1)^2+b^2<1$$

Example 2.4.2 Let $L = L_{ar} = \{0, S, +, \cdot, <\}$ and $\mathcal{N} = \langle \mathbb{N}, 0, S, +, \cdot, <\}$, and consider the formulas:

$$A(x) : \exists y(x = y\cdot y)$$
$$GR^{\mathcal{N}}_{A}(a) \text{ iff } a \text{ is a square.}$$

$$B(x,y) : \exists w\exists v[S(S(w)) = x \wedge S(S(v)) = y]\wedge$$
$$\forall z[\exists p(z\cdot p = x) \wedge \exists q(z\cdot q = y) \Rightarrow z = S(0))]$$

$GR^{\mathcal{N}}_{B}(a,b)$ iff $a, b \geq 2$ and a, b are prime to each other (i.e., $\gcd(a,b)=1$).

It can be shown that there is a formula $P(x,y)$ in L_{ar} such that $GR^{\mathcal{N}}_{P}(a,b)$ iff a is the bth prime number, and also a formula $R(x,y,z)$ such that

$$GR^{\mathcal{N}}_{R}(a,b,c) \text{ iff } a^b = c.$$

This follows from a clever idea due to K. Gödel that allows us to *code* in an appropriate, first-order definable sense, finite sequences of numbers by numbers.

Definition 2.4.3 A relation $R \subseteq M^n$ is *first-order definable* if there is a formula $A(x_1,\ldots,x_n)$ such that $R = GR^{\mathcal{M}}_{A}$, i.e., for any $a_1,\ldots,a_n \in M$

$$R(a_1,\ldots,a_n) \text{ iff } \mathcal{M} \models A[a_1,\ldots,a_n].$$

A function $f : M^n \to M$ is *first-order definable* if its graph

$$\text{graph}(f) \subseteq M^{n+1}$$

is first-order definable, i.e. there is a formula $A(x_1,\ldots,x_n,x_{n+1})$ such that for all $a_1,\ldots,a_n,a_{n+1}$ in M,

$$f(a_1,\ldots,a_n) = a_{n+1} \text{ iff } M \models A[a_1,\ldots,a_n,a_{n+1}].$$

(If $n = 0$, an element $a \in A$ is *first-order definable* if there is a formula $A(x)$ such that a is the unique element of M with

$$\mathcal{M} \models A[a].)$$

Examples 2.4.4

(i) If f is a function symbol in L, then $f^{\mathcal{M}}$ is definable by the formula

$$f(x_1, \ldots, x_n) = x_{n+1}.$$

If R is a relation symbol in L, then $R^{\mathcal{M}}$ is definable by the formula

$$R(x_1, \ldots, x_n).$$

If t is a term, say t is $t(x_1, \ldots, x_n)$, then $t^{\mathcal{M}}$ is definable by the formula

$$t = x_{n+1}.$$

(ii) The definable relations in $\mathcal{M}$ form a *Boolean algebra*, i.e., are closed under $\sim, \cup, \cap$, where if $R \subseteq M^n$,

$$\sim R \subseteq M^n \text{ and } \sim R = M^n \setminus R,$$

and if $R, S \subseteq M^n$, then $R \cap S$, $R \cup S$ have their usual meaning. This is clear, since if R is defined by A, then $\sim R$ is defined by $\neg A$ and if R, S are defined by B, C, resp., then $R \cap S$, $R \cup S$ are defined by $B \wedge C$, $B \vee C$, resp.

(iii) The definable relations in $\mathcal{M}$ are closed under *projection*. If $R \subseteq M^{n+1}$ is defined by $A(x_1, \ldots x_n, x_{n+1})$, then $\text{proj}(R) = \{(a_1, \ldots, a_n)\colon$ for some a_{n+1}, $R(a_1, \ldots, a_{n+1})\}$ is defined by $\exists x_{n+1} A(x_1, \ldots, x_n, x_{n+1})$.

Example 2.4.5 Let $L = \{+, \cdot, <\}$ and $\mathcal{M} = \langle \mathbb{R}, +, \cdot, <\}$.

First we check, by induction, that every integer $n \in \mathbb{Z}$ is definable. The most important integers, of course, are 0 and 1:

$$\begin{array}{ll} 0 : x + x = x & (A_0(x)) \\ 1 : x \cdot x = x \wedge \neg x + x = x(A_1(x)) & \end{array}$$

Now assume a formula $A_n(x)$ defines n (in $\mathcal{M}$). Then the formula $\exists y \exists z (x = y + z \wedge A_n(y) \wedge A_1(z))$ $(A_{n+1}(x))$ defines $n + 1$. So by induction, any $n \geq 0$ is definable by a formula $A_n(x)$. Consider now the formula

$$A_{-n}(x) : \exists y \exists z (A_0(z) \wedge A_n(y) \wedge x + y = z)$$

This defines $-n$. So every integer can be defined.

One can then show that every rational number is definable, and (though this is a bit trickier) so is every algebraic number (solution of polynomial with integer coefficients).

Using this, it follows that all finite unions of intervals with algebraic endpoints are definable. It turns out that this is it, i.e. every definable unary relation (set) is a finite union of intervals with algebraic endpoints.

For binary, ternary, or higher order relations things are more complicated, but it turns out that every definable relation is a Boolean combination of relations that can be defined using polynomial equations or inequalities with integer coefficients, as for example:

$$(x \cdot x + y \cdot y < z \wedge x \cdot x \cdot x = y \cdot y + z) \vee x \cdot y + y \cdot z < x \cdot y \cdot z.$$

Every polynomial function with integer coefficients is definable (but there are more complex definable functions than polynomials). For example, the polynomial function $f(a, b, c) = 2a^2 - b + c^2$ is definable by the formula

$$A(x, y, z, w) : \exists p \exists q (A_2(p) \wedge A_{-1}(q) \wedge w = p \cdot x \cdot x + q \cdot y + z \cdot z).$$

Another formula that defines the graph of f is

$$B(x, y, z, w) : y + w = x \cdot x + x \cdot x + z \cdot z.$$

Example 2.4.6 Let $L = L_{ar} = \{0, S, +, \cdot, <\}$ and $\mathcal{N} = \langle \mathbb{N}, 0, S, +, \cdot < \rangle$. Again any polynomial function with coefficients in $\mathbb{N}$ is definable, but it turns out that much more complicated functions are definable, e.g. m^n, $i \mapsto p_i =$ the ith prime, etc. In fact, every standard function one studies in number theory is definable. (This is by no means obvious—try to define the exponential function, as an example.) Similarly all the ordinary relations, e.g. "n is prime", are definable. In fact "n is prime" is definable by the following formula:

$$A(x) : S(0) < x \wedge \forall y \forall z (y \cdot z = x \Rightarrow y = S(0) \vee z = S(0)).$$

Remark. Assuming that the language L has only countably many symbols (i.e. the non-logical symbols in L can be enumerated in a sequence), there are only countably many definable relations and functions in each given structure $\mathcal{M}$ for L. Since in any structure whose universe is infinite, there are uncountably many possible relations and functions, the definable ones form a very small subset of these. But so far, we have not seen even one example of a *non*-definable relation.

Definition 2.4.7 Let L be a first-order language and $\mathcal{M} = \langle M, \dots \rangle$ a structure for L. A relation $R \subseteq M^n$ is *definable with parameters* if there is a formula $A(x_1, \dots, x_n, x_{n+1}, \dots, x_{n+m})$, $m \geq 0$, and fixed $p_1, \dots, p_m \in \mathcal{M}$ (the *parameters*), such that

$$R(a_1, \dots, a_n) \text{ iff } \mathcal{M} \models A[a_1, \dots, a_n, p_1, \dots, p_m].$$

Example 2.4.8 Let $L = \{+, \cdot, <\}$ and $\mathcal{M} = \langle \mathbb{R}, +, \cdot, < \rangle$. Consider an interval $(b, c) \subseteq \mathbb{R}$. For arbitary b, c this may not be definable, but it is always definable with parameters:

$$A(x, y, z) : y < x \wedge x < z$$
$$a \in (b, c) \text{ iff } \mathcal{M} \models A[a, \underbrace{b, c}_{\text{parameters}}]$$

Similarly, a function $f : M^n \to M$ is *definable with parameters* whenever its graph is definable with parameters. Notice also that every element of M is definable with parameters; one parameter suffices (that element).

Example 2.4.9 In example 2.4.5, *every* polynomial function is definable with parameters.

2.4.B Isomorphisms

We will next discuss a method for showing that certain relations are *not* definable.

Definition 2.4.10 Let L be a first order language and $\mathcal{M}_1 = \langle M_1, \dots \rangle$ and $\mathcal{M}_2 = \langle M_2, \dots \rangle$ be two structures of L. An *isomorphism* of $\mathcal{M}_1$ with $\mathcal{M}_2$ is a map $\pi : M_1 \to M_2$ which is one-to-one and onto (i.e. a bijection of M_1, M_2) and such that for each n-ary function symbol f in L,

$$\pi(f^{\mathcal{M}_1}(a_1, \dots, a_n)) = f^{\mathcal{M}_2}(\pi(a_1), \dots, \pi(a_n)),$$

for all $a_1, \dots, a_n \in M_1$, and for any m-ary relation symbol R in L,

$$R^{\mathcal{M}_1}(a_1, \dots, a_m) \text{ iff } R^{\mathcal{M}_2}(\pi(a_1), \dots, \pi(a_m)),$$

for all $a_1, \dots, a_m \in M_1$.

We denote this by

$$\pi : \mathcal{M}_1 \cong \mathcal{M}_2$$

Clearly also

$$\pi^{-1} : \mathcal{M}_2 \cong \mathcal{M}_1.$$

Example 2.4.11 Let $L = \{f; R\}$ with f a binary function symbol and R a binary relation symbol, and consider the structures

$$\begin{aligned} \mathcal{M}_1 &= \langle \mathbb{R}^+, \cdot, < \rangle \quad (\text{i.e. } \cdot = f^{\mathcal{M}_1},\ < = R^{\mathcal{M}_1}) \\ \mathcal{M}_2 &= \langle \mathbb{R}, +, < \rangle \quad (\text{i.e., } + = f^{\mathcal{M}_2},\ < = R^{\mathcal{M}_2}) \end{aligned}$$

where $\mathbb{R}^+ = \{\alpha \in \mathbb{R} : \alpha > 0\}$. Let $\pi : \mathbb{R}^+ \to \mathbb{R}$ be given by $\pi(a) = \ln a$. Then $\pi : \mathcal{M}_1 \cong \mathcal{M}_2$ and $\pi^{-1}(a) = e^a$.

We now have the following basic fact.

Theorem 2.4.12 *If $\pi : \mathcal{M}_1 \cong \mathcal{M}_2$ and $A(x_1, \ldots, x_n)$ is any formula, then for any $a_1, \ldots, a_n \in M_1$,*

$$\mathcal{M}_1 \models A[a_1, \ldots, a_n] \textit{ iff } \mathcal{M}_2 \models A[\pi(a_1), \ldots, \pi(a_n)].$$

Proof. First show by induction on the construction of the term t, that if t is $t(x_1, \ldots, x_n)$, then

$$\pi(t^{\mathcal{M}_1}(a_1, \ldots, a_n)) = t^{\mathcal{M}_2}(\pi(a_1), \ldots, \pi(a_n))$$

for any $a_1, \ldots, a_n \in M_1$. Then the theorem can be proved by induction on the construction of A. ⊣

Definition 2.4.13 If $\mathcal{M}_1 = \mathcal{M}_2 = \mathcal{M}$, we call any isomorphism $\pi : \mathcal{M} \cong \mathcal{M}$ an *automorphism* of $\mathcal{M}$.

Example 2.4.14 If $\mathcal{M} = \langle \mathbb{Q}, +, < \rangle$, then $\pi(a) = 3a$ is an automorphism of $\mathcal{M}$.

In particular, if π is any automorphism of $\mathcal{M}$, the preceding theorem says that for any formula $A(x_1, \ldots, x_n)$ and any $a_1, \ldots, a_n \in M$

$$\mathcal{M} \models A[a_1, \ldots, a_n] \quad \text{iff} \quad \mathcal{M} \models A[\pi(a_1), \ldots, \pi(a_n)]$$

i.e. every definable relation is invariant under automorphisms.

Similarly, any definable relation with parameters $p_1, \dots, p_m$ is invariant under any automorphism π that fixes the parameters, i.e. $\pi(p_i) = p_i$, $i = 1, \dots, m$.

We can use this fact to prove that certain relations are not definable: just exhibit an automorphism of the structure under which they are not invariant. Here are some examples.

Application. Consider the language $L = \emptyset$ with no nonlogical symbols and any structure $\mathcal{M} = \langle M \rangle$ for it. What are the definable subsets (i.e. unary relations) on M?

Clearly $\emptyset, M$ are definable (by $x \neq x$, $x = x$, resp.). We claim that these are the only definable subsets of M. Let A be a subset of M, $A \neq \emptyset$, $A \neq M$, and let $a \in A$, $b \notin A$. Let $\pi : M \to M$ be the bijection such that $\pi(a) = b$, $\pi(b) = a$ and $\pi(c) = c$ if $c \notin \{a, b\}$. Then π is an automorphism of M but A is not invariant under π, so A is not definable.

What are the definable with parameters subsets of M? Clearly any finite subset $\{a_1, \dots, a_n\} \subseteq \mathcal{M}$ is definable by the formula

$$x = x_1 \vee \cdots \vee x = x_n$$

and parameters $a_1, \dots, a_n$ (for $x_1, \dots, x_n$). But then every co-finite (complement of finite) subset of M is definable with parameters. We claim these are the only ones. Let $A \subseteq M$ be neither finite or co-finite. Assume it is definable with parameters $p_1, \dots, p_n$. Since A is neither finite nor co-finite, there is some element of A, say a, distinct from all $p_1, \dots, p_n$ and some element of $M \setminus A$, say b, distinct from $p_1, \dots, p_n$. Let $\pi(p_i) = p_i$, $\pi(a) = b$, $\pi(b) = a$, $\pi(c) = c$ if $c \notin \{p_1, \dots, p_n, a, b\}$. Then π is an automorphism of $\mathcal{M}$ that fixes the parameters, but does not leave A invariant, a contradiction.

Application. $\mathbb{N}$ is not definable in

$$\mathcal{M} = \langle \mathbb{R}, 0, 1, \cdot, < \rangle$$

Indeed, $\pi(a) = a^3$, is an automorphism of $\mathcal{M}$ but it does not leave $\mathbb{N}$ invariant, i.e., $a \in \mathbb{N}$ iff $\pi(a) \in \mathbb{N}$ clearly fails. (It turns out that $\mathbb{N}$ is not even definable in $\langle \mathbb{R}, 0, 1, +, \cdot, < \}$, even though each integer is.)

Remark. This method of automorphisms does not always apply. For example, there are structures which have no automorphisms except the identity (these are called *rigid*). An example is

$$\langle \mathbb{N}, 0, S \rangle.$$

Another is

$$\langle \mathbb{R}, 0, 1, +, \cdot, < \rangle$$

So although, for example, it turns out that $\cdot$ is not definable in $\langle \mathbb{N}, 0, S, +, < \rangle$, this cannot be shown by the automorphism method.

Remark. By the way, $+$ is definable in $\langle \mathbb{N}, 0, S, \cdot \rangle$ since:

$$a + b = c \text{ iff } [(a+1)(c+1)+1][b(c+1)+1] = (c+1)^2[(a+1)b+1]+1$$
$$\text{i.e. } S(S(a) \cdot S(c)) \cdot S(b \cdot S(c)) = S((S(c) \cdot S(c)) \cdot S(S(a) \cdot b))$$

2.5 Prenex Normal Forms and Games

We will now discuss a "normal form" for first-order wffs which makes many proofs easier, and follow it with an application to game theory.

2.5.A Prenex Normal Forms

Definition 2.5.1 A formula A (in a fixed language L) is in *prenex normal form* (pnf) if A has the form

$$Q_1 y_1 Q_2 y_2 \ldots Q_n y_n B,$$

where y_i are distinct variables, each Q_i is either $\exists$ or $\forall$ and B is *quantifier-free* (has no quantifers), i.e. it is built from atomic formulas using propositional connectives only.

We call $Q_1 y_1 Q_2 y_2 \ldots Q_n y_n$ the *prefix* of A and B the *matrix* of A.

Example 2.5.2

$$\overbrace{\forall x \forall y \exists z \exists w}^{\text{prefix}} \overbrace{[P(x,z) \Rightarrow \neg Q(y,w)]}^{\text{matrix}}$$
$$\text{(empty prefix) } \underbrace{P(x,y) \vee S(f(x), h(y))}_{\text{matrix}}$$

are in pnf, but

$$\neg \exists x [P(x) \wedge \exists y (Q(x,y) \vee \neg \exists z R(x,z))]$$

is *not* in pnf.

Theorem 2.5.3 *For each formula A, we can find a formula A^* in pnf, logically equivalent to A, i.e., $A \equiv A^*$.*

Proof. We will use the following equivalences. If $Q = \forall$ (resp. $\exists$), let $\check{Q} = \exists$ (resp. $\forall$); we call $\check{Q}$ the *dual* of Q.

(a) $\neg QxA \equiv \check{Q}x\neg A$

(b) $A \vee (QxB) \equiv Qx(A \vee B)$, if x if not free in A

(c) $A \wedge (QxB) \equiv Qx(A \wedge B)$, if x is not free in A

(d) $(A \Rightarrow QxB) \equiv Qx(A \Rightarrow B)$, if x is not free in A

(e) $(QxA \Rightarrow B) \equiv \check{Q}x(A \Rightarrow B)$, if x is not free in B.

(f) $Q_1y_1Q_2y_2\ldots Q_ny_nA \equiv Q_1y'_1Q_2y'_2\ldots Q_ny'_nA[y_1/y'_1,\ldots,y_n/y'_n]$, where A is quantifier-free and $y'_1\ldots y'_n$ are any distinct variables not appearing in $Q_1y_1Q_2y_2\ldots Q_ny_nA$ (it is assumed that $y_1,\ldots,y_n$ are distinct here.)

(g) If in $Q_1y_1Q_2y_2\ldots Q_ny_nA$, A quantifier free, a variable y_i appears more than once and Q_my_i $(m \leq n)$ is the rightmost occurrence of y_i in the prefix, then if we eliminate all $Q_{m'}y_i$ for $m' < m$, we obtain a formula B equivalent to $Q_1y_1\ldots Q_ny_nA$.

We now show how to construct A^* by recursion.

- If A is atomic, we simply take $A^* = A$.

- Assume A^*, B^* have been constructed, and consider $\neg A,\ A \wedge B,\ A \vee B,\ A \Rightarrow B,\ A \Leftrightarrow B$.

  We can eliminate the last case, $A \Leftrightarrow B$, simply by replacing it by $(A \Rightarrow B) \wedge (B \Rightarrow A)$. We do this because there is no simple equivalence corresponding to a–e for $\Leftrightarrow$, so it would be rather complicated to consider separately. (We could also eliminate $\Rightarrow$ and one of $\wedge$ and $\vee$ as well, but this does not seem to offer any great advantage.) We will construct, therefore, $(\neg A)^*$, $(A \wedge B)^*$, $(A \vee B)^*$, and $(A \Rightarrow B)^*$.

Let

$$\begin{aligned} A^* &: Q_1 y_1 \ldots Q_n y_n C \\ B^* &: Q'_1 z_1 \ldots Q'_m z_m D. \end{aligned}$$

Then, by applying (a) n times we see that

$$\begin{aligned} \neg A &\equiv \neg A^* \\ &\equiv \check{Q}_1 y_1 \check{Q}_2 y_2 \ldots \check{Q}_n y_n \neg C \end{aligned}$$

so $(\neg A)^* = \check{Q}_1 y_1 \ldots \check{Q}_n y_n \neg C$.

Also

$$\begin{aligned} A \wedge B &\equiv A^* \wedge B^* \\ &\equiv Q_1 y_1 \ldots Q_n y_n C \wedge Q'_1 z_1 \ldots Q'_m z_m D. \end{aligned}$$

Using (f) we have that

$$Q'_1 z_1 \ldots Q'_m z_m D \equiv Q'_1 u_1 \ldots Q'_m u_m D[z_1/u_1, \ldots z_m/u_1]$$

where $u_1, \ldots, u_m$ are variables distinct from all of the variables in $Q_1 y_1 \ldots Q_m y_m C$ and $Q'_1 z_1 \ldots Q'_m z_m D$. Thus

$$\begin{aligned} A \wedge B &\equiv Q_1 y_1 \ldots Q_n y_n C \wedge Q'_1 u_1 \ldots Q'_m u_m D[z_1/u_1 \ldots z_m/u_m] \\ &\equiv Q'_1 u_1 \ldots Q'_m u_m (Q_1 y_1 \ldots Q_n y_n C \wedge D[z_1/u_1, \ldots, z_m/u_m]) \end{aligned}$$

by repeatedly using (c). Now let $v_1, \ldots, v_n$ be distinct variables not appearing in $Q_1 y_1 \ldots Q_n y_n C$ and $D[z_1/u_1, \ldots, z_m/u_m]$. Then by using (f) and (c) (and the fact that $P \wedge Q \equiv Q \wedge P$) we have

$$\begin{aligned} A \wedge B \equiv\ & Q'_1 u_1 \ldots Q'_m u_m (Q_1 v_1 \ldots Q_n v_n C[y_1/v_1, \ldots, y_n/v_n] \wedge \\ & D[z_1/u_1, \ldots, z_m/u_m]) \\ \equiv\ & Q'_1 u_1 \ldots Q'_m u_m Q_1 v_1 \ldots Q_n v_n (C[y_1/v_1, \ldots, y_n/v_n] \wedge \\ & D[z_1/u_1, \ldots, z_m/u_m]), \end{aligned}$$

and this last formula is by definition $(A \wedge B)^*$.

The cases of $(A \vee B)^*$ and $(A \Rightarrow B)^*$ are similar, using (b)-(e).

- Finally, assume A^* is defined and consider $\exists xA$ and $\forall xA$. If $A^* = Q_1y_1 \ldots Q_ny_nC$, then

$$\begin{aligned}\exists xA &\equiv \exists xA^* \\ &\equiv \exists xQ_1y_1 \ldots Q_ny_nC.\end{aligned}$$

If x is different from all $y_1, \ldots, y_n$ we simply take

$$(\exists xA)^* = \exists xQ_1y_1 \ldots Q_ny_nC.$$

Otherwise, using (g), we take

$$(\exists xA)^* = Q_1y_1 \ldots Q_ny_nC$$

The case of $(\forall xA)^*$ is similar. ⊣

Example 2.5.4 Let A be the formula:

$$\neg\exists x(P(x) \Rightarrow \exists y(Q(z,y) \vee \neg\exists zR(x,z))).$$

The following sequence of steps transforms it in pnf.

$$\begin{gathered}\neg\exists x(P(x) \Rightarrow \exists y(Q(z,y) \vee \neg\exists zR(x,z))) \\ \neg\exists x(P(x) \Rightarrow \exists y(Q(z,y) \vee \forall z\neg R(x,z))) \\ \neg\exists x(P(x) \Rightarrow \exists y\forall z'(Q(z,y) \vee \neg R(x,z'))) \\ \neg\exists x\exists y\forall z'(P(x) \Rightarrow (Q(z,y) \vee \neg R(x,z'))) \\ \forall x\forall y\exists z'\neg(P(x) \Rightarrow (Q(z,y) \vee \neg R(x,z')))\end{gathered}$$

Remark. Of course the matrix of a formula is pnf can always be replaced, up to $\equiv$, with one which is in disjunctive or conjunctive normal form (on atomic formulas).

If $A = Q_1y_1 \ldots Q_ny_nB$, is a formula in pnf we can group together consecutive similar quantifiers and write it in one of the forms

$$\exists\bar{z_1}\forall\bar{z_2}\cdots Q\bar{z_n}B \quad \text{or} \quad \forall\bar{z_1}\exists\bar{z_2}\ldots Q\bar{z_n}B,$$

where $\bar{z_i} = z^i_1, \ldots, z^i_{k_i}$ is a string of variables and $Q\bar{z_i}$ is an abbreviation for $Qz^i_1Qz^i_2 \ldots Qz^i_{k_i}$ (where Q is $\exists$ or $\forall$).

Example 2.5.5

$$\exists y_1 \exists y_2 \forall y_3 \exists y_4 \exists y_5 B(y_1, \dots, y_5)$$

will be written as

$$\exists y_1, y_2 \forall y_3, y_4 \exists y_5 B(y_1, \dots, y_5)$$

and

$$\forall y_1 \exists y_2 \exists y_3 \exists y_4 \forall y_5 \forall y_6 B(y_1, \dots, y_6)$$

as

$$\forall y_1 \exists y_2, y_3, y_4 \forall y_5, y_6 B(y_1, \dots y_6).$$

2.5.B Games and Strategies

Let $\mathcal{M}$ be a structure for the language L in which A is a sentence in prenex normal form. We will reformulate the statement $\mathcal{M} \models A$ in terms of a game. For simplicity, we will assume that A is a sentence of the form

$$A : \exists z_1 \ \forall z_2 \exists z_3 \dots \forall z_{2n} B(z_1, \dots, z_{2n}).$$

Obvious modifications can be made to handle the general case.

We associate with $\mathcal{M}, A$ a game $G_A^{\mathcal{M}}$ played as follows: We have two players called $\exists$ and $\forall$. $\exists$ plays first an arbitrary element $a_1 \in M$. Then $\forall$ plays an arbitrary element $a_2 \in M$. $\exists$ plays then $a_3 \in M$, $\forall$ plays $a_4 \in M$, etc., for a total of $2n$ moves. We say that $\exists$ wins this run of the game if

$$\mathcal{M} \models B[z_1 \mapsto a_1, \dots, z_{2n} \mapsto a_{2n}]$$

and $\forall$ wins this run of the game if

$$\mathcal{M} \models \neg B[z_1 \mapsto a_1, \dots, z_{2n} \mapsto a_{2n}].$$

$\exists$	$\forall$
a_1	
	a_2
a_3	
	a_4
$\vdots$	$\vdots$
a_{2n-1}	
	a_{2n}

It is assumed in this game that each player can see the opponent's previous moves (i.e. when $\forall$ plays a_2, he knows a_1 etc.).

Fact 2.5.6 *(i) $\mathcal{M} \models A$ iff $\exists$ has a winning strategy in $G_A^{\mathcal{M}}$.*

(ii) $\mathcal{M} \models \neg A$ iff $\forall$ has a winning strategy in $G_A^{\mathcal{M}}$.

Proof. Assume $\mathcal{M} \models A$, i.e. $\mathcal{M} \models \exists z_1 \forall z_2 \ldots \forall z_{2n} B(z_1, \ldots, z_{2n})$. Then there is $a_1 \in M$ such that

$$\mathcal{M} \models \forall z_2 \exists z_3 \ldots \forall z_{2n} B(z_1, \ldots, z_{2n})[z_1 \mapsto a_1].$$

$\exists$ starts by playing a fixed such a_1. Then for any $a_2 \in M$ that $\forall$ could play in his next move, we have

$$\mathcal{M} \models \exists z_3 \forall z_4 \ldots \forall z_{2n} B(z_1, \ldots, z_{2n})[z_1 \mapsto a_1,\ z_2 \mapsto a_2].$$

Thus $\exists$ can respond by playing some a_3 so that

$$\mathcal{M} \models \forall z_4 \ldots \forall z_{2n} B(z_1, \ldots, z_{2n})[z_1 \mapsto a_1,\ z_2 \mapsto a_2,\ z_3 \mapsto a_3],$$

and so on. By induction, if $\exists$ follows this strategy, once $a_1, a_2, \ldots, a_{2n}$ have been played, then $\mathcal{M} \models B[z_1 \mapsto a_1, \ldots, z_{2n} \mapsto a_{2n}]$, i.e. $\exists$ has won.

If on the other hand $\mathcal{M} \models \neg A$, then since

$$\begin{aligned} \neg A &= \neg \exists z_1 \forall z_2 \ldots \forall z_{2n} B(z_1, \ldots, z_{2n}) \\ &\equiv \forall z_1 \exists z_2 \ldots \exists z_{2n} \neg B(z_1, \ldots, z_{2n}), \end{aligned}$$

we have

$$\mathcal{M} \models \forall z_1 \exists z_2 \ldots \exists z_{2n} \neg B(z_1, \ldots, z_{2n}).$$

So assume now $\exists$ starts with an arbitrary $a_1 \in M$. Then we have

$$\mathcal{M} \models \exists z_2 \forall z_3 \ldots \exists z_{2n} \neg B(z_1, \ldots, z_{2n})[z_1 \mapsto a_1].$$

So $\forall$ can respond by playing some a_2 such that

$$\mathcal{M} \models \forall z_3 \exists z_4 \ldots \exists z_{2n} \neg B(z_1, \ldots, z_{2n})[z_1 \mapsto a_1,\ z_2 \mapsto a_2]$$

etc., as before. If $\forall$ follows this strategy, at the end of this run we will have $a_1, a_2, \ldots, a_{2n}$ such that

$$\mathcal{M} \models \neg B[z_1 \mapsto a_1, \ldots, z_{2n} \mapsto a_{2n}],$$

so $\forall$ has won.

Since it is clear that it cannot be that *both* $\exists$ and $\forall$ have winning strategies in $G_A^{\mathcal{M}}$ (otherwise they can play their winning strategies against each other and then they both will win, which is impossible), it follows that if $\exists$ has a winning strategy in $G_A^{\mathcal{M}}$, then we cannot have $\mathcal{M} \models \neg A$, so we must have $\mathcal{M} \models A$. Similarly, in case $\forall$ has a winning strategy, we have $\mathcal{M} \models \neg A$, so we are done. $\dashv$

Example 2.5.7 Consider the sentence A:

$$\exists x_1 \forall x_2 \exists x_3 (x_3 + x_3 = x_2 \vee x_3 + x_3 = x_2 + x_1)$$

Let $\mathcal{M} = \langle \mathbb{N}, + \rangle$. Then the game $G_A^{\mathcal{M}}$ is as follows

$\exists$	$\forall$
a_1	
	a_2
a_3	

where $\exists$ wins if $2a_3 = a_2 \vee 2a_3 = a_1 + a_2$, otherwise $\forall$ wins. Then $\exists$ has the following winning strategy (and hence $\mathcal{M} \models A$):

$\exists$ starts with $a_1 = 1$. Then after $\forall$ plays an arbitrary a_2, $\exists$ responds by playing $a_3 = \frac{a_2}{2}$, if a_2 is even, and by $a_3 = \frac{a_2+1}{2}$, if a_2 is odd.

Using this interpretation, we can give an application of first-order logic to game theory.

Definition 2.5.8 A *finite game* is determined by a set M and a relation $R \subseteq M^k$ ($k = 1, 2, \dots$). Take for simplicity $k = 2n$ to be even. The game is played as follows: We have two players, I, II which take turns in playing $a_1, a_2, \dots, a_{2n-1}, a_{2n}$ in M. I wins this run of the game if $R(a_1, \dots, a_{2n})$. Otherwise II wins (i.e. if $R(a_1, \dots, a_{2n})$ fails).

I	II
a_1	
	a_2
a_3	
	a_4
$\vdots$	$\vdots$
a_{2n-1}	
	a_{2n}

We can use fact 2.5.6 to prove the following:

Theorem 2.5.9 (Zermelo, von Neumann) *Every finite game is* determined*, i.e. one of the two players has a winning strategy.*

Proof. Consider the language with one $2n$-ary relation symbol $\bar{R}$ and let $\mathcal{M}$ be its structure

$$\mathcal{M} = \langle M, R \rangle,$$

(i.e. $R = \bar{R}^{\mathcal{M}}$). Let A be the sentence

$$\exists x_1 \forall x_2 \dots \exists x_{2n-1} \forall x_{2n} \bar{R}(x_1, \dots, x_{2n}).$$

Clearly $G_A^{\mathcal{M}}$ is the above finite game, with I=$\exists$ and II=$\forall$, so by fact 2.5.6, if $\mathcal{M} \models A$, I has a winning strategy, while if $\mathcal{M} \models \neg A$, II has a winning strategy, and since either $\mathcal{M} \models A$ or $\mathcal{M} \models \neg A$, one of the two must have a winning strategy. $\dashv$

Remark. One could also define *basically finite* games. Here we have a set M and for each n a (possibly empty) relation $R_n \subseteq M^n$. Let $R = \bigcup_n R_n$. Players I and II alternate playing elements a_i of M as before. In order for the game to be basically finite, two conditions need to hold: (1) At each *move*, a player has only finitely many possibilities, and (2) the game always ends after finitely many stages, when the output $(a_1, a_2, \dots, a_n)$ thus produced does not belong to R. The last player to have played loses. An easy application of König's lemma shows that if a game if basically finite then there is an n such that each play of the game lasts at most n moves. It then follows from the argument of the Zermelo-von Neumann theorem that basically finite games are also determined.

One can continue this line of argument and analyze infinite determined games, but we will not pursue this road here.

2.6 Theories

Let S be a set of sentences in a fixed first-order language L. We denote by $\mathrm{Con}(S)$ the set of all sentences which are logical consequences of S, i.e.

$$\mathrm{Con}(S) = \{A : S \models A, A \text{ a sentence}\}.$$

Notice that $\mathrm{Con}(\mathrm{Con}(S)) = \mathrm{Con}(S)$, i.e. $\mathrm{Con}(S)$ is closed under logical consequences.

Definition 2.6.1 A set of sentences T is called a *theory* if it is closed under logical consequences, i.e. iff $T = \mathrm{Con}(T)$.

Definition 2.6.2 If $\mathrm{Con}(S) = T$, we say that S is a *set of axioms* for the theory T. Note that in this case S and T have exactly the same models.

Example 2.6.3 Consider $L = \{<\}$, $<$ a binary relation symbol. The theory of *partial order* has as axioms:

(1) $\forall x \forall y (x < y \Rightarrow \neg y < x)$ (antisymmetry)
(2) $\forall x \forall y \forall z (x < y \wedge y < z \Rightarrow x < z)$ (transitivity)

A model $\langle M, < \rangle$ of this theory is called a *partially ordered set* (or just *partial order*). If we add the axiom:

(3) $\forall x \forall y (x < y \vee x = y \vee y < x)$ (linearity)

we obtain the theory of *linear order*, whose models are called *linearly ordered sets* (or just *linear orders*).

The theory of *dense linear order* is obtained by adding two more axioms:

(4) $\exists x \exists y (x \neq y)$
(5) $\forall x \forall y (x < y \Rightarrow \exists z (x < z \wedge z < y))$ (density)

Examples of models of this theory (i.e., *dense linear orders*) are $\langle \mathbb{Q}, < \rangle$, $\langle \mathbb{R}, < \rangle$. Examples of linear orders which are not dense are $\langle \mathbb{Z}, < \rangle$ and $\langle \mathbb{N}, < \rangle$.

Finally the theory of *dense linear orders without endpoints* is obtained by adding the axioms:

(6) $\forall x \exists y (x < y)$
(7) $\forall x \exists y (y < x)$.

Models of these are again $\langle \mathbb{R}, < \rangle$, $\langle \mathbb{Q}, < \rangle$, but not $\langle [0,1], < \rangle$.

Remark. A *well ordered set* (or *wellorder*) is a linear order $\langle M, < \rangle$ with the following additional property: Every non-empty set $S \subseteq M$ has a least element, e.g. $\langle \mathbb{N}, < \rangle$. It turns out that one cannot express this by axioms in first-order logic, i.e. there is no set of axioms S in the first-order language $\{<\}$ whose models are exactly the wellorders. This will follow from the Compactness Theorem, that we will prove later on.

Example 2.6.4 Consider $L = \{E\}$, E a binary relation symbol. The theory of *(undirected) graphs* has as axioms

(1) $\forall x \neg (xEx)$
(2) $\forall x \forall y (xEy \Rightarrow yEx)$

A model of this theory is a *graph*.

Remark. It turns out again that there is no set of axioms in first-order logic whose models are exactly the connected graphs. Similarly for trees, i.e. connected acyclic graphs.

Example 2.6.5 Let $L = \{e, \cdot\}$, e a constant symbol and $\cdot$ a binary relation symbol. The *theory of groups* has the following axioms:

(1) $\forall x \forall y \forall z (x \cdot (y \cdot z) = (x \cdot y) \cdot z)$ (associativity)
(2) $\forall x (x \cdot e = x \wedge e \cdot x = x)$ (identity)
(3) $\forall x \exists y (x \cdot y = e \wedge y \cdot x = e)$ (existence of inverses)

A model of this theory is called a *group*, e.g. $\langle \mathbb{Z}, 0, + \rangle$, or $\langle \mathbb{R}^+, 1, \cdot \rangle$, where

$$\mathbb{R}^+ = \{x \in \mathbb{R} : x > 0\}.$$

The theory of *abelian groups* has one more axiom:
(4) $\forall x \forall y (x \cdot y = y \cdot x)$ (commutativity)
Examples of abelian groups are $\langle \mathbb{Z}, 0, + \rangle$, $\langle \mathbb{Z}_m, 0, + \rangle$, etc. Examples of groups which are not abelian are $\langle \mathcal{C}([0,1]) \circ \rangle$, (functions under composition) and $\langle GL_n(\mathbb{R}), \cdot \rangle$ ($n \times n$ matrices with matrix multiplication).

Example 2.6.6 Let $L = \{0, 1, +, \cdot\}$, 0, 1 constant symbols , $+, \cdot$ binary function symbols. The theory of *commutative rings with identity* has the following axioms:

(1)-(4) Same as the four axioms in example 2.6.5 with $\cdot$ replaced by $+$ and e by 0 (i.e. the axioms for abelian groups for $0, +$)
(5) $\forall x \forall y \forall z (x \cdot (y \cdot z) = (x \cdot y) \cdot z)$ (associativity for $\cdot$)
(6) $\forall x \forall y (x \cdot y = y \cdot x)$ (commutativity for $\cdot$)
(7) $\forall x (x \cdot 1 = x)$ (identity for $\cdot$)
(8) $\forall x \forall y \forall z (x \cdot (y + z) = x \cdot y + x \cdot z)$ (distributivity)

Examples of models of this theory (i.e. *commututive rings with identity*) are $\langle \mathbb{Z}, 0, 1, +, \cdot \rangle$, $\langle \mathbb{Q}, 0, 1, +, \cdot \rangle$, $\langle \mathbb{Z}_m, 0, 1, +, \cdot \rangle$, etc.
The *theory of integral domains* is obtained by adding the axiom
(9) $\forall x \forall y (x \cdot y = 0 \Rightarrow x = 0 \vee y = 0)$ (no zero divisors)
Examples of integral domains are $\langle \mathbb{Q}, 0, 1, +, \cdot \rangle$, $\langle \mathbb{Z}, 0, 1, +, \cdot \rangle$, $\langle \mathbb{Z}_p, 0, 1, +, \cdot \rangle$ (p a prime), but not $\langle \mathbb{Z}_6, 0, 1, +, \cdot \rangle$.
The theory of *fields* is obtained by adding the axioms:
(10) $0 \neq 1$
(11) $\forall x (x \neq 0 \Rightarrow \exists y (x \cdot y = 1))$ (existence of inverses for $\cdot$)

For example, $\langle \mathbb{Q}, 0, 1, +, \cdot \rangle$, $\langle \mathbb{R}, 0, 1, +, \cdot \rangle$ are fields, but $\langle \mathbb{Z}, 0, 1, +, \cdot \rangle$ is not. The theory of *fields of characteristic zero* is obtained by adding the infinite list of axioms:

$(12)_2$ $1 + 1 \neq 0$

$(12)_3$ $1 + 1 + 1 \neq 0$

$\vdots$

$(12)_n$ $\underbrace{1 + 1 + \cdots + 1}_{n} \neq 0$

for each $n \geq 2$. Examples of models of this theory, i.e. fields of characteristic zero, are $\langle \mathbb{Q}, 0, 1, +, \cdot \rangle$, $\langle \mathbb{R}, 0, 1, +, \cdot \rangle$, $\langle \mathbb{C}, 0, 1, +, \cdot \rangle$ but not $\langle \mathbb{Z}_p, 0, 1, +, \cdot \rangle$, p a prime.

Remark. Notice that the theory of fields of characteristic zero has infinitely many axioms. It turns out that it is not *finitely axiomatizable*, i.e. it cannot have a finite set of axioms.

Another way a theory can arise in practice is as follows. Let L be a first-order language and $\mathcal{M}$ as structure for L. The *theory* of $\mathcal{M}$, $\mathrm{Th}(\mathcal{M})$, is defined by

$$\mathrm{Th}(\mathcal{M}) = \{A : A \text{ a sentence and } \mathcal{M} \models A\}.$$

(It is indeed easy to check that $\mathrm{Th}(\mathcal{M})$ is a theory.)

Notice that here $\mathrm{Th}(\mathcal{M})$ is not given in any meaningful sense in terms of axioms (except in the trivial way, namely taking these axioms to be the theory itself). It is often an important problem to find a reasonable set of axioms for the theory of $\mathcal{M}$, for various structures of mathematical interest. In other words, we are looking for some explicit list of sentences true about $\mathcal{M}$, so that any other one follows logically from them.

Example 2.6.7 Consider $L = \{0, 1, +, \cdot\}$ and $\mathcal{M} = \langle \mathbb{R}, 0, 1, +, \cdot \rangle$. It turns out that the following is a set of axioms for $\mathrm{Th}(\mathcal{M})$:

(1)-(11) the axioms for fields

(12) $\forall x \exists y (x = y^2 \vee x + y^2 = 0)$

(here y^2 abbreviates $y \cdot y$; similarly y^n abbreviates $\underbrace{y \cdot y \cdot \cdots \cdot y}_{n}$)

(13) $\forall x \forall y \exists z (x^2 + y^2 = z^2)$

(14) $\forall x (x^2 \neq -1)$

$(15)_n$ $\forall x_0 \dots \forall x_n (x_n \neq 0 \Rightarrow \exists y (x_n \cdot y^n + x_{n-1} \cdot y^{n-1} + \cdots + x_1 \cdot y + x_0 = 0))$,

where $n = 1, 3, 5 \dots$ ranges over all *odd* numbers.

Example 2.6.8 For the same L and $\mathcal{M} = \langle \mathbb{C}, 0, 1, +, \cdot \rangle$, it turns out that a set of axioms consists of the axioms for fields of characteristic 0 plus all the axioms $(15)_n$ or above but now for $n = 1, 2, 3, \ldots$ ranging over *all* numbers.

Example 2.6.9 Let now $L_{ar} = \{0, S, +, \cdot, <\}$ be the language of arithmetic and $\mathcal{N}$ be its standard structure $\langle \mathbb{N}, 0, S, +, \cdot, <\}$. Consider the following set of axioms, called the *first-order Peano axioms* (PA):

(1) $\forall x(S(x) \neq 0)$
(2) $\forall x \forall y(x \neq y \Rightarrow S(x) \neq S(y))$
(3) $\forall x(x + 0 = x)$
(4) $\forall x \forall y(x + S(y) = S(x + y))$
(5) $\forall x(x \cdot 0 = 0)$
(6) $\forall x \forall y(x \cdot S(y) = x \cdot y + x)$
$(7)_A$ $\forall y_1 \cdots \forall y_n[(A(0, y_1, \ldots, y_n) \wedge \forall x(A(x, y_1, \ldots, y_n)$
$\Rightarrow A(S(x), y_1 \ldots y_n))) \Rightarrow \forall x A(x, y_1, \ldots, y_n)]$

for any formula $A(x, y_1, \ldots, y_n)$. (So this is an infinite set of axioms.)

Practically every known fact of number theory is a logical consequence of these axioms. However this set of axioms is not sufficient as a set of axioms for Th($\mathcal{N}$), i.e., there are some sentences A true in $\mathcal{N}$ which cannot be derived logically from these axioms. In fact, *there is no way to find a reasonable set of axioms for* Th($\mathcal{N}$) (contrast this with the previous examples 2.6.7 and 2.6.8). This is a consequence of the celebrated *Gödel First Incompleteness Theorem.*

Example 2.6.10 Consider the language $L = \{\in\}$, where $\in$ is a binary relation symbol. This is called the *language of set theory* and its intended meaning is to view $\in$ as representing membership of sets and the variables are ranging over all sets. One can formulate in this language a list of axioms called the *Zermelo-Fraenkel Axioms with the Axiom of Choice* (ZFC), from which one can logically derive practically all of the present day ordinary mathematics, which, as it is well understood today, can be founded on the theory of sets. Here are a few of the ZFC axioms:

(1) $\forall x \forall y(x = y \Leftrightarrow \forall z(z \in x \Leftrightarrow z \in y))$ (Extensionality axiom)
(2) $\forall x \forall y \exists z \forall w(w \in z \Leftrightarrow w = x \vee w = y)$ (Pairing axiom)
(3) $\forall x \exists y \forall z(z \in y \Leftrightarrow \forall t(t \in z \Rightarrow t \in x))$ (Powerset axiom)
$\vdots$

Thus the theory Con(ZFC) can be viewed as a global theory encompassing most present day mathematics. (But not quite all – by the same Gödel Incompleteness Theorem no reasonable axiomatic theory can be *complete* in the sense that there will always be, for any fixed reasonable set of axioms S, a sentence A such that $S \not\models A$ and $S \not\models \neg A$, provided S is powerful enough to include some very elementary number theory. *Reasonable* here means that there is a algorithm to decide whether a sentence belongs to S or not.)

2.7 A Proof System for First-Order Logic

Just as we did for propositional logic in section 1.11, we will now discuss a proof system for first-order logic, i.e. a way to write formal proofs of statements from axioms and rules of inference. Our main goal, as before, is to prove the *Gödel Completeness Theorem*:

$$S \vdash A \text{ iff } S \models, A.$$

but we will not get to this until section 2.8.

2.7.A Formal Proofs

As in section 1.11 it will be convenient to consider as basic symbols in the language of first-order logic the following

$$\neg, \Rightarrow,), (, x_1, x_2, \ldots, =, \forall$$

and view $(A \wedge B)$ as an abbreviation of $\neg(A \Rightarrow \neg B)$, $(A \vee B)$ as an abbreviation of $(\neg A \Rightarrow B)$, $(A \Leftrightarrow B)$ as an abbreviation of $\neg((A \Rightarrow B) \Rightarrow \neg(B \Rightarrow A))$ and

$$\exists x A \text{ as an abbreviation of } \neg\forall x \neg A.$$

Definition 2.7.1 If A is a formula, then we call any formula of the form

$$\forall y_1 \forall y_2 \ldots \forall y_n A$$

a *generalization* of A.

Example 2.7.2 $\forall x \forall y(P(x) \Rightarrow \exists z Q(x, y, z))$ is a generalization of $(P(x) \Rightarrow \exists z Q(x, y, z))$.

We will now describe a Hilbert-type proof system for first-order logic (this particular system is essentially coming from H. Enderton, *A Mathematical Introduction to Logic*). At this point, we fix a first-order language L and consider only formulas in L from now on.

A *(logical) axiom* is any formula which is a *generalization* of a formula of the following form:

(a) (i) $A \Rightarrow (B \Rightarrow A)$

(ii) $((A \Rightarrow (B \Rightarrow C)) \Rightarrow ((A \Rightarrow B) \Rightarrow (A \Rightarrow C))$

(iii) $((\neg B \Rightarrow \neg A) \Rightarrow ((\neg B \Rightarrow A) \Rightarrow B))$

(b) $(\forall x(A \Rightarrow B) \Rightarrow (\forall x A \Rightarrow \forall x B))$

(c) $(A \Rightarrow \forall x A)$, provided x is not free in A

(d) $\forall x A \Rightarrow A[x/t]$, provided t is a term *substitutable* for x in A, where this proviso will be explained shortly.

(e) (i) $x = x$

(ii) $(x = y \wedge y = z) \Rightarrow (x = z)$

(iii) $x = y \Rightarrow y = x$

(iv) $(y_1 = z_1 \wedge \cdots \wedge y_n = z_n) \Rightarrow (f(y_1, \ldots, y_n) = f(z_1, \ldots, z_n))$

(v) $(y_1 = z_1 \wedge \cdots \wedge y_m = z_m) \Rightarrow (R(y_1, \ldots, y_m) \Leftrightarrow R(z_1, \ldots, z_m))$

(Here A, B, C are arbitrary formulas, $x, y, z, y_1, \ldots y_n, z_1, \ldots z_n$ arbitrary variables, f arbitrary n-ary function symbols and R arbitrary m-ary relation symbols.)

We also have a rule of inference, namely *modus ponens* (MP).

From A, $(A \Rightarrow B)$ derive B.

Definition 2.7.3 If S is any set of formulas, a *formal proof* from S is a sequence $A_1, \ldots, A_n$ of formulas such that each A_i is either a logical axiom, belongs to S or comes by applying modus ponens to some A_j, A_k with $j, k < i$. We call this sequence a *formal proof of* A_n *from* S.

If there is a formal proof of a formula A from S, we say that A is a *formal theorem* of S and write

$$S \vdash A.$$