Chapter II. Mathematical Structures and their Descriptions in First Order Logic.

In this chapter we will look at a few well-known examples of first order theories. These examples are important in their own right, i.e. as formalisations of structures which arise in certain branches of mathematics and other scientific domains. But they will also serve as illustrations of certain general logical issues and we shall use them as opportunities to introduce and discuss those.

The kinds of structures which we will discuss fall into four main classes:

(i) orderings

(ii) certain classes of algebraic structures such as boolean and non-boolean lattices and groups

(iii) the structure of the natural numbers and that of the real numbers with their familiar arithmetical operations + and .

(iv) feature structures

The first order theories of these structures and structure classes we will present will serve as anchors for the discussion of such general issues as: incompleteness, completeness and categoricity of theories; theory extensions and Lindenbaum algebras; quantifier elimination; independence; implicit and explicit definability; equational logic as a subsystem of first order logic; and feature logic as an alternative to first order logic.

2.1 Orderings.

Our first examples concern the concept of order. Mathematically, order is most naturally represented in the form of a binary relation - either a *strict* ordering relation $<$ or a *non-strict* ordering relation $\leq$. (Strict ordering relations are irreflexive and non-strict orderings reflexive. Given a strict ordering $<$ we can define a corresponding non-strict ordering $\leq$ by: $a \leq b$ iff $a < b \vee a = b$; conversely, from a non-strict ordering $\leq$ we get a strict ordering $<$ via: $a < b$ iff $a \leq b \,\&\, a \neq b$.) Orderings can be classified in terms of the properties of $<$ (or, equivalently, of $\leq$). First, there is the distinction between *linear orders* and *partial orders*. In a linear order of a domain D any two distinct

elements a, b of D are ordered in the sense that either a stands in the ordering relation to b or b else stands in the relation to a. In partial orders this condition is in gene4ral not satisfied. (Thus the notion of a partial order is the more general one; linear orders are a special kind of partial order.)

A second important distinction is that between denseness and discreteness. In a dense order there is for any a and b such that $a < b$ an element c such that $a < c < b$; in a discrete order there exists for any a and b such that $a < b$ a c with the properties that (i) $a < c \leq b$ and (ii) there is no d such that $a < d < c$; and, similarly, if $b < a$ then there is a c such that (i) $b \leq c < a$ and (ii) there is no d such that $c < d < a$. (The element c in question is called the *immediate successor* (*predecessor*) *of* a *in the direction of* b.) It should be emphasised that denseness and discreteness are mutually exclusive (in the sense that no non-trivial ordering - i. e. no ordering which holds between at least two different elements - can be dense and discrete at the same time), but that they are not jointly exhaustive: An ordering may be neither dense nor discrete, for instance because it consists of one part which is dense and another which is discrete.

Here we will look at two distinct kinds of ordering structures:

(a) certain linear orders, among them the ordering of the rational numbers, that of real numbers (both dense orderings), that of the narural numbers and that of the integers (both discrete orderings);

(b) partial orders which have the additional property of being *lattices*. A *lattice* is a partial order in which for any two elements a and b there is a "smallest element above both of them" - i.e. an element c such that $a \leq c$ and $b \leq c$ and which is least with regard to this condition, i.e. if c' is any other element such that $a \leq c'$ and $b \leq c'$, then $c \leq c'$ - and, dually, there exists for any a and b a "greatest element that is $\leq$ a and $\leq$ b".)

Lattice-like orders have the important property that they c an be described not only in terms of their orderings, but also interms of the lattice operations $\cup$ and $\cap$, which can be defined in terms of the ordering ($a \cup b$ is the least element above a and b and $a \cap b$ the greatest element below a and b) and which in their turn allow definition of the

ordering relation (e.g. via the definition: a $\leq$ b iff a $\cup$ b = b). Thus lattices can also be viewed as *algebraic structures* or *algebras* - that is, structures consisting of a universe together with a number of functions defined on that universe. (In other words, an algebraic structure is a model for a language L all of whose non-logical constants are function constants.)

Of particular importance among the lattices that we will discuss are the *boolean lattices* (or *boolean algebras*, the term that is used to refer to them when they are pressented as structures involving functions). The logical importance of boolean algebras will no doubt be familiar: classical propositional logic with the connectives & and v has the structure of a boolean algebra.

The order in which we proceed in this section is as follows. We begin with the ordering theory T_{rat} of the rational numbers, presenting the conceptually and historically important theorem of Cantor's according to which any denumerable model of T_{rat} is isomorphic to the ordering structure of the rationals. This will be the basis for introducing the notion of a theory being categorical in a certain cardinality κ. Cantor's Theorem shows that T_{rat} is categorical in the cardinality of the denumerably infinite sets, but as it turns out not in any other infinite cardinality. The subsection closes with a brief discussion of Morley's Categoricity Theorem.

Next, in subsection 2.1.2, we proceed to lattices. We begin with axiomatic characterisations of the class of all lattices, first from the ordering perspective (i.e. formulating our axioms in the first order language $\{\leq\}$ whose only non-logical constant is the 2-place relation $\leq$, and then form the algebraic perspective, using the language $\{\cup,\cap\}$. We show that each of these two theories is definable within the other. We then extend these axiomatisations to obtain theories for the class of all boolean lattices and for that of all boolen algebras, respectively, theories that are again definable within each other. Section 2.1.3 is concerned with the variety of boolean algebras. It presents some particular boolean algebras and some properties in terms of which arbitrary boolean algebras can be classified. 2.1.4 presents the Cech-Stone Representation Theorem, according to which every boolean algebra is isomorphic to (and thus 'can be represented as') a set algebra - a boolean algebra consisting of sets, with the set inclusion relation as the partial order of the lattice. Representation theorems, which assert that every structure with certain abstract properties can be 'represented', or 'realised' as a structure of some more specific

kind, are of great importance in many areas in mathematics; the Cech-Stone Theorem can be regarded as the classical paradigm of theorems of this general form.

The theory of boolean algebras is incomplete, since among its models are boolean algebras that can be distinguished from each other by properties expressible in the language of the theory itself. Even more obvious is the incompleteness of the theory of all lattices, since among its models are on the one hand the boolean lattices, which are also models of the theory of boolean lattices, and on the other hand non-boolean lattices, which are not models of that theory. (Thus the theory of boolean lattices is a proper extension of the general theory of lattices, which proves the latter's incompleteness.) In Section 2.2 we look at incomplete theories from a more general and systematic perspective. The structure consisting of all theories of a given language l., and more generally that consisting of all theories of L which extend a given theory T, are both lattices (though in general not boolean lattices). Thus the study of these structures provides with a further application of lattice theory, as well as giving more insight in the structure of first order logic.

The lattice consisting of all extensions of a given theory T as well as a certain boolean sublattice of this structure, the so-called Lindenbaum algebra of T, are studied in 2.2.1. 2.2.2 contains a discussion of almost complete theories. here we return to linear orderings comparing theories of dense orderings with certain theories of discrete orderings.

2.1.1. The Theory of Dense Linear Orders without End Points.

We choose as our first task in this chapter that of formulating a first order theory that captures all truths about the ordering of the rational numbers. To this end we choose as our language, in which the theory will be formulated, the language $\{<\}$, whose only non-logical constant is the two-place predicate $<$. We will refer to $\{<\}$ also as $L_<$. Our task is thus to state a theory of $L_<$ whose theorems are all and only the truths expressible in $L_<$ about the structure $\langle Q,<_Q\rangle$, where Q is the set of rational numbers and $<_Q$ is the standard ordering of the rationals.

Here is our proposal: Let T_{rat} be the theory consisting of all logical consequences of the following set of axioms:

Def. 1 (Axioms of T_{rat})

L1. $(\forall x)(\forall y)\ (x < y \rightarrow \neg\ (y < x))$
L2. $(\forall x)(\forall y)(\forall z)\ ((x < y\ \&\ y < z) \rightarrow x < z)$
L3. $(\forall x)(\forall y)\ (x < y\ \vee\ x = y\ \vee\ y < x)$
L4 $(\forall x)(\forall y)\ (x < y \rightarrow (\exists z)\ (x < z\ \&\ z < y))$
L5 $(\forall x)(\exists y)\ (x < y)$
L6. $(\forall x)(\exists y)\ (y < x)$

T_{rat} is also known as the *theory of dense linear orders without endpoints*. The subtheory of T_{rat} that is axiomatised by L0-L3 is known as the *theory of linear orders* and that axiomatised by L0-L2 as the *theory of partial orders*.[1] WE will refer to the first as T_{lin} and to the second as T_{par}.

Some the properties of T_{rat} are stated in Theorem 10.

Theorem 1. (1) Every model of T_{rat} is infinite:
(2) (Cantor) Every two denumerably infinite models of T_{rat} are isomorphic.
(3) T_{den} is complete.

Proof.

(1) Note that because L0 is an axiom of T_{rat} any model of T_{rat} must have at least 2 elements. Secondly, suppose that $M = <U_M, <_M>$ is a finite model of T_{rat}, i.e. that U_M consists of elements $a_1, \ldots, a_n$, where n is some natural number. As just observed, n must be at least 2. Furthermore, since $<_M$ is a linear order, there must be among the elements $a_1, \ldots, a_n$ at the very least one pair of elements (a_i, a_j) such that $a_i < a_j$ and for no a_k, $a_i < a_k\ \&\ a_k < a_j$. But this contradicts L5.[2]

---

[1] Often axiom L0 is not included in axiomatisations of the theories of linear or partial orderings. leaving it out has the effect that among the models of the hteory one also includes structures of the form $<\{a\}, \varnothing>$, where $\{a\}$ is any singleton set and < is interpreted as the empty relation $\varnothing$. Whether such structures are included or not makes no real difference to what the theory says about the structures which really matter, viz. those in which the universe contains more than one element. In hte present context it has proved to be a little more convenient to exclude them from the start, and thus to include L0 among the axioms.

[2] Strictly speaking the existence of a pair (a_i, a_j) as just stated should be proved. In fact it is easy to prove, by induction on n, that every model of the theory of linear orders whose universe consists of n elements contains such a pair: Suppose this holds for n and let M be a model with universe $\{a_1, \ldots, a_n, a_{n+1}\}$.

(2) Let M, M' be denumerable models of T_{rat} with universes $U_M = \{a_1, a_2, \ldots\}$ and $U_{M'} = \{b_1, b_2, \ldots\}$. We refer to the interpretations $<_M$ and $<_{M'}$ in respectively M and M' of the predicate < as $<$ and $<'$. We construct, by induction on n, partial isomorphisms h_n from M to M' with domains $\{a^1, \ldots, a^n\}$ and ranges $\{b^1, \ldots, b^n\}$. In this notation we assume that $a^1 < \ldots < a^n$ and $b^1 <' \ldots <' b^n$ (and thus that h_n is defined by: $h_n(a^i) = b^i$, for i = 1, .., n. Moreover, the h_n will be constructed in such a way that, putting $h = \bigcup_n h_n$, h is an isomorphism from M to M'.

We proceed as follows. Suppose that the elements $a^1, \ldots, a^n$ and $b^1, \ldots, b^n$ have alredsy been chosen. We distinguish between the case where n is odd and that where n is even.

(a) Suppose n is odd. Then we pick the first element a_i from the enumeration $\{a_1, a_2, \ldots\}$ which does not occur among $\{a^1, \ldots, a^n\}$. For the position of a_i with respect to the $a^1, \ldots, a^n$ there are three possibilities:

(i) $a_i < a^1$;
(ii) $a^n < a_i$;
(iii) $a^k < a_i < a^{k+1}$, for some $k < n$.

(i) Because M' is a model of T_{rat} and T_{rat} contains L4, we know that there is a b among $\{b_1, b_2, \ldots\}$ such that $b <' b^1$. Let b_j be the first such b and let $h_{n+1} = h_n \cup \{<a_i, b_j>\}$. Then h_{n+1} is an isomorphism with $DOM(h_{n+1}) = \{a_i, a^1, \ldots, a^n\}$ and $RAN(h_{n+1}) = \{b_j, b^1, \ldots, b^n\}$.

(ii) This case is just like (i): We know that there is a b in $\{b_1, b_2, \ldots\}$ such that $b <' b^n$, etc.

(iii) This time we make use of L5. Because T_{rat} contains L5 that we may infer that $\{b_1, b_2, \ldots\}$ contains a b such that $b^k <_{M'} b <' b^{k+1}$. Again we let b_j be the first such b. Putting, as before, $h_{n+1} = h_n \cup$

Then consider the restriction M' of M to the set $\{a_1, \ldots, a_n\}$, i.e. the model with universe $\{a_1, \ldots, a_n\}$ in which the interpretation of < is the restriction of the interpretation of < in M to $\{a_1, \ldots, a_n\}$. Since M' has n elements, there is by assumption a pair (a_i, a_j) $(i,j \leq n)$ such that there is no a_k in M' with $a_i < a_k$ & $a_k < a_j$. If it is not the case that $a_i < a_{n+1}$ & $a_{n+1} < a_j$ then (a_i, a_j) is a pair for M of the required kind. If $a_i < a_{n+1}$ & $a_{n+1} < a_j$ then a pair of the desired kind is (a_i, a_{n+1}).

$\{<a_i,b_j>\}$, we conclude that h_{n+1} is an isomorphism with Domain $\{a^1, \ldots, a^k, a_i, a^{k+1}, \ldots, a^n\}$ and Range $\{b^1, \ldots, b^k, b_j, b^{k+1}, \ldots, b^n\}$.

(b) n is even. In this case, let b_i be the first element from the enumeration $\{b_1, b_2, \ldots\}$ which does not occur among $\{b^1, \ldots, b^n\}$ and find, in each of the cases (i) - (iii), an a_j in M which is "similarly situated" with respect to $\{a^1, \ldots, a^n\}$. We put $h_{n+1} = h_n \cup \{<a_j,b_i>\}$.

It is not hard to verify that the union h of all the h_n has for its Domain all of $\{a_1, a_2, \ldots\}$ (because of the steps in the construction for odd n) and that it has for its Range all of $\{b_1, b_2, \ldots\}$ (because of the steps for n even). Moreover, it is obvious from the construction that if a, a' are elements of M and
$a < a'$, then $h(a) <' h(a')$. From linearity (Axiom L3!) it then follows that for all a, a' from M, $a < a'$ iff $h(a) <' h(a')$.

(3) This follows almost directly from (2). Note that if T_{rat} were not complete, then there would be a sentence A from the language $L_<$ such that $\neg(A \in T_{rat})$ and $\neg(\neg A \in T_{rat})$. So it follows that $T_{rat} \cup \{A\}$ and $T_{rat} \cup \{\neg A\}$ are both consistent and thus each of them has a model. Let M_1 be a model of $T_{rat} \cup \{A\}$ and M_2 a model of $T_{rat} \cup \{\neg A\}$. By (i) both models are infinite. So by the downward Skolem-Löwenheim Theorem there are denumerably infinite models M'_1 and M'_2 such that $M'_1 \equiv M_1$ and $M'_2 \equiv M_2$. So A is true in M'_1 and false in M'_2. But by (ii) $M'_1 \equiv M'_2$: contradiction. We conclude that T_{rat} is complete.

q.e.d.

The centre piece of Theorem 1 is part (2). This result is generally known as 'Cantor's Theorem' (or more fully 'Cantor's Theorem about Dense Linear Orders', in order to distinguish this theorem from the equally famous theorem of Cantor that the cardinality of the power set of a given set X always exceeds that of X). The proof of this theorem has, like Cantor's proof of his power set theorem, been a milestone in the development of our understanding of what constitutes valid mathematical reasoning. At first, many mathematicians were very sceptical with regard to the soundness of these proofs. Precisely because their initially controversial status, Cantor's arguments were a major input to the debates over the foundations of mathematics that became a vital concern in the second half of the nineteenth Century and which in its turn provided much of the impetus to the development of formal logic as a fool-proof framework for doing mathematics. (Recall the interlude on Set Theory in Chapter I.)

As opposed to part (2) of Theorem 10, which is specifically about dense linear orderings. the purport of part (3) is much more general. The general statement. known as 'Vaught's test', is this:

Prop. 1 (Vaught's Test)

> Whenever T is a theory which (i) has only infinite models and (ii) is such that for some infinite cardinaltiy κ any two models of T of caridnality κ are isomorphic, then T is complete.

Complete theories are the closest we can get to characterising the properties of a given mathematical structure, when we want to do this by dscribing them within some logical language L. We have already seen some general limits to what can be achieved along these lines, viz. those imposed by the Skolem-Löwenheim Theorems presented in Chapter I. But in fact, for many structures, the best that can be achieved is even farther from the ideal (characterisation of the structure up to isomorphism) than the Skolem-Löwenheim Theorems would in principle allow for.

Let us be more exact. In order that a theory T of a first order language L can be considered a characterisation of some given structure *A*, two conditios must be satisfied. First, all the structural properties of *A* must be expressible in L. That is, we must be able to represent *A* as a model $M_O = <U_O,F_O>$ of L such that each relation that is relevant to the structure of *A* is either given as the interpretation $F_O(\alpha)$ of some non-logical constant α of L or else must be definable in terms of one or more relations $F_O(\alpha)$ with $\alpha \varepsilon$ L. (For a general discussion of notions of definability see Section 2.3.) Second, all sentences of L that are true in M_O must be iderivable from T as theorems (and thus, because T is closed under logical consequence,, must be members of T).

Assume that we have succeeded in choosing a language L such that the first condition is fulfilled - i.e. that we can represent *A* as some particular model M_O of L. In that case there exists - trivially - a unique theory T of L which verifies all and only the sentences of L that are true in M_O, viz, $Th(M_O)$. That the set $Th(M_O)$ always exists follows from general principles of set-theory (which will be spelled out in Ch. 3). But from the general principles which guarantee the existence of $Th(M_O)$ nothing follows that has anything to do in particular with the structure *A* whose properties $Th(M_O)$ describes. What we eally want is a non-trivial characterisation of $Th(M_O)$ that reveals some of the special

properties of $Th(M_o)$, and that ideally gives us some insight into them that might have eluded us without them. A natural way to go about this is to try to find 'axioms' for $Th(M_o)$ - sentences belonging to the theory which on the one hand can be readily verified as true in M_o and on the other as entailiing all other sentences that are true in M_o. It seems particularly desirable form this perspective to find a finite set of axioms for the theory. As we saw in Chapter I, this is always possible whenn *A*, and therewith M_o, are finite. But for infinite structures *A* the situation is very different. For instance, it is an interesting and surprising consequence of Gödel's Incompleteness Theorems that for many infinite structures *A* no finite axiomatisation of $Th(M_o)$ exists. (In fact, the situation is even worse in that there isn't even an infinite recursively enumerable set of axioms for $Th(M_o)$; for 'recursively enumerable' see Ch. ??.)

These negative results hold in spite of the fact that iby requiring only that our theory captures all the truths about *A* that are expressible in L we haven't pitched our aims necessarily very high. There is also another, stricter sense in which one can define complete characterisation of *A* by T

Any model M of T is isomorphic to M_o

(where again M_o is represented as model for the language L of T) Again, when *A* is finite, then, as established by Thm. 6 in Chapter I, a theory T satisfying this requirement can always be found (and when L is also finite, then this theory is finitely axiomatisable, e.g. by the single axiom described in th proof of Thm. 6). But the Skolem-Löwenheim Theorems tell us that this desideratum is never met iwhen *A* is infinite. For as soon as *A* is infinite, $Th(M_o)$ will have models of different infinite cardinalities and these can never be isomorphic to each other. The best we can hope for is that models of $Th(M_o)$ are isomorphic to each other so long as they are of the same cardinality. But even this weaker condition is only seldomly fulfilled and holds only for rather uninteresting structures *A*, with largely trivial structural properties.

In fact, even for the ordering structure $<Q,<_Q>$ of the rationals this weaker requirement is not fulfilled, Cantor's Theorem notwithstanding. For while , as the Thm states, any two *denumerable* models of $Th(<Q,<_Q>)$ (= T_{rat}) are isomorphic, this is not so for non-denumerable models - see Exercise ??.

For easier formulations during the remainder of this section we introduce some further terminology.

Def. 2 A theory T in a first order language L is called *categorical in* a cardinality κ, or also κ-*categorical*, iff any two models of T of cardinality κ are ismorphic.

Using this defintion we can restate what has just been said about T_{rat}) as:

(i) T_{rat} is ω-categorical (where ω is the cardinality of the denumerable sets and structures)

(ii) For any non-denumerable cardinality κ, T_{rat} is not κ-categorical.

Another way to describe these two facts makes use of the notion of the *categoricity spectrum* of a (complete) theory T. By the *categoricity spectrum of* T, CS(T), we understand that function which maps an infinite cardinality κ to 1 iff any two models of T of cardinality κ are ismorphic, and otherwise maps κ to 0. In terms of categoricity spectra the characterisation of T_{rat} is as follows:

(i) $CS(T_{rat})(\omega) = 1$;
(ii) $CS(T_{rat})(\kappa) = 0$, if κ non-denumerable.

From what little has been said so far, we should be prepared for all sorts of categoricity spectra - functions CS(T) according to which the collection of infinite cardinalities κ such that $CS(T)(\kappa) = 1$can take a wide variety of different forms. But as a matter of fact this is not so. It was shown in the early sixties by Morley - arguably the first truly deep result in general model theory - that for categoricity spectra CS(T) there are altogether only four possibilities: :

i. $CS(T)(\kappa) = 1$ for all infinite cardinalities κ;
ii. $CS(T)(\omega) = 1$; $CS(T)(\kappa) = 0$ for κ non-denumerable;
iii. $CS(T)(\omega) = 0$; $CS(T)(\kappa) = 1$ for κ non-denumerable;
iv. $CS(T)(\kappa) = 0$ for all infinite cardinalities κ.

As indicated above, case (i) turns out to be very rare and arises only for essentially trivial structures. (An example is the theory T_{inf} of the language {} which says that there are infinitely many indviduals.) An example of case (ii) is, as we have seen, our theory T_{rat}, but there

aren't many other interesting examples in this category, involving structures that are familiar on independent grounds. Examples of case (iii) are also rare; one - very surprising - example is the first order theory of the arithmetic operations + and . on the real numbers (see Section 2.4.2).

The bulk of mathematically important structures gives rise to theories falling under (iv). Among these structures there are in particular all those which contain the arithmetical structure of the natural numbers (i.e. the natural numbers with the operations of + and .)as a definable substructure. (Trivially, this includes in particular to the arithmetical structure of the natural numbers itself. For that structure contains itself as an (improper) substructure, definable by means of identity maps.)

All these negative results are indications of the limits of first order logic as a toll for characterising non-trivial mathematical structure.

Morley's Theorem is usually stated in the following form[3]:

Theorem 2 (Morley).

> Suppose that T is a theory of some first order language L and that T is κ-categorical for some non-denumerably infinite cardinality κ. Then T is κ-categorical for all non-denumerably infinite cardinalities κ.

2. 1.2 Lattices, as Partial Orders and as Algebras.

We noted in 2.1 that lattices can be viewed in two different ways. On the one hand they can be described as partial orderings with certain special properties (any two elements a and b have a least element above them (the *supremum* of a and b) and a greatest element below them (the *infimum* of a and b). But they can also be described as algebraic structures, characterised by two binary operations $\cup$ and $\cap$, which

---

[3] We do not prove Morley's theorem in these Notes. The proof is of this theorem is hard (much harder than any proof presented in these Notes) and would detain us for far too long. Proofs can be found in several textbooks on model theory, for instance in Chang & Keisler, *Model Theory*. or Hodges *Model Theory*.

assign to any pair of elements a, b their supremum $a \cup b$ and their infimum
$a \cap b$.

We first present lattices as partial orders with the mentioned properties; that is, we formulate an axiomatic theory T_{lato} ('lato' stands for 'lattice order') in the language L_{lato} (the language whose only non-logical constant is the 2-place predicate $\leq$ and for which the canonical reference would be '{$\leq$}') whose models are all and only the partial ordering that are lattices. We then show how the operations $\cup$ and $\cap$ can be defined in this theory and form a new theory T'_{lato} in the language $\{\leq,\cup,\cap\}$ by adding the proposed definitions of $\cup$ and $\cap$ to the given axioms of T_{lato}. From the axioms of T'_{lato} (which, remember, include the definitions of $\cup$ and $\cap$ in terms of $\leq$) we derive a certain set of theorems which are phrased strictly in terms of $\cup$ and $\cap$ (and thus do not contain $\leq$). These theorems can serve in their turn as axioms of a theory T_{lata} in the language $L_{lata} = \{\cup,\cap\}$. In this theory it is now possible to define $\leq$ (either in terms of just $\cup$ or in terms of just $\cap$). And these definitions are the reverse of the definitions of $\cup$ and $\cap$ in terms of $\leq$ in that adding them to the axioms of T'_{lata} yields a theory T'_{lato}:

(1) $\quad T'_{lata} = T'_{lato}$

Equation (1) captures the ultimate equivalence of the two directions from which lattice structure can be approached.

After having obtained this result we proceed to the theories of boolean lattices and boolean algebras. These theories - T_{bl} and T_{ba} (for 'boolean lattices' and boolean algebras', respectively) - are obtained by adding further axioms to T_{lato} and T_{lata}. It is easy to show that T_{lato} and T_{lata} stand in the same relation of definitional equivalence as T_{lato} and T_{lata}.

As implied by what was said in the introductory remarks to this section, it is convenient to o axiomatise the theory of lattice-like partial orderings using as primitive relation not the strict ordering relation <

but rather the corresponding weak ordering relation $\leq$.[4] In other words we start with the language $L_{lato} = \{\leq\}$. Let T_{lato} be the theory axiomatised by the following sentences of this language.

Des. 3 (Axioms for T_{lato})

$Ax_{lato}.1$ $(\forall x)(\forall y)(x \leq y \ \& \ y \leq x \leftrightarrow x = y)$
$Ax_{lato}.2$ $(\forall x)(\forall y)\forall z)(x \leq y \ \& \ y \leq z \rightarrow x \leq z)$
$Ax_{lato}.3$ $(\forall x)(\forall y)((\exists z)(x \leq z \ \& \ y \leq z \ \& \ (\forall u)(x \leq u \ \& \ y \leq u \rightarrow z \leq u))$
$Ax_{lato}.4$ $(\forall x)(\forall y)((\exists w)(w \leq x \ \& \ w \leq y \ \& \ (\forall u)(u \leq x \ \& \ u \leq y \rightarrow u \leq w))$

Note that $Ax_{lato}.1$ says that $\leq$ is both reflexive and antisymmetric. Thus $Ax_{lato}.1$ and the transitivity axiom $Ax_{lato}.2$ together state that $\leq$ is a partial ordering. $Ax_{lato}.3$ and $Ax_{lato}.4$ assert the existence of suprema and infima.

Our first task is to show that the suprema and infima whose existence is asserted by $Ax_{lato}.3$ and $Ax_{lato}.4$ are unique. We will argue the case for suprema; the case of infima is analogous.

We argue informally. (Here as elsewhere the argument could be turned without a formal derivation without any real difficulties, but such formal derivations tend to be lengthy and cumbersome and to obscure the idea of the argument.) Let x and y be any elements. Suppose that z and z' have the properties stated in (2) and (3)

(2) $(x \leq z \ \& \ y \leq z) \ \& \ (\forall u)((x \leq u \ \& \ y \leq u) \rightarrow z \leq u)$
(3) $(x \leq z' \ \& \ y \leq z') \ \& \ (\forall u)((x \leq u \ \& \ y \leq u) \rightarrow z' \leq u)$

Then we have, instantiating u to z' in (2),

(4) $(x \leq z' \ \& \ y \leq z') \rightarrow z \leq z'$

Since the antecedent of (3) is a conjunct of (2), we get $z \leq z'$ by MP. In the same way we get $z' \leq z$ by instantiating u to z in (2). From $Ax_{lato}.1$ we then get $z = z'$.

---

[4] As noted in the opening remarks to this Chapter the choice between $<$ and $\leq$ is strictly one of convenience. If we choose $<$ as primitive, then we can define $\leq$ in terms of it via $x \leq y \equiv_{df} x < y \vee x = y$; and if we choose $\leq$, then $<$ can be defined via $x < y \equiv_{df} x \leq y \ \& \ x \neq y$.

Exercise: Derive the sentence

$$(\forall x)(\forall y)(\forall z)(\forall z')((x \leq z \,\&\, y \leq z \,\&\, (\forall u)(x \leq u \,\&\, y \leq u \rightarrow z \leq u)) \,\&\, x \leq z' \,\&\, y \leq z' \,\&\, (\forall u)(x \leq u \,\&\, y \leq u \rightarrow z' \leq u))) \rightarrow z = z')$$

from T_{lato}. (The easiest way to do this is to construct a Semantic Tableau. Constructing a derivation in some system of Natural Deduction is also quite doable. An axiomatic derivation is (here as in most other cases) much harder.)

Given that T_{lato} entails the existence and uniqueness of suprema and infima, we can define the operations $\cup$ and $\cap$ in T_{lato} in terms of $\leq$ as in Def($\cup$,{$\leq$}) and Def($\cap$,{$\leq$})below. These definitions correctly determine the interpretations of $\cup$ and $\cap$ in any model of T_{lato}. Also, they can be added to T_{lato} without undesirable 'side effects', i.e. without adding new theorems that can be expressed in the language L_{lato} of T_{lato}.[5]

Def($\cup$,{$\leq$}) $(\forall x)(\forall y)(\forall z)(x \cup y = z \leftrightarrow (x \leq z \,\&\, y \leq z \,\&\, (\forall u)(x \leq u \,\&\, y \leq u \rightarrow z \leq u)))$

Def($\cap$,{$\leq$}) $(\forall x)(\forall y)(\forall z)(x \cap y = z \leftrightarrow (z \leq x \,\&\, z \leq y \,\&\, (\forall u)(u \leq x \,\&\, u \leq y \rightarrow u \leq z)))$

Let, as already indicated in the introduction to this section, T'_{lato} be the theory in the language $L_{lat} = \{\leq, \cup, \cap\}$ that is obtained by adding the definitions Def($\cup$,{$\leq$}) and Def($\cap$,{$\leq$})as new axioms to the axiom set {$Ax_{lato}.1$-$Ax_{lato}.4$} of T_{lato}. It is not hard to show that the following sentences are all theorems of T'_{lato}:

---

[5] If existence and/or uniqueness could not be proved fromT_{lato}, then adding the definitions would also add the non-derivable statement or statements of T_{lato} which expressing existence and uniqueness, respectively to the theory. The reason is that the left hand sides of the biconditionals in the definitions Def($\cup$,{$\leq$}) and Def($\cap$,{$\leq$}) (e.g. $x \cup y = z$ for the first of these) entail existence and uniqueness of z simply because that is part of the general logical properties of function constants. The fact that T_{lato} entails the existence and uniqueness conditions associated with the right hand sides of the biconditionals guarantees that addition of the two definitions is what is called a *conservative* extension of T_{lato}, i. e. an extension which has exactly the same theorems as T_{lato} in its original language L_{lato}. For more on conservativity and other properties of formal definitions see Section 2.3.

$Th_{lata}.1$ $(\forall x)\ x \cup x = x$

$Th_{lata}.2$ $(\forall x)x \cap x = x$

$Th_{lata}.3$ $(\forall x)(\forall y)\ x \cup y = y \cup x$

$Th_{lata}.4$ $(\forall x)(\forall y)\ x \cap y = y \cap x$

$Th_{lata}.5$ $(\forall x)(\forall y)(\forall z)\ (x \cup y) \cup z = x \cup (y \cup z)$

$Th_{lata}.6$ $(\forall x)(\forall y)(\forall z)\ (x \cap y) \cap z = x \cap (y \cap z)$

$Th_{lata}.7$ $(\forall x)(\forall y)\ (x \cup y) \cap x = x$

$Th_{lata}.8$ $(\forall x)(\forall y)\ (x \cap y) \cup x = x$

Exercise: Show that these are theorems of T'_{lato}.

The theorems $Th_{lat}.1$ - $Th_{lat}.8$ can now be used in their turn as axioms of a theory T_{lata} of the language $L_{lata} = \{\cup, \cap\}$. In this new capacity we refer to them as $Ax_{lata}.1$ - $Ax_{lata}.8$. T_{lata} allows us to define $\leq$ in terms of the non-logical constants $\cup$ and $\cap$ of its language L_{lata}. In fact, as adumbrated earlier, we need only one of $\cup$ and $\cap$ in such a definition. Two such definitions, one in terms of $\cup$ and one in terms of $\cap$, are given below as $Def(\leq,\{\cup\})$ and $Def(\leq,\{\cap\})$.

$Def(\leq,\{\cup\})$ $(\forall x)(\forall y)(x \leq y \leftrightarrow x \cup y = y)$

$Def(\leq,\{\cap\})$ $(\forall x)(\forall y)(x \leq y \leftrightarrow x \cap y = x)$

Adding either $Def(\leq,\{\cup\})$ or $Def(\leq,\{\cap\})$ as a new axiom to the set $\{Ax_{lata}.1,.., Ax_{lata}.8\}$ of axioms of T_{lata} yields an extension in the language L_{lat} from which the our original axioms $Ax_{lato}.1$ - $Ax_{lato}.4$ can be derived in their turn. For the sake of definiteness let us assume that the definition that is added is $Def(\leq,\{\cup\})$ and that the resulting extension of T_{lata} is the theory T'_{lata} of the language L_{lat}. As we noted in the introduction, it turns out that this theory is the very same theory as the theory T'_{lata} which we obtained by approaching the characterisation of lattice structure from the perspective of partial orderings. That is, we have the equality (1).

(1) $T'_{lata} = T'_{lato}$.

Exercise: Show the equality (1) is true. This requires showing -in addition to what has already been asked of the reader in earlier exercises from this section:

(i) $T'_{lato} \models Def(\leq,\{\cup\})$;

(ii) $T'_{lata} \models Ax_{lato}.i$ for i = 1,..,4;

(iii) $T'_{lata} \models Def(\cup,\{\leq\})$ and $Def(\cap,\{\leq\})$.

2. 1.3 Lattices based on sets and Boolean Lattices

Prominent among the models of T_{lato} are *power set inclusion structures*. These are models of the form $<P(X), \subseteq>$, where $P(X)$ is the power set of some set X and $\subseteq$ is the set inclusion relation (restricted to $P(X)$). Similarly a prominent subclass of the models of T_{lata} is that consisting of models of the form $<P(X), \cup, \cap>$, where $\cup$ and $\cap$ are the operations of set-theoretic union and intersection, again restricted to $P(X)$. What we have seen in general terms in the last section - viz. that $\cup$ and $\cap$ are definable in terms of $\leq$ and that $\leq$ is conversely definable in terms of $\cup$ or $\cap$ - is reflected by the well-known fact that set-theoretic union and intersection are definable in terms of $\subseteq$ and conversely. In fact, for any given X we can combine the structures $<P(X), \subseteq>$ and $<P(X), \cup, \cap>$ into a single structure $<P(X), \subseteq, \cup, \cap>$), which is a model of the theory which we have denoted either as T'_{lato} or as T'_{lata}.

But models of this kind are speial not only in that they are based on set-theoretic relations and operations. They are also special in that they all verify some additional conditions, which are expressible in the languages of our theories but are not derivable from those theories.

Among these conditions are in particular the so-called *distribution laws* for $\cup$ and $\cap$. Formulations of these laws are given in BA9 and BA10.

DISTR.1 $(\forall x)(\forall y)(\forall z)\ (x \cup y) \cap z = (x \cap z) \cup (y \cap z)$

DISTR.2 $(\forall x)(\forall y)(\forall z)\ (x \cap y) \cup z = (x \cup z) \cap (y \cup z)$

It follows from the results in the last section that DISTR.1 and DISTR.2 can be expressed in the language $\{\leq\}$. (In fact, one way to obtain such

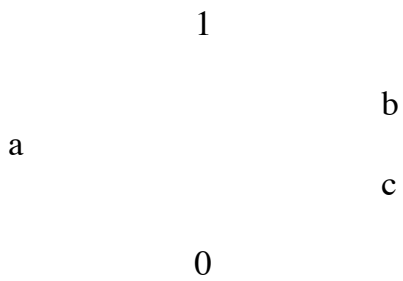
formulations is to translate DISTR.1 and DISTR.2 into formulas of L_{lat} using definitions Def($\cup$,{$\leq$}) and Def($\cap$,{$\leq$}) of $\cup$ and $\cap$ in terms of $\leq$.)

Exercise: Carry out this translation for DISTR.1 and prove that $Cl(T'_{lata} \cup \{DISTR.1\}) = Cl(T'_{lato} \cup \{DISTR'.1\})$, where DISTR'.1 is the translation of DISTR.1.

Lattices satisfying DISTR.1and DISTR.2 (or, what comes to the same thing, satisfying their translations into L_{lato}) are called *distributive* lattices. The following simple example shows that not all lattices are distributive. Let M be hte following model for the language L_{lato}:

> M = <{0,a,b,c,1}, $\leq$>, where $\leq$ is the following set of ordered pairs: {<0,0>,<0,a>,<0,b>,<0,c>,<0,1>,<a,a>,<a,1>, <c,c>,<c,b>, <c,1>,<b,b>,<b,1>,<1,1>}.

More perspicuously, M can be represented as the following directed graph[6]:



In this lattice we have: a $\cup$ c = a $\cup$ b = 1 and a $\cap$ c = a $\cap$ b = 0. So (a $\cap$ b) $\cup$ c = 0 $\cup$ c = c and (a $\cup$ c) $\cap$ (b $\cup$ c) = 1 $\cup$ b = b, falsifying DISTR.2.

Exercise: Show that the structure M described above also falsifies DISTR.1.

---

6 A directed graph G is a structure <U,R> where U is a set (the *nodes* of the graph G) and R is some binary relation on U. The pairs (a,b) of elements of U that belong to R are the *(directed) edges* of G. The edge (a,b) *goes from* a *to* b. Certain directed graphs, in which R is antisymmetric and either reflexive or irreflexive, can be used to represent partial orderings. When a graph G is used in this way, its node set represents the universe of the ordering,while the ordering relation itself is the transitive closure of R. Thus the ordering relation holds between two nodes a and b iff there is a path (i.e. a chain of edges) from a to b.

When a lattice is finite, it always has a smallest element - keep taking infima of pairs of elements - first taking the infimum c of two arbitrarily chosen elements a an b, then taking the infimum of c and some element d chosen arbitrarily from the elements not yet considered, and so on - until you have used up all elements of the lattice's finite universe - and a largest element (obtainable by taking suprema until the universe has been exhausted). Infinite lattices - i.e. infinite models of our theory T_{lato} - do not necessarily a smallest element (an element a such that for all other elements b in the lattice a $\leq$ b - or a largest element. (A counterexample is any unbounded linear order, such as, for instance, the orderings of the integers, the rationals or the reals.[7]) For the remainder of this section, however, we will focus on lattices which do have a smallest and a largest element.[8] We will refer to these as *the 0 of* the lattice and *the 1 of* the lattice, respectively. We will also use '0' and '1' as individual constants to denote these elements. We further limit our attention to distributive lattices. Thus - stated in terms of the language L_{lato} - we will be dealing with models of the theory $T_{d,0,1}$, whose axioms are, besides those of T_{lato}, translations into L_{lato} of the axioms DISTR.1 and DISTR.2 as well as the following two axioms, which assert the existence of a smallest and a largest element:

ExO $(\exists z)(\forall u)\ z \leq u$

Ex1 $(\exists z)(\forall u)\ u \leq z$

It is easy to see that $T_{d,0,1}$ entails that both the smallest and the largest element are uniquw. (This follows from ExO and Ex1, respectively, together with the fact that the models of T_{lato} are partial orderings.) This means that we can, for the same reason that this was possible earlier for $\cup$ and $\cap$, and following the same procedure, introduce individual constants 0 and 1 into the language L_{lato} by definitions obtained from the existence axioms ExO and Ex1. For the sake of explicitness the two definitions are given below.

Def(0,{$\leq$}) $(\forall z)(0 = z \leftrightarrow (\forall u)\ z \leq u)$

Def(1,{$\leq$}) $(\forall z)(1 = z \leftrightarrow (\forall u)\ u \leq z)$

---

7 Every linear order is a lattice. Exerice: prove that this is so.

8 The notion of a lattice is sometimes d*efined* as including the existence of a smallest and a largest element. This is not the practice we have adopted here.

Note that all set inclusion algebras are distributive lattices with a 0 and 1. On the other hand, as we already noted, linear orderings are distributive lattices, but they need not have a 0 or 1.

From here on it will be convenient to work in a language which contains all the constants considered so far - the 2-place predicate $\leq$, the two 2-place operations $\cup$ and $\cap$ and the individual constants 0 and 1. For the moment this is the language which contains just these five constants, i.e. $\{\leq,\cup,\cap,0,1\}$. Let $T'_{d,0,1}$ be the theory of this language whose axioms are:

(i) $Ax_{lato}.1$ - $Ax_{lato}.4$,
(ii) $Def(\cup,\{\leq\})$ and $Def(\cap,\{\leq\})$,
(iii) DISTR.1 and DISTR.2
(iv) $Def(0,\{\leq\})$ and $Def(1,\{\leq\})$

The theory $T'_{d,0,1}$ provides a suitable basis for the introduction of yet another operation, the 1-place operation of *complement*. To pave the way for the introduction of this operation we proceed once as we did before in the case of $\cup$, $\cap$, 0 and 1, viz. by first adopting a new axiom which asserts the existence of suitable values for the operation, then proving that these values are unique, and then, on the basis of this result introducing hte operation by means of a definition that is derived directly from the existence axiom.

Our existence axiom, COMP, asserts that for every element x there is an element y such that (a) the supremum of x and y is the 1 of the lattice and (b) the infimum of x and y is the 0 of the lattice:

COMP $\quad (\forall x)(\exists y)(x \cup y = 1 \ \& \ x \cap y = 0)$

From the combination of $T'_{d,0,1}$ and COMP it is possible to prove that the element y mentioned in COMP is uniquely determined in relation to x. We argue as follows. First we observe that the sentences (i) and (ii) are theorems of $T'_{d,0,1}$. (The proof of this is left to the reader.)

(i) $\quad (\forall u)\ u \cap 1 = u$

(ii) $\quad (\forall u)(u \cup 0 = u$

Assume that y_1 and y_2 both satisfy the matrix (= the quantifier-free part) of COMP for some given x, i.e. that

(a) $x \cup y_1 = 1$ (b) $x \cup y_2 = 1$
(c) $x \cap y_1 = 0$ (d) $x \cap y_2 = 0$

Then, since $x \cup y_1 = 1$, $(x \cup y_1) \cap y_2 = 1 \cap y_2 = y_2$, by (i). By DISTR.1 $(x \cup y_1) \cap y_2 = (x \cap y_2) \cup (y_1 \cap y_2)$ and $(x \cap y_2) \cup (y_1 \cap y_2) = 0 \cup (y_1 \cap y_2) = y_1 \cap y_2$, by (ii) and assumption (d). So $y_1 \cap y_2 = y_2$. Similarly we show that $y_2 \cap y_1 = y_1$. So $y_1 = y_2 \cap y_1 = y_1 \cap y_2 = y_2$.

The definitions Def($\cup$,{$\leq$}) and Def($\cap$,{$\leq$}) enable us to translate the axioms DISTR.1, DISTR.2 and COMP into sentences DISTR.1($\leq$), DISTR.2($\leq$) and COMP($\leq$) of the language {$\leq$}. Consider the theory T_{bl} that we obtain when these translations to the theory T_{lato}. (The subscript 'bl' stands for 'boolean lattice'.) The models of T_{bl} are called *boolean lattices*. In view of the existence and uniqueness of complements in such models we can, in the same way in which we extended the theory of lattice orderings with definitions for the supremum and infimum functions and those for the '0-place functions' 1 and 0, now add a definition of the complement function. We denote this function as '-'. (That is, -x is the complement of x.)

The definition Def(-,{$\cup$,$\cap$}) of - can, as we already said, be directly obtained form the corresponding existence axiom COMP.

Def(-,{$\cup$,$\cap$}) $(\forall x)(\forall y)(y = -x \leftrightarrow (x \cup y = 1 \ \& \ x \cap y = 0))$

'-' and Def(-,{$\cup$,$\cap$}) are our final additions to language and theory, respectively. Let L_{bla} be the language {$\leq$,$\cup$,$\cap$,0,1,-} and let T_{bla} be the extension of $T'_{d,0,1}$ with COMP and Def(-,{$\cup$,$\cap$}). The models of T_{bla} are on the one hand, because of the properties of their partial ordering relation, boolean lattices, while on the other hand they have, because of the properties of their operations $\cup$,$\cap$,0,1 and -, the structure of *boolean algebras*.

To amplify this last statement: We have seen that the theory of lattices can be formulated in terms of the operations $\cup$ and $\cap$. (This was the theory T_{bla}.) We can extend this theory with existence axioms and

definitions for 0,1 and - all couched in terms of $\cup$ and $\cap$. It is not hard to show that the theory that we obtain this way, and which belongs to the language $\{\cup,\cap,0,1,-\}$ is identical with the restriction of T_{bla} to the sentences of this language. This theory is known as the theory of boolean algebras and its models as *boolean algebras*. So as to fit in with this terminology we refer to the language $\{\cup,\cap,0,1,-\}$ as L_{ba} and to the theory of this language which we have just described as T_{ba}.

For further reference we list once more the set of axioms for T_{ba} which has emerged in the course of this discussion. In this list we have combined the existence axioms which guarantee the legitimacy of the corresponding definitions we used to introduce the new operation symbols into single axioms, in which the operation symbols take the place of the existentially quantified variables in the existence axioms.

Def. 4 (Axioms for the theory T_{ba} of boolean algebras.[9])

$Ax_{ba}.1$ $(\forall x)\ x \cup x = x$

$Ax_{ba}.2$ $(\forall x) x \cap x = x$

$Ax_{ba}.3$ $(\forall x)(\forall y)\ x \cup y = y \cup x$

$Ax_{ba}.4$ $(\forall x)(\forall y)\ x \cap y = y \cap x$

$Ax_{ba}.5$ $(\forall x)(\forall y)(\forall z)\ (x \cup y) \cup z = x \cup (y \cup z)$

$Ax_{ba}.6$ $(\forall x)(\forall y)(\forall z)\ (x \cap y) \cap z = x \cap (y \cap z)$

$Ax_{ba}.7$ $(\forall x)(\forall y)\ (x \cup y) \cap x = x$

$Ax_{ba}.8$ $(\forall x)(\forall y)\ (x \cap y) \cup x = x$

$Ax_{ba}.9$ $(\forall x)(\forall y)(\forall z)\ (x \cup y) \cap z = (x \cap z) \cup (y \cap z)$

$Ax_{ba}.10$ $(\forall x)(\forall y)(\forall z)\ (x \cap y) \cup z = (x \cup z) \cap (y \cup z)$

$Ax_{ba}.11$ $(\forall u)\ u \cap 1 = u$

$Ax_{ba}.12$ $(\forall u)(u \cup 0 = u$

$Ax_{ba}.13$ $(\forall x)(x \cup -x = 1 \ \&\ x \cap -x = 0))$

---

[9] Here $Ax_{ba}.1$ - $Ax_{ba}.8$ are the theorems $Th_{lat}.1$ - $Th_{lat}.8$ of Section 2.1.2; $Ax_{ba}.9$ and $Ax_{ba}.10$ are the earlier DISTR.1 and DISTR.2; $Ax_{ba}.11$ and $Ax_{ba}.12$ - mentioned earlier as (i) and (ii) in the proof of uniqueness of complements, are the results of combining the existence axioms Ex0 and Ex1 for the lattice One and the latice Zero with the definitions of the individual constants 0 and 1 in terms of $\cup$ and $\cap$ - these definitions we did not actually give, but they can be obtained from the defintions Def(0,{$\leq$}) and Def(1,{$\leq$}) we did give by translating them into sentences of L_{ba} using the definition of $\leq$ in terms of $\cup$; $Ax_{ba}.13$ results form combining the axiom COMP with Definition Def(-,{$\cup,\cap$}).

This concludes our general account of lattices, lattice algebras, boolean lattices and boolean algebras. The route we have followed, with all the switching back and forth between partial orderings and operations, may appear rather round-about and hard to follow, certainly on a first reading. But I believe that this is a price worth paying. The central methodological point of the last two sections has been to show, by means of the example that lattices and the corresponding algebras provide, how two at first sight very different perspectives on structure - here that of structure in the form of partial order and structure in the form of a number of connected operations - can nevertheless prove to be concerned with what is essentially the same structure after all. In order to bring out how and why this convergence arises in the case in question, switching between the two perspectives was essential. That does of course require a greater effort, both on the part of the presenter and that of the reader, than a simple presentation of lattices *just* as ordered structures or of lattice algebras and boolean algebras *just* in terms of their operations.

There is also a practical spin-off to the presentation of lattices as being describable either as partial orders or as algebras: Now that we have explored the nature of this correspondence thoroughly, we can, with the benefits of that investigation, join the wide-spread practice of switching betgween the two perspectives in discussions of such structures if and when this proves convenient. We will make use of this freedom in particualr in the next sections.

In the next two sections we focus exclusively on boolean algebras. 2.1.4 presents a number of distinct types of boolean algebras and defines certain properties in terms of which they can be distinguished from each other and classified. 2.1.5 is devoted to the Stone Cech Theorem, according to which every boolean algebra is isomorphic to a structure in which the operations $\cup$, $\cap$ and - are set-theoretic union, intersection and subtraction, respectively.

2.1.4 Some Examples of Boolean Algebras.

As compared with lattices in general, boolean lattices form a quite special category. But even so there is much variety even within this special domain. One important subtype is that identified by the power set inclusion lattices $<P(X),\subseteq>$ that were mentioned earlier. These are distinguished by two properties: they are (i) *atomic* and (ii) *complete*.

Before we define these two properties, first, in Prop. 2, an obvious observation about the power set lattices, viz that they are determined up to isomorphism by their carrier sets X:

Prop. 2 If $|X| = |X'|$, then $<P(X),\subseteq> \cong <P(X'),\subseteq>$.

Proof: It suffices to note that a bijection between X and X' induces a bijection between $P(X)$ and $P(X')$ and carries the inclusion relation restricted to $P(X)$ into the inclusion relation restricted to $P(X')$.

Next the definitions of atomicity and completeness. The first of these presupposes the notion of an element being an *atom*, which is important in its own right.

Def. 5 (a) Let $BL = <U,\leq>$ be a boolean lattice, b an element of BL. b is an *atom* of BA iff

(i) $a \neq 0$ and

(ii) there is no c in BL such that $0 < c < a$ (where $<$ is the strict partial order corresponding to the lattice ordering $\leq$.

(b) BL is *atomic* iff for every b in BL there is an atom a of BA such that $a \leq b$.

Def. 6 A boolean lattice $BL = <U,\leq>$ is *complete* iff for every subset V of U there is a least element c in U such that for all $v \in V$, $v \leq c$. More formally:

For each $V \subseteq U$ there is a c in U such that

(i) $(\forall v \in V)\ v \leq c$, and

(ii) $(\forall c')((\forall v \in V \rightarrow v \leq c') \rightarrow c \leq c')$

To show that power set inclusion lattices are atomic and complete is once again very easy to show and we record the fact as another proposition.

Prop. 3 Every power set inclusion lattice is atomic and complete.

Proof: Let PSIL = $<P(X),\subseteq>$ be any 0power set inclusion lattice. Note that the 0 of PSIL is the empty set $\varnothing$. So the atoms of PSIL are the singleton sets {x}, where x is any element of X. Suppose that Y is any element of PSIL distinct from 0. Then Y is a subset of X and $Y \neq \varnothing$. So Y contains at least one element $x \in X$. But then we have $\{x\} \subseteq Y$, i.e. the lattice ordering relation holds between the atom {x} and Y. Thus PSIL is atomic.

To see that PSIL is complete, let V be any subset of $P(X)$. Then $\cup V$ is a subset of X and thus a member of $P(X)$. It is easy to verify (i) that for all $V \in V$, $V \subseteq \cup V$ and (ii) if W is any other element of $P(X)$ such that for all $V \in V$, $V \subseteq W$, then $\cup V \subseteq W$. Thus PSIL is complete.

But not all boolean lattices are either atomic or complete, In fact, there are boolean lattices that are the extreme opposite of atomic in that they have no atoms at all. And there are also boolean lattices that are the extreme opposite of complete in that they have the following property:

> Every set V of elements is either *essentially finite* or else V does not have a supremum.

Here by *essentially finite* we mean the following: V is *essentially finite* iff there is a finite subset W of V such that $(\forall v \in V)(\exists w \in W)\ v \leq w$. (Note that in this case the supremum of W, which must exist since W is finite, is also the supremum of V.)

But besides boolean lattices which occupy the opposite end of the spectrum from the power set inclusion lattices with regard to either atomicity or completeness or both, there are also many which display less extreme forms of non-atomicity or incompleteness. For instance there are boolean lattices which do contain some atoms but which nevertheless do not have enough of them to make them atomic.

Our first example of a boolean lattice that is not like the power set inclusion lattices, BL1, differs from them in being not complete,

although it shares with them the property of being atomic. The example also illustrates another important fact, the true significance of which will become clear when we turn to the Stone Cech Representation Theorem in the next section. This is because it is a boolean lattice whose ordering relation is, just like it is for the power set inclusion lattices, set-theoretic inclusion. The only, but crucial difference with the power set inclusion lattices is that in our example the universe is no longer a full power set $P(X)$, but rather some proper subset of such a power set. (The Stone Cech Theorem says that just by varying the universes of inclusion lattices all possible properties of boolean lattices can be exemplified.)

In the case of BL1 the universe is defined as the set of all fintie and all cofinite subsets of the set N of natural numbers. Here a cofinite subset of N is a subset Y of N such that N \ Y is finite. In other words, if U is the set of all finite and cofinite subsets of N, then BL1= $\langle U,\subseteq\rangle$, where is the relation of set-theoretic inclusion restricted to U.

Before we show that BA1 has the mentioned properties, i.e. that it is atomic but not complete, we first have to show that it is a boolean lattice- in other words, that it is a lattice and that is it boolean. To this end we make use of the possibility of switching back and forth between lattices and the corresponding algebras. To start, note that the restriction of $\subseteq$ to any set of sets will always be a partial order. To show that in the case at hand this order is a lattice we note that U is closed under the set-theoretic operations $\cup, \cap$. To see that the union X $\cup$ Y of two subsets X and Y of U belongs to U, we have to distinguish between two cases: (i) if X, Y are both finite, then X $\cup$ Y is finite and thus in U; (ii) if at least one of X, Y is cofinite, then X $\cup$ Y is cofinite and thus also in U. In the same way one shows that U is closed under $\cap$. From the fact that U is closed under $\cup$ and $\cap$ it follows that $\langle U,\subseteq\rangle$ is a lattice. For if, say, the union of the sets X and Y from U is again a member of U, then it will be the supremum of X and Y in $\langle U,\subseteq\rangle$; likewise, since the intersection of subsets X and Y of U again belongs to U it must be the infimum of X and Y. Thus $\langle U,\subseteq\rangle$ is a lattice.

Furthermore, $\varnothing$ and N both belong to $\langle U,\subseteq\rangle$, since is a finite and N a cofinite subset of N. But then it is obvious that these are the smallest and largest element, respectively, of $\langle U,\subseteq\rangle$. So $\langle U,\subseteq\rangle$ has a 0 and a 1. We also note that set-theoretic union and intersection satisfy the distributivity laws DISRT1 and DISTR2. So $\langle U,\subseteq\rangle$ is a distributive lattice

with a 0 and a 1. To see that <U,⊆> is boolean, we note that U is closed under the operation of complementation relative to N (that is, the operation of subtracting a given X from N, denoted as N \ X). For the relative complement of a finite subset of N is a cofinite subset and vice versa. Using the same reasoning as above, we conclude that the relative complement is the operation we obtain when we apply Def(-,{∪,∩}) (see section 1.2.3) to the supremum and infimum operations of <U,⊆>, which, as we have already shown, are nothing but the operations of set-theoretic union and intersection. Moreover the relative complement operation of set theory does satisfy, in conjunction with union and intersection, the laws $Ax_{ba}.11$ and $Ax_{ba}.12$. So <U,⊆> is a boolean lattice.

We next show that BA1 is atomic. This is easy. All singleton sets {n}, where the n ε N, are finite and thus belong to U. Clearly they are the atoms of <U,⊆>. And if X is a member of U that is different from the 0 of U, i.e. X ≠ ∅, then there must be some n such that n ε X and thus {n} ⊆ X; so there is an atom between 0 and X.

Finally BA1 is not complete. For let A be a subset of N such that both A and N\A are infinite. (For instance, we could take for A the set of even numbers.) Let *A* be the set {{n}: n ε A}. Then *A* has no supremum in <U,⊆>. For if Y is any element in U with the property that (∀Z)(Z e *A* → Z ⊆ Y), then A ⊆ Y. The only elements of U with this property are the cofinite subsets of N which include A. But among these there is no smallest element: Take any such Y. Then Y\A is non-empty (in fact it is infinite). Let m ε Y\A and Y' = Y\{m}. Then Y' ε U, A ⊆ Y' and Y' is a proper subset of Y. So there is no smallest member of U which includes all members of *A*.

Our second example, BA2, is presented as a boolean algebra. And it is not a set algebra. Once again the set N of natural numbers is our starting point. But this time we begin by defining an equivalence relation on the subset of N:

$$X \equiv Y \text{ iff}_{\text{def.}} X - Y \text{ is finite.}$$

Here "X - Y" denotes the *symmetric difference* between X and Y, i.e. X - Y = (X\Y) ∪ (Y\X).

The first thing to observe is that $\equiv$ is a *congruence relation* with respect to the set-theoretic operations $\cup, \cap$ and $\setminus$. That is, if the arguments of the operations stand to each other in the relation $\equiv$, then so are the results of those operations. For instance, suppose that $X \equiv X'$ and that $Y \equiv Y'$. Then also $(X \cup Y) \equiv (X' \cup Y')$. That this must be so is not hard to see. On the one hand $(X \cup Y) \setminus (X' \cup Y') \subseteq (X \setminus X') \cup (Y \setminus Y')$. This entails that if the term on the right of $\subseteq$ is finite, so is the one on the left. Analogously $(X' \cup Y') \setminus (X \cup Y)$ is finite. So $(X \cup Y) \setminus (X' \cup Y')$ is finite. It follows that $(X \cup Y) \equiv (X' \cup Y')$. Likewise for the other two operations.

Let V be the set $\{[X]_{\equiv}: X \subseteq N\}$. (N.B. during the remainder of our discussion of BA2 we will leave out the subscript $\equiv$.) The congruence of $\equiv$ w.r.t. $\cup, \cap$ and $\setminus$ entails that we can define the following operations on V:

Def. For arbitrary $X, Y \subseteq N$,

(i) $[X] \underline{\cup} [Y] = [X \cup Y]$
(ii) $[X] \underline{\cap} [Y] = [X \cap Y]$
(iii) $\underline{\neg} [X] = [N \setminus X]$

Now let BA2 be the structure $\langle V, \underline{\cup}, \underline{\cap}, \underline{\neg}, [\varnothing], [N] \rangle$. That this is a Boolean algebra follows straightforwardly from the Boolean nature of the set-theoretical operations $\cup, \cap$ and $\setminus$, in terms of which we have defined the operations $\underline{\cup}, \underline{\cap}, \underline{\neg}$. Note that the lattice ordering $\underline{\leq}$ of this structure holds between any two members [X] and [Y] of V iff $X \setminus Y$ is finite. To see this, recall that $\underline{\leq}$ can be defined in terms of $\underline{\cup}$ by: $[X] \underline{\leq} [Y]$ iff $[X] \underline{\cup} [Y] = [Y]$. This entails that $(X \cup Y) \setminus Y$ is finite. But $(X \cup Y) \setminus Y$ is the same set as $X \setminus Y$.

We first observe that BA2 is atomless, and thus not atomic. Suppose that $[X] \neq [\varnothing]$. Then X is infinite. But then we can split X into two infinite subsets Y and $X \setminus Y$ But in that case we have $\varnothing < [Y] < [X]$, where $<$ is the strict order in relation corresponding to the lattice ordering $\leq$ of BA2. So [X] is not an atom.

BA2 is also not complete. To see this, let A be a denumerably infinite set of infinite mutually disjoint subsets of N whose union is N. (That is, if $X \in A$, then X is infinite, if $X, Y \in A$ and $X \neq Y$, then $X \cap Y = \emptyset$ and $\cup A$ = N.) Then there is no element in V which is the supremum of A. For suppose that [Z] were the supremum of A. Then for each $X \in A$, $[X] \leqq [Z]$. So, by the remark at the end of the penultimate paragraph $X \setminus Z$ is finite. Consequently, since X infinite and $X = (X \setminus Z) \cup (X \cap Z)$, $Z \cap X$ must be infinite and thus $\neq \emptyset$. So we can for each X in A pick an element n_X from $X \cap Z$. Note that if $Y \in A$ and $Y \neq X$, then by assumption Y is disjoint from X and therefore n_X is not an element of Y. So each n_X belongs to exactly one element of A. That is, if $B = \{n_X : X \in A\}$, then for each $X \in A$, $X \cap B = \{n_X\}$. Now let $Z' = Z \setminus B$. Since $B \subseteq Z$, $Z \setminus Z' = B$ and thus $Z \setminus Z'$ is infinite. So $[Z'] \leq [Z]$. On the other hand, for any $X \in A$, $Z' \setminus X = (Z \setminus X) \cup \{n_X\}$, and this set is finite, since $Z \setminus X$ is finite. So, by the remark at the end of the one-but-last paragraph, $[X] \leqq [Z']$. It follows that [Z] is not the supremum of A.

There are also boolean lattices that are complete but not atomic. [An example of such a lattice can be found in the exercises.]

2.1.5 The Stone-Cech representation Theorem

One of the most famous and most fundamental results in the theory of boolean algebras is the *Stone-Cech Representation Theorem,* which says that every boolean algebra is isomorphic to (and thus 'can be represented as') a set algebra; that is, it is isomorphic to a structure $\langle U, \cup, \cap, -, 0, 1 \rangle$ in which the elements of U are subsets of some set X, the operations $\cup, \cap, -$, are set-theoretic union, intersection and complementation relative to X, 0 is the empty set and 1 the set X. (Once more, note well that U will in general not consist of all subsets of X.)

The proof of the Stone Cech Theorem involves the notion of an *ideal* of a boolean lattice, or, alternatively, that of a *filter*. So we begin by defining these notions as well as a few others connected with them.

Def. 7. Let BL = <U,≤> be a boolean lattice.

1. A subset V of U is an *ideal of* BL iff (i) V ≠ ∅; (ii) V ≠ U; (iii) if b ε V and a ≤ b, then a ε V; and (iv) if a, b ε V, then a ∪ b ε V.

2. A subset V of U is a *filter of* BL iff (i) V ≠ U; (ii) V ≠ ∅; (iii) if b ε V and b ≤ a, then a ε V; and (iii) if a, b ε V, then a ∩ b ε V.

3. Let b ε BL, b ≠ 1. The *ideal* of BL *generated by* b is the set {a ε U: a ≤ b}
An ideal is called a *principal* ideal if it is generated by some b ε U such that b ≠ 1.

   Likewise, if b ≠ 0, the *filter* of BL *generated by* b is the set {a ε U: b ≤ a}; and a filter is called a *principal* filter if it is generated by some b ε U such that b ≠ 0.

4. An ideal V of BL is called a *prime* or *maximal* ideal of BL iff for each b ε U either b ε V or else -b ε V.

   Likewise, a filter V of BL is called a *prime* or *maximal* filter of BA iff for each b ε U either b ε V or else -b ε V.

Prop. 4. 1. If V is an ideal of a boolean lattice BL = <U,≤>, then -V = {-b: b ε V} is a filter of BL, and conversely.

2. If V is a principal ideal {a ε U: a ≤ b} of BL, then -V is the principal filter {a ε U: -b ≤ a} of BL, and conversely.

3. If V is a prime ideal of BL, then -V is a prime filter of BL, and conversely.

Lemma. 1. (Boolean Prime Ideal Theorem for Boolean Lattices)

Let V be an ideal of some BL <U,≤>. Then there exists a prime ideal V' of BL such that V ⊆ V'.

A general proof of the Prime Ideal Theorem, which applies to lattices of arbitrary cardinality, is not possible at this stage, since it requires set-theoretic assumptions and methods that are not available to us as yet. We can only prove the theorem for BA's which are at most infinitely denumerable. For this case the argument goes as follows.

If $<U,\leq>$ is denumerable, then we can assume an enumeration $u_1, u_2, \ldots$ of all elements of U and extend V stepwise, first with u_1 or $-u_1$, then with u_2 or $-u_2$, and so on, obtaining in the limit an extension of V which is a prime ideal. We just sketch the first step, in which V is extended with either u_1 or $-u_1$. (The other steps are completely analogous.)

With regard to V and u_1 we distinguish two cases:

(a) For all finite $W \subseteq V$, $\sup(W) \cup u_1 \neq 1$.
(b) For some finite $W \subseteq V$, $\sup(W) \cup u_1 = 1$.

In case (a) $V_1 = \{u \in U: (\exists W)(W \subseteq V \ \& \ W \text{ finite} \ \& \ u \leq \sup(W) \cup \{u_1\}\}$;
in case (b) $V_1 = \{u \in U: (\exists W)(W \subseteq V \ \& \ W \text{ finite} \ \& \ u \leq \sup(W) \cup \{-u_1\}\}$

We begin by showing that in case (a) V_1 is an ideal. First, note that $u_1 \in V_1$. This is so since the empty set $\varnothing$ is a subset of V and $u_1 \leq 0 \cup u_1 = \sup(\varnothing) \cup u_1$. Second, suppose $b \in V_1$ and $a \leq b$. Then there is a finite $W \subseteq V$ such that $b \leq \sup(W) \cup u_1$. But then also $a \leq \sup(W) \cup u_1$, so $a \in V_1$. Third, suppose that $V_1 = U$. Then $1 \in V_1$. This means that there is a finite $W \subseteq V$ such that $1 \leq \sup(W) \cup u_1$, which is equivalent to: $\sup(W) \cup u_1 = 1$. This contradicts the assumption of case (a) and we conclude that $V_1 \neq U$. Lastly, let $a, b \in V_1$. Then there are finite subsets W_a, W_b of V such that $a \leq \sup(W_a) \cup u_1$ and $b \leq \sup(W_b) \cup u_1$. If W_a and W_b are both finite subsets of V, then so is $W_a \cup W_b$. Also, $\sup(W_a) \leq \sup(W_a \cup W_b)$, so $a \leq \sup(W_a \cup W_b) \cup u_1$. Similarly, $b \leq \sup(W_a \cup W_b) \cup u_1$. So $a \cup b \leq \sup(W_a \cup W_b) \cup u_1$. Our final observation is that $V \subseteq V_1$. Suppose that $u \in V$. Then $u \leq \sup(\{u\}) \cup u_1$, with $\{u\}$ a finite subset of V. So $u \in V_1$.

From all this we conclude for case (a): V_1 is an ideal which extends V and contains one of u_1 and $-u_1$.

Now consider case (b). We show:

(*) for all finite $W' \subseteq V$, $\sup(W') \cup -u_1 \neq 1$.

Suppose this is not so. Then there is a finite $W' \subseteq V$ such that $\sup(W') \cup -u_1 = 1$. By assumption of case (b) there also is a finite $W \subseteq V$ such that $\sup(W) \cup u_1 = 1$. Let $W'' = W \cup W'$. Then W'' is a finite subset of V. Let $w = \sup(W'')$. Then $w \in V$, so, since V is an ideal, $w \neq 1$ (for otherwise we would have that $V = U$). Furthermore, $\sup(W) \leq w$. So, since $\sup(W) \cup u_1 = 1$, $w \cup u_1 = 1$. Similarly $w \cup -u_1 = 1$. So $(w \cup u_1) \cap (w \cup -u_1) = 1 \cap 1 = 1$. But $(w \cup u_1) \cap (w \cup -u_1) = ((w \cup u_1) \cap w) \cup ((w \cup u_1) \cap -u_1) = w \cup ((w \cap -u_1) \cup (u_1 \cap -u_1)) = w \cup ((w \cap -u_1) \cup 0) = w \cup (w \cap -u_1) = w$. So $w = 1$, contrary to what we established above. This proves (*).

We can now show as in case (a) that V_1 is an ideal which extends V and contains $-u_1$. So it follows in either case that V_1 is an ideal which extends V and contains one of u_1 and $-u_1$.

In this way we coinstruct a denumerable sequence $V_1, V_2, \ldots$ of ideals extending V such that for each n V_n will contain, for $i = 1, \ldots, n$, one of u_i and $-u_i$.

Now let $V' = \bigcup_n V_n$. Then it is easy to show that V' is an ideal. (In particular V' does not cootain 1. For is it did than 1 would be an element of some V_n, contradicting the already established fact that V_n is an ideal. From the construction of V' it is also clear that V' is maximal.

q.e.d.

Corollary. Let u be an element of some BL $\langle U, \leq \rangle$ such that $u \neq 1$. Then there exists a prime ideal V' of BL such that $u \in V'$.

Proof. Suppose that u is as described in the statement. Then $V_u = \{v \in U: v \leq u\}$ is an ideal. (Show this. N. B. ideals of this form, which consist of all elements $\leq$ some given element, are called *principal ideals*.) So, accroding to Lemma 3 there is a prime ideal V such that $V_u \subseteq V$. Clearly u e V.

q.e.d.

We now turn to the Stone-Cech Theorem itself.

Theorem. 3 (Stone-Cech Theorem for Boolean Lattices)

Let $M = \langle U, \leq \rangle$ be any boolean lattice. Then there is a set inclusion lattice M* - i.e. a structure $\langle U^*, \subseteq \rangle$ in which U* is a subset of some power set $P(X)$ and $\subseteq$ is the set-theoretic inclusion relation on U* - such that $M \cong M^*$.

Proof. For any $u \in U$ let u* be the set consisting of all maximal ideals V of M such that not $u \in V$: $u^* = \{V: V \text{ is a prime ideal of M and not } u \in V\}$. Let $U^* = \{u^*: u \in U\}$. Then $U^* \subseteq P(P(U))$; so $U^* \subseteq P(X)$ for some X. We show that * is 1-1 map from U onto U*. That * is onto follows from the definition of U*. To show that * is 1-1 we argue as follows. First suppose that $u, u' \in U$ and that $u \neq u'$. Then either not $u \leq u'$ or not $u' \leq u$. Assume that not $u \leq u'$. (The other case is analogous.) Then $u' \cup -u \neq 1$. For if $u' \cup -u = 1$, then $u \cap u' = u' \cap u = (u' \cap u) \cup 0 = (u' \cap u) \cup (-u \cap u) = (u' \cup -u) \cap u = 1 \cap u = u$, so $u \leq u'$, contrary to assumption. Since $u' \cup -u \neq 1$, there is according to the Corollary to Lemma 3 a maximal ideal V containing $u' \cup -u$. Since $-u \in V$ it is not the case that $u \in V'$, For otherwise $u, -u \in V'$, so $u \cup -u \in V'$ and V' wouldn't be an ideal. So by teh definition of *, $V' \in u^*$. On the other hand, $u' \in V'$. So it is not the case that $V' \in u'^*$. So $u^* \neq u'^*$.

Next we prove that $u \leq u'$ iff $u^* \subseteq u'^*$. First assume $u \leq u'$. Then, as can easily be shown, $-u' \leq -u$. Let V be any maximal ideal in u*. Then, since not $u \in V$, $-u \in V$. So, since $-u' \leq -u$, $-u' \in V$. So it is not the case that $u \in V$, and therefore $V \in u'^*$. So $u^* \subseteq u'^*$. Conversely assume that $u^* \subseteq u'^*$. Suppose it is not the case that $u \leq u'$. Then, as we saw above, there is a maximal ideal V' such that $u' \in V'$ but not $u \in V'$. So $V' \in u^*$, but not $V' \in u'^*$, contrary to the assumption that $u^* \subseteq u'^*$.

q.e.d.

The Stone-Cech Representation Theorem for Boolean Algebras is a paradigm for a type of result that has proved of great value in mathematics and logic in a number of distinct contexts. Results of this type are generally called 'representation theorems'. Informally speaking, a *representation theorem for* a theory T of a language L is a statement to the effect that a certain class *M'* of models for some

extension L' of L is representative for T's models. Putting the matter more formally, representation theorems take the following general form:

> Let T be a theory in some language L. Let *M* be the class of all models of T. Let L' be an extension of L and let *M'* be a class of models for L' such that for each M ε *M'* the reduction M⌈L of M to L is a member of *M*. Then *M' is representative of the models of* T iff for each M ε *M* there is an M' ε *M'* such that M ≅ M'⌈L.

The use and importance of representation theorems is in most cases that they provide a clearer view of the range of variation among the models of a given theory T and/or a way of studying this variation. In order to obtain a picture of the different (isomorphism) types of models of T it is enough to study the variation within the representing class *M'*. And in many cases this latter investigation is helped by the fact that the models within this class are of a special kind, e.g. in that they have additional properties which do not apply to models of T in general. (Normally this is because these properties are not expressible within the language L of T, but only in the extended language L' of the models in *M'*.

The Stone Representation Theorem for boolean lattices is a good example of this: There are ways to explore the possible structure of set inclusion lattices which are not directly available for arbitrary boolean lattices. On the other hand, however, the very fact that the Stone Theorem is true is an indication of how much variation can be found among set inclusion lattices. To take just one example, our algebra BAII was not a set inclusion lattice as we defined it. Stone's Theorem tells us that there is a set inclusion lattice isomorphic to BAII, and also gives us a method for how to construct such a lattice. But the resulting lattice is not a set inclusion structure that one would easily have thought of off the bat. Should one have expected that a set inclusion lattice with this structure actually exists? That would of course depend on our general knowledge of set theory, but at the very least the answer is not obviously 'yes'.

In fact, one way to look at the Stone-Cech Theorem is as a statement telling us how much variation can be obtained by starting from the narrowly circumscribed notion of a power set lattice $<P(X), \subseteq>$ - recall: any such lattice is atomic and complete and it is fully determined by the cardinality of the carrier set X - and then to broaden this notion by allowing for variation in just one respect: the universe U

need not be all of $P(X)$, but may also be some proper subset of it. All variety, in other words, can be located in the choice of U.

2.1.6 Booelan Algebra and Logic.

We noted at the outset of this Chapter that boolean algebras are of particular importance for logic; some the most prominent structures that are studied in formal logic have the properties of such algebras.[10] The simplest (and arguably most central) example is the 'algebra of propositions', in which the disjunction p v q of two propositions p and q interpreted as the supremum of p and q, their conjunction p & q as their infimum and the negation -p of p as its complement. Exactly what boolean algebra this will give us depends on how we decide to characterise propositions. When we identify 'propositions' with the Fregean denotations of sentences - 'the True', or '1', and 'the False', or '0' - then we get a boolean algebra whose universe is the two-element truth value space {0,1}, in which the boolean operations are as follows:

(i) $1 \cup 1 = 0 \cup 1 = 1 \cup 0 = 1, 0 \cup 0 = 0;$
(ii) $1 \cap 0 = 0 \cap 1 = 0 \cap 0 = 0, 1 \cap 1 = 1;$
(iii) $-1 = 0, - 0 = 1.$

Note that this algebra results as the image of any language L of propositional logic under any classical valuation Suppose that V is a classical valuation of the set of propositional letters of L (classical in the sense that it assigns each letter one of the classical truth values 0 and 1). Then V will map each formula of L into {0,1} according to the familiar truth table rules:

---

10 Boolean algebras and lattices owe their name to one of the founders of modern logic, the 19-th century mathematician George Boole (1815-1864). Boole was together with his compatriot Augustus de Morgan, the first to look at logic from an algebraic perspective, according ot which the logical connectives &, v, ¬, etc. are seen as operators, or functors, which can be used to obtain propositions out of other propositions (e.g. the conjunction 'A & B' from the propoisitions A and B). Boole tried to formulate the laws of logic (his 'Laws of Thought') in algebraic terms, i.e. as equations that express logical equivalences that hold between propositions in virtue of their logical structure, such as e.g.

(i) A & B = B & A
(ii) (A & B) & C = A & (B & C)

to express the commutativity and associativity of conhunction. Eventually such equations became the axiomatic foundation of the definition of the concept of a boolean algrebra, see our axioms $Ax_{ba}.1$ - 13 on p. 21 of this Chapter.

(i) $V(A \vee B) = V(A) \cup V(B)$;

(ii) $V(A \& B) = V(A) \cap V(B)$;

(iii) $V(\neg A) = 1$ if $V(A) = 0$ and $V(\neg A) = 0$ if $V(A) = 1$.

More interesting is the kind of algebra that we get when propositions are characterised *intensionally*, viz. as sets of possible worlds. Let W be the set of all possible worlds. Then each proposition p determines a subset of W, consisting of those worlds in which p is true. According to the *intensional* theory of propositions this set - or, if one prefers, the division of W into two parts that comes with it, the part of those possible worlds in which p is true and those in which it is false - fully identifies the proposition p; in other words, propositions *are* sets of possible worlds; and on the assumption that the set of all possible worlds is W, they are subsets of W. The logical operations of disjunction, conjunction and negation now turn into set-theoretic operations. For instance, the conjunction p & q of the propositions (i.e. subsets of W) p and q is the set of worlds of W in which both p and q are true, i.e. the worlds which belong both to the subset p of W and to the subset q of W. Thus p & q is the set-theoretic intersection of p and q. Similarly, p v q becomes the union of p and q and $\neg$p the set-theoretic difference W\p. Furthermore, the 0 of the proposition algebra thus defined is the empty set $\varnothing$ ('the contradictory proposition') and its 1 the entire set W ('the tautologous proposition').

This 'intensional' proposition algebra is the model-theoretic fundament of the currently most popular developments of modal and intensional logic, in which logical relations are defined in terms of a 'Kripkean'[11] model-theoretic semantics, propositions are interpreted as sets of possible worlds and modalities are analysed in terms of relations between such worlds. It is also the model-theoretic foundation of the system of Higher Order Intensional Logic, the logical formalism that was introduced by Montague[12] in his seminal work on the semantics of natural languages - work that, in various guises has served as the formal basis for the formal semantics of natural languages since the early seventies.

---

[11] Saul Kripke (1940 -) is the founder of modern modal logic. He did his astoundingwork in this area at the astoishingly young age of 16, while still in high school.

[12] Richard Montague (1930 - 1971). Founder of the model-theoretic approach to the analysis of meaning of natural languages. Montague was the first to see that it was possible and illuminating to apply the model-theoretic methods developed by his teacher Tarski for the formal languages of mathematical logic, such as, in particular, the predicate calculus..

There is a further connection between the semi-formal ideas expressed above and Montague's conception of the semantics of (formal and natural) languages. Montague thought of the way in which the syntactic structure of a sentence determines its meaning as generally taking the form of a *homomorpism* from syntactic structures to meanings (or 'semantic values'). In the context of the present discussion of boolean structure this idea can be explained rather succinctly. Doing so, moreover, will give an opportunity to introduce the general notion of a homomorphism and its systematic connections to the already familiar notions of an equivalence relation and that of one relation being congruence relation wrt. some other relation. And finally it throws an illuminating light on the ideas that Boole and De Morgan were after but that can be stated fully transparently only now that we know how to draw a clear distinction between sentences of a language as symbol strings with a syntactic structure and the samentic values ('propositions') they denote.

In the more formal discussion that follows we focus on first order languages as we have been doing hitherto. This will also allow for a natural transition to the topic of the next two sections.

Central to the discussion will be the language of boolean algebra, i.e. the language L_{ba} whose logical constants are $\cup$,$\cap$,0,1and -. Let L be any first order language. We can use the set S_L consisting of the sentences of L to define the following model M_L for L_{ba}: the universe is S_L and the interpretations of the non-logical constants of L_{ba} are given by the following function F_L:

$F_L(\cup)(A,B) = (A \vee B)$; $F_L(\cap)(A,B) = (A \& B)$; $F_L(0) = \neg(\forall v_1)\ v_1 = v_1$; $F_L(1) = (\forall v_1)\ v_1 = v_1$; & B); $F_L(-)(A) = \neg A$, where A and B are arbitrary sentences of L.

In other words, the 'boolean' operator symbols $\cup$ etc. are interpreted in as *syntactic* operations of the sentences of L. For instance, $\cap_L$ operates on arbitrary sentences (that is, arbitrary well-formed symbol strings) A and B of L and maps such a pair to the symbol string (A & B).

Now let $\mathbb{M}$ be some class of models for L. Then each sentence A of L can be said to express wrt $\mathbb{M}$ a 'proposition' $[[A]]^{\mathbb{M}}$, consisting of those models M in $\mathbb{M}$ for which $M \models A$: $[[A]]^{\mathbb{M}} = \{M \in \mathbb{M}: M \models A\}$. (It is reasonable to refer to $[[A]]^{\mathbb{M}}$ as the 'proposition expressed by A wrt. $\mathbb{M}$' insofar as $[[A]]^{\mathbb{M}}$ tells us for each $M \in \mathbb{M}$, and thus for each of the

'possible worlds' described by models of $\mathbb{M}$, whether or not A (or the proposition A expresses) is true in that world or model.)

From the propositions $[[A]]^{\mathbb{M}}$ expressed by A wrt. $\mathbb{M}$ it is possible to construct another model for L_{ba}, to which we refer as $M_{\mathbb{M}}$. The universe of this model is the set $\{[[A]]^{\mathbb{M}}$: A is a sentence of L} anf its interpretation function $F_{\mathbb{M}}$ is defined by:

$$F_L(\cup)([[A]]^{\mathbb{M}},[[B]]^{\mathbb{M}}) = ([[A]]^{\mathbb{M}} \mathbf{\cup} [[B]]^{\mathbb{M}};$$
$$F_L(\cap)([[A]]^{\mathbb{M}},[[B]]^{\mathbb{M}}) = ([[A]]^{\mathbb{M}} \mathbf{\cap} [[B]]^{\mathbb{M}};$$
$$F_L(0) = \varnothing;\ F_L(1) = \mathbb{M};\ F_L(-)([[A]]^{\mathbb{M}}) = \mathbb{M}\backslash[[A]]^{\mathbb{M}}.$$

(Here we have used bold face $\mathbf{\cup}$ and $\mathbf{\cap}$ to distinguish the set-theoretical union and intersection from the function constants $\cup$ and $\cap$ of the language L_{ba}.)

It follows directly from what we seen in the last section that $M_{\mathbb{M}}$ is a boolean algebra. On the other hand the model M_L is not, for one thing because syntactic disjunction and conjunction, the functions which interpret the function constants $\cup$ and $\cap$ in M_L, are not commutative. (For instance, in general, (A&B) is not the same string as (B & A); in particular, $(\neg(\forall v_1)\ v_1 = v_1\ \&\ (\forall v_1)\ v_1 = v_1)$ is not the same string as $((\forall v_1)\ v_1 = v_1\ \&\ \neg(\forall v_1)\ v_1 = v_1)$; and so on.) This means that the function $[[\ ]]^{\mathbb{M}}$ maps the non-boolean model for L_{ba} onto the boolean model $M_{\mathbb{M}}$.

Given a first order language L many different classes of models $\mathbb{M}$ are possible and for each such choice we get a different function $[[\ ]]^{\mathbb{M}}$. The possible choices of $\mathbb{M}$ are bounded on the one side by the smallest such choices- those where $\mathbb{M}$ is a singleton set {M} - and on the other side by the maximal choice, where $\mathbb{M}$ is the class of all models for L. When $\mathbb{M}$ = {M}, then the universe of the model $M_{\mathbb{M}}$ consists of just two elements, the set {M} itself and the empty set $\varnothing$. We can think of these two elements as 'true in $M_{\mathbb{M}}$' and 'false in $M_{\mathbb{M}}$' and replace them by 1 and 0. This gives us the 2-element boolean algebra, whose universe is the set {0,1} and whose operations are the familiar connectives of classical propositional logic, given by the classical truth tables. (For example, the interpretation of $\cap$ in this model is the 2-place function $\underline{\&}$ defined by $\underline{\&}(1,1) = 1$; $\underline{\&}(1,0) = \underline{\&}(0,1) = \underline{\&}(0,0) = 0$, and so on.) In this case the

notion of a 'proposition wrt $M_{\mathbb{M}}$' reduces to that of a mere truth value. $[[\ ]]^{\mathbb{M}}$ throws together any two sentences that have the same truth value in M and we end up with just two 'bags' one for the sentences of L that are true in M and one for the false sentences.

At the other extreme, where $\mathbb{M}$ is the class of all models for L, we get a maximal diversity of bags. Now two sentences A and B end up in the same bag only iff tghey are logically equivalent: $[[A]]^{\mathbb{M}} = [[B]]^{\mathbb{M}}$ iff for every model M for L, $M \models A$ iff $M \models B$.

The function $[[\ ]]^{\mathbb{M}}$ is an example of a *homomorphism*. Homomorphisms are maps from one structure into another which are structure-preserving. In general such maps are not 1-1. And that is true also for $[[\ ]]^{\mathbb{M}}$, since any two different logically equivalent sentences will be mapped onto the same value. For instance, we have for any sentences A and B that $[[(A \ \& \ B)]]^{\mathbb{M}} = [[(B \ \& \ A)]]^{\mathbb{M}}$, even though the two conjunctions (A & B) and (B & A) are, as we have just observed, in general distinct. In fact, the point of a homomorphism is often that it isolates those aspects of a given type of structure that are relevant from a certain perspective while abstracting from all remaining features. It does this by 'throwing into the same bag' any two elements for which the structural features that are relevant from the given perspective are the same and that thus only differ in respects that do not matter. Thus $[[\ ]]^{\mathbb{M}}$ identifies, by mapping them onto the same value, any two sentences whose structure guarantees that they have the same truth value in all models of $\mathbb{M}$.

We will define the notion of a homomorphism only for algebraic structures - that is, for models of algebraic languages. (There is a way of generalising the notion to arbitrary first order languages, some or all of whose non-logical constants are predicates, but since we won't need this generalisation here or later, we will limit ourselves to the case of algebraic languages only.

Def. 8

a. Let L be any algebraic language, M, M' models for L, h a function from U_M into $U_{M'}$. h is a *homomorphism from* M *into* M' iff for every non-logical constant f^n of L and every n elements $d_1, ..., d_n$ from U_M:

$$h(f^n{}_M(d_1,..,d_n)) = f^n{}_{M'}(h(d_1),..,h(d_n))$$

Special cases are those where a homomorphism h from M into M'is (i) onto M' and (ii) where h is 1-1. It is immediate that if h is both 1-1 and onto, then it is an isomorphism from M onto M'. We already noted that $[[\ ]]^{\mathbb{M}}$ is onto $M_{\mathbb{M}}$. Usually $M_{\mathbb{M}}$ is a proper submodel of the set inclusion lattice $M' = \langle P(\mathbb{M}), \subseteq \rangle$, and when that is so, $[[\ ]]^{\mathbb{M}}$ is a homomorphism into, but not onto, M'. (Exercise: For which combinations of of a first order language L and a class $\mathbb{M}$ of models for L is $M_{\mathbb{M}}$ a proper submodel of M'?)

There is an important general connection between homomorphisms and congruence relations. Again we use our 'syntax-semantics interface function' $[[\ ]]^{\mathbb{M}}$ to illustrate the matter. As a preliminary recall that there is a general correlation between functions and equivalence relations: (i) Let f be a function defined on some set X. Then f induces an equivalence relation $\sim$ on X, defined by:

(1) for any $x,y \in X$, $x \sim y$ iff $f(x) = f(y)$.

Conversely, any $\sim$ equivalence relation on a set X induces a function on X which maps each $x \in X$ onto the equivalence class $[x]_{\sim}$ it generates under $\sim$. Moreover, when (1) is applied to this function, it gets us back to the relation $\sim$.

This correlation holds in particular for functions that are homomorphisms. In particular, when h is a homomorphism from one structure M into another structure M', then there will be a corresponding equivalence relation $\sim$ on U_M induced by h via (1). In this case, however, $\sim$ has additional properties, which reflect the fact that h is a homomorphisms (and not just any function): is a *congruence relation wrt* each of the operations of M. We recall the notion of a congruence relation: Suppose that f is an n-place function defined on some set X, i.e. both the arguments and the values of f belong to X, and that $\sim$ is a binary relation on X. Then $\sim$ is a *congruence relation wrt* f iff for any $x_1, ..., x_n, x'_1, ..., x'_n$ from X such that $x_1 \sim x'_1, ..., x_n \sim x'_n$, $f(x_1,.,x_n) \sim f(x'_1,.,x'_n)$.

It is easily verified that when h is a homomorphism from a model M for an algebraic language L into some other model M' for L, then the relation $\sim$ induced by h via (1) is congruence relation wrt. all interpretations in M of function constants of L. Moreover, the converse

also holds in this case: If $\sim$ is an equivalence relation on U_M which is a congruence relation on the interpretations in m of all the non-logical constants of L, then the function which maps any element d of U_M onto its equivalence class $[d]_\sim$ is a homomorphism from M into the model M' whose universe is the set of equivalence classes $[d]_\sim$ and which interprets each n-place function constant f of L via the definition:

$$f_{M'} = \{<[d_1]_\sim, ..,[d_n]_\sim,[d]_\sim>: d_1, ..,d_n,d \varepsilon U_M \ \& \ f(d_1,..,d_n) = d\}$$

(This definition is legitimate because $\sim$ is a congruence relation wrt f.)

Returning to $[[\ ]]^{\mathbb{M}}$ we recall that this function is a homomorphism in that this function preserves the interpretations of all the function constants of L_{ba}. (For instance, $[[\ ]]^{\mathbb{M}}$ converts the syntactic conjunction operation & into the 'propositional conjunction' which maps the model sets $[[A]]^{\mathbb{M}}$ and $[[B]]^{\mathbb{M}}$ onto their intersection.) It follows from the general connection between homomorphisms and congruence relations we have described above that the relation which holds between sentences A and B iff they have the same truth values in each of the models of $\mathbb{M}$ is a congruence relation wrt to the syntactic operations that interpret the function constants of L_{ba} in $M_{L_{ba}}$. This is the formal justification for looking at the connectives of classical propositional logic as algebraic operations on 'sentence meanings'.

As noted in footnote **??**, the conception of the way in which meaning is determined by form as a homomorphism that maps syntactic strings onto meanings, thereby identifying any two strings whose structures make them identical in meaning, is a central assumption in the approach to meaning in natural languages developed by Montague in the late sixties and early seventies and now generally known as 'Montague Grammar'. The idea is that the syntax of any language - natural languages no less than the formal languages of logic and computer science (including in particular the first order languages that are the topic of these Notes) - can always be characterised by a set of syntactic operations which build complex expressions from constituents, and that to each such syntactic operation corresponds a rule which combines the semantic values of the constituents into the semantic value of the expression that is the output of the syntactic operation. It became clear soon that (except for very restricted fragments) the strictest implementation of this conception comes at a cost of assumptions about the syntax of natural languages that are quite artificial, and are ill supported by intrinsically syntactic evidence, of the

kind that linguists do, and should, take seriously. Nevertheless, the attempt to develop a syntax-semantics interface that is based on an independently plausible syntax and yet keeps as closely to Montague's original conception has proved a principle of immense methodological value in the development of semantics over the past 40 years.

The model $M_{\mathbb{M}}$ for L_{ba} that we obtain when $\mathbb{M}$ is the class of all models for L is known as the *Lindenbaum algebra of* L. Lindenbaum algebras will play an important part in the next section, be it in the different guise of structures whose elements are the finitely axiomatisable deductive theories of a given first order language L. (There is an obvious 1-1 correspondence between the finitely axomatisable theories of L and the classes $[[A]]^{\mathbb{M}}$ into which $[[\ ]]^{\mathbb{M}}$ partitions the set of all sentences of L and that make up the universe of $M_{\mathbb{M}}$ when $\mathbb{M}$ contains all models for L. For on the one hand, if t is a finitely axiomatisable theory of L, then there is a single sentence A of L such that $T = Cl_L(\{A\})$. On the other hand, when two sentences A and A' belong to the same class, i.e. if $[[A]]^{\mathbb{M}} = [[A']]^{\mathbb{M}}$, then A and A' are logically equivalent and thus axiomatise the same theory: $Cl_L(\{A\}) = Cl_L(\{A'\})$. Thus each finitely axiomatisable theory of L corresponds to exactly one element of the universe of $M_{\mathbb{M}}$.)

2.2 Incomplete Theories and their Extensions.

In section 2.1 we saw that complete theories do not always do what one might have expected of them, and for which they are often designed: describe a given structure uniquely up to isomorphism. A complete theory always succeeds in doing this, we observed, when the structure it is meant to describe is finite. (See Thm. 6 of Ch.1.) But for theories with infinite models the picture is much more complicated. We know that if a complete theory has an infinite model, then all its models are infinite (see exercise ??). But the differences between these infinite models may still be considerable. Not only will the theory always have models that are not isomorphic for the simple reason that their universes are of different cardinality - recall that the Skolem-Löwenheim Theorems tell us that theories with infinite models always have models of every possible infinite cardinality -, there exist complete theories that have non-isomorphic models even within the same cardinality. Though Morley's Theorem indicates that the range of possibilities us much more limited than one might have thought, there nevertheless remains considerable room for variation. For suppose a theory T has non-isomorphic models in some infinite cardinality κ. Then there is the further question how *wide* the variety of models of T of cardinality κ is. To answer this question a much finer - and much deeper - analysis of complete first order theories is needed than anything presented in these notes. Such an analysis exists. it is known as Stability Theory, a subject of considerable complexity, developed and brought to conclusion almost single-handedly by the Israeli mathematician and logician Saradon Shelah **[ref. to Shelah]**

When we move from complete to incomplete theories we find much wider ranges of possible models. Now the models of a theory T can be given a first classification in terms of the sentences they verify, in other words, in terms of those of the complete extensions of T which they verify. So the range of models of an incomplete theory T can be studied from two complementary perspectives, first the set of complete extensions of T, and second, for each of these complete extensions the range of models for that extension.

So far we have encountered examples of complete as well as of incomplete theories. But we haven't looked much at the structure of the entire field of theories in a given language L, including both its complete and its ncomplete theories. It is this issue that we will pursie in the present section.

2.2.1 Lattices of Theories.

Let L be a first order Language and let $\mathbb{T}_L$ be the set of all theories of L. The structure $\mathcal{T}_L = <\mathbb{T}_L, \subseteq>$, where $\subseteq$ is the relation of set-theoretical inclusion restricted to $\mathbb{T}_L$, is called the *Lattice of Theories of L.* We will also refer to it as the *Tarski Lattice of L* in honour of the logician A. Tarski, who was the first to study these structures.

Our first task is to show that $\mathcal{T}_L$ is a distributive lattice with 0 and 1. We already noted in the precious sections that any restriction $\subseteq_V$ of $\subseteq$ to some set V of sets is a weak partial order on V. To show that when V = $\mathbb{T}_L$ this partial ordering is a lattice, we must show that for each pair of theories T_1 and T_2 of L $\subseteq \mathbb{T}_L$ yields an infimum and a supremum with. First, note that $T_1 \cap T_2$ (where $\cap$ is set-theoretic intersection) is a theory of L. For suppose B is any sentence of L such that $T_1 \cap T_2 \models B$. Then $T_1 \models B$ and $T_2 \models B$, So, since T_1 and T_2 are theories, B ε T_1 and B ε T_2. So B ε $T_1 \cap T_2$. Since this holds for arbitrary B, $T_1 \cap T_2$ is a theory. It now follows almost directly that $T_1 \cap T_2$ is the infimum of T_1 and T_2 in $\mathcal{T}_L$. For if T is any theory of L such that T $\subseteq$ T_1 and T $\subseteq$ T_2, then T $\subseteq$ $T_1 \cap T_2$.

The case of $\cup$ is different because $T_1 \cup T_2$ is in general not a theory. (It is a theory only if $T_1 \subseteq T_2$ or $T_2 \subseteq T_1$, (See Exercise 20.ii of Ch.1) But T_1 and T_2 do have a supremum in $\mathcal{T}_L$ nevertheless, viz. the theory $Cl_L(T_1 \cup T_2)$.
To see this, observe that $T_1 \subseteq Cl_L(T_1 \cup T_2)$ and $T_2 \subseteq Cl_L(T_1 \cup T_2)$. Now suppose that T' is any theory of L such that $T_1 \subseteq$ T' and $T_2 \subseteq$ T'. Let B be any sentence from $Cl_L(T_1 \cup T_2)$. Then $T_1 \cup T_2 \models B$. So by the Completeness Theorem $T_1 \cup T_2 \vdash B$. From this is can easily be inferred that there must be a single sentence C ε T_1 and a single sentence D ε T_2, such that C & D $\vdash$ B. Since $T_1 \subseteq$ T', C ε $T_1 \cup T_2$ and thus C ε T'. Similarly D ε T'. So, T' $\models$ C & D and so since T' is a theory, C & D ε T'. So since C & D $\models$ B, also B ε T'.

Having shown that the supremum and the infimum of any two members of $\mathcal{T}_L$ exist, we facilitate further discussion by introducing the symbols $\cup_L$ and $\cap_L$ for these operations:

(1) (i) $T_1 \cup_L T_2 =_{df} Cl_L(T_1 \cup T_2)$
(ii) $T_1 \cap_L T_2 =_{df} T_1 \cap T_2$

That $\mathbb{T}_L$ has a 0 and a 1 is obvious. Its 0 is the theory $0_L = \{A \in L: \vdash A\}$ and its 1 the contradictory L-theory 1_L consisting of all sentences of L. That is distributive requires an argument. We show that the distributive law DISTR.2 holds in $\mathbb{T}_L$.[13] (The validity of the other law is shown in much the same way.)

DISTR.2 $T_1 \cup_L (T_2 \cap_L T_3) = (T_1 \cup_L T_2) \cap_L (T_1 \cup_L T_3)$

To show the inclusion of the left hand side in the right hand side is straightforward. (In fact this inclusion holds in all lattices.) To show inclusion in the opposite direction, let $B \in (T_1 \cup_L T_2) \cap_L (T_1 \cup_L T_3)$. Then $B \in (T_1 \cup_L T_2)$ and $B \in (T_1 \cup_L T_3)$. Since $B \in (T_1 \cup_L T_2)$, there are $C' \in T_1$ and $D \in T_2$ such that $C' \& D \vdash B$. Similarly, since $B \in (T_1 \cup_L T_3)$, there are $C'' \in T_1$ and $E \in T_3$ such that $C'' \& E \vdash B$. Putting $C =_{df} C' \& C''$, we have $C \& D \vdash B$ and $C \& E \vdash B$. So $C \& (D \vee E) \vdash B$. But $D \vee E \in T_2$ and $D \vee E \in T_3$. So $D \vee E \in T_2 \cap_T T_3$. So $T_1 \cup_L (T_2 \cap_L T_3) \vdash B$. So $B \in T_1 \cup_L (T_2 \cap_L T_3)$.

q.e.d.

While $\mathbb{T}_L$ is always a distributive lattice, it is never a boolean lattice. The reason is that if T is a theory of a first order language L which is not finitely axiomatisable, then there is no theory T' of L such that $T \cup_L T' = 1_L$ and $T \cap_L T' = 0_L$. And every first order language has theories that are not finitely axiomatisable. We record this fact as Theorem 4.

Thm. 4 For no first order language L is $\mathbb{T}_L$ a boolean lattice.

We postpone the proof of Thm. 4 till later in this section.

While $\mathbb{T}_L$ is never a boolean lattice, each $\mathbb{T}_L$ has a certain sublattice which invariably is boolean. This is the so-called *Lindenbaum algebra of* L.[14] It consists of all finitely axiomatisable theories of L, i.e. all

---

13 See Section 2.1.2. Note that here we have omitted the universal quantifiers binding T_1, T_2 and T_3.

14 Speaking on the one hand of 'Tarski lattices and on the other of Lindenbaum algebras will seem incoherent. The term 'Lindenbaum *algebra* has'

theories T of L such that for some finite set *A* of L-sentences $T = Cl_L(A)$. We denote the Lindenbaum Algebra *of* L as $\mathbb{L}_L$.

To show that $\mathbb{L}_L$ is a boolean lattice, we recall that a theory T is finitely axiomatisable iff there is a single sentence A such that $T = Cl_L(\{A\})$ - see Exercise 12.a of Ch. 1. (For easier reading we write 'T_A' instead of '$Cl_L(\{A\})$'.) It is straightforward to verify that if T_1 and T_2 are finitely axiomatisable theories of L and $T_1 = T_A$ and $T_2 = T_B$, then the following two conditions hold (Exercise: Show this.)

(1) (i) $T_1 \cup_L T_2 = T_{A\&B}$
(ii) $T_1 \cap_L T_2 = T_{A \vee B}$

Now let T be any finitely axiomatisable theory of L and suppose that $T = T_A$. Let $T' = T_{\neg A}$. Then according to (3.i,ii) $T_A \cup_L T_{\neg A} = T_{A\&\neg A}$ and $T_A \cap_L T_{\neg A} = T_{A \vee \neg A}$. But $T_{A\&\neg A} = Cl_L(\{A\&\neg A\}) = 1_L$ and $T_{A\vee\neg A} = Cl_L(\{A \vee \neg A\}) = 0_L$. So T' is the complement of T, in that the two satisfy the characteristic equations, repeated in (2).

(2) (i) $T \cup_L T' = 1_L$
(ii) $T \cap_L T' = 0_L$

Since for each member T of $\mathbb{L}_L$ there is a complement T' in $\mathbb{L}_L$ such that (2.i,ii) are satisfied, $\mathbb{L}_L$ is boolean. q.e.d.

As noted in the remarks leading up to Thm. 4, theories that are not finitely axiomatisable do not have boolean complements. However, it is possible to define an operation on arbitrary theories that (a) satisfies at least one of the conditions in (2), viz. (2.ii), (b) is the largest element satisfying this condition and (c) coincides with the boolean complement of any finitely axiomatisable theory. One definition of this operation is given in Def. 9.

It is possible to define a complement operation on theories of L which acts as a boolean complement when the theory in question is a theory

---

been adopted because of its general use in the literature - few people if anyone speak of the Lindenbaum *lattice* of L. Because of the equivalence between lattices and algebras nothing much hangs on this terminological issue. In fact we might just as well speak of Lindnbaum latticesa as of Lindenbaum algebras, and likewise, speaking of Tarski algebras is just as legitimate as talking about Tarski lattices.

of $\mathbb{L}_L$. The definition we will give is such that it can be applied to arbitrary theories. But only when the theory is finitely axiomatisable, will the theory and its complement stand in the relations that are distinctive of boolean algebras.

Def. 9 Let T be an element of $\mathbb{T}_L$. The *pseudocomplement of* T i*n* $\mathbb{T}_L$, $-_L T$, is defined by: $-_L T = \cup \{T' \in \mathbb{T}_L : T \cap_L T' = 0_L\}$[15]

Prop. 5 (i) $-_L T$ is the largest theory T' of L such that $T \cap_L T' = 0_L$.
(ii) Suppose that $T = T_A$. Then $-_L T = T_{\neg A}$.

Proof.

(i) Let $\underline{T} = Cl_L(-_L T)$. Suppose that $B \in T \cap \underline{T}$. Then $B \in T$ and there is a $C \in -_L T$ such that $C \models B$. But if $C \in -_L T$, then there is some theory T' such that $T \cap_L T' = 0_L$ and $C \in T'$. Since $C \in T'$ and $B \in T$, $C \vee B \in 0_L$. On the other hand, since $C \models B$ and $B \models B$. $C \vee B \models B$. So, since 0_L is a theory, $B \in 0_L$. This establishes that $\underline{T}$ is a theory T' such that $T \cap_L T' = 0_L$. Therefore $Cl_L(-_L T) = \underline{T} \subseteq -_L T$. So $-_L T = Cl_L(-_L T)$. That is, $-_L T$ is a theory It now follows directly from Def. 8 that it is the largest theory T' such that $T \cap_L T' = 0_L$.

(ii) Suppose that $T = T_A$. Then, as we have already seen, $T_{\neg A}$ is a theory T' such that $T \cap_L T' = 0_L$. So $T_{\neg A} \subseteq -_L T$. Now let T' be any theory such that $T \cap_L T' = 0_L$. Suppose that $B \in T'$. Then, since $A \in T$, $A \vee B \in 0_L$; that is, $\models A \vee B$. But $A \vee B$ is logically equivalent to $\neg A \rightarrow B$. So $\models \neg A \rightarrow B$, and therefore $\neg A \models B$. So $B \in T_{\neg A}$. This establishes that $-_L T \subseteq T_{\neg A}$. So $T_{\neg A} = -_L T$.

q.e.d.

---

15 Tarski lattices are thus structuresc which, according to a well-esrablished terminology are called *pseudo-complemented lattices*. A pseudo-complemented lattice is a lattice with an additional 1-place operation - with the properties that for all x, -x is the largest element such that $x \cap -x = 0$. Tarski-lattices have additional properties, one of which is that they are distributive. In fact, most of the well-known examples of pseudo-complemented lattices that are not Boolean algebras are distributive. However, the existence of a pseudo-complement does not entail distributivity. For instance, the 5-element lattice of Section 2.1.3 is pseudo-complemented (-1 = 0, -0 = 1, - a = b, -b = -c = a), but as we saw it is not distributive. Sometimes the pseudo-complement of x is defined as the smallest element y such that $x \cup y = 1$. From a formal poitn of view this comes in last analysis to the same ting because of the duality of $\cup$ and $\cap$.

We now proceed to the proof of Thm. 4.

Proof of Thm. 4

(a) Suppose that T is a theory of some first order language L and that $T \cup_L {}^{-L}T = 1_L$. Then there are a sentence A from T and a sentence B from ^{-L}T such that $A \& B \models \bot$. This entails that $B \models \neg A$. So we have $\neg A \in {}^{-L}T$. We show that $T = T_A$. Suppose that $C \in T$. Then, since $\neg A \in {}^{-L}T$, $C \vee \neg A \in 0_L$. So $\models C \vee \neg A$, which is equivalent to: $A \models C$. So $C \in T_A$. So we have shown that $T \subseteq T_A$. On the other hand, since $A \in T$, $T_A \subseteq T$. So $T = T_A$.

(b) We observe that the following infinite set of sentences $\{D_n\}_{n = 2,3,..}$ is strictly increasing in that for all n, $D_{n+1} \models D_n$ but not $D_n \models D_{n+1}$:

D_2: $(\exists v_1)(\exists v_2)\ v_1 \neq v_2$
D_3: $(\exists v_1)(\exists v_2)(\exists v_2)\ (v_1 \neq v_2 \& v_1 \neq v_3 \& v_2 \neq v_3)$
.
.
(D_n says that there are at least n different elements in the universe.)

Let L be any first order language and let $T_{inf,L}$ be the theory axiomatised by the sentences , i.e. $T_{inf,L} = Cl_L(\{D_n\}_{n = 2,3,..})$. (Note that the sentences only use logical vocabulary and thus belong to any first order language whatever.) Then according to Exercise 7.b of Ch. 1 $T_{inf,L}$ is not finitely axiomatisable. So $T_{inf,L}$ has no complement in L satisfying both of the two conditions (2.i,ii).
It follows that for no L is $\mathbb{T}_L$, the Tarski lattice for L, a boolean lattice.
q.e.d.

So far we have considered the Tarski lattices $\mathbb{T}_L$ of first order languages and just one type of substructure of those, the Lindenbaum algebras. But of course we could in principle study many other sublattices of the $\mathbb{T}_L$s. Of special importance among those sublattices are certain lattices whose bottom element is not 0_L, but rather some theory T of L. More particularly, it has proved useful in a variety of contexts to study (i) the lattice consisting of all extensions of T, and (ii) the lattice consisting of the *finitely axiomatisable* extensions of T (those extensions T' of T for

which there is a sentence A of L such that T' = $Cl_L(T \cup \{A\})$.[16] We call these the T*arski lattice of* L *generated by* T and the *Lindenbaum algebra of* L *generated by* T, respectively, and denote them as $\mathbb{T}_{L,T}$ and $\mathbb{L}_{L,T}$.

Def. 10 Let L be a language, T a theory of L.

a. The *Tarski lattice of* L *generated by* T is the structure $\mathbb{T}_{L,T} = <\mathbb{T}_{L,T}, \subseteq>$, where $\mathbb{T}_{L,T}$ is the set of all L-extensions of T and $\subseteq$ is the relation of set-theoretic inclusion restricted to $\mathbb{T}_{L,T}$.

b. The *Lindenbaum algebra of* L *generated by* T is the structure $\mathbb{L}_{L,T} = <\mathbb{L}_{L,T}, \subseteq>$, where $\mathbb{L}_{L,T}$ is the set of all L-extensions of T which are finitely axiomatisable over T - that is. All those L-extensions T' of T for which there is an L-sentence A such that T' = $Cl_L(T \cup \{A\})$) and $\subseteq$ is the inclusion relation on $\mathbb{L}_{L,T}$.

Like $\mathbb{T}_L$, $\mathbb{T}_{L,T}$ is always a distributive lattice with 0 and 1. This can be shown in just the same way as we did for $\mathbb{T}_L$. The argument that $\mathbb{L}_{L,T}$ is always boolean also goes as before. So far, then, there is no difference between the more general cases of $\mathbb{T}_{L,T}$ and $\mathbb{L}_{L,T}$ and the more specific cases of $\mathbb{T}_L$ and $\mathbb{L}_L$, in which the bottom element is 0_L. But there is nevertheless one difference, viz. that among the lattices $\mathbb{T}_{L,T}$ we now find many that are boolean (while, as we have seen, this is never so for the lattices $\mathbb{T}_L$). It can be inferred from what has already been established in this section that this happens only when the Tarski lattice generated by T and the Lindenbaum algebra generated by T coincide, i.e. when all extensions of T are finitely axiomatisable over T. In the next section we will see a number of comparatively simple examples of this situation.

Besides the lattices $\mathbb{T}_{L,T}$ and $\mathbb{L}_{L,T}$ other sublattices of $\mathbb{T}_L$ are worth consideration as well. Among these are in particular the lattice of all subtheories of a given theory T and the lattice consisting of all its finitely axiomatisable subtheories. (Exercise: prove that the former is again a distributive lattice with 0 and 1, where the set of tautologies of L is the 0 and T is the 1, and that the latter is a boolean lattice.) Even

---

16 Often the lattice $\mathbb{T}_{L,T}$ provides us with certain insights into the nature of T. For by telling us something aqbout the range of possible extensions of T it also tells us something about the range of its possible models, or true interpretations. and with that of the range of variability among the models of T.

more generally, we can, for any pair of L-theories T and T' such that $T \subseteq T'$, consider the Tarski lattice and Lindenbaum algebra consisting of those L-theories (or finitely axiomatisable L-theories, respectively) that lie between T and T', - in other words, at the sublattices of $\mathbb{T}_L$ whose 0 is T and whose 1 is T'. None of these, however, will be further considered in these Notes.

We have already observed that $\mathbb{T}_L$ is never boolean - not even for the simplest language {}. This is not so for the lattices $\mathbb{T}_{L,T}$. These can be boolean. Among them is the trivial lattice $\mathbb{T}_{L,\perp_L}$, whose only element is $\perp_L$, and all two element lattices $\mathbb{T}_{L,T}$, for T a consistent and complete theory of L, lattices whose only elements are $\perp_L$ and T.

In general, lattices of the form $\mathbb{T}_{L,T}$ are always both atomic and complete. More precisely, this is so for any such lattice with more than two elements. (If a lattice has ≤ 2 elements, then there are no atoms and the concept of atomicity is not applicable.) To see that $\mathbb{T}_{L,T}$ is atomic, assume that $\mathbb{T}_{L,T}$ has > 2 elements and observe that the complete consistent extension of T are the 'anti-atoms' of $\mathbb{T}_{L,T}$: they are those theories different from the inconsistent theory of L such that there is no theory between them and the inconsistent theory. It is easy to show - Exercise: do this! - that the atoms of $\mathbb{T}_{L,T}$ are precisely the theories $\text{-}_L T'$ where T' is any complete and consistent extension of T. With this in mind it is easy to see that $\mathbb{T}_{L,T}$ is atomic. For let T' be any proper extension of T (i.e. any extension of T that is different from T). Let A be any consistent sentence in T'\T - there will be such sentences if $\mathbb{T}_{L,T}$ has > 2 elements - and let T'' be any complete and consistent extension of $\text{Cl}(\{\neg A\})$. Then $\text{-}_L T''$ is an atom below T'. (Exercise: prove this!)

That $\mathbb{T}_{L,T}$ is a complete lattice is straightforward. Let $\mathcal{T}$ be any set of extensions of T. It is easy to show that $\text{Cl}_L(\cup \mathcal{T})$ is the supremum of $\mathcal{T}$.

We already know that $\mathbb{T}_{L,T}$ is not always a boolean lattice. (In particular, this is never so when T is 0_L.) For some L and T, however, $\mathbb{T}_{L,T}$ is boolean. Trivial examples are those where T is the inconsistent theory of L, in which case $\mathbb{T}_{L,T}$ is the trivial boolean algebra consisting of just one element and the case we already considered, where T is a complete consistent theory, in which case $\mathbb{T}_{L,T}$ consists of two elements, T and the inconsistent theory of L. There are also many examples of boolean $\mathbb{T}_{L,T}$ of more than 2 elements. However, *all*

boolean lattices $\mathbb{T}_{L,T}$ are finite. Note that this does not simply follow from the fact that such lattices are atomic and complete. For there exist infinite atomic and complete boolean lattices, viz. the power set inclusion structures $\langle P(X),\subseteq\rangle$ in which X is infinite.

The fact that boolean lattices of the form $\mathbb{T}_{L,T}$ are always finite thus has to do with the special properties of theory lattices. Since we have already established that $\mathbb{T}_{L,T}$ is always atomic and complete, the argument is quite simple. It goes as follows. First we observe the following general property of complete atomic boolean lattices $\mathbb{L}$:

(1) Let $\mathbb{L}$ be a complete atomic boolean lattice and let A_1 and A_2 be two distinct sets of atoms of $\mathbb{L}$. Then the suprema in $\mathbb{L}$ of these two sets, $\sup(A_1)$ and $\sup(A_2)$, are distinct.

We prove (1) by making use of (2), which we leave as an exercise:

(2) Let $\mathbb{L}$ be a boolean lattice and let a, a' be distinct atoms of $\mathbb{L}$. Then $a \leq -a'$.

Proof of (1): Let A_1 and A_2 be two distinct sets of atoms of $\mathbb{L}$. Then there is an $a \in A_1 \setminus A_2$ or there is an $a \in A_2 \setminus A_1$. Assume that $a \in A_1 \setminus A_2$. Then by (2) for each $a' \in A_2$, $a' \leq -a$. So, $\sup(A_2) \leq -a$. On the other hand $a \leq \sup(A_1)$. So it is not the case that $\sup(A_1) \leq -a$; for that would mean that $a \leq -a$, which is obviously impossible, as it would entail that $-a = 1_{\mathbb{L}}$, which evidently it isn't. (If it were, then $a = --a = 0_{\mathbb{L}}$, and thus a would not be an atom.)

We next observe (3)

(3) Any complete, atomic boolean lattice $\mathbb{L} = \langle U,\leq\rangle$ with atom set A is isomorphic to the power set inclusion lattice $\langle P(A),\subseteq\rangle$.

(3) follows from (1) and (4), the proposition that in a complete atomic boolean lattice $\mathbb{L}$ each element other than $0_{\mathbb{L}}$ is the supremum of the set of all atoms below it.

(4) Let $\mathbb{L}$ be a complete atomic boolean lattice with atom set A and let b be any element of $\mathbb{L}$ such that $b \neq 0_{\mathbb{L}}$. Let A_b be the set of atoms below b: $A_b = \{a \in A: a \leq b\}$. Then $b = \sup(A_b)$.

The proof of (4) is left as an exercise. (See **Exercise ?? at the end of this Chapter.**)

In view of (1) and (4) we can define the following map h from $\mathbb{L}$ to $P(A),\subseteq>$: for b ε U such that $b \neq 0_{\mathbb{L}}$ $h(b) = \sup(A_b)$; and $h(0_{\mathbb{L}}) = \varnothing$. It is then easy to see that h is onto an that it transfers $\leq$ into the inclusion relation on $P(A)$.

Suppose now that $\mathbb{T}_{L,T}$ is infinite. Then because of (3) its atom set A must be infinite. Now let A' be any proper infinite subset of A. Since each element a of is finitely axiomatisable we can choose for each such a a single sentence A_a which axiomatises a. Let $T(A')$ be the theory of L which is the supremum of A' in $\mathbb{T}_{L,T}$. Then, since A' is a proper subset of A, there is at least one atom a that does not belong to A'. Then, as we have seen, $T(A') \leq -a$, so $T(A')$ consistent. But then $T(A')$ is not finitely axiomatisable. The argument is like that of Exercise 12 of Ch. 1. Let a_1, a_2, ... be an enumeration of all members of A'. Note that A' is denumerable. (Why?). Furthermore, let the sentences B_n be defined as follows: (i) $B_1 = A_{a_1}$; $B_{n+1} = B_n \,\&\, A_{a_{n+1}}$. Then it is easily verified (i) that the B_n are strictly increasing in logical strength - i. e. we have for all n that $B_{n+1} \models B_n$, but not $B_n \models B_{n+1}$ - and (ii) that $T(A') = Cl_L(\{B_n\}_n = 1,2, ..)$. So we can argue as in Exercise 12 of CH.1 that $T(A')$ is not finitely axiomatisable. But then, as shown in Exercise 21 of CH. 1, $T(A') \sqcup -T(A') \neq 1$. So $\mathbb{T}_{L,T}$ is not boolean.

This concludes the proof of our claim that when a lattice $\mathbb{T}_{L,T}$ is boolean, it must be finite. We record this claim once more, as part of the following more elaborate Theorem 5, which gives three additional equivalent conditions.

Thm. 5 Let T be a theory in some first order language L
Then the following five statements are equivalent:

(i) $\mathbb{T}_{L,T}$ is boolean.
(ii) T has finitely many complete extensions.
(iii) T has finitely many extensions. (i.e. $\mathbb{T}_{L,T}$ is finite.)
(iv) All of T's complete extensions are finitely axiomatisable over T.
(v) All of T's extensions are finitely axiomatisable over T.

The main work of the proof of Theorem 5 has been done above. What remains is left as an exercise.

Theorem 5 entails that boolean lattices of the form $\mathbb{T}_{L,T}$ are comparatively rare. They are found only 'at the upper end' of the set of all lattices $\mathbb{T}_L$, i.e. when T is cose to being complete. (The cases we have already mentioned, i.e. the lattices $\mathbb{T}_{L,T}$ where T is itself a complete theory, are the extreme examples of this.) In the next section we will look at some simple cases of boolean lattices of the form $\mathbb{T}_{L,T}$.

To get a clear picture of the structure of the lattices $\mathbb{T}_L$ for different languages L turns out to be a far from trivial problem. Only for the very simplest languages is it possible to describe the structure of $\mathbb{T}_L$ in fairly straightforward and readily understandable terms. This is so in particular for the language without any non-logical constants, {}. Already for the language {P} whose only non-logical constant is the 1-place predicate P, ta complete description proves to be considerably more involved. But a much higher degree of complexity is reached when the language contains predicates of 2 or more places or function constants whose arity is ≥ 1. There are all sorts of questions that can be asked here, for instance:

(a) What is the full range of isomorphism types of lattices $\mathbb{T}_L$ for various first order languages?

(b) How does the structure of $\mathbb{T}_L$ depend on L?

(c) Call two languages L_1 and L_2 *isomorphic* iff they have essentially the same signature; that is, if there is a bijection h of the set NLC_1 of non-logical constants of L_1 onto the set NLC_2 of non-logical constants of L_2 which preserves signature in that for any $\alpha \in NLC_1$, $L_1(\alpha) = L_2(h(\alpha))$.

Question: Are there (finite) non-isomorphic languages for which the corresponding theory lattices are isomorphic nevertheless? And if so, for which language pairs is this so?

To none of these questions do I have answers, and I do not know whether answers to them exist.

2.2.2. Tarski Lattices of some almost complete Theories

In this section we look at two examples of Tarski lattices $\mathbb{T}_{L,T}$ which are comparatively simple and tractable.

In the first example the theory T is the theory T_{den} of arbitrary dense linear orderings. One of the extensions of this the theory T_{rat} of the ordering of the rationals (or, what comes to the same thing: the theory of all dense linear orderings without beginning or end point) which we investigated in Section 2.1.1. Of T_{rat} we showed that it is ω-categorical, und thus, since it also has the property that all its models are infinite, complete.

T_{den} is axiomatised by the following axioms $T_{den}.0$ - $T_{den}.4$. $T_{den}.1$ - $T_{den}.4$ are from our earlier axiomatisation of T_{rat}; $T_{den}.0$ has been added in order to eliminate the degenerate order which consists of just one element. (In the case of T_{rat} this possibility was excluded by the presence of axioms L5 and L6, repeated below, which assert that there is no beginning and no end point, respectively-)

$T_{den}.0$ $\quad (\exists x)(\exists y)\ (x \neq y)$

$T_{den}.1$ $\quad (\forall x)(\forall y)\ (x < y \rightarrow \neg\ (y < x))$

$T_{den}.2$ $\quad (\forall x)(\forall y)(\forall z)\ ((x < y\ \&\ y < z) \rightarrow x < z)$

$T_{den}.3$ $\quad (\forall x)(\forall y)\ (x < y\ \ v\ \ x = y\ \ v\ \ y < x\ )$

$T_{den}.4$ $\quad (\forall x)(\forall y)\ (x < y \rightarrow (\exists z)\ (x < z\ \ \&\ \ z < y))$

L5. $\quad (\forall x)(\exists y)\ (x < y)$

L6. $\quad (\forall x)(\exists y)\ (y < x)$

Unlike T_{rat} T_{den} is of course not complete. But it is not far removed from that. It has a total of no more than four complete extensions. One of these is T_{rat}, which we get by adding the axioms L5 and L6. The other three are obtained by adding the other boolean combinations of these two axioms: (i) {¬ L5, L6}, (ii) {L5, ¬L6}, (iii) {¬L5, ¬L6}.

We denote the four extensions of T_{den} as(i) $T_{den}(+,+)$, (ii) $T_{den}(+,-)$, (iii) $T_{den}(-,+)$and (iv) $T_{den}(-,-)$. The + and - signs indicate the presence or absence of a first or last point. For instance, if the first sign is a plus, then the models of the theory all have a beginning point, and if it is - then all models don't. In other words, $T_{den}(+,+)$ is the theory we get by adding to T_{den} the axioms ¬ L5 and ¬ L6, and so on, In particular $T_{den}(-,-) = T_{rat}$.

That each of the theories $T_{den}(+,+)$, $T_{den}(+,-)$ and $T_{den}(-,+)$ is consistent and complete can be shown in the same way as we did this for T_{rat} in Section 2.1.1. In fact, since the rational interval (0,1) is one of the models of T_{rat} (Exercise: show this!), it follows from what was shown in Section 2.1.1 that every denumerable model of T_{rat} is isomorphic to (0,1). Using the same method we can also prove that [0,1), (0,1] and [0,1] are models of $T_{den}(+,-)$, $T_{den}(-,+)$ and $T_{den}(+,+)$, respectively, and that they are the only denumerable models of these theories up to isomorphism. So since each of the theories only has infinite models (Exercise: show this!), they are all complete as well as consistent.[17]

It is also easy to show that these are all the complete and consistent extensions of T_{den}. For suppose that T is any complete extension of T_{den} and that M is a model of T. M will either have or fail to have a first point and likewise it will either have or fail to have a last point. This gives a total of four possibilities, corresponding to the four boolean combinations of L5 and L6 mentioned above. In each case T is identical with the theory we get by adding this boolean combination to T_{den}. For instance, suppose that M has both a first and a last point. Then it will verify both ¬ L5 and ¬ L6. So these sentences are consistent with T, and so, since by assumption T is complete, they must belong to T. So T is the theory $T_{den}(+,+)$. Likewise for the other three possibilities.

This shows that the lattice $\mathbb{T}_{L,T_{den}}$ has exactly four 'anti-atoms'. So it also has exactly four atoms, which means that it consists of 2^4 theories altogether. Exercise: give explicit axiomatisations for each of the theories that make up $\mathbb{T}_{L,T_{den}}$!

2.2.3 Quantifier Elimination

---

[17] The same is true for the other three complete extensions of T_{den}. Consider for instance $T_{den}(+,-)$. The only complication which we have to deal with, when constructing martching tuples $<a_1, ..., a_n>$, $<b_1, ..., b_n>$ from two models M_1, M_2 of $T_{den}(+,-)$ is that if $<a_1, ..., a_n>$ contains the first element of M_1, and more precisely, if this first element is a_i, then b_i must be the first element of M_2, and conversely. That that is the only additional precaution we need to take in constructing the finite sequences $<a_1, ..., a_n>$, $<b_1, ..., b_n>$ and the isomorphiisms between them rests on the fact that all elements of M_1 (casu quo M_2) which are distinct from its first element are "infinitely far away from it" in the sense that there are infinitely many points between any such point and the first point (just as there are infinitely many points between any two distinct points of any model of T_{den},)

Our second example concerns the theory of discrete linear orderings. We will explore the Tarski lattice $\mathbb{T}_{L,T_{dis}}$, where T_{dis} is the theory defined below.

This exploration will be more involved than that of $\mathbb{T}_{L,T_{den}}$ in the last section, and that for two distinct reasons. First, $\mathbb{T}_{L,T_{dis}}$ is a more complex lattice than $\mathbb{T}_{L,T_{den}}$, although its complexity is still quite modest when compared with most Tarski lattices. But also - and this will be the bigger hurdle we will encounter - proving that the structure of the lattice is indeed what we will claim it to be, will prove a good deal more involved than tit was in the case of $\mathbb{T}_{L,T_{den}}$ and it will require a fundamentally different method. This is the method of quantifier elimination mentioned in the title to this section.

The base theory of our lattice, T_{dis}, is once more a theory of the language $L = \{<\}$. T_{dis} is axiomatised by the axioms $T_{dis}.0$ - $T_{dis}.5$. Not surprisingly there is a considerable overlap with the axioms of T_{dis}. For after all both theories deal with linear orderings. Consequently the first four axioms are the same, and divergence from T_{den} comes only with the discreteness axioms $T_{dis}.4$ and $T_{dis}.5$.

$T_{dis}.0$ $\quad (\exists x)(\exists y)\ (x \neq y)$

$T_{dis}.1$ $\quad (\forall x)(\forall y)\ (x < y \rightarrow \neg (y < x))$

$T_{dis}.2$ $\quad (\forall x)(\forall y)(\forall z)\ ((x < y\ \&\ y < z) \rightarrow x < z)$

$T_{dis}.3$ $\quad (\forall x)(\forall y)\ (x < y\ \vee\ x = y\ \vee\ y < x)$

$T_{dis}.4$ $\quad (\forall x)((\exists y)\ (x < y \rightarrow ((\exists y)\ (x < y\ \&\ \neg (\exists z)\ (x < z\ \&\ z < y))$

$T_{dis}.5$ $\quad (\forall x)((\exists y)\ (y < x \rightarrow ((\exists y)\ (y < x\ \&\ \neg (\exists z)\ (y < z\ \&\ z < x))$

T_{dis} is not complete and for much the same reasons as T_{den}: Nothing is said about the existence or non-existence of beginning or end points. Using the same notation that we resorted to in our discussion of T_{den}, we define the theories $T_{dis}(+,+)$, $T_{dis}(+,-)$, $T_{dis}(-,+)$ and $T_{dis}(-,-)$ to be those which we get by adding the boolean combinations of L5 and L6 described in the last section. (Thus $T_{dis}(+,+)$ is obtained by adding $\neg$ L5 and $\neg$ L6, etc.) All of these have, like the corresponding extensions of T_{den}, infinite models. In particular, $T_{dis}(+,-)$is satisfied by the ordering of the natural numbers, $T_{dis}(-,+)$ by the order of the negative integers,$T_{den}(-,-)$ by the order of the positive and negative integers and

$T_{dis}(+,+)$ by the structure which we get when we put the negative integers "behind" the natural numbers.[18]

There is however an important difference between $T_{dis}(+,+)$ and the other three: while the latter only have infinite models, $T_{dis}(+,+)$ has finite models as well. In fact, $T_{dis}(+,+)$ has models of cardinality n for all finite $n \geq 2$: any linearly ordered set of n elements will be a model of $T_{dis}(+,+)$.[19] On the other hand it is also clear that for each finite cardinality n there is essentially just one model for T_{dis} of that cardinality: Any two linearly ordered sets of n elements are (obviously) order-isomorphic; we can define, in the obvious way, an order-preserving correspondence between them. This means that if we add to $T_{dis}(+,+)$ a sentence which states that there are exactly n elements, then the resulting theory will have for its only models the linear orders of n elements. And since any two such orders are isomorphic, it follows that all these theories are complete.

In the spirit of the notation which we have been using, let us denote as $T_{dis}(+,+,n)$ the theories obtained by adding to $T_{dis}(+,+)$ a sentence saying that there are exactly n elements; and let us denote as $T_{dis}(+,+,\infty)$ the theory obtained by adding to $T_{dis}(+,+)$ the infinitely many sentences $D\geq n$ which say that there are at least n elements.

What can we say about the theories $T_{dis}(-,-)$, $T_{dis}(+,-)$, $T_{dis}(-,+)$ and $T_{dis}(+,+,\infty)$? The first pertinent observation is that unlike what we found for the corresponding extensions of T_{den}, these theories are not ω-categorical. Let us focus on $T_{dis}(+,-)$. One of its denumerably infinite models, we noted, is the set of the natural numbers with their natural order. But there are other denumerably infinite models too, and

---

18 More presicely, we can define this structure as the ordered disjoint union of these two structures, viz as the set of all pairs <0,n>, with $n \varepsilon N$ and all pairs <1,-n> with $n \varepsilon N$, with the ordering relation < defined by:

(i) $<0,n> < <0,m>$ iff $n <_N m$
(ii) $<1,-n> < <1,-m>$ iff $m <_N n$
(iii) $<0,n> < <1,'m>$ for arbitrary n, m

19 The requirement that $n \geq 2$ comes from $T_{dis}.0$, which we have retained from our axiomatisation of T_{den}. We could have dropped this axiom without changing much to the structure of $\mathbb{T}_{L,T_{dis}}$. The only effect would have been that the degenerate, one point ordering would have been included among the possible models of T_{dis}. This would have meant that in addition to the complete extensions of T_{dis} we are in the process of describing there would have been the extension which says that there is exactly one point.

as a rule these will not be isomorphic to the natural number structure. The simplest model of $T_{dis}(+,-)$ which is not isomorphic to the natural numbers is the structure that we obtain when we put a copy of Z (the negative and positive integers) behind a copy of the natural numbers. We can make this precise in the same way as we did for the infinite model we considered for $T_{dis}(+,+)$ described in footnote 16. That is we let M be the model $< U_M, <_M >$, where

(a) $U_M = \{<0,n>: n \in N\} \cup \{<1,z>: z \in Z\}$

(b) $<_M = \{<<0,n>:, <0,m>>: n <_N m\} \cup \{<<1,z>:, <1,y>>: z < y\}$
$\cup \{<<0,n>:, <1,z>>\}$

It is obvious that M is not isomorphic to the set N of natural numbers with their standard order. Just try to construct an isomorphism between N and M, starting with the 0 of N, 0_N. Obviously there is only one element of M on which an order isomorphism h from N to M could map 0, viz. M's first point <0,0>. In other words, it is necessarily the case that $h(0_N) =<0,0>$. Likewise the number 1 of N, 1_N, which is the immediate successor of 0 in N, can only be mapped onto the immediate successor <0,1> of $<0_M$ in M. That is, we must have $h(1_N) = <0,1>$. In the same way the structure of N and M fixes the images under h of all the other elements of N. This means that, when N has been exhausted - i.e. h has been defined for all of N - only the "N-part" of M (consisting of the pairs of the form <0,n>) has been covered in the range of h.

The non-isomorphism of N and M entails that the completeness of $T_{dis}(+,-)$ cannot be established by the simple technique which we used to prove Cantor's theorem (the ω-categoricity of $T_{den}(-,-)$) in Section 2.1.1 and which would also be applied to the three other extensions of T_{den} which we considered in the last section. Nevertheless, $T_{dis}(+,-)$ is complete and the same is true of the remaining three extensions of T_{dis} which have infinite models, $T_{dis}(-,-)$, $T_{dis}(-,+)$ and $T_{dis}(+,+,\infty)$. But the proof that they are complete is harder than the Cantor-type proofs for the corresponding extensions of T_{den}. We will give the proof for the case of $T_{dis}(+,-)$. The proofs that the three other theories are complete are virtually identical.

In presenting the proof that $T_{dis}(+,-)$ is complete we will proceed as follows. We first focus on the concrete task before us. We show that any two models of $T_{dis}(+,-)$ are elementary equivalent. This argument will reveal the general features of the method used (that of quantifier

elimination). Un the next section we will then describe and discuss the method of quantifier elimination in general.

Recall the basic architecture of Cantor's proof. We considered two models $M_1 = < U_1, <_1 >$ and $M_2 = < U_2, <_2 >$ of $T_{den}(-,-)$ and constructed, by going back and forth between the universes U_1 and U_2, ever longer matching n-tuples $<a_1, ..., a_n>$ of elements from U_1 and $<b_1, ..., b_n>$ of elements from U_2, which were order-isomorphic. Because of the special properties of dense linear orderings it proved to be always possible to match a new element a_{n+1} chosen from U_1 by a new element b_{n+1} from U_2 which stood in exactly the same order relations to each of the b_i ($i = 1, ..., n$) as a_{n+1} stood to each of the a_i; and conversely. For models of theories of discrete orderings - among them the models for $T_{dis}(+,-)$ - the situation is different. Here the "distance" between two points - i.e. the number of points between them - can be either finite or infinite; and the distance could involve any finite number n of intermediate points. The model N is special among the models of $T_{den}(+,-)$ in that the distance between two of its elements is always finite. But in this respect it is unique. Any model of $T_{den}(+,-)$ which is not isomorphic to N will have points that are infinitely far from each other. (This is true in particular for the model we considered above, in which a copy of N is followed by a copy of Z. In this model there is an infinite distance between any two elements $<0,n>$ and $<1,z>$.)

A consequence of this is that when we consider a formula A of our language and two tuples $<a_1, ..., a_n>$, $<b_1, ..., b_n>$ belonging to two models M_1, M_2 and ask whether A gets the same truth value in M_2 under the assignment provided by $<b_1, ..., b_n>$ that it gets in M_2 under the assignment provided by $<a_1, ..., a_n>$, then we will have to take into account the quantifier complexity of A: It will depend on this complexity how similar $<a_1, ..., a_n>$ and $<b_1, ..., b_n>$ will have to be in order that we can be certain that they confer upon A the same truth value in their respective models M_1 and M_2. A few simple examples will illustrate this.

First consider a quantifier-free formula, e.g. $v_1 < v_2$. Let M_1, M_2 be models of $T_{dis}(+,-)$ and let $<a_1,a_2>$, $<b_1,b_2>$ be ordered pairs of elements of M_1 and M_2 which are order-isomorphic to each other, i. e. $a_1 <_{M_1} a_2$ iff $b_1 <_{M_2} b_2$. Then clearly $M_1 \models (v_1 < v_2)[a_1,a_2]$ iff

$M_2 \models (v_1 < v_2)[b_1,b_2]$. The same holds for any other quantifier-free formulas such as $v_1 < v_2 \,\&\, v_2 < v_3$, $v_1 < v_2 \,\&\, \neg(v_2 < v_3)$, etc, etc. This is just as in the case of dense orderings.

As soon as A contains quantifiers, however, the mere order isomorphism between $<a_1, ..., a_n>$ and $<b_1, ..., b_n>$ will no longer suffice. For example, let A be the formula $(\exists v_2)(v_1 < v_2 \,\&\, v_2 < v_3)$. Suppose that M_1 and M_2 are both the natural number structure N and that $<a_1, a_2> = <4,7>$ and $<b_1,b_2> = <8,9>$. Then $<a_1,a_2>$ and $<b_1,b_2>$ are order-isomorphic; yet $N \models A[a_1,a_2]$, while on the other hand not $N \models A[b_1,b_2]$. The source of the problem is obvious. A says something about the distance between the points represented by its free variables v_1 and v_3, viz. that there is at least one point between them. This is a condition which a mere order isomorphism need not preserve. And that is precisely what we see in our example: $<a_1,a_2>$ and $<b_1,b_2>$ are both order-isomorphic, but the point pair $<a_1,a_2>$ satisfies the condition that there is at least one point between them whereas the pair $<b_1,b_2>$ does not. In other words, in order to be sure that two pairs $<a_1,a_2>$ and $<b_1,b_2>$ confer upon A the same truth value, they must not just be order-isomorphic, but stand in some tighter relationship, which also involves information about how many points there are between them.

As we move to formulas A more quantifiers even stronger similarity relations must hold between $<a_1,a_2>$ and $<b_1,b_2>$ to guarantee that a_1 and a_2 satisfy A in M_1 iff b_1 and b_2 satisfy A in M_2. This is because with more quantifiers we can say more about the number of points between two given points a_1 and a_2. For instance, with two quantifiers, but not with just one, it is possible to say that there are at least two points between a_1 and a_2; and so on. And the same goes, more generally, for formulas A with free variables $x_1, ..., x_n$: Ever stronger relations must hold between an n-tuple $<a_1, ..., a_n>$ of elements from M_1 and an n-tuple $<b_1, ..., b_n>$ of elements from M_2 in order to guarantee that $<a_1, ..., a_n>$ satisfies A in M_1 iff $<b_1, ..., b_n>$ satisfies A in M_2.

It would be convenient iff we could define a relation between tuples $<a_1, ..., a_n>$ and $<b_1, ..., b_n>$ such that any two tuples standing in this relation will confer the same truth value on all formulas. But often - this is true of our present problem but also for many others - there is no direct way of defining such a single relation; all that can be done is

to define a hierarchy $\equiv_1, \equiv_2$, ... of ever tighter relations between n-tuples so that whenever $\langle a_1, ..., a_n\rangle \equiv_k \langle b_1, ..., b_n\rangle$, then $\langle a_1, ..., a_n\rangle$ and $\langle b_1, ..., b_n\rangle$ confer the same truth values on all formulas whose *quantifier depth* is $\leq$ k. By the *quantifier depth* of a formula A we understand the maximal degree of nesting of quantifiers in A. There is no particular difficulty in defining this notion for arbitrary formulas. But it is somewhat more convenient to limit our attention to prenex formulas. For a formula A in prenex form the *quantiifer depth* of A is simply the number of quantifiers in its quantifier prefix. Since every formula is logically equivalent to a formula in prenex normal form, satisfaction preservation of all prenex formulas will entail preservation of all other formulas.

For the argument below it will be also convenient to asssume a slightly different form for prenex formulas, one in which the prefix contains only existential quantifiers but no universal ones. We can obtain such a prefix from a standard prefix by replacing every universal quantifier $(\forall v_i)$ in the standard prefix by the equivalent combination $\neg(\exists v_i)\neg$. So the formulas with which we will be concerned will always begin with a (possibly empty) prefix consisting of existential quantifiers and negation signs, followed by a quantifier-free formula. The *quantifier depth* of such a formula is then the number of existential quantifiers in its prefix.

In (1) we repeat for further reference the basic requirement we have already stated on the relations $\equiv_k$.

(1) Let M_1 and M_2 be models of $T_{dis}(+,-)$. And let A be any formula of quantifier depth $\leq$ k whose free variables are among $x_1, ..., x_n$. Then for any n-tuples $\langle a_1, ..., a_n\rangle$ and $\langle b_1, ..., b_n\rangle$ of elements chosen from M_1 and M_2, respectively, such that $\langle a_1, ..., a_n\rangle \equiv_k \langle b_1, ..., b_n\rangle$, $M_1 \models A[a_1,..,a_n]$ iff $M_2 \models A[b_1,.., b_n]$.

We already know what is required of the relation $\equiv_o$, which according to (1) should guarantee that if $\langle a_1, ..., a_n\rangle \equiv_o \langle b_1, ..., b_n\rangle$, then $\langle a_1, ..., a_n\rangle$ and $\langle b_1, ..., b_n\rangle$ satisfy the same formulas of quantifier depth 0 (i.e. the same quantifier-free formulas). This requires that the function h, given by the condition: $h(a_i) = b_i$, is an order isomorphism between (the submodels of M_1 and M_2 determined by) $\{a_1, ..., a_n\}$ and $\{b_1, ..., b_n\}$, respectively. We define $\equiv_o$ accordingly:

(2) Let $<a_1, ..., a_n>$ and $<b_1, ..., b_n>$ be n-tuples of elements chosen from models M_1 and M_2, respectively. Then

$<a_1, ..., a_n> \equiv_0 <b_1, ..., b_n>$ iff

the function h given by: "for i = 1,2, .., n, $h(a_i) = b_i$" is an order isomorphism between the submodels of M_1 and M_2 whose universes are $\}a_1, ..., a_n\}$ and $\{b_1, ..., b_n\}$.

A second requirement on the relations $\equiv_k$, which is imposed by the strategy we will follow to show that two models M_1 and M_2 of $T_{dis}(+,-)$ are elementarily equivalent, is that successive relations $\equiv_k$ and $\equiv_{k+1}$ stand in the following relation:

(3) Suppose that M_1 and M_2 are as under (1), that, for arbitrary number n, $<a_1,..,a_n>$, $<b_1,..,b_n>$ are n-tuples of elements of M_1 and elements of M_2, respectively and that $<a_1,..,a_n> \equiv_{k+1} <b_1,..,b_n>$. Then

i. if a is any element of M_1, then there is an element b of M_2, such that $<a_1,..,a_n,a> \equiv_k <b_1,..,b_n,b>$.

ii. if b is any element of M_2, then there is an element a of M_1, such that $<a_1,..,a_n,a> \equiv_k <b_1,..,b_n,b>$.

From (2) and (3) we can derive that the condition (1) holds for all formulas of the special prenex form described above, in which a quantifier-free part is preceded by a string of existential quantifiers and negations. We repeat this restricted version of (1) as (1') below. Since every formula can be transformed into a logically equivalent formula of this special form, (1') entails (1).

(1') Let M_1 and M_2 be models of $T_{dis}(+,-)$. And let A be any prenex formula with a prefix consisting of existential quantifiers and negation signs, that A has quantifier depth $\leq$ k and that its free variables are among $x_1, ..., x_n$. Then for any n-tuples $<a_1, ..., a_n>$ and $<b_1, ..., b_n>$ of elements chosen from M_1 and M_2, respectively, such that $<a_1, ..., a_n> \equiv_k <b_1, ..., b_n>$,

$$M_1 \models A[a_1,..,a_n] \text{ iff } M_2 \models A[b_1,.., b_n].$$

To derive (1') from (2) and (3) we argue by induction on the complexity of such formulas. The base case is constituted by quantifier free formulas. Strictly speaking, this reuires an inductive proof in its own right: First, when $\langle a_1, ..., a_n \rangle \equiv_o \langle b_1, ..., b_n \rangle$ and A is an atomic formula - that is, A is either of the form "$v_i = v_j$" or of the form "$v_i < v_j$", then obviously A is satisfied by $\langle a_1, ..., a_n \rangle$ in M_1 iff it is satisfied by $\langle b_1, ..., b_n \rangle$ in M_2. The inductive step then consists in showing that the condition in (1') holds for B and for C then it holds for $\neg$ B, B & C, and likewise for the other sentence connectives. But this is trivial.

The inductive step makes use of (3). Suppose that (1') holds for formulas of quantifier depth $\leq$ k and that A is a formula in our special prenex form and is of quantifier depth k + 1. If A begins with a negation sign - i.e. A is of the form $\neg$ B, where B too has our special prenex form, then the result will hold for A provided it holds for B. Let us assume therefore that A begins with an existential quantifier, i.e. A is of the form $(\exists x)B$. Suppose then that the free variables of B are among $v_1, ..., v_n$ and that $\langle a_1, ..., a_n \rangle \equiv_{k+1} \langle b_1, ..., b_n \rangle$. Without loss of generality we may assume x is the variable v_{n+1}. (We do not really need this assumption, but it simplifies notation.) Assume that $M_1 \models A[a_1,..,a_n]$. Then there is an element a of M_1 such that $M_1 \models B[a_1,..,a_n,a]$. Given (3) we can find an element b in M_2 such that $\langle a_1, ..., a_n,a \rangle \equiv_k \langle b_1, ..., b_n,b \rangle$. By induction hypothesis $M_2 \models B[b_1,..,b_n,b]$. So it follows that $M_2 \models (\exists x)B[b_1,..,b_n]$. In the same way we show that if $M_2 \models A[b_1,..,b_n]$, then $M_1 \models A[a_1,..,a_n]$.

This concludes the argument that (1) provided that we can define a sequence of relations $\equiv_o, \equiv_1, \equiv_2,..$ such that $\equiv_o$ is the relation defined in (2) and successive relations $\equiv_k, \equiv_{k+1}$ satisfy (3). In the present case . the one concerning the theory $T_{dis}(+,-)$ - the relations can be given by independent explicit definitions. (In other applications of the quantifier elimination method their definition may be more complicat4ed and require itself an induction on k.) The definitions are given in (4)

(4) Let M_1 and M_2 be models of $T_{dis}(+,-)$. Let a_{beg} be the first element of M_1in the sense of its order relation $<_{M_1}$ - there must be a unique such element since M_1 is a model of $T_{dis}(+,-)$ - and let b_{beg} be the first element of M_2. Let $\langle a_1, ..., a_n \rangle$ and $\langle b_1, ..., b_n \rangle$ be n-tuples from M_1 and M_2, respectively.

Then $\langle a_1, ..., a_n \rangle \equiv_k \langle b_1, ..., b_n \rangle$ iff the following conditions are fulfilled:

(i) $\langle a_1, ..., a_n \rangle$ and $\langle b_1, ..., b_n \rangle$ are order-isomorphic.
(For simplicity we assume, as we have been all along, that their elements have been arranged "in order of magnitude", i.e. $a_1 <_{M_1} a_2$, etc. and similarly for the elements of $\langle b_1, ..., b_n \rangle$)

(ii) For any pair of successive elements a_i, a_{i+1} from the first tuple and corresponding pair b_i, b_{i+1} from the second we have either (a) or (b):

(a) the number of elements) between a_i and a_{i+1} in M_1 and that between b_i and b_{i+1} in M_2 are both $< 2^k$ and they are identical;

(b) the number of elements between a_i and a_{i+1} in M_1 and that between b_i and b_{i+1} in M_2 are both $\geq 2^k$.

(iii) For the elements a_1 and b_1 we have either (c) or (d):

(c) the number of elements between a_1 and a_{beg} and that between b_1 and b_{beg} are both $< 2^k$ and they are identical;

(d) number of elements between a_1 and a_{beg} and that between b_1 and b_{beg} are both $\geq 2^k$.

N.B. For the case where $k = 0$ condition (ii) is vacuous, since the first possibility they mention - of the distances between a_i and a_{i+1} and between b_i and b_{i+1} being $< 2^o$ - cannot arise. Similarly condition (ii) is vacuous, So only (i) matters and thus the specification that (4) provides of $\equiv_o$ coincides with that given in (2).

It remains to show that the relations of (4) satisfy (3). Suppose that $\langle a_1, ..., a_n \rangle \equiv_{k+1} \langle b_1, ..., b_n \rangle$. We have to show that for any element a of M_1 there is an element b of M_2 such that $\langle a_1, ..., a_n, a \rangle \equiv_{k+1} \langle b_1, ..., b_n, b \rangle$ and conversely. we only consider the first half. Let a be any element of U_{M1}. There are three possibilities to be considered:

(i) $a <_{M_1} a_1$;
(ii) $a_i <_{M_1} a <_{M_1} a_{i+1}$ for some $i < r$

(iii) $a_n <_{M_1} a$.

Assume (i). Let $D(a_{beg},a_1)$ be the number of elements between a_{beg} and a_1. Then either (c) $D(a_{beg},a_1) < 2^{k+1}$ or (d) $D(a_{beg},a_1) \geq 2^{k+1}$. First suppose (c). Since the number of elements in M_2 between b_{beg} and b_1, $D(b_{beg},b_1)$, is the same as $D(a_{beg},a_1)$, we can pick as the b required by (3) that element of M_2 which lies just as many elements before b_1 in M_2 as a lies before a_1 in M_1. Then the distance between b and b_1 is the same as that between a and a_1 and the same is true for the distance between the b and b_{beg} and the distance between a and a_{beg}. So $<a_1,..,a_n,a> \equiv_k <b_1,..,b_n,b>$.

Now suppose that both $D(a_{beg},a_1)$ and $D(b_{beg},b_1)$ are $\geq 2^{k+1}$. First suppose that the distance between a and a_1 is $< 2^k$. Then we pick from M_2 the element b which lies before b_1 at just the same distance that a lies before a_1 in M_1. This guarantees that there are as many elements between a and a_1 in M_1 as there are between b and in b_1 in M_2. Moreover, since by assumption the distance between a_{beg} and a_1 and that between b_{beg} and b_1 are both $\geq 2^{k+1}$, the distance between a_{beg} and a and that between b_{beg} and b will be both $\geq 2^k$. So again we have that $<a_1,..,a_n,a> \equiv_k <b_1,..,b_n,b>$.

The second possibility to be considered is that where $D(a_{beg},a) < 2^k$. Then we pick the element b of M_2 which lies at that same distance from b_{beg}. This time $D(a_{beg},a_1)$ and $D(b_{beg},b_1)$ are both $\geq 2^k$. So again $<a_1,..,a_n,a> \equiv_k <b_1,..,b_n,b>$.

The third possibility we must consider for the position of a before a_1 is that where both $D(a_{beg},a)$ and $D(a,a_1)$ are $\geq 2^k$. In this case the fact that $D(a_{beg},a_1)$ is $\geq 2^{k+1}$ guarantees that we can pick an element b from M_2 such that $D(b_{beg},b)$ and $D(b,b_1)$ are both $\geq 2^k$. Again $<a_1,..,a_n,a> \equiv_k <b_1,..,b_n,b>$.

This completes case (i), in which a lies before a_1 in M_1. We leave the other two cases - that where a lies between a_i and a_{i+1} for some $i < 1$ and that where a lies beyond a - to the reader, and thus reach the end of the argument that if $<a_1, ..., a_n> \equiv_{k+1} <b_1, ..., b_n>$, then for any choice of an element a from M_1 we can make a matching choice of an

element d from M_2 such that $<a_1,..,a_n,a> \equiv_k <b_1,..,b_n,b>$[20] and therewith the proof that definition (4) entails (3).

We have now proved (1'). One step remains towards the conclusion that M_1 and M_2 are elementarily equivalent. But this is straightforward. Suppose that A is any sentence and that A' is formula in our prenex form that is logically equivalent to A. Then A' may be assumed to also be a sentence. Suppose that A' has quantifier depth k. In order that $M_1 \models A'$ iff $M_2 \models A'$ we need to show that the empty sequence $< >$ of elements of M_1 satisfies A' in M_1 iff the empty sequence $< >$ of elements of M_2 satisfies A' in M_2. According to (1') this will be the case, provided these two sequences stand in the relation $\equiv_k$. But it is obvious from def. (4) that the empty sequences of elements of M_1 and M_2 trivially satisfy this requirement.

q.e.d.

We have now proved that any two infinite models of $T_{dis}(+,-)$ are elementarily equivalent. Since $T_{dis}(+,-)$ only has infinite models, $T_{dis}(+,-)$ is complete. The argument is much like the one justifying Vaught's Test. (See Ch. 1, **Theorem ??**.) Suppose that $T_{dis}(+,-)$ were not complete. Then there would be a sentence A such that neither A nor $\neg$ A belong to $T_{dis}(+,-)$. So both $T_{dis}(+,-) \cup \{A\}$ and $T_{dis}(+,-) \cup \{\neg A\}$ are consistent. So each of them has a model. Both models must be infinite. So, because of the Downward Skolem-Löwenheim theorem, we may assume that they are both denumerably infinite. So, since they are both models of $T_{dis}(+,-)$, it follows from what we have just proved that they are elementarily equivalent. This contradicts the assumption that the first model verifies A and the second $\neg$A.

By the same method that we have used to prove that $T_{dis}(+,-)$ is complete we can also prove completeness for the three remaining theories, $T_{dis}(-,+)$, $T_{dis}(-,-)$ and $T_{dis}(+,+,\infty)$. This rounds off our survey of the complete consistent extensions of T_{dis}: There are four extensions whose models are infinite and denumerably many - the theories $T_{dis}(+,+,n)$ - whose models are of cardinality n. These latter

---

[20] N. B. the tuples $<a_1,..,a_n,a>$ and $<b_1,..,b_n,b>$ are not necessarily arranged in order of magnitude, even if this was true for the tuples $<a_1,..,a_n>$ and $<b_1,..,b_n>$, since the new elements a and b But of course we can rearrange the elements of $a_1,..,a_n,a>$ and $<b_1,..,b_n,b>$ so that their order in hte tuples reflects there order in the sense of M_1 and M_2.

theories are absolutely categorical - any two models of $T_{dis}(+,+,n)$ are isomorphic - whereas the first four are complete but not ω-categorical.

Since $\mathbb{T}_{L,T_{dis}}$ is infinite, it follows from Thm. 5 that it is not boolean. The trouble maker is $T_{dis}(+,+,\infty)$. All other complete extensions of T_{dis} are finitely axiomatisable over T_{dis} (and in fact, since T_{dis} is finitely axiomatisable itself, finitely axiomatisable simpliciter). From this and the infinity of $\mathbb{T}_{L,T_{dis}}$ we can conclude that the one remaining complete theory of $\mathbb{T}_{L,T_{dis}}$, viz. $T_{dis}(+,+,\infty)$, is not finitely axiomatisable. (This is a result that we can also easily derive directly, making use of the particular axioms - those of $T_{dis}(+,+)$ together with the difference axioms D_n - which we have given, but we get it from Thm. 5 "for free".

Exercise. Determine which extension of T_{dis} is the complement $-T_{dis}(+,+,\infty)$ of $T_{dis}(+,+,\infty)$ relative to T_{dis}. (In particular, give an explicit axiomatisation for $-T_{dis}(+,+,\infty)$.)

The purpose of this section has been two-fold. On the one hand it is meant as counterpoint to our investigation of the much simpler lattice $\mathbb{T}_{L,T_{den}}$ in Section 2.2.2. As we noted earlier, the lattice $\mathbb{T}_{L,T_{dis}}$ of this section is still of modest complexity when compared with the Tarski lattices for most languages and theories. But it is nevertheless significantly more complex than $\mathbb{T}_{L,T_{den}}$. Crucially, $\mathbb{T}_{L,T_{den}}$ is boolean while $\mathbb{T}_{L,T_{dis}}$ is not.

However, the section also has served a second, more general purpose, that of introducing the method of Quantifier Elimination. The general method is contained in the argument we have given for the inductive step in the proof of (1') from condition (3). This argument is fully general in that it makes no use of any special properties of the models for T_{dis}. To turn that arument into a proof that any two models of T_{dis} are elementarily equivalent we needed in addition (i) a definition of the relations $\equiv_k$ together with (ii) a proof that that the relations thus defined satisfy (3) and (ii) a proof that $\equiv_o$ satisfies condition (2) for quantifier free formulas. In each application of the method of Quantifier Elimination (i)-(iii) have to be dealt with anew, in a way which reflects the special properties of the problem to which it is being applied. But the general architecture is always the same. The next section contains some further general reflections about this method and some remarks about its history.

One final remark on the nature of our investigations in the last three sections (Sections 2.2.1 -2.2.3). On the one hand these invesitgations can be seen as a continuation of the exploration of first order theories of boolean and other lattices which we started in Section 2.1.2. From this pointt of view there is no fundamental difference between our exploration of Tarski and Lindenbaum lattices in the last four sections and, say, our look at the two boolean algebras of Section 2.1.4. But there is also another point of view from which what we have been doing from Section 2.2 onwards is importantly different from what precedes it. In these last sections we have been applying the formal tools of analysis - that of investigating structures as models of first order theories - to the structure of those tools themselves. In other words, here we have one example of the situation described informally in Sections 1.3.2 and 1.3.3 of Ch. 1: the possibility and potential usefulness of applying the tools of formal logic to the structures of formal logic - its expressions, languages and theories - themselves. As announced in Ch. 1 we will another instance of this in Ch. 3 when we develop set theory as a first order theory. While there are many important differences between what we will do in Ch. 3 and the explorations of the last three sectons, they nevertheless have in common that both show the methods of formal logic can be made into their own topic.

2.2.4 Why "Quantifier Elimination"?

N.B. The following section - is mostly of historical interest and can be skipped without any loss to the substance of these Notes.

The term "quantifier elimination" refers originally to a method which it describes perfectly: To show that all sentences A of a given language L have a certain semantic property which involves truth in certain Models or classes of models, show that in relation to the models M in question every sentence A is equivalent to a quantifier-free sentence A', in the sense that for each such model M we have $M \models A$ iff $M \models A'$. In the simplest cases where quantifier elimination is possible in this sense, the quantifier-free formulas A' are formulas of the very language L one starts out with. but very often the method isn't applicable in this simple form. Quantifier-free equivalents for sentences with quantifiers can be found, but only in some extension L´ of L. Typically L' is that where is a *definitional extension of* L in the following sense. Each new non-logical constant α of L' is defined by a formula ϕ_α of L, with as many free variables as α has arguments. Thus, if α is an n-place

predicate, then ϕ_α has the free variables $v_1,..., v_n$. (Function constants present an additional complication, which is not directly relevant here. So we leave them out of consideration. If necessary, n-place function constants can always be "simulated" as n+1-place predicates.) The definitions of the new constants of L' provide us with a way of expanding any model for L to a model for L': If M = <U,I> is a model for L, α a new n-place predicate of L' and ϕ_α the definition of α, then the interpretation function I' of the expansion M' of M will assign α the set of all n-tuples $<a_1,..., a_n>$ of elements of M such that $M \models \phi_\alpha[a_1,..., a_n]$. This transforms in particular each of the models which determine the notion of equivalence relevant to the given application into corresponding L'-models.

The defining formulas ϕ_α will often contain quantifiers. When this is so, the term "quantifier elimination" for the existence, for each sentence A of L, of a quantifier-free formula in L' is easily somewhat misleading. for by permitting in the "quantifier-free" formula A' of L' that is equivalent to A predicates that are defined by quantified formulas of L we allow quantification to sneak back in as it were, and A' should be considered as "quantifier-free" only in an attenuated sense. In fact, when we translate A' back into L by replacing all occurrences in it of new predicates by their definitions in L, then we will in general get a formula A'' which does contain quantifiers. The point of the method in these cases is that while A'' does contain quantifiers, it contains them only in quite special configurations, and this is what makes it (or, equivalently, the formula A' from which A'' is obtained) behave in ways that are relevantly similar to the behaviour of the quantifier-free formulas of L. In particular - this is the crucial point here - A'' ought to behave much like a quantifier-free formula with regard to the questions of the form: "Does $M \models A''[a_1,.., a_n]$?", where M is one of the relevant L-models and $<a_1,..., a_n>$ an n-tuple of elements from M (assuming that the free variables of A'' are among $v_1,.., v_n$). For instance, when the issue is to show that two such models M_1, M_2 are elementarily equivalent, then it should be true that $M_1 \models A''[a_1,.., a_n]$ iff $M_2 \models A''[b_1,.., b_n]$, where $a_1,.., a_n$, $b_1,.., b_n$ are from M_1, M_2, respectively, and $<a_1,.., a_n> \equiv_0 <b_1,.., b_n>$ $\equiv_0$ is some relation of moderate complexity, and we should be able to prove that.

In fact, the use of quantifier elimination in this sense for the purpose of proving elementary equivalence may involve much more complicated arguments than the one that was needed in the proof above to establish the truth of condition (3).

The method of quantifier elimination in this form becomes particularly involved in those cases where it is not only necessary to extend the language L with which one starts to a larger language L', but where L must be extended with infinitely many new predicates, The definitions of these predicates will necessarily be of increasing complexity, and in particular of increasing quantifier complexity.[21]

About the simplest illustration of quantifier elimination in the literal sense of the term concerns the theory T_{rat}, to which we applied the method of Cantor's proof in Section 2.1. The simplicity of the proof that any two denumerable models of this theory are isomorphic is directly reflected in the ease with which the quantifier elimination method is applied in this instance. In particular, it is not necessary in this case to extend the language {<} of the theory to a larger language.

We begin by considering quantifier-free formulas of L in the variables $v_1, ..., v_n$. We think of these variables as designating points of some dense linear order. Among formulas of this kind there are in particular those which fully describe the order relations between these points, and also say which variables are to be seen as designating the same point. Any formula A of this kind can be written in a form which is the conjunction of three conjunctions A_1, A_2, A_3, which can be dscribed as follows.

(i) A_1 is a conjunction of equations of the form $v_i = v_j$ $(i < j \leq n)$. These give us all combinations of variables v_i, v_j which, according to the situation described by A, designate the same point.

(ii) A_2 is the conjunction of all formulas of the form $v_i \neq v_j$ $(i < j \leq n)$ such that $v_i = v_j$ is not a conjunct of A_1.

(iii) Let $x_1, ..., x_m$ $(m \leq n)$ be all those variables v_j from $\{v_1, ..., v_n\}$ such that A_1 contains no equation of the form $v_i = v_j$. Then A_1 is a conjunction of formulas $x_i < x_j$ which completely fixes a linear order between the x's.

It is easy to see (a) that any such conjunction A is consistent with

---

[21] If L is finite (i.e. has only finitely many non-logical constants), then only finitely many non-equivalent predicates of a given arity can be defined in L if we only consider defining formulas whose quantifier depth does not exceed some given fintite number k.

T_{rat} in that we can find a model M for L and objects a_1, ..., a_n of M such that $M \models T_{rat}$ and $M \models A[a_1, ..., a_n]$, and (b) A is maximal in the sense that if we take any other quantifier-free formula B of L in v_1, ..., v_n, such that B is consistent with $T_{den}(-,-)$, then either $T_{rat} \models (\forall v_1)...(\forall v_n)(A \rightarrow B)$ or $T_{rat} \models (\forall v_1)...(\forall v_n)(A \rightarrow \neg B)$; and, finally, (c) any quantifier-free formula of B L in v_1, ..., v_n that is maximal consistent in the sense above is equivalent to an A of the kind described, i.e. there is an A as described such that $T_{rat} \models (\forall v_1)...(\forall v_n)(A \leftrightarrow B)$.

(a), (b) and (c) together entail that any quantifier-free formula B of L in v_1, ..., v_n which is consistent with T_{rat} is equivalent modulo T to some disjunction $\bigvee_i A_i$ of conjunctions A_i of the described kind:

$$T_{rat} \models (\forall v_1)...(\forall v_n)(B \leftrightarrow \bigvee_i A_i).$$

We can generalise to the case of inconsistent formulae B by stipulating that they are equivalent to some fixed logical contradiction $\perp$, identifying $\perp$ with the "empty disjunction" of formulas.

Now let A be an arbitrary sentence of L in the kind of prenex form used in Section 2.2.3 - i.e. one whose prefix consists of existential quantifiers and negations - and let us assume that the matrix B of A is given as a disjunction $\bigvee_i A_i$ of maximal conjunctions A_i of the kind we have described. Without loss of generality we may assume that the matrix is immediately preceded by an existential quantifier $(\exists v_n)$. (In case the last element of the prefix is a negation sign, this negation can be moved towards the inside of the matrix formula and the resulting formula rewritten once more as a disjunction $\bigvee_i A_i$.) We first observe that $(\exists v_n)(\bigvee_i A_i)$ is logically equivalent to $\bigvee_i (\exists v_n)A_i$. Now consider any one of the disjuncts A_i. Let A'_i be the formula which we obtain from A_i by eliminating from it all conjuncts which contain v_n.
Claim: $T_{rat} \models (\exists v_n)A_i \leftrightarrow A'_i$. First the implication from left to right. This is a theorem of predicate logic. For (i) $\models A_i \rightarrow A'_i$, since in going from A_i to A'_i we have only thrown out conjuncts; (ii) since v_n does not occur in A'_i, (i) entails that A'_i also follows logically from the existential quantification $(\exists v_n)A_i$ of A_i. For the opposite direction we have to distinguish between several cases. First, suppose that v_n occurs in A_i in a conjunct of the form $v_j = v_n$. Then v_n will occur in A_i only in conjuncts that have the form of equations. So in this case, adding these

conjuncts again to A'_i and then quantifying existentially over v_n yields a formula which is entailed by A'_i, and this formula is (obviously equivalent to) $(\exists v_n)A_i$. Second suppose that v_n does not occur in A_i in a conjunct of the form $v_j = v_n$. Then v_n will occur in at least one conjunct involving $<$. There are three cases to be considered here:

(i) v_n occurs only in conjuncts of the form $v_n < v_j$. Then A_i describes v_n as the first element among its "points". In particular, v_n is described as lying before the point which is described by A'_i as the first of the points designated by $v_1, ..., v_{n-1}$. Let v_j be the variable (or one of the variables) designating this first point of the order described by A'_i. Since $T_{rat} \models (\forall v_j)(\exists v_n)\ v_n < v_j$, we also have that $T_{rat} \models A'_i \rightarrow (\exists v_n)A_i$.

(ii) The second possibility is that v_n occurs in A_i both in conjuncts of the form $v_n < v_j$ and in conjuncts of the form $v_j < v_n$. In that case there will be two variables v_j and v_k such that A_i entails that v_j, v_n and v_k are adjacent in the order it describes. This time we make use of the fact that $T_{rat} \models (\forall v_j)(\forall v_k)(v_j < v_k \rightarrow (\exists v_n)(v_j < v_n \ \& \ v_n < v_k))$ to see that $T_{rat} \models A'_i \rightarrow (\exists v_n)A_i$.

(iii) The third case is that where A_i only contains conjuncts of the form $v_j < v_n$. This case is just like case (i).

This competes the argument that

(7) $T_{rat} \models (\exists v_n)A_i \leftrightarrow A'_i$.

(7) entails that when we replace $(\exists v_n)A_i$ by A'_i in A, we obtain a sentence which is equivalent to A, but in which the quantifier $(\exists v_n)$ no longer occurs. In an analogous way we can eliminate all quantifiers of A but one. At this point we have a sentence C equivalent to A modulo T_{rat} which contains one quantifier $(\exists x)$, with or without a negation sign in front of it and some quantifier-free formula D following it in which the only variable is x. It is easy to verify by checking the small number of different forms that D can take that either $T_{rat} \models (\exists x)D$ or $T_{rat} \models \neg (\exists x)D$. Then we also have: $T_{den(-,-)} \models C$ or $T_{rat} \models \neg C$. So in particular we have $T_{rat} \models A$ or $T_{rat} \models \neg A$. This shows the completeness of $T_{den(-,-)}$ and by the same token the fact that modulo it every formula is equivalent to either a theorem of the theory or a

contradiction.

q.e.d.

Evidently this has been a rather fussy proof, with lots of little details that had to be checked along the way, and far lengthier than Cantor's proof of the same result presented in Section 2.1. For more complicated cases, where Cantor's proof can't work, the method outlined is also much fussier than the one we described in connection with the extensions of T_{dis}. Let us briefly look at the case of $T_{dis}(+,-)$ in connection with the present method. This time we must, as indicated above, extend L to a larger language L', and in fact to one with infinitely new predicates. The following 2-place predicates $D\geq_r$ for r = 1, 2, ...will fit the bill. Intuitively, $D\geq_r(x,y)$ says that x lies before y and that there are at least r points between them. It is left to the reader to define these predicates in L. (That is, to find formulas $E_r(x,y)$ of L with x and y as free variables whose extension in any model of $T_{dis}(+,-)$ consists exactly of the pairs <a,b> such that a and b stand in the relation $D\geq_r$. With the help of the predicates $D\geq_r$ we can also define predicates $D=_r$ which say that between x and y there are exactly r points. Evidently $D=_r(x,y)$ holds iff $D\geq_r(x,y)$ & $\neg$ $D\geq_{r+1}(x,y)$. For k = 1,2, ... let L_k be the extension of L with the predicates $D=_r$ for r = 1, ..., 2^k together with the predicate $D\geq_{2k+1}$. Suppose that B is a quantifier-free formula in v_1, ..., v_n of $L_{k'}$ and that $k' \leq k$. Then B is equivalent modulo $T_{dis}(+,-)$ to a disjunction of conjunctions of literals from L_k.

Now let A be a sentence of L and assume that A is in prenex form with a prefix consisting of existential quantifiers and negations. Consider the innermost quantifier $(\exists v_n)$ of A. Rewrite the matrix of A as a disjunction $\bigvee_i A_i$ of maximal consistent formulas of L. Again $(\exists v_n)\bigvee_i A_i$ is logically equivalent to $\bigvee_i (\exists v_n)A_i$. Consider $(\exists v_n)A_i$. A_i is equivalent to a disjunction $\bigvee_j A_{ij}$ of maximal consistent formulas of L_1. Let A'_{ij} be the result of eliminating all conjuncts containing v_n from A_{ij}. It is not hard to see that $T_{dis}(+,-) \models (\exists v_n)A_i \leftrightarrow \bigvee_j A'_{ij}$. So we can replace the part inside A beginning with $(\exists v_n)$ by a quantifier-free formula from L_1 in which v_n no longer occurs and which is equivalent to this part modulo $T_{dis}(+,-)$. In this way we can remove all the quantifiers from A. Note, however, that each time we remove a new quantifier the matrix formula which we remove together with it will belong to one of the languages L_k and the disjunction replacing it will then belong to the next language L_{k+1}. This recursion is the direct counterpart of the one

which in our earlier proof of this result made use of the hierarchy of relations $\{+_k\}$.

Not very nice proofs. But they do explain how our earlier, nicer, method came to its name.

2.3 More about Algebraic Theories

Our only encounter with algebraic languages and theories so far was with the languages and theories of lattice algebras and boolean algebras (L_{lata}, L_{ba}, T_{lata}, T_{ba}; see Sections 2.1.2 and 2.1.3). One of the points we stressed about those structures, all of which are lattices, was that they can be characterised alternatively as algebraic structures, involving a number of operations with certain equationally definable properties, or as structures that involve a partial ordering with special properties. As a matter of fact this kind of duality between an algebraic and a relational conception of structure is quite rare, of which the case of lattices is arguably the most striking example in mathematics and logic as they are known today. For most types of relational structures there seem to exist no algebraic alternatives that provide a significantly different perspective; and, similarly, no significantly different relational formulations seem possible for most algebraic structures that play a prominent part in mathematics.

It should be stressed that these are informal assessments, which it would be hard to turn into hard-nosed formal claims that it would be possible to prove or conclusively refute. For what is it for an alternative characterisation of a type of structure to be 'significantly' different? That seems rather a matter of taste, for which it would be difficult to find a convincing formal definition. And that significance is the crucial notion here follows from the fact that some way of redefining relational structures in algebraic terms is almost trivially possible. And the same holds for, conversely, redefining algebraic structures in relational terms. As regards the redefinition of algebraic structure in relational terms we refer to Exercise EA2 at the end of the Appendix to Ch. 1, where it was shown how each n-place function constant can be replaced by a corresponding n+1-place relation constant together with an axiom stating that the relation denoted by the new constant is functional in its last argument; and further, how each

formula couched in the original functional vocabulary is to be translated into a formula couched in the new relational one.

The converse reformulation is slightly more involved. We know from set theory that the extension of any n-place relation R - i.e. any set of n-toples of objects drawn from some domain U - can be turned into the corresponding characteristic function f_R which maps the tuples belonging to the extension to one of two special objects - the one which intuitively speaking signifies 'yes' - and maps the other n-tuples to the other special object, which intuitively means 'no'. Usually the two objects chosen for this purpose are the numbers 1 and 0, but of course that is not essential for the reduction - any two objects will do, provided that they can be kept suitably distinct from the objects in U. There are various ways in which distinctness can be secured. One of these makes use of a simple technique that has proved useful in formal logic elsewhere too is to extend the universes of the algebraic structures M that are to be redescribed in relational terms with a pair of new objects 1_M and 0_M which serve as the 'yes' and the 'no' in the context of M. Some care has to be taken to make sure that the relational translations of the sentences of the original algebraic language are true in the new extended models $M[0_M,1_M]$ iff the original sentences were true in the non-extended models M. But these matters are essentially trivial. For details see Exercise **??** of this Chapter.

The types of algebraic structures to which we turn now, groups and semi-groups, conform to what appears to be the rule in that no significantly different relational characterisations of these types seem to exist. They are also typical of algebraic structures more generally in that they can be characterised by axioms all of which have the form of universally quantified equations, just as we found this to be possible in the case of lattices, distributive lattices and boolean algebras. In Universal Algebra - the branch of mathematics which studies algebraic structures from a general and abstract point of view - types of structure (i.e. classes of models) that are defined by sets of such equational axioms are known as *varieties*. It is important to keep the distinction between this notion and the more general one of an axiomatically definable type of algebraic structure firmly in mind. In general axiomatic characterisations of types of algebraic structures may involve axioms that can be any sentences from the first order language for which the structures are models. The equational axiomatisations that make the characterised model class into a variety constitute a comparatively small special subclass from the range of all possible first order axiomatisations. (Note in this connection that equational axioms are (i) purely universal sentences, but in addition (ii) even among the

purely universal sentences they form a specialised subclass.) It seems safe to infer that the class of varieties is a correspondingly small subclass of the class of all axiomatisable structure types.)

We start, mostly as a preamble to our discussion of the Theory of Groups, with a brief introduction to the Theory of Semi-Groups. The notion of a semi-group is simpler and more fundamental than that of a group, although, as the terminology suggests, the notion of a group came first. This is comparable to what can be observed in connection with orderings, where the notion of a linear ordering was well understood before the general notion of a partial ordering was properly articulated and made into the topic of the exploration of a theory - the Theory of Partial Orders - which subsumes the Theory of Linear Orders as one of several specialisations (Lattice Theory being another).

2.3.1 The Theory of Semi-Groups

The language of the theory of semi-groups consists of a single 2-place function constant. We follow the widely accepted convention of denoting this constant as a full stop and of writing the terms involving it in 'infix notation', just as with ordinary multiplication. So the language, L_{sg}, is {.}, and the term we get when applying. to, say, the variables x and y is written as 'x.y'.

The Theory of Semi-groups, T_{sg}, is nothing more or less than the theory of an associative operation. Thus it consists of all consequences of the single axiom ASS:

ASS $\quad (\forall x)(\forall y)(\forall z)\ x.(y.z) = (x.y).z$

Associative operations can be found in all kinds of contexts and they come in a variety of very different forms. Three salient categories are:

(i) 'arithmetical operations like addition and multiplication, as operations on a range of different domains: natural numbers, integers, rational numbers, real numbers, complex numbers.

(ii) fairly closely related to these, set-theoretical union and intersection, and more generally supremum and infimum operations in lattice-like structures.

(iii) 'function application', in the widest sense of the word. In a sense this is just one operation. But it is found in such a wide variety of

contexts that its instances provide a quite diverse spectrum of different semi-groups, both conceptually and as regards their further formal properties.

In (iii) the basic idea is that of a succession of operations which transform objects of a certain sort into other objects of that sort. the objects can be numbers, geometrical figures, linguistic expressions, computer files or documents, .. - in fact, they can be data structures of and kind. And similarly, the operations can be of any kind too, provided that they return objects of the same sort that they take as input. All that is required is that these operations can be carried out in succession, but that is in essence guaranteed by the fact that their outputs are such that they can serve again as inputs to further applications of the operations.

Under these conditions it is possible to form complex operations by combining two operations O_1 and O_2 into a complex operation $O_1.O_2$ which consists in first executing O_1 and then applying O_2 to the output that the first operation produced. That is, for any input x we have $(O_1.O_2)(x) = O_2(O_1(x))$. It should be obvious that the 'second order operation' (= operation on operations), will always be associative: First executing $O_1.O_2$ and then O_3 obviously amounts to the same thing as first executing O_1 and then $O_2.O_3$; in both cases we get a succession of first executing O_1, then executing O_2 and finally executing O_3.

More 'mathematically' the second order operator, can be identified with the operation $\circ$ of function composition: Let U be any set of 1-place functions from an 'object set' X into itself. Then for any two functions f and g from U, we can form the function $f \circ g$ which maps each object x from X to g(f(x)). Evidently this is again a function from X into X. That $\circ$ is associative follows for the obvious reasons spelled out above.

The three types of associative operations listed above are distinguished by additional formal properties. Arithmetical operations are typically commutative, ie. they satisfy the commutativity axiom COM.

COM $(\forall x)(\forall y)\ x.y = y.x$

Function composition, in contrast, is in general not commutative. Consider for instance the functions f(x) = x +1 and g(x) = 2x on the natural numbers. Then $(f \circ g)(1) = g(f(1)) = 2(1+1) = 4$, but $(g \circ f)(1) = f(g(1)) = 2 + 1 = 3$. However, while non-commutativity is the rule for function composition, there do exist (naturally arising) function spaces

on which composition is commutative. An example is the set U of all functions (say, on the natural numbers, but other number sets will do too here) that map each number onto a certain multiple of it. That is, $U = \{\lambda x.nx: n \varepsilon \omega\}$ (where $\lambda x.nx$ is that function f which for any number y as argument returns the number n.y as value). On the other hand, in modern mathematics one studies number systems ('skew number fields') in which addition and/or multiplication are not commutative. So commutativity is a property that *tends* to hold for semi-groups of types (i) and (ii) and not to hold for semi-groups of type (iii), but this is only a matter of tendencies.

A distinction between semi-groups of types (i) and (ii) is that those of the second type typically satisfy the law of idempotency, given as IDP below, while those of type normally do not:

IDP $(\forall x)\ x.x = x$

This does not mean that in semi-groups of the first type there are no elements at all which satisfy the equation x.x = x. More often than not such semi-groups have some element that satisfies the equation. But these elements are, in case they exist at all, rare, and often they are unique. For instance, the additive semi-groups of the natural numbers, the integers and the reals (i.e. the operation of addition on the natural numbers, the integers or the reals, respectively) all have exactly one such element, viz. the number 0. In the multiplicative groups of (among other number systems) the reals and the rationals (i.e. the multiplication operation on the reals and the rationals) there are two such elements, viz. 0 and 1.

That semi-groups of the first kind contain such elements is closely connected with another property that singles out a certain subclass of semi-groups. This is the property of having an *identity*. An identity of semi-group is an element e such that for any element x of the semi-group x.e = e.x = x. In additive groups this is the unique element that satisfies the equation x.x = x, i.e. 0: for any number x, 0 + x = x + 0 = x. (That an identity satisfies x.x = x follows logically from the definition.) In the case of multiplicative semi-groups the identity is not 0 but 1.

The existence of an identity is quite common among semi-groups of each of the three types. Thus among the salient examples of semi-groups of type (ii), structures of the form $\langle U,\cup\rangle$, where U is some set of sets and $\cup$ is set-theoretic union, have an identity iff U contains a bottom element wrt. set-theoretic inclusion, i.e. an element b that is

included in all other elements of U. For then it will be the case for all x in U that $b \cup x = x \cup b = x$. (A common way for this condition to be satisfied is when u contains the empty set $\varnothing$, which will always be the bottom element so long as it is present.)

Among semi-groups of type (iii) the existence of an identity is also a common occurrence. This will be so in particular when the universe U of a given semi-group contains the identity function I_X on the associated object set X, i.e. the function whose domain is X and which maps each x in X to x. Obviously we have for any function f in U that $I_X \circ f = f \circ I_X = f$.

The existence of an identity is our first property of semi-groups that cannot be expressed by means of an equational axiom - evidently so, for we are not dealing with a general condition that all elements of the structure must satisfy, but an existence claim, to the effect that there is at least one element that satisfies a certain equational condition. As stated this has the form of an $\exists\forall$-formula; and indeed, in the language {.} there seems to be no simpler way of stating it. For the sake of explicitness we give the $\exists\forall$ formula:

IDE $\quad (\exists y)(\forall x)(y.x = x \ \& \ x.y = x)$

One might wonder if this formulation isn't somewhat redundant. Do we really need the conjunction of the two equations $y.x = x$ and $x.y = x$? Wouldn't one of those be enough? The answer to this question is negative. But there are some slight subtleties to the matter, so we will dwell on it a little. Let us, just as we have called an element that satisfies the condition $(\forall x)(y.x = x \ \& \ x.y = x)$ an identity, use the terms *left identity* and *right identity* for elements that satisfy the conditions $(\forall x)y.x = x$ and $(\forall x)\ x.y = x$, respectively; and let us call the statements that a left, resp. right identity exists, IDEL and IDER:

IDEL $\quad (\exists y)(\forall x)\ y.x = x$

IDER $\quad (\exists y)(\forall x)\ x.y = x$

Evidently an identity is both a left identity and a right identity. But we will see in Section ?? that in general a left identity need not be a right identity (and thus not be an identity) and conversely. Nor does the existence of a left identity entail that there is some other element that is a right identity or vice versa. That is, in general neither of IDEL and

IDER entails the other, and so a fortiori neither entails IDE.[22] On the other hand, when a semi-group has both a left identity and a right identity, then these two elements must be identical, and this element will thus be an identity. Similarly, any two left identities and any two right identities must be identical (and so any two identities must be identical). But of course the identity of two left or two right identities doesn't entail that they will be identities.

Exercise. a. Suppose that e_l and e_r are a left an right identity of some semi-group <U,.>. Show that $e_l = e_r$.

b. Suppose that e_1 and e_2 are both left identities of <U,.>. Show that $e_1 = e_2$.

Some semi-groups with an identity are distinguished by a further property, which makes them into *groups*. A *group* is a semi-group with an identity e in which each element x as an *inverse*, i.e. an element z such that x.z = z.x = e. Expressing this property in our language of semi-groups, {.}, is cumbersome, since it must incorporate the assertion that there exists an identity within it.

INV $(\exists y)(\forall x)(y.x = x \ \& \ x.y = x \ \& \ (\forall x)(\exists z)(z.x = y \ \& \ x.z = y))$

Once again the question arises whether we need ther conjunction of the two conditions in the scope of $(\exists z)$. This time the immediate answer is negative. But here too there are subtleties that deserve to be pointed out, and which will emerge in the next section. So once again we distinguish, so that we will be in a better position to discuss those when we come to them, between a l*eft inverse* z_l *of* an element x, which has the property that $z_l .x = e$ and a *right inverse* z_r *of* x, which has the property that $x.z_r = e$.

The answer to the question above is negative in the following precise sense. Suppose that a semi-group M = <U,.> has an identity e. Then if every element of M has a left inverse it is also the case that every element has a right inverse; and conversely, if every element has a right inverse, then every element has a left inverse. Moreover, in either case the left and right inverse of any element will coincide,.

---

22 When we say that (e.g.) IDEL does not 'entail' IDER, what is meant is that IDEL doesn't entail IDER within the Theory of Semi-Groups, T_{sg}. That is, IDER does not follow logically from the conjunction of IDEL and T_{sg}'s only axiom ASS.

Consequently if every element has a left inverse, then every element x has an inverse in the sense of INV (i.e. an element z such that x.z = e & z.x = e)

The proof of these different claims is not complicated. First suppose that every element of M has a left inverse. Let x be any element of M, let z be a left inverse of x, i.e. z.x = e. We must show that x has a rght inverse. Let u be a left inverse of z, i.e. u.z = e. Then x = e.x = (u.z).x = u.(z.x) = u.e = u. But then x.z = u.z = e, so z is right inverse of x. This establishes not only that every element of M has a right inverse, but that for each x there is an element that is both left and right inverse. A parallel argument shows that this conclusion follows equally from the assumption that every element of M has a right inverse.

We can summarise the upshot of this by observing that relative to the Theory of Semi-Groups INV is equivalent to each of the two following sentences INVL and INVR.

INVL $(\exists y)(\forall x)(y.x = x \ \& \ x.y = x \ \& \ (\forall x)(\exists z)\ z.x = y\ )$

INVR $(\exists y)(\forall x)(y.x = x \ \& \ x.y = x \ \& \ (\forall x)(\exists z)\ x.z = y\ )$

In thee next section we look at the Theory of Groups. As we have seen this theory can be axiomatised in the language of semi-groups we have been using in this section (the language L_{sg}, or {.}), e.g. by the axioms ASS and INV. But the second of these is not in equational form, and it seems that it cannot be converted into such a form, or be replaced by one more others of such form that yield the same theorems in conjunction with ASS - at least not when we stick with the language L_{sg}. As we have seen this theory can be axiomatised in the language of semi-groups we have been using in this section (the language L_{sg}. (We are not giving an actual proof that such a replacement is impossible, and as far as we know such a proof this not all that easy.)

However, we will see in the next section that it does become possible to axiomatise the Theory of Groups in equational form if we extend L_{sg} with additional non-logical constants.

2.3.2 The Theory of Groups

We have already given one formulation of the first order theory of groups and thus specified what groups are like. But, as in the case of lattices, there are other ways of formalizing the notion, even if in the

present case the differences aren't quite as dramatic. As we already said, the main advantage of the alternative formulation we present below is that it enables us to state all the axioms as equations. The comparison between this new axiomatisation and the one given in the last section is interesting from a general methodological point of view in that it shows a trade-off of a kind not yet encountered: That between a parsimonious choice of primitive notions (our language {.} with its one 2-place function constant) but axioms of a more complicated structure and on the other hand a richer set of primitives with a corresponding gain in simplicity as far as teh axioms are concerned.

The section serves to focus on two other issues of general significance. The first is the question of independence as applied to axiom systems, or sets of sentences. Usually when we specify a set of axioms as a way of characterising a given formal theory, we try to avoid redundancies: none of the axioms in the set should follow logically from the rest. However, proving that this desideratum has in fact been satisfied can be very tricky. And when there are many axioms, there is a lot of work to be done, since each axiom requires its own independence proof. For the axiomatisations of group theory that are considered in this section this problem is manageable since there are few axioms to deal with. But the independence arguments we wil give for them should provide a clear impression of the general nature of independence proofs and also give a little taste of why such proofs can be difficult.

The third point of general significance that the section seeks to illustrate was already brought up in the last section, when we drew attention to the wide conceptual and formal diversity of semi-groups. This is also true of groups, and here the value of extracting what is common to a great diversity of structures by describing them as models of a single formal theory that covers them all has been of great importance in the history and current practice of pure and applied mathematics.

A fourth point concerns the special properties of 'equations', that is of those purely universal sentences in which the quantifier prefix is followed by a single equation. Equations, in this sense of the word, form a kind of closed subsystem of the set of sentences of a given language L, with their own proof theory and its own special model-theoretic properties. This subsystem is known as Equational Logic. A separate section (Section ??) will be devoted to it.

The axiomatisation of the Theory of Groups we gave in the last section had to resort to axioms that were not of equational form. These axioms

contain existential quantifiers that are needed to express that groups contain entities with special properties: (i) an identity and (ii) for each element x an inverse of x. However, we saw that if such entities exist a all, then they are unique. This means that we can also proceed as follows: We introduce constants in our language to denote these entities and then give axioms stating that the denotations of those constants have the required properties. The constants we need are (i) a o-place function constant e to denote the group identity and (ii) a 1-place function constant $^{-1}$ to denote a function that maps each element to its inverse.

Thus we are led to the language $\{\cdot, ^{-1}, e\}$, to which we will also refer as L_{G1}. $\{\cdot, ^{-1}, e\}$ is the group-theoretic vocabulary that is usually treated as basic in discussions of groups.)

In L_{G1} the Theory of Groups can be axiomatised with the axioms T_{G1}.A1-T_{G1}.A3, which we present both in the standard notation of first order predicate logic and also in the abridged notation of equational logic, in which the universal quantifiers are implicit

T_{G1}.A1	$(\forall x)\ (\forall y)(\forall z)\ \ (x\cdot y)\cdot z = x\cdot(y\cdot z)$	$(x\cdot y)\cdot z = x\cdot(y\cdot z)$
T_{G1}.A2	$(\forall x)\ x\cdot x^{-1} = e$	$x\cdot x^{-1} = e$
T_{G1}.A3	$(\forall x)\ x\cdot e = x$	$x\cdot e = x$

But whether we explicitly write the quantifiers of these axioms or not, they are there, and they are meant as axioms of a theory consisting of all sentences of L_{G1} that logically follow from them, and not just those that are universally quantified equations themselves. We will see this presently when we go through a few simple theorems of this theory and proofs of those from the axioms: some of these theorems do have the form of equations, but not all of them.

The proofs of the equational theorems that follow make use of notation that is familiar from the way arguments in universal algebra are often presented, where all mention of quantifiers is suppressed. (Where both premises and conclusions of an argument are in equational form this is very natural, and hardly needs a justification. Nevertheless, it is an interesting, and as it turns out non-trivial, logical question exactly how this form of derivation relates to standard methods of logical deduction like those discussed in Ch. 1. In Section ??, which is devoted to Equational Logic as an alternative to predicate logic, we will go into this question in detail.)

T_{G1}.T1 $x^{-1} \cdot x = e$

Proof. $x \cdot e = x \cdot (x^{-1} \cdot (x^{-1})^{-1}) = (x \cdot x^{-1}) \cdot (x^{-1})^{-1} = e \cdot (x^{-1})^{-1}$.
Therefore:

$x^{-1} \cdot x = (x^{-1} \cdot x) \cdot e = x^{-1} \cdot (x \cdot e) = x^{-1} \cdot (e \cdot (x^{-1})^{-1}) =$

$(x^{-1} \cdot e) \cdot (x^{-1})^{-1} = x^{-1} \cdot (x^{-1})^{-1} = e.$

T_{G1}.T2 $e \cdot x = x$

Proof. $e \cdot x = (x \cdot x^{-1}) \cdot x = x \cdot (x^{-1} \cdot x) =_{(TG1.T1)} x \cdot e = x$

T_{G1}.T3 $(x^{-1})^{-1} = x$

Proof. Combine T_{G1}.T2 and the first line of the proof of T_{G1}.T1.

Exercise. Turn the proofs of T_{G1}.T1 - T_{G1}.T3 into predicate logic derivations in the formal sense of the definition on p. 5.

Given what was said about groups in the last section, theorems T_{G1}.T1 and T_{G1}.T2 are a natural complement to axioms T_{G1}.A1 - T_{G1}.A3. In fact, when one looks at these axioms without the hindsight that these theorems provide, the suspicion might easily arise that the axioms are too weak. For T_{G1}.A2 only asserts that x^{-1} ia a right inverse of x, and T_{G1}.A3 only that e is a right identity. Is that enough to guarantee that e is also a left identity and x^{-1} also a left inverse? Theorems T_{G1}.T1 and T_{G1}.T2 tell us that they are. But that this is so has to do with a subtle interaction between T_{G1}.A2 and T_{G1}.A3. We will see in the next section that when one of T_{G1}.A2 and T_{G1}.A3 is changed into its opposite (i.e. T_{G1}.A2 into the axiom which says that e is a left identity), then the axiom system does become too weak.

Exercise. Prove the following theorems of G1 from its axioms:

(i) $(x.y)^{-1} = y^{-1}.x^{-1}$

(ii) $x.y = y.x \leftrightarrow y^{-1}.x.y = x \leftrightarrow y.x.y^{-1} = x \leftrightarrow x.y.x^{-1} = y \leftrightarrow x^{-1}.y.x = y$

(Here "A $\leftrightarrow$ B $\leftrightarrow$ C $\leftrightarrow$.." is used as shorthand for

"(A $\leftrightarrow$ B) & (B $\leftrightarrow$ C) & (C $\leftrightarrow$ .. ")

Exercise. Let "x/y" be short for "$x.y^{-1}$". Show:

(i) $e = x/x$
(ii) $x^{-1} = (x/x)/x$
(iii) $x.y = x/((y/y)/y)$

The next theorems do not have the form of equations:

T_{G1}.T4 $(\forall x)(\forall y)(\forall z)(\ x\cdot y = z \leftrightarrow z\cdot y^{-1} = x)$

Proof. First suppose that $x\cdot y = z$. Then $z\cdot y^{-1} = (x\cdot y)\cdot y^{-1} = x\cdot(y\cdot y^{-1}) = x\cdot e = x$. Conversely, if $z\cdot y^{-1} = x$, then $x\cdot y = (z\cdot y^{-1})\cdot y = z\cdot(y^{-1}\cdot y) = z\cdot e = z$.

T_{G1}.T5 $(\forall x)(\forall y)(\ x\cdot y = e \rightarrow y = x^{-1})$

Proof. Suppose $x\cdot y = e$. Then $x^{-1} = x^{-1}\cdot e = x^{-1}\cdot(x\cdot y) = (x^{-1}\cdot x)\cdot y = e\cdot y = y$.

We have now seen two formalisations of the Theory of Groups, one in the language L_{sg} and involving the axioms ASS and INV, and one in the language L_{G1} and involving the axioms T_{G1}.A1-T_{G1}.A3. The move from L_{sg} to L_{G1} was motivated by the observation that the existence statements made by INV provide to be of elements that turn out to be uniquely characterised by the conditions that IV specifies. This means that we could also have proceeded in the same way as we did when extending the theory of lattices T_{lato} in the language $\{\leq\}$ to the theory in which we have constants to refer to the operations $\cup$ and $\cap$ that T_{lato} enables us to define in terms of $\leq$. That is, we can (i) extend L_{sg} to L_{G1} (as we have done), and (ii) extend the theory $Cl_{L_{sg}}(\{ASS,INV\})$ to a theory in L_{G1} by adding the following two definitions of e and $^{-1}$ as axioms:

(Def.e) $(\forall y)(e = y \leftrightarrow (\forall z)\ z.y = z)$
(Def.$^{-1}$) $(\forall x)(\forall y)(x^{-1} = y \leftrightarrow x.y = e)$

It is not hard to show that this is the same theory as T_{G1}.

Exercise: Prove this.

The difference with the situation we found to obtain in the case of lattices is that this time the converse route is not possible: We cannot formulate the Theory of Groups in the language whose non-logical constants are just the ones that we added when passing from L_{sg} to L_{G1}; no axiomatsaton of the Theory of Groups is possible within the language $\{e,^{-1}\}$.

Exercise: Prove this. (Hint: there is no way to define the two place operation, with the help of just the 0-place function e and the 1-place function $^{-1}$.)

These formalisations of the Theory of Groups are by no means the only ones possible. As a matter of fact, in a strict formal sense the number of possible formalisations of a theory is always infinite; for any one formalisation there twill always be infinitely many alternatives, although as a rule most of these will be uninteresting variants which it is as pointless to present as they are easy to construct. But often genuinely different alternatives exist, which cast a different light on what is being formalised. The alternative formalisation of lattices as orderings and as algebras was a particularly striking example of this. Nothing quite like that compares with it in the case of groups. But there is one alternative that is worth mentioning, at least because it answers a certain formal question that naturally arises in connection with what we have said above about our two axiomatisations in the languages L_{sg} and L_{G1}. The choice between those was presented as a kind of trade-off between (i) having just the single function constant, and (ii) having only axioms in equational form. The alternative that is discussed in the following exercise can be seen as combining the advantages of both. It uses a single 2-place function constant / and it only needs equational axioms. The function / is the 'division operator' of Group Theory, which can be defined in terms of . and $^{-1}$ as: $x/y = x.(y^{-1})$.

Exercise. Give a complete axiomatisation, all axioms of which are equations, of the Theory of Groups in the language {/}, where / is the 2-place operation of group-theoretical division: More precisely, provide equational axioms A/.1,..,A/.n (for some number n) such that the theories T_1 and T_2 defined below are identical.

Definition of T_1 and T_2:

Let T/ = Cl{/}({A/.1,..,A/.n}). Let L' be the language {/, . ,$^{-1}$,e}.

(a) T_1 is the theory of L' that is obtained by adding to the axioms of T/ the following definitions of e, $^{-1}$ and . in terms of /:

(i) $(\forall x)\ e = x/x$

(ii) $(\forall x)\ x^{-1} = (x/x)/x$

(iii) $(\forall x)(\forall y)\ x.y = x/((y/y)/y)$

(b) T_2 is the theory of L' that is obtained by adding to the axioms of T_{G1} the following definition of / in terms of . and $^{-1}$:

(iv) $(\forall x)(\forall y)\ x/y = x.y^{-1}$

(Solution. One solution is the following set of axioms A/.1,.,A/.4:

A/.1	y/y = x/x	
A/.2	y/(y/y) = y	y/e = y
A/.3	(y/y)/(x/y) = y/x = x/x	e/(x/y) = y/x
A/.4	x/(y/z) = (x/((z/z)/z))/y	x/(y/z) = (x/(e/z))/y

In the formulations of A/.2-A/.4 on the right, subterms of the form α/α have been abbreviated as 'e', in accordance with A/.1.)

2.3.3 Independence

In the introduction to this section we mentioned the question of the *independence* of the members of a given axiom set. As indicated, it is generally considered a matter of logical hygiene that the sets of axioms used to formalise a given structure or concept contain no *redundant* axioms. That is, if G is any such set and A ε G, then it should not be the case that (G\{A} $\models$ A. If this is not the case, then we say that A is *independent in* G; and if all members of G are independent, G is called an *independent set* of axioms.

As a matter of fact, all axiom sets presented so far in this chapter have been independent in the sense just defined. Showing that this is so, however, is not trivial. In general, proving that an axiom set is independent tends to be not only a fair bit of work - to show that the set A1, ..., An is independent requires n separate proofs, one for each Ai - some independence questions can be a real challenge. Also independence proofs may provide real insight into what precisely is

contributed by a given axiom to the given characterisation of the intended class cof structures that is not contributed by the other axioms. More about this towards the end of this section.

Here we consider only two of the three independence questions connected with the axiom set $\{T_{G1}.A1, T_{G1}.A2, T_{G1}.A3\}$. We show the independence of $T_{G1}.A3$ from the remaining two axioms explicitly, and provide a hint for establishing the independence of $T_{G1}.A2$. As regards $T_{G1}.A1$, the reader is on his own (see Exercise **??**).

First $T_{G1}.A3$. Consider the following model $M = \langle U,F \rangle$ for L_{G1}:

(i) U = the set of all pairs $\langle i,n \rangle$, where $i \in \mathbb{Z}$ (the set of integers) and $n \in \mathbb{N}$ (the set of natural numbers).

(ii) $F(\cdot)$ = the function f such that for any $\langle i,n \rangle, \langle j,m \rangle \in U$, $f(\langle i,n \rangle, \langle j,m \rangle) = \langle i+j,m \rangle$

(iii) $F(e) = \langle 0,0 \rangle$

(iv) $F(^{-1})$ = the function g such that for any $\langle i,n \rangle \in U$, $g(\langle i,n \rangle) = \langle -i,0 \rangle$

Then it is straightforward to verify that $T_{G1}.A1$ and $T_{G1}.A2$ hold in M. But $T_{G1}.A3$ does not hold, since e.g. $\langle 1,1 \rangle \cdot e = \langle 1,1 \rangle \cdot \langle 0,0 \rangle = \langle 1,0) \neq \langle 1,1 \rangle$.

It is easy to turn this construction into a demonstration that the second axiom is independent of the other two by changing the definition of $F(\cdot)$ into

(ii') $F'(\cdot)$ = the function f' such that for any $\langle i,n \rangle, \langle j,m \rangle \in U$, $f'(\langle i,n \rangle, \langle j,m \rangle) = \langle i+j,n \rangle$

It is worth noting that while M falsifies $T_{G1}.A3$ it verifies the superficially similar sentence

$T_{G1}.A3'$ $\quad (\forall x)\ e \cdot x = x$

Recall that $T_{G1}.A3'$ is nothing other than $T_{G1}.T2$. So we have also shown that $T_{G1}.A3$ cannot be derived from $T_{G1}.A1$, $T_{G1}.A2$ and $_{G1}.A3'$. Apparently, then, this sentence is, given $T_{G1}.A1$ and $T_{G1}.A2$, genuinely weaker than $T_{G1}.A3$, and replacing $T_{G1}.A3$ by $T_{G1}.A3'$ in the axiomatisation of T_{G1} would yield a different, weaker theory. In the same vein it can be observed that the modified model $M' = \langle U,F' \rangle$ verifies the sentence

$T_{G1}.A2'$ $(\forall x)\ x^{-1} \cdot x = e$

So replacing $T_{G1}.A2$ by $T_{G1}.A2'$ while leaving $T_{G1}.A1$ and $T_{G1}.A3$ the same would also lead to a weakening of deductive power. On the other hand it is easy to verify that if we replace both $T_{G1}.A2$ and $T_{G1}.A3$ by $T_{G1}.A2'$ and $T_{G1}.A3'$ the result is a theory that is equivalent to T_{G1}.

Exercise: Show this.

Exercise: Show that the associativity axiom $T_{G1}.A1$ is independent of the axioms $T_{G1}.A2$ and $T_{G1}.A3$.

Hint: 1. Consider the model M = <U,F>, where U = the set of the rational numbers without 0 and let $F(\cdot)(r,s) = r/s$. Then $T_{G1}.A1$ evidently fails. Choose $F(^{-1})$ and F(e) so that M verifies $T_{G1}.A2$ and $T_{G1}.A3$.

Other solution. Here is another possibility. U is the set {0,1,2, ..,n-1}. F(e) = 0, $F(^{-1})(k)$ is the unique number m from U such that k + m = 0(mod n) and F(.) is defined as follows: (i) F(.)(k,k) = k; (ii) if $k \neq m$, then F(.)(k,m) = k + m (mod n). Then it is easily verified that (writing "." instead of "F(.)" and using infix notation) 0.k = k.0 = k and that $k.k^{-1} = k^{-1},k = 0$. But in general F(.) will not be associative. For instance, if n = 4, then (2.2).3 = 2.3 = 5 (mod 4) = 1, but 2.(2.3) = 2. (5 (mod 4)) = 2.1 = 3. Note that in this example F(.) is commutative and that (because of this) not only the axioms $T_{G1}.A2$ and $T_{G1}.A3$ are verified, but also the formulas which we get by switching the arguments of the left hand term around, i.e. $(\forall x)$ e.x = x and $(\forall x)\ x \cdot x^{-1}.x = e$.

[End Exercise]

The three independence arguments presented here are comparatively simple. They do give insight why each of the three axioms contributes something that the others do not, but precisely because models that satisfy all but one of the axioms are comparatively easy to find, the insight gained from any one such models (and thus from the independence proof it provides) are limited: Other models might give additional insights in the contributions of the different axioms in the set and quite possibly more important ones.

But in this regard our examples are not representative. In the history of mathematics and logic certain independence questions have had an enormous impact. Their solution have led to the discovery of structures that have proved of lasting importance and to methods of mathematical reasoning and mathematical construction that subsequently found many additional applications. Even some attempts at finding a solution to an independence question that did not answer the question that they were meant to have led to significant progress in other areas.

Perhaps the most famous example from mathematics id the parallel postulate from Euclid's axiomatisation of plane geometry, the statement that for every point p that is not on a straight line l there is exactly one straight line m that goes through p and is parallel to l. Ever since Euclid it was felt that this postulate was less self-evident than Euclid's other postulates. Since it was widely thought that Euclidean geometry described a structure that was in some sense necessary - space just couldn't have been different from what it is! - and since it was thought also that since the properties of the structure of space were necessary, they should be directly accessible to intellectual judgement, the lacking self-evidence of the parallel postulate was seen as an imperfection of Euclid's system, and an imperfection that could be removed only by either finding a more intuitive replacement for it or - even better - to derive it from Euclid's other postulates. In the course of the many centuries during which this was an open question an enormous amount of mathematical energy and ingenuity must have gone into the project of deriving the parallel postulate from the other postulates. Eventually, in the second half of the 18-th century it dawned on some mathematicians that the persistent failure to find a proof of the parallel postulate from the others might have a very simple explanation, viz that there is no such proof, in other words, that the parallel postulate was independent from the other postulates. This led to the new and contrary effort to demonstrate the independence of the parallel postulate, or, what comes to the same thing, the consistency of the other postulates with the negation of the parallel postulate. (It no longer needs to be said here that being a model in which postulates A_1, ..., A_{n-1} hold and A_n doesn't is the same as being a model in which A_1, ..., A_{n-1} and $\neg A_n$ hold together.) The models of the negation of the parallel postulate jointly with the other Euclidean postulates - as described in the work of the Hungarian mathematician Janos Bolyai (1802-1860), the Russian mathematician Lobachewski (1792-1856) and the German mathematicians Gauss (1777-1855) and Riemann (1826-1866) - have done more than anything else to revolutionarise geometry as mathematical discipline as it in the course of the 19-th century. And it has also deeply affected our understanding of the distinction between

necessary and contingent truth as well as the distinction between geometry as a conceptual structure (along the lines it was seen by, for instance, Kant) and geometry as part of the structure of the physical world.[23]

A second independence problem, which was specific to the development of mathematical logic in the 20-th century, concerns the Continuum Hypothesis in Set Theory, the Hypothesis that there are no sets whose cardinality is intermediate between that of the natural numbers (the smallest infinite cardinality) and that of the set of real numbers, which is the same as that of the power set of the set of natural numbers). As we noted earlier, the Continuum Hypothesis was formulated by Cantor, the founder of set theory. Cantor is said to have worked desperately on a proof of the Continuum Hypothesis from other set-theoretical principles, whose validity he did not consider in doubt, and the effort is supposed to have seriously affected his health. His unsuccessful efforts were followed by those of many others, and among these efforts were in particular those to derive the Continuum Hypothesis from the other established set-theoretical axioms e.g. those of Zermelo-Fraenkel (see Ch. 3). But in this case too eventually the suspicion arose that no such derivation could be given, since the Continuum Hypothesis was in fact independent from the other, uncontroversial, axioms of set theory. And independence was finally proved in 1963 by the American mathematician Paul Cohen. In this case too the method used to establish independence has proved immensely fruitful, leading in particular to a series of further independence results within the realm of set theory.

There is an interesting similarity between these tow cases - the parallel postulate in geometry and the Continuum Hypothesis in set theory - in that in both cases a conception of the subject matter as involving necessary and therefore presumably ultimately self-evident truths drove scholars to persistent efforts to decide what seemed not self-

---

[23] The first to have clearly understood this second disrtinction appears to have been Gauss , who engaged as early as the first half of the nineteenth century in a large scale prohect of geodetical measurments in order to determine whether the physical geometry whose straight lines are the paths of light rays is in fact Euclidean or not. (i.e. if light rays conform to the parallel postulate.) Gauss suspician of non-Euclidenan character of the genoatryx of light rays was confirmed only when in the forst quarter of the 20-th century physisicts and astronomers looked for a experimental confirmation of one of the implications of Einstein's gedneral Theory of Relativity, which is that gravitation 'bends' the paths of light rays, so that the geometry they define is - in the presence of graviational fields, which is always the case in our actual cosmos - non-Eucildean. Einstein's Theory of General Relativity, it has been sad would not have been possible without the work of Riemann.

evident on the bassi of those principles that were considered self-evident. One of the general lessons that has been learned from both efforts is that the line between necessity and contingency is much more difficult to draw than people seem to have realised through most of the history of philosophy (and than, in the depth of our hearts, many of us would still like to believe today); and, connected with that, that we should not set too much store by our intuitions on what is 'self-evident' and what is not.

2.3.4 The Theory of Groups and Group Theory

1. What has been called the (first order) Theory of Groups here should not be confused with what is normally understood by 'Group Theory'. First, the 'mini-theorems' of the Theory of Groups of which we have given a few examples here bear no comparison with the theorems about groups that mathematicians find interesting. But more fundamentally, those results can as a rule not even be stated within the first order languages we have been using. For instance, many results in Group Theory have to do with characterisations of groups in terms of the kinds of subgroups they have - that is, in our terminology, in terms of their submodels. (Note that a submodel of a structure that satisfies the axioms T_{G1}.A1-T_{G1}.A3 will automatically be itself a model of these axioms and thus again a group. (Exercise: Prove this and/or Section ?? below.) To state such a characterisations of a group we need to quantify over its subgroups and thus over subsets of its universe, and to do that we need second, not first order logic. So at a minimum we will need the second order extension of one of our first order languages {.} or {,.,$^{-1}$,e}. Also, there are many theorems of Group Theory which involve reference to natural numbers (e.g. to describe the possible size(s) of finite groups with certain properties, and/or the sizes of certain parts of them. The proofs of such theorems often make use of quite complicated facts of combinatorial number theory,. In these cases formalisation requires a logical vocabulary that includes number-theoretic notions as well as the group-theoretic ones that are the only non-logical constants of the language we have used here, and for a formalisation of the proofs of these statements we will need an axiomatisation of number theory as well.

All this goes to say that Group Theory as it is practiced by algebraists involves far more than our 'bare bones' languages provide. Even if such a language suffices to characterise the general notion of a group, it falls far short of what is needed to state and proof what a mathematician wants to know. This is a somewhat sobering comment on the power of

first order formalisations, not only of the structures that are the subject of group Theory, but of most kinds of mathematically interesting structures generally.

2. It was pointed out more than once in this Chapter that the point of many algebraic theories is that their models cover a wide range of different structures. This is true in particular of the theory of groups. The class of all groups shows a great deal of diversity, in the sense that it contains structures which vary substantially either in their conception or in their formal properties or both.

The value of an algebraic theory with such coverage is, we have noted, that the theorems that can be derived from the general theory are applicable to all the different structures that are among its models. This is as true of the Theory of Groups as it is of other theories with wide structure coverage. But on the other hand the diversity among the different types of groups is such, and certain types of groups are so important, that these types have become the subject of a separate branch of mathematical investigation. A prominent example of this is the class of *Abelian* or *commutative* groups, in which the group operation $\cdot$ is commutative (i.e. $x \cdot y = y \cdot x$ holds for all elements x,y of the group).

In this particular case the additional property that singles out the given class of groups, viz. commutativity of $\cdot$, can be expressed by a first order axiom. But for many other properties that define important subtypes of groups this is not so. An example is the notion of a *simple group*, i.e. group that doesn't contain any proper subgroups (i.e. for which there are no properly included submodels which consist of more than one element); the notion of a finite group - finiteness cannot be expressed by a first order axiom -; or the class of all *permutation* groups, a notion which will be explained below.

To give an impression of how different certain models of the Theory of Groups can be from each other in origin and/or appearance we remind the reader of the two types of examples that were mentioned briefly in the introduction to Section 2.2.1. The first type, it may be recalled, consists of structures in which the group operation $\cdot$ is one of the familiar arithmetical operations of addition or multiplication, or some variant thereof. One example of this type of groups are: the integers with the binary operation of addition, the 1-place operation of sign inversion (i.e. n^{-1} is the number -n) and the number 0 as e constitute a group. Similar examples are provided by the rational numbers and the

real numbers, each with the same operations of addition, sign inversion and 0. Closely related examples are the *additive groups modulo n*, consisting of the numbers {0,1,..., n} with "+ mod(ulo) n" for the operation $\cdot$ (where i+j (mod n) is the remainder of i+j after division by n), "sign inversion modulo n" for the operation $^{-1}$ (i.e. $i^{-1} = n - i$) and again 0 as e. Besides these additive groups there are also multiplicative groups, in which $\cdot$ is multiplication. One example we have already encountered: the rational numbers without 0, with multiplication for $\cdot$, $1/r$ for r^{-1}, and 1 for e. Yet another example is provided by the real numbers (also without 0) with the usual operations of times, multiplicative inverse and 1. There are many more examples of this general sort, involving either some variant of addition or multiplication and/or the use of some alternative notion of "number" (complex numbers, quaternions, etc.).

As a rule groups of this type are commutative, since operations of addition and multiplication tend to be commutative (though there are exceptions).

The second type of group to be mentioned here is that where the elements of the group are functions, $\cdot$ is the operation of function composition, $^{-1}$ is function inverse and e is the identity map. In order that these notions are defined for all elements of the structure it is necessary that all elements (i.e. all functions) have one and the same domain and range. Moreover, the requirement that the inverse operation be everywhere defined entails that all functions are injections. Thus a group of this kind will consist of a set of bijections from some given set X to itself. Such bijections from X to X are also known as *permutations of* X.

It is easy to verify that any set of permutations from X to X which includes the identity map on X and is closed under inverses and function composition forms a group. (Exercise: Show this.) Such groups are called *permutation groups*. Within the class of permutation groups we still find a remarkable spectrum of variety. Among the simplest examples are those groups which consist of all permutations of some finite set $\{a_1,..., a_n\}$. Evidently, the properties of any such group are determined entirely by the cardinality of the set - the group of all permutations of $\{a_1,..., a_n\}$ and the group of all permutations of $\{b_1,..., b_m\}$ are isomorphic iff n = m. So it is possible to confine attention to the full permutation groups of {1,..., n} for the different natural numbers n.

Function composition is usually not a commutative operation. So, contrary to the groups based on arithmetical operations permutation groups are hardly ever commutative.

Exercise. i. Show this, by defining a permutation group in which the commutativity law $x \cdot y = y \cdot x$ is invalid.

ii. What is the smallest number n such that the full permutation group on {1,..., n} is not commutative?

[To be added to the list of exercises at the end pf Ch. 2]

Exercise: In Section 2.2.1.1 it was shown that the axiom $T_{G1}.A1$ is independent of the axioms $T_{G1}.A2$ and $T_{G1}.A3$. The model discussed in that exercise did not establish the following stronger independence result, according to which $T_{G1}.A1$ is not entailed by the set consisting of $T_{G1}.A2$ and $T_{G1}.A3$ and their "converses" $T_{G1}.A2'$ and $T_{G1}.A3'$:

$T_{G1}.A2'$ $\quad x^{-1} \cdot x = e$

$T_{G1}.A3'$ $\quad e \cdot x = x$

One way to get this stronger result is to make use of permutation models. Let M = <U,F>, where U is the set of permutations of the set {1,2, ..., n}, for some n > 2. $F(^{-1})$ and F(e) are defined for permutation groups, i.e. $F(^{-1})(f)$ is the inverse f^{-1} of f and F(e) is the identity map.

But we now define $F(\cdot)$ by: $F(\cdot)(f,g) = g^{-1} \circ f$. Show that in this model $T_{G1}.A2$, $T_{G1}.A3$, $T_{G1}.A2'$ and $T_{G1}.A3'$ all hold, but that $T_{G1}.A1$ fails.

Exercise: Missing from the independence proof for the axiom set {$T_{G1}.A1$, $T_{G1}.A2$, $T_{G1}.A3$} in Section 2.2.1.1 was the independence of $T_{G1}.A2$.

To show independence of this axiom from the other two is very easy, because it is the only axiom that contains the opersation $^{-1}$.

a. Why? Prove the independence of $T_{G1}.A2$.

More interesting is the independence from $T_{G1}.A1$ and $T_{G1}.A3$ of the weaker principles (i) that there is for each x an element y such that

$x \cdot y = e$ and (ii) that, for ay $x_,$, any two elements y and y' such that $x \cdot y = e$ and $x \cdot y = e$ are identical:

(i) $(\forall x)(\exists y)\ x \cdot y = e$

(ii) $(\forall x)(\forall y)(\forall y')(x \cdot y = e\ \&\ x \cdot y' = e \rightarrow y = y')$

b. Prove the independence of (i) and of (ii) from T_{G1}.A1 and T_{G1}.A3.

2.4 Equational Logic.

Equations - purely universal sentences whose matrices are of the form $\sigma = \tau$, where σ and τ are terms - have special properties. First, they allow for a special method of deduction: if an equation B follows from equations A_1, ..., A_n, then this can be shown by deriving B from A_1, ..., A_n via special rules, which are designed to fit the special form that equations have.

Secondly, equations are characterised by special model-theoretic properties. These of course include the properties that are shared by all purely universal sentences (see Ch. 1, Sn 1.5.2). But equations are distinguished from universal sentences in general by some additional properties. As for purely universal sentences in general this fact can be cast in the mould of a preservation theorem, a theorem first stated and proved by the American algebraist G. Birkhoff.

These then are the topics of this section. We will first present the special deduction system for equations and prove its soundness and completeness, and then present and prove Birkhoff's Theorem.

Let $L = \{f_1, ..., f_k\}$ be an algebraic language, where, for i = 1, ..., k, f_i is an n(i)-ary function constant. By an *identity of* L we understand any purely universal sentence of the form $(\forall x_1)...(\forall x_m)\ s = t$, where s and t are terms of L and $x_1,...,x_m$ are the variables that have occurrences in at least one of s and t. We denote the identity $(\forall x_1)...(\forall x_m)\ s = t$ also as "$s \equiv t$".

There is a sense in which the identities of L form a "self-contained" subsystem of the set of all formulae of L:

Suppose $\Gamma \models E$, where E is an identity and Γ is set of identities. Then it is possible to derive E from Γ by means of a set of five inference rules $RE_{ref.}$,..., $RE_{repl.}$, each of which only involves identities. That is, there always exists in such a case a derivation of E from Γ which consists of identities only (and in which each line is either a premise from Γ or comes from earlier lines by application of one of the rules).

Here are the rules:

$RE_{refl.}$ $t = t$ (that is: each identity of the form "t = t", where t is any term, may be written down as a new line; thus this rule functions as an axiom.)

$RE_{sym.}$
$\underline{s = t}$
$t = s$

$RE_{trans.}$
$\underline{r = s,\ s = t}$
$r = t$

$RE_{subst.}$ Suppose that x_1 ,..., x_m are the free variables occurring in the identity s = t and that r_1,..., r_m are terms. Let s' be the result of simultaneously substituting the terms r_1,..., r_m for the variables x_1 ,..., x_m in s; and likewise for t' and t. Then

$\underline{s = t}$
$s' = t'$

N.B. This rule also covers the case where we substitute terms for only some of the free variables in s = t (and in particular the case where we do this for only one variable). In such cases we choose for each variable x_i that we want to "leave alone" that variable itself as term r_i.

$RE_{repl.}$ Suppose that s has an occurrence as a subterm in t and that t' results from t by replacing this occurrence of s in t by the term s'. Then

$$\frac{s = s'}{t = t'}$$

A *EL Derivation (Equational Logic derivation) from* a set of equations Γ *in* an algebraic language L is a sequence $<E_1, ..., E_p>$ of identities of L in which each line E_i either (i) is a member of Γ, or (ii) results from an application of $RE_{refl.}$, or (iii) comes from one or more earlier lines by an application of one of the rules RE_{sym} - $RE_{repl.}$.

Exercise. Show that the proofs of T_{G1}.T1 - T_{G1}.T3 from T_{G1} can be turned into derivations of Equational Logic.

Theorem 12 (Completeness Theorem for Equational Logic).

Suppose that L is an algebraic language and that $\Gamma \models E$, where E is an identity of L and Γ is set of identities of L. Then $\Gamma \vdash_{eq} E$
(That is, there is a derivation in Equational Logic of E from Γ in L.)

Proof: As in the completeness proof for the first order predicate logic we proceed by contraposition. Suppose that it is not the case that $\Gamma \vdash_{eq} s_o = t_o$. We construct a model M such that $M \models \Gamma$ but not $M \models s_o = t_o$. (Recall in this connection that the identities are really universally quantified formulas. Thus $M \models \gamma$ means that for all possible value assignments **a** to the variables of γ $[[\gamma]]_{M,\mathbf{a}} = 1$. On the other hand, in order to show that not $M \models s_o = t_o$ it suffices to find one assignment **b** such that $[[s_o = t_o]]_{M,\mathbf{b}} \neq 1$.)

Informally, we proceed as follows: We identify all terms s, t for which the identity s = t is derivable from Γ. The (equivalence) classes [s], [t], ... obtained in this way will be the elements of the universe of M. We can then define on this universe the interpretations of the function constants of L so that the identities in Γ are all universally satisfied in M. Since it is not the case that $M \models s_o = t_o$, s_o and t_o will not belong to the same equivalence class; hence if **b** assigns to each of the free variables of $s_o = t_o$ its own equivalence class, then $[[s_o]]_{M,\mathbf{b}} \neq [[t_o]]_{M,\mathbf{b}}$.

Formally: Let the relation $\sim_\Gamma$ on the terms of L be defined by:

(1) $s \sim_\Gamma t$ iff $\Gamma \vdash_{eq} s = t$.

Because of the rules $RE_{refl.}$, $RE_{sym.}$ and $RE_{trans.}$ $\sim_\Gamma$ is an equivalence relation. So we can form the corresponding equivalence classes $[t]_{\sim_\Gamma}$. Let $U_M = \{[s]_{\sim_\Gamma}: s \text{ a term of L}\}$. Furthermore, in virtue of $RE_{repl.}$, $\sim_\Gamma$ is a *congruence relation with respect to* each function constant f^n of L, that is:

(2) when for i = 1, .., n, $s_i \sim_\Gamma t_i$, then $f(s_1, ..., s_n) \sim_\Gamma f(t_1, ..., t_n)$.

This means that the following definition of the interpretation f_M of f^n in M is coherent and defines a total function on U_M:

(3) $\quad <[t_1]_{\sim_\Gamma}, .., [t_n]_{\sim_\Gamma},[t]_{\sim_\Gamma}> \varepsilon f_M$ iff $\Gamma \vdash_{eq} f(_1, .., t_n) = t$

(As regards totality of f_M: EQ1 guarantees that there is at least one term t such that "$f(t_1, .., t_n) = t$" is derivable, viz. $f(_1, .., t_n)$.)

This completes the definition of M. To show that M is a countermodel to the claim that $\Gamma \models s = t$ we first establish the following:

(4) Let r be any term of L with variables $x_1,.., x_n$ and let **a** be an assignment in M such that for j = 1, ..., n, $\mathbf{a}(x_j) = [x_j]_{\sim_\Gamma}$. Then $[[r]]_{M,\mathbf{a}} = [r]_{\sim_\Gamma}$.

(4) is proved by a simple induction on the complexity of r.

We now show that for each $E_i \varepsilon \Gamma$, E_i is true in M. Suppose that E_i is the equation $s_i = t_i$. Recall that "equations" are really *sentences*, which are obtained from the bare equations by universally quantifying over all the variables occurring in them. So in order that the equation $s_i = t_i$ is true in M it is necessary and sufficient to show that for arbitrary assignments **a** in M, $[s_i = t_i]_{M,\mathbf{a}} = 1$.

Assume that $x_1,.., x_n$ are the variables occurring in $s_i = t_i$. Let **a** be any assignment in M. Suppose that for j = 1, ..., r, $\mathbf{a}(x_j) = [r_j]_{\sim_\Gamma}$. Let s_i' be the term $s_i[r_1/x_1,..., r_n/x_n]$ - i.e. s' is the result of simultaneously substituting the terms r_j for the variables x_j in s - and similarly for t_i and t_i'.

Since $s_i = t_i \varepsilon \Gamma$, we have, trivially, $\Gamma \vdash_{eq} s_i = t_i$. So, by the rule $RE_{subst.}$ it follows that we also have $\Gamma \vdash_{eq} s_i' = t_i'$. So

(5) $[s_i']_{\sim\Gamma} = [t_i']_{\sim\Gamma}$.

Let $y_1,.., y_m$ be all the variables occurring in $s_i' = t_i'$ and let $\mathbf{a}'$ be an assignment such that for h = 1, ..., m, $\mathbf{a}'(y_h]) = [y_h]_{\sim\Gamma}$. From Lemma 3, established in connection with the Completeness Proof for Predicate Logic in Ch. I, we know that:

(6) $[[s_i']]_{M,a'} = [[s_i]]_{M,\mathbf{a}''}$,

where $\mathbf{a}'' = \mathbf{a}'[\ [[r_1]]_{M,\mathbf{a}'}/x_1,..., [[r_n]]_{M,\mathbf{a}'}/x_n\ ]$.

By (4) we get (i)

(7) $[[s_i']]_{M,\mathbf{a}'} = [s_i']_{\sim\Gamma}$ and $[[t_i]]_{M,\mathbf{a}'} = [t_i']_{\sim\Gamma}$.

and (ii)

(8) $[[r_j]]_{M,\mathbf{a}'} = [r_j]_{\sim\Gamma}$, for j = 1,..., n.

From (8) it follows that $\mathbf{a}'' = \mathbf{a}'[\ [r_1]_{\sim\Gamma}/x_1,..., [r_n]_{\sim\Gamma}/x_n\ ]$. Thus $\mathbf{a}''$ and $\mathbf{a}$ coincide on the variables $x_1,.., x_n$. Therefore, since $x_1,.., x_n$ are all the (free) variables of $s_i = t_i$, it follows by Lemma 1 from Part I that

(9) $[[s_i]]_{M,\mathbf{a}} = [s_i]]_{M,\mathbf{a}''}$ and $[[t_i]]_{M,\mathbf{a}} = [t_i]]_{M,\mathbf{a}''}$.

From (5), (6) and (9) we get:

$[s_i]]_{M,\mathbf{a}} = [s_i]]_{M,\mathbf{a}''} = [[s_i']]_{M,\mathbf{a}'} = [s_i']_{\sim\Gamma} = = [t_i']_{\sim\Gamma} = [[t_i']]_{M,\mathbf{a}'} = [t_i]]_{M,\mathbf{a}''} = [t_i]]_{M,\mathbf{a}}$.

This establishes that $M \models \Gamma$.

To see that not $M \models s = t$, it suffices to note that it follows from (4) above that $[[s]]_{M,\mathbf{b}} \neq [[t]]_{M,\mathbf{b}}$, where $\mathbf{b}$ is an assignment such that $\mathbf{b}(w_i) = [w_i]_{\sim\Gamma}$, for i = 1,..., h, where w_1, ..., w_h are all the variables occurring in s = t. The existence of such assignments entails that the equational sentence s = t is false in M.

q.e.d.

It is striking how much simpler this proof is than the Completeness Proof we gave in Part I. In a way this should not come as a surprise since we are dealing with formulas of a comparatively simple logical structure. Still, it is to be notes that while the present result is weaker than the full completeness proof precisely in that it deals with a small subclass of formulas, it is stronger in that it shows that when G and E stand in the consequence relation then a proof can be found of a very special and simple form. The following Corollary makes this a little more explicit.

Corollary. If L is an algebraic language and $\Gamma \vdash E$, where, as above, E is an identity of L and Γ is set of equations of L and $\vdash$ is the proof relation of full first order logic, then $\Gamma \vdash_{eq} E$.

This Corollary follows immediately from the Theorem and the soundness of the proof relation $\vdash$. The result is interesting in its own right insofar as it gives a certain normal form for proofs whose premises and conclusion all have the simple form of a universally quantified equation.

(To turn a derivation within Equational Logic into a "simple" proof of the universal generalisation of the conclusion from the universal generalisations of the premises is not completely trivial but very nearly so. In particular, a little reflection makes clear that one can turn the proof into (i) a series of applications of UI to the needed premises and to the identiy axioms; (ii) a series of steps involving MP corresponding to the successive steps of the given Equational Logic proof; and (iii) UG on the variables of the conclusion.)

Note also that the present proof yields like the completeness proof we presented for the full predicate calculus the additional information that a countermodel never need be more than denumerable in size. From the proof we have just gone through this follows from the fact that for any of the languages L we consider in this script the set of terms is denumerable. So a model whose universe consists of equivalence classes of such terms can be at most denumerable.

It should be noted, though, that in the case of equational logic the counter models constructed in the completeness proof as we have presented it here are almost always denumerably infinite. The reason is simple and relates to equations of the form $v_i = v_j$, with variables on both sides of =. If any such equation is entailed by a given set of equations Γ, then this will be true for all of them. For it is easy to see

that any one entails any other. So we have only two possibilities as regards such equations: (i) for all i,j such that $i \neq j$, $[v_i]_{\sim_\Gamma} \neq [v_j]_{\sim_\Gamma}$, in which case the model $M_{\sim_\Gamma}$ will be infinite; or (ii) for some i,j such that $i \neq j$, $[v_i]_{\sim_\Gamma} = [v_j]_{\sim_\Gamma}$, in which case we have $[s]_{\sim_\Gamma} = [t]_{\sim_\Gamma}$ for all terms s, t. In this second case the model $M_{\sim_\Gamma}$ will have a universe consisting of only one element, viz. the set of all terms of L.

Identities (i.e. equational sentences) differ from purely universal sentences in general in that they have special preservation properties. More precisely, we have a preservation theorem for conjunctions of identities: A sentence of L is logically equivalent to a conjunction of identities iff it is preserved under (i) submodels; (ii) homomorphic images; and (iii) direct products.

Of the three model-theoretic relations that are involved in these preservation properties the first two -that of a model M being a submodel oof some other model M' and that of h being a homomorphism of a model M into a model M' have laready been defined (the first in Ch. 1 Sn. 1.5.2, Def. 20 and the second in this Chapter, Sn. 2.1.6, Def. 8).

The *direct product* $M_1 \otimes M_2$ of two models $M_1 = \langle U_1, F_1 \rangle$ and $M_2 = \langle U_2, F_2 \rangle$ of L is defined as follows: The universe U of the product is the set of all ordered pairs $\langle a,b \rangle$ with $a \in U_1$ and $b \in U_2$; and for any n-place function constant f, the interpretation F of f is the function defined as follows:

$$F(f)(\langle a_1,b_1 \rangle,.., \langle a_n,b_n \rangle) = \langle F_1(f)(a_1,.., a_n), F_2(f)(b_1,.., b_n) \rangle.$$

Def. 11 Let $M_1 = \langle U_1, F_1 \rangle$ and $M_2 = \langle U_2, F_2 \rangle$ be models for the algebraic language L. The *direct product of* M_1 and M_2 is the model $M = \langle U, F \rangle$, where:

(i) $U = \{\langle a,b \rangle : a \in U_1 \ \& \ b \in U_2\}$
(ii) $F(f) = \{\langle\langle a_1,b_1 \rangle,..,\langle a_n,b_n \rangle, \langle F_1(f)(a_1,.., a_n), F_2(f)(b_1,.., b_n) \rangle : a_1,.., a_n \in U_1 \ \& \ b_1,.., b_n \in U_2\}$

The direct product of M_1 and M_2 is denoted as $M_1 \otimes M_2$.

Exercise: Show that if E is an equation of L, M is the direct product $M_1 \otimes M_2$ of two models M_1 and M_2 for L, $M_1 \models E$ and $M_2 \models E$, then

$M \models E$.

Hint. First show, by induction on the complexity of terms t of L, that for any assignments **a** in M_1 and **b** in M_2, the product assignment $\mathbf{a} \otimes \mathbf{b}$ in $M_1 \otimes M_2$ assigns to t in $M_1 \otimes M_2$ the value $<[[t]]_{M_1,\mathbf{a}}, [[t]]_{M_2,\mathbf{b}}>$. Here $\mathbf{a} \otimes \mathbf{b}$ is the assignment which assigns to each variable v_i the element $<\mathbf{a}(v_i),\mathbf{b}(v_i)>$ of $M_1 \otimes M_2$.

Before we turn to the exact formulation and proof of the preservation result for conjunctions of identities, it will be useful to first make a general observation about a special type of model for algebraic languages. These are the so-called *term models*. We encountered an example of such a model in the Completeness Proof for Equational Logic just given, where we constructed a counter example to the consequence claim $\Gamma \models E$ in the form of a model M whose elements were equivalence classes of terms. In general, a term model for an algebraic language L is a model whose universe consists of equivalence classes of the terms of L, where these equivalence classes are generated by equivalence relations which are also congruence relations with respect to all the function constants of L.

More specifically, given a congruence relation $\sim$ of the set Te_L of all terms of L, the corresponding model $M_\sim$ will have for its universe the set $\{[t]: t \varepsilon Te_L\}$, where Te_L, and as interpretation for any n-place function constant f of L the function defined by:

$$f_{M\sim}([t_1]_\sim, ..., t[t_n]_\sim) = [f(t_1, ..., t_n)]_\sim$$

The term models for a given algebraic language L are situated between two extremes. At the one and of the spectrum we find the so-called *free algebra for* the language L. This is the model generated by the identity relation on the set Te_L. Obviously this is an equivalence relation and congruence relation wrt to all function constants of L. Its equivalence classes are all the singleton sets {t}, where $t \varepsilon Te_L$. We denote this model as $M_{fr(L)}$. Clearly any other term model $M_\sim$ for L, generated by some congruence relation $\sim$, is a homomorphic image of $M_{fr(L)}$. For it easy to see that the map $\{t\} \Rightarrow [t]_\sim$ is is homomorpism from $M_{fr(L)}$ onto $M_\sim$. At the other end of the spectrum we find the model generated by the universal relation U_{Te_L} on Te_L. Again, this

relation is an equivalence relation and congruence relation wrt. the function constants of L. The model generated by this relation has for its universe the singleton set $\{Te_L\}$, and the interpretation of the function constants are, of necessity functions which map the one tuple all of whose members are the one element of this universe to this element. Since every congruence relation $\sim$ is a refinement of U_{Te_L}, the model just described is a homomorphic image of the model $M_\sim$.

More generally, if $\sim_1$ and $\sim_2$ are equivalence and congruence relations on Te_L and $\sim_1 \subseteq \sim_2$, then $M_{\sim_2}$ is a homomorphic image of $M_{\sim_1}$. For the map h which maps each element $[t]_{\sim_1}$ of the universe of $M_{\sim_1}$ onto $[t]_{\sim_2}$ is a homomorphism from $M_{\sim_1}$ onto $M_{\sim_2}$. At the opposite and from we find the one element term algebra $<Te_L, F>$, where for any $f^n \varepsilon$ L, $F(f) = \{<Te_L, .., Te_L, Te_L>\}$ (with $<Te_L, .., Te_L, Te_L>$ the n+1-tuple all of whose members are Te_L).

Given any model M for L we can associate a term model with M in several ways. First, we can form the equivalence relation $\sim_M$ on the set of terms of L defined by: $s \sim_M t$ iff $M \models s \equiv t$. Evidently, the resulting term model $M_{\sim_M}$ will verify exactly the same equations as M. But beyond that it is not so easy to say how M and $M_{\sim_M}$ are related. A second method goes as follows. We extend L to a language L^+ with names for each of the objects in U_M. (i.e. $L^+ = L \cup \{c_a: a \varepsilon U_M\}$; cf. the definition of the diagram of M in Ch. 1.) Let M^+ be the expansion of M in L^+, i.e. $c_{a_{M+}} = a$ for $c_a \varepsilon L^+ \backslash L$ and otherwise M^+ is like M. Now let $\sim_{M+}$ be the relation between terms of L^+ defined by

$$s \sim_{M+} t \text{ iff } M^+ \models s \equiv t$$

$\sim_{M+}$ is an equivalence relation on Te_{L+} and a congruence relation wrt all function constants of L^+. Thus $M_{\sim_{M+}}$ is a well-defined model for L^+. In this case too an equation of L will be true in the derived term model iff it is true in the original model M. Moreover, since for distinct objects a and b in U_M, $M^+ \models c_a \neq c_b$, $[c_a]_{\sim_{M+}} \neq [c_b]_{\sim_{M+}}$. So the map $a \Rightarrow [c_a]_{\sim_{M+}}$ is a 1-1 map into the universe of $M_{\sim_{M+}}$. It is easy to verify that this map is an isomorphism between M and a submodel of $M_{\sim_{M+}}$, but in general this will be a proper submodel of $M_{\sim_{M+}}$. A third possibility is to form a model $M'_{\sim_{M+}}$, whose universe consists of the

equivalence classes under $\sim_{M+}$ of all the *closed* terms of L^+. Here, the map $a \Rightarrow [c_a]_{\sim_{M+}}$ is a 1-1 map *onto* the universe of $M'_{\sim_{M+}}$ and thus an isomorphism from M to $M'_{\sim_{M+}}$.

To conclude these remarks on term models, we recall an important property concerning the values of terms in term models which we established and made use of in the Completeness Proof above:

(*) Let $M_\sim$ be a term model for the language L based on the congruence relation $\sim$, let t be a term of L, let $x_1, ..., x_n$ be the variables occurring in t and let a be an assignment in $M_\sim$ such that for i = 1,.., n, $a(x_i) = [x_i]_\sim$. Then $[t]_{M_\sim, a} = [t]_\sim$.

As we have seen, (6) can be proved by a simple induction on the complexity of terms.

We are now ready to prove the mentioned preservation theorem for equations:

Theorem 13 (Birkhoff)

Let L be an algebraic language. A sentence A of L is logically equivalent to a conjunction of identities of L iff (a) A is satisfiable and (b) A is preserved under (i) submodels; (ii) homomorphic images; and (iii) direct products.

Proof

$\Rightarrow$ The direction from left to right is straightforward. Clearly each identity is preserved by taking submodels (since identities are purely universal sentences), direct products (since the matrix of an identity is an atomic formula); and homomorphic images (since the matrix has the form of an equation "s = t"). And since the individual identities satisfy these conditions, the same is obviously true of their conjunctions. Finally, if the truth of each such conjunction is preserved under the model relations in question, then the same will be true for any sentence that is logically equivalent to such a conjunction.

$\Leftarrow$ The hard part is (as always with preservation theorems) the direction from right to left. Suppose that A is a sentence that is

preserved under taking submodels, direct products and homomorphisms. Let $\Gamma = \{E: E \text{ is an identity such that } A \vdash E\}$. First we show that if $M{\sim}_\Gamma \models A$, then $\Gamma \models A$.

To show that $\Gamma \models A$, we have to show that if M is any model of Γ, then $M \models A$. In view of the Completeness Proof we know that it suffices to show this for denumerable models. So let M be a denumerable model of Γ. Let g be an assignment in M which maps the set of variables onto U_M. We extend g to the set of all terms of L by letting $g(t) = [[t]]_{M,g}$. Suppose that s and t are two terms of L such that $s \sim_\Gamma t$. Then $\Gamma \models s = t$. So, since $M \models \Gamma$, $M \models s = t$. So $[[s = t]]_{M,g} = 1$. So $[[s]]_{M,g} = [[t]]_{M,g}$. So the map g from terms t to elements $[[t]]_{M,g}$ induces a map from the equivalence classes $[t]_{\sim_\Gamma}$ onto the elements of M. It is also easily verified that this map is a homorphism. So, since A is preserved by homomorphisms and by assumption $M{\sim}_\Gamma \models A$, it follows that $M \models A$.

So we conclude that $\Gamma \models A$. But then there is a finite set of $E_1, .., E_n$ in Γ such that $E_1 \& .. \& E_n \vdash A$. So, since on the other hand $A \vdash E_i$ for all i ($1 \leq i \leq n$), $\vdash A \leftrightarrow (E_1 \& .. \& E_n)$.

It remains to show that $M{\sim}_\Gamma \models A$. Suppose not. Then $M{\sim}_\Gamma \models \neg A$. Let $(M{\sim}_\Gamma)^+$ be the expansion of $M{\sim}_\Gamma$ in some language $L^+ = L \cup \{c_a: a \in U_{M{\sim}_\Gamma}\}$ and let $D((M{\sim}_\Gamma)^+)$ be the set of (a) all equations s = t with s, t constant terms of L^+ and (b) all negations of such sentences. Then $D((M{\sim}_\Gamma)^+) \cup \{A\}$ is inconsistent. For if not, then $D((M{\sim}_\Gamma)^+) \cup \{A\}$ has a model. But this model will be (isomorphic to) an extension of $(M{\sim}_\Gamma)^+$. So $(M{\sim}_\Gamma)^+$ will be a submodel of this model and consequently, because A is preserved by taking submodels, $(M{\sim}_\Gamma)^+ \models A$. This contradicts the assumption that $M{\sim}_\Gamma \models \neg A$. Since $D((M{\sim}_\Gamma)^+) \cup \{A\}$ is inconsistent, there are $E_1, .., E_k, D_1, .., D_n$ in $D((M{\sim}_\Gamma)^+)$, where the E_i are of type (a) and the D_j of type (b) (see the def. of $D((M{\sim}_\Gamma)^+)$) and

(1) $A \vdash \neg (E_1 \& .. \& E_k \& D_1 \& . \& D_m)$

Since A does not contain any of the constants $\{c_a: a \in U_{M\sim\Gamma}\}$,

(2) $A \vdash (\forall x_1)..(\forall x_r) \neg (E'_1 \& .. \& E'_k \& D'_1 \& . \& D'_m)$

where (i) $c_{a_1}, ..., c_{a_r}$ are all the new constants occurring in $E_1, .. , E_k, D_1, .., D_m$, (ii) $x_1, ..., x_r$ are r new variables (i.e. variables not occurring in A or $E_1, .. , E_k, D_1, .., D_m$) and (iii) the E'_i and D'_j are the result of replacing in the E_i and D_j the constants c_{a_h} by the variables x_h.

First assume that k = 0 (i.e. all the conjuncts on the right hand side in (1) are of type (b)):

(3) $A \vdash (\forall x_1)..(\forall x_r) \neg (D'_1 \& . \& D'_m)$

Consider D'_1. Suppose D'_1 is the inequality $s_1 \neq t_1$. We know that the elements $a_1, ..., a_r$ of $U_{M\sim\Gamma}$ satisfy $s_1 \neq t_1$ in $M\sim_\Gamma$. This means that the identity $s_1 = t_1$ does not belong to Γ, for if it did it would be satisfied in $M\sim_\Gamma$ by all possible combinations of elements of $U_{M\sim\Gamma}$. So it is not the case that $A \vdash s_1 = t_1$. That is, A is consistent with $(\exists x_1)..(\exists x_r)\ s_1 \neq t_1$. So there is a model M_1 of $\{A\} \cup \{(\exists x_1)..(\exists x_r)\ s_1 \neq t_1\}$. So there are objects $a_{11}, ..., a_{1r}$ in U_{M_1} which satisfy $s_1 \neq t_1$ in M_1. In the same way we can find models M_j of $\{A\} \cup \{(\exists x_1)..(\exists x_r)\ s_j \neq t_j\}$ and sequences of objects $a_{j1}, ..., a_{jr}$ in their universes which satisfy $s_j \neq t_j$, for each of the remaining disjuncts D'_j. Let M be the direct product of the models M_j and let for i = 1, ..., r $b_i = \langle a_{1,i}, ..., a_{m,i}\rangle$. Then (i) since A is preserved by direct products, A holds in M and (ii) the sequence $\langle b_1, ..., b_r\rangle$ simultaneously satisfies all inequalities $s_1 \neq t_1, .., s_m \neq t_m$ in M. But the existence of such a model contradicts (3).

Now assume that k > 0. Consider E'_1. E_1 is of the form $s_1 = t_1$. Since $(M\sim_\Gamma)^+ \models E_1$, the elements $a_1, ..., a_r$ of $M\sim_\Gamma$ satisfy the equation $s'_1 = t'_1$ in $M\sim_\Gamma$. Now let $q_1, ..., q_r$, be terms of L such that for i = 1, .., r, $q_i \in a_i$. Then we have, for i = 1, .., r, $a_i = [q_i]\sim_\Gamma$. Let $z_1, ..., z_s$ be all the variables occurring in $q_1, ..., q_r$, and let b be an assignment such that for h = 1, ..., s, $b(z_h) = [z_h]\sim_\Gamma$. Then according to (*), we have for i = 1, ..., r that

(4) $[[q_i]]_{M\sim\Gamma},b = [q_i]_{\sim\Gamma}$.

Let s''_1 be the result of substituting the terms q_i for the variables x_i in s'_1; in the same way we obtain t''_1 from t'_1. By Lemma 3 of Ch. 1,

(5) $[[s''1]]_{M\sim\Gamma},b = [[s'1]]_{M\sim\Gamma},b'$,

where b' is the assignment which is like b except that for i = 1, ..., r, $b'(x_i) = [[q_i]]_{M\sim\Gamma},b$; and similarly for t''_1 and t'_1. But according to (4), $[[q_i]]_{M\sim\Gamma},b = [q_i]_{\sim\Gamma} = a_i$. So $[[s'1]]_{M\sim\Gamma},b'$ is the value of s'_1 in $M\sim_\Gamma$ under any assignment which assigns the a_i to the x_i, and the same is true for $[[t'1]]_{M\sim\Gamma},b'$. Since $M\sim_\Gamma \models s'_1 = t'_1\ [a_1, ..., a_r]$, it thus follows that

(6) $[[s''1]]_{M\sim\Gamma},b = [[t''1]]_{M\sim\Gamma},b$.

Now note that the variables in s''_1 and t''_1 are $z_1, ..., z_s$. So we can apply (*) once more, obtaining that $[[s''1]]_{M\sim\Gamma},b = [s''1]_{\sim\Gamma}$ and similarly for t''_1. So from (6) we conclude that $[s''1]_{\sim\Gamma} = [t''1]_{\sim\Gamma}$, that is:

(7) $s''_1 \sim_\Gamma t''_1$.

But this means that

(8) $\Gamma \vdash s''_1 = t''_1$.

Since $A \vdash \Gamma$, $A \vdash s''_1 = t''_1$, that is

(9) $A \vdash (\forall z_1)..(\forall z_s)(s'[qi/xi] = t'[qi/xi])$

Now substitute the terms $q_1,..,q_r$ for the corresponding variables $x_1,..,x_r$ throughout the matrix of the formula on the right of $\vdash$ in (2). This will turn the conjuncts E'_i, D'_j into new conjuncts E''_i, D''_j which are substitution instances of the E'_i and D'_j. From (2) we infer that

(10) $A \vdash (\forall z_1)..(\forall z_s) \neg (E''_1 \,\&\, .. \,\&\, E''_k \,\&\, D''_1 \,\&.\, \&\, D''_m)$

Note further that the argument we have just given for E'_1 applies equally to each of the other E'_j (if any) and that the choice of the terms q_i can be the same in each case (i.e. irrespective of which E'_j we consider. In other words we have:

(11) $A \vdash s''_j \equiv t''_j$, for j = 1, .., k.

Because of (9) we can eliminate the disjunct E''_1 from the negated conjunction. This reduces (10) to (12)

(12) $A \vdash (\forall z_1)..(\forall z_s) \neg (E''_2 \,\&\, .. \,\&\, E''_k \,\&\, D''_1 \,\&.\, \&\, D''_m)$

But because of (11), the same argument applies to each of the other E''_i (i = 2,.., k). So each of these conjuncts can be removed from (12) and we end up with a formula of the form (4) with each of the conjuncts satisfiable in $M_{\sim\Gamma}$. We have already seen that this leads to a contradiction.

q.e.d

Exercise. Let L be the algebraic language consisting of two 1-place function constants f and g. Let Γ be the pair of equations {f(x) = x, g(x) = x}. Show: there is no single equation E of L which is logically equivalent to the conjunction $(\forall x)(f(x) = x) \,\&\, (\forall x)(g(x) = x)$.

We conclude this section with the comment which we promised in the introduction to Section 2.2. There we noted that formulas that contain function constants may seem to carry, because of those function constants, additional quantificational information other than what is directly visible from the quantifiers that are overtly displayed. This extra information becomes explicit, when the formula is translated into one in which the function constants are replaced by predicates. In particular, this translation will normally convert a purely universal formula into one that is AE. In the light of this observation it might seem surprising that the preservation theorem for purely universal formulas which we proved towards the end of Ch. 1 applies not only to languages that only have predicates, but also to those some or all of whose non-logical constants are function constants. If it is true, one might ask, that in general a purely universal sentence with function constants has the force of an AE sentence, how then can it be that such formulas obey the same model-theoretic restrictions as the "genuinely purely universal" sentences which consist of a purely universal prefix

followed by a quantifier-free matrix in which there are no function symbols?

The explanation of this apparent paradox is that when we are dealing with a language L which has function constants, the submodel relation between models for L is subject to restrictions which do not play a role when we deal with models for languages which only have predicates. Whenever M = <U, F> is any model for a language without function constants and U' is a subset of U, then there is always a unique submodel M' = <U',F'> of M, in which F' assigns to each predicate P of the language the restriction to U' of the interpretation F(P) assigned to P in M. When the language L contains function constants, this no longer holds in general. Suppose for instance that L contains the 1-place function constant f and let M be any model for <U, F> and U' a subset of U. In order that there be a submodel M' = <U',F'> of M whose universe is U' it should be the case that the restriction of F(f) to U' satisfies the requirements for interpretations of 1-place function constants, viz that the interpretation is a function from the universe into itself. In general this won't be the case, for there may well be elements a ε U' such that F(f)(a) belongs to U \U'. In that case the pair <a,F(f)(a)> will not belong to F'(f), F'(f) will thus only be a partial but not a total function form U' into U' and thus unsuitable as interpretation for f.

The upshot of this is that when L contains function constants, then the submodel relation is much harder to satisfy than it is for pure predicate languages. Consequently truth preservation under arbitrary submodels is a condition that is easier to satisfy for such languages than for pure predicate languages - since there are fewer submodels, it is easier for a sentence to have the property that whenever it is true in a given model it is also true in all its submodels. In fact, the general validity of preservation theorem of Ch. shows that the extra quantificational complexity that formulas may seem to have because of containing complex terms is "matched" by the special constraints which function constants impose on the submodel relation.

Arguably this comment would have been more appropriate after the proof of the preservation theorem in Ch. 1. But since the general issue that prompted it was raised only in this chapter, this seemed the next best place to make the comment. For the preservation properties of universally quantified equations are, as Birkhoff's Theorem asserts, even stricter than those for purely universal formulas - preservation under formation of submodels being one (but only one) of the properties that distinguish sentences that are equivalent to a universally

quantified equation. Since universally quantified equations are preserved under submodel formation and since they too will usually produce additional existential quantifiers when translated into formulas with predicates, they too give rise to the apparent paradox of which we have spoken.

2.4.1 Unification

A very different conception and use of equations is found in connection with *unification*. Here equations are understood as constraints on a structure consisting of (presumably) connected objects which are represented by the variables of a given set $\mathbb{E}$ of equations. Thus the equations in the set $\mathbb{E}$ are not understood as universally true - i.e. as universally quantified sentences - but as "locally true" - i.e. as true of the particular objects which the variables occurring in $\mathbb{E}$ represent. What one is after is a particular set of values for the variables for which all the equations are satisfied.

In certain situations one moreover wants the simultaneous solution to $\mathbb{E}$ to be "provably correct". More specifically, what one is looking for is a way of specifying the values so that the fact that they form a solution to the equations becomes a fact of pure logic. There is one salient and natural way in which this may be accomplished, and it is this: Let L be the language of the equations in $\mathbb{E}$ and let $M_{fr(L)}$ be the free algebra for L. Suppose that $x_1, ..., x_n$ are the variables occurring in $\mathbb{E}$ and that a is an assignment in $M_{fr(L)}$ such that $[[E]]^{M_{fr(L)},a} = 1$ for all $E \varepsilon \mathbb{E}$. Suppose that for $i = 1, ..., n$, $a(x_i) = [r_i] = \{r_i\}$. It is easy to see that, supposing that E is the equation $s = t$, s' is the result of replacing $x_1, ..., x_n$ in s by $r_1, ..., r_n$ and likewise for t' and t, $[[E]]^{M_{fr(L)},a} = 1$ implies that the equation $s' = t'$ is a tautology, i.e. s' is the very same term as t'; and thus that $s' = t'$ is a (trivial) theorem of pure logic.

The problem of finding such a "logically valid" simultaneous solution to the equations in a given set $\mathbb{E}$ in the free algebra for L is known as the problem of (*term*)*unification*.

The problem of unification is usually stated as the question whether a set of equations has a *unifier* (or *unifying substitution*). Let us begin by introducing the relevant notions.

Def. 12 Let L be an algebraic language, X a set of variables.

i. A *substitution on* X *in* L is a function σ with domain X, which assigns each variable x_i in X a term r_i of L.

ii. Suppose σ is a substitution on X. There is a standard extension σ' of σ to the set of all variables, defined by

$\sigma'(v_j) = \sigma(v_j)$, if $v_j \in X$
$\sigma'(v_j) = v_j$ otherwise

Since there is an obvious 1-1 correspondence between substitutions on subsets X of the set of all variables and their extensions as just defined, we won't distinguish between them, using "σ" both to refer to the substitution σ on X itself and to its extension σ'.

iii. Let σ, τ be two substitutions. By $\sigma \circ \tau$, the *composition of* σ and τ, we understand the substitution ρ which assigns to each variable v_j the term $\rho(v_j)$ which we obtain by simultaneously substituting for the variables v_k occurring in $\sigma(v_j)$ the terms $\tau(v_k)$.

iv. Suppose that $\mathbb{E}$ is a set of equations of L and that σ is a substitution in L. Then σ is called a *unifier of* $\mathbb{E}$ iff for each $E \in \mathbb{E}$, $\models E[\sigma]$, where $E[\sigma]$ is the result of simultaneously substituting the terms $\sigma(v_j)$ for the variables v_j which have free occurrences in E.

The main result about unification is that for finite sets of equations the problem whether a unifier exists is decidable: There exists an algorithm (due to Martelli & Montanari), which will find a unifier in a finite number of steps if one exists, and will return a negative answer to the question, when there is no simultaneous solution. Moreover, the algorithm returns, in those cases where there is a solution, a so-called "most general unifier" for the given equation set.

Def. 13 Let $\mathbb{E}$ be a set of equations of L and σ a substitution in L. Then σ is called a *most general unifier of* $\mathbb{E}$ iff (i) σ is a unifier of $\mathbb{E}$; and (ii) for any unifier ρ of $\mathbb{E}$ there is a substitution τ such that $\rho = \sigma \circ \tau$.

Thm. 14

i. There exists an algorithm which (i) returns for any finite set of equations $\mathbb{E}$ of any algebraic language L in finitely many steps either a unifier σ for $\mathbb{E}$ or else the answer that no unifier of $\mathbb{E}$ exists.

ii. The unifier σ which the algorithm returns when $\mathbb{E}$ does admit of a simultaneous solution is a most general unifier for $\mathbb{E}$.

[Ref. ??]

N.B. 1. Note that when $\mathbb{E} = \{E_1, ..., E_n\}$, then, if σ is unifier of ,

$$\models (\forall x_1)...(\forall x_k)(E_1[\sigma] \,\&\, ... \,\&\, E_n[\sigma]),$$

where $x_1, ..., x_k$ are all the variables occurring in $(E_1[\sigma], ..., E_n[\sigma]$. This formulation is especially apt to show how strong a claim unifiability really is.

2. The unification problem is special in that it asks for a substitution which turns all equations in the set into tautologies. There are many situations where such a result is stronger than one really needs. Rather, what is wanted is a substitution which turns all equations into theorems of a given theory T:

$$\text{For all } E \in \mathbb{E},\ T \models E[\sigma]$$

It should be stressed that with each different T the corresponding unification problem one is dealing with is a different one; and as a rule the problems are very different indeed, involving very different combinatorics, as a function of the axiomatic principles that T includes. This is so in particular in certain cases where T is itself an equational theory. For a few simple examples of such equational theories T the unification problem has been showed to be undecidable - which is one indication of how different the problem may become when a non-tautological theory T is brought into play.

(Two references on Unification: (i) Martelli, A. & U. Montanari. An Efficient Unification Algorithm. ACM Transactrions on Programming Languages and Systems, April 1982. (ii) Lloyd, J.W., *Foundations of Logic Programming*. Springer, 1984)

2.5 Definitions.

It is common practice to extend given scientific theories by adding new notions via definitions. Sometimes the point of a definition is strictly one of notational convenience: the defined concept abbreviates a complicated expression in the "primitive" vocabulary of the theory (that is, of the vocabulary in which the theory is given initially) and thus allows simplification of statements which contain this expression as a part. In other cases the defined notion has a conceptual significance of its own, which will make it easier to understand and handle statements in which it is represented as a unit - i.e. by a single symbol or term - than they would be if the concept were circumscribed in the theory's primitive vocabulary. And in yet other cases the defined concept may be one that is directly accessible to empirical observation, and deserve to be made explicit by a separate definition for that reason. In fact, the method of introducing concepts by definition is so general and of such methodological importance that most textbooks on logic and/or scientific methodology devote a separate chapter to it.

Here we will look at issues connected with definitions within the specific context of theories formalised within first order logic. That somewhat limits the range of issues that the theory and pactice of definition give rise to in general. Nevertheless, there remain a number of useful things to be said and these we will address. (Something that does not fit within the setting we adopt here is the conceptually important question of (non-)circularity of definitions. We will have a few observations about this notion towards the end of the section.)

In relation to first order theories questions of definition arise in two different settings. The first is that implicit in what was said in the opening paragraph: We have a theory T of some first order language L and want to extend T by adding some notion by definition. Formally this will consist in (i) choosing a new symbol α for the notion that is to be added to it, (ii) extending the language L to the language $L' = L \cup \{\alpha\}$ and then (iii) extending T to the theory T' of L' which is obtained by adding the definition of α to T and then closing under logical consequence in L'. This is what might be called the *external perspective* on definition.

But questions of definition can also be raised from a theory-i*nternal* prespective. Suppose again that T is a theory of L but now α is a non-logical constant of L. We can then ask the question whether α could not be defined within T in terms of its remaining vocabulary: Is there a definition D of α in terms of the remaining vocabulary which (i) is a

theorem of T and (ii) will give us back all of T when combined with the reduction T' of T to the language L' = L\{α} (i.e. the theory which consists of all theorems of T that belong to L')? Or - to put the question a little more informally - could we not eliminate all statements involving α from T and then restore them again to T by adding D?

In order to state this second question with the necessary precision we need to first have a clearer notion of what a "definition" is. We just spoke of "adding a definition of α" to some first order theory. That implies that the defintion in question must be a first order sentence, which we can add to a theory as an additional axiom. But which sentences should qualify as possible definitions of some non-logical constant α? What do we, or should we, expect of a sentence that is to serve as a definition? There are two criteria that, as the result of discussions of the purpose and form of definitions that stretched over centuries, have emerged as the central functional requirements. These are:

(i) *conservativity*

and

(ii) *determination*.

(i) Conservativity is a notion that does not only arise in connection with definitions. Its general context is that of a theory T of some language L and an extension T' of T whose language is some extension L' of L. T' is called a *conservative extension of* T iff T' coincides with T as far as L is concerned: if A is a sentence of L, then A is a theorem of T' iff it is a theorem of T.

The notion of conservativity as definability constraint involves a straightforward application of the "conservative extension" relation. Intuitively, the constraint is that adding a definition D of a new notion α to a theory T should not introduce new information that is expressible in the primitive vocabulary L of T. The formal expression of this requirement is as follows: every sentence A of L that is a theorem of the theory $T' = Cl_{L'}(T \cup \{D\})$ (where as before L' is the language $L \cup \{\alpha\}$) is already a theorem of L; or, put in terms of the notion just introduced: T' is a conservative extension of T.

(ii) Determination is the principle that a definition D of α should fully determine the extension of α when the extensions of the notions in

terms of which D defines α are given. The formal characterisation of this condition is model-theoretic: Let T and T' be as under (i) and let M = <U,F> be a model for L that is a model of T. Then there should be one and only one way to expand M to a model M' = <U,F'> for the language L' = L $\cup$ {α} that is a model of M'. That is, there ought to be only one way of extending F to an interpretation function F' of the non-logical constants of L', i.e. only one way of adding an interpretation F'(α) for α which verifies all the additional theorems of T' (including, in particular, the new "axiom" D)

Of these two criteria determination is the stronger one; it entails conservativity. For supppose that T, T', L and L' are as above and that D satisfies determination of α in relation to T. That is:

(3) For every model M of T there is one and only one expansion M' of M to L' which is a model of T'.

To show that T' = $Cl_{L'}$(T $\cup$ {D}) is a conservative extension of T assume that (3) holds and that A is a sentence of L such that T' $\models$ A. We must show that T $\models$ A. Suppose that it is not the case that T $\models$ A, Then T $\cup$ {$\neg$ A} consistent. Let M be a model of T $\cup$ {$\neg$ A}. Then there will be no expansion M' of M that is a model of T'. For every such expansion will verify $\neg$ A, while A is a theorem of T'.[24]

With this we are now in a position to address the question what form a definition should have in order that the mentioned criteria are satisfied. Since determination entails conservativity, it suffices to consider just determination.

Within formal logic we find two different forms of definitions which both satisfy determination. For the first of these, known as *explicit definition*, this is almost trivial. For the second, *definition by recursion*, - also called "definition by induction", or "recursive definition" or "inductive definition" - determination isn't quite as obvious, but even for this type of definition it is relatively easy to see that all the familiar

---

24 It is natural to ask whether conservativity in its turn entails determination. As it stands, I so not know the answer to this question. (I suspect the answer must be known but i haven't done the extensive literatire check need to find out whether this is so.) my hunch is that the entailment in this direction does not hold. it may fold under certain restrictions, but I have no clear idea what these might be either.

instances do satisfy determination. In this section we will only consider explicit definitions.[25]

Explicit definitions are universally quantified biconditionals in which an atomic formula involving the symbol that is being defined stands to the left of $\leftrightarrow$ and its definition - some formula A of the language in which the new symbol is being defined - to its right. (The left hand side and the right hand side are often referred to as the *definiendum* and the *definiens* of the given definition.) Exactly what this comes to still depends on what type of symbol α - or, more accurately: what type of non-logical constant α - is being defined. If α is an n-place predicate P, then an *explicit definition for* α *in* a language L has the form specified in (4)

(4) $(\forall x_1)...(\forall x_n)(P(x_1,.., x_n) \leftrightarrow A(x_1,.., x_n))$,

where $x_1,.., x_n$ are n distinct variables and A is a formula of L not containing P whose only free variabes are $x_1,.., x_n$.

Explicit definitions of function constants are essentially of the same form, except that the atomic formula on the left reflects the fact that we are dealing with a function constant rather than a predicate constant. The form of an explicit definition for an n-place function constant is given in (5).

(5) $(\forall x_1)...(\forall x_n)(\forall x_{n+1})(f(x,.,x_n) = x_{n+1} \leftrightarrow A(x_1,.., x_n,x_{n+1}))$,

where $x_1,.,x_n, x_{n+1}$ are n + 1 distinct variables and A is a formula of L not containing f whose only free variabes are $x_1,.., x_{n+1}$.

It is easy to see that sentences of the form (4) sastisfy determination. Suppose again that T is a theory of L, that P does not belong to L and that we form the theory $T' = Cl_{L'}(T \cup \{D\})$ of the language $L' = L \cup \{P\}$, where D has the form given in (4). Let $M = \langle U,F \rangle$ be a model of T. The right hand side A of D has for its extension the set $[[A]]^M$ in M, where $[[A]]^M = \{\langle u_1,.., u_n\rangle$: for i = 1,..,n, $u_i \varepsilon U$ & $[[A]]^M[u_1,.,u_n] = 1\}$.
Let $M' = \langle U,F' \rangle$ be the expansion of M to L' defined by F' =

---

[25] Examples of recursive definitions will be encountered in the next section, where we deal with the axiomatisation of natural number arithmetic. In chapter 3 recusive definitions will be discussed in greater depth; there we will in particular look at the systematic connections that exist between recursive and explicit definitions.

$F \cup \{<P,[[A]]^M>\}$. It is easily verified that $M' \models T'$. (This follows from the fact that on the one hand $M' \models T$, while on the other the choice of F'(P) guarantees that $M' \models D$.) This establishes that there is at least one expansion of M which verifies T'. Secondly, suppose that $M'' = <U,F''>$ is another expansion of M such that $M'' \models T'$. Then in particular $M'' \models D$. This means that for every n-tuple $<u_1,..,u_n>$ of elements of U, $[[P(x_1,.,x_n)]]^{M''}[u_1,..,u_n] = 1$ iff $[[A]]^{M''}[u_1,..,u_n] = 1$.

In other words, $[[P(x_1,.,x_n)]]^{M''} = [[A]]^{M''}$, where

$$[[P(x_1,.,x_n)]]^{M''} = \{<u_1,..,u_n>: u_i..u_n \in U \ \& \ [[P(x_1,.,x_n)]]^M[u_1,..,u_n] = 1\}$$

and

$$[[A]]^{M''} = \{<u_1,..,u_n>: u_i \in U \text{ for } i = 1,..,n \ \& \ [[A]]^M[u_1,..,u_n] = 1\}.$$

But the first of these two sets is nothing other than F''(P) and the second set equals $[[A]]^M$. This entails that $[[P(x_1,.,x_n)]]^{M''} = [[A]]^M = [[P(x_1,.,x_n)]]^{M'}$ and thus that $M'' = M'$.

The case of (5) is a little more complicated. A definition D of the form (5) does not automatically guarantee determination, because the form of D imposes certain constraints on the semantics of its definiens A. D says that $A(x_1,..,x_n,x_{n+1})$ is equivalent to a statement of the form "$f(x_1,..,x_n) = x_{n+1}$". This means that in any model M' of D there will have to be for any n-tuple $<u_1,.,u_n>$ of elements of the unviverse exactly one u_{n+1} such that $[[A]]^{M'}[u_1,..,u_n,u_{n+1}] = 1$. This means that the correspnding "unique value" condition (6) for A will be a theorem of T', whether or not it is a theorem of T.

$$(6) \quad (\forall x_1)...(\forall x_n)(\exists y)(A(x_1,..,x_n,y) \ \& \ (\forall y')A(x_1,..,x_n,y) \rightarrow y' = y),$$

So if T' is to be a conservative extension of T, then (6) should be a theorem of T to begin with.

The upshot of this is that an explicit definition D of a function constant is acceptable as an addition to a theory T only if T already entails the corresponding unique value condition (6) for its definiens A. For only then will the addition of D be conservative. However, when this condition is fulfilled, then the addition of D will not only satisfy conservativity but also determination. (The argument is the same as for explicit definitions of predicates.)

The general moral of this discussion is that sentences of the form (4) and, with the qualifications just noted, also those of the form (5) satisfy the requirements we laid down for good definitions. This is consistent with the almost universal practice to cast definitions of new symbols in these particular forms.[26]

This concludes our discussion of the external perspective on the question what constitutes a proper definition, and we now turn to the internal perspective. In discussing the questions that this perspective gives rise to we follow the tradition in that we assume the notion of an explicit definition, as specified in (4) and (5), as our syntactic characterisation of proper definitions.

Suppose that T is a theory of the language L and that α is a non-logical constant of L. We already stated what it means for α to count as definable within T: there has to be some definition D of α in the language $L' = L\backslash\{\alpha\}$) such that $T = Cl_L(T' \cup \{D\})$, where $T' = T \cap \{A: A \in L'\}$. Now that we have adopted a specific syntactic characterisation of definitions we can turn this notion of definability into a strictly formal characterisation:

(7) Let T be a theory of a first order language L and α a non-logical constant of L. Let $L' = L\backslash\{\alpha\}$ and $T' = T \cap \{A: A \text{ is a sentence of } L'\}$. Then α is *explicitly definable in* T iff there exists an explicit definition D of α in L' such that $T = Cl_L(T' \cup \{D\})$.

We have already seen that when α is explicitly definable in T, then α is also *implicitly definable in* T, where implicit definability is characterised model-theoretically as in (8).

(8) Let T, L, α, L' and T' as in (7). Then α is *implicitly definable in* T iff the following condition holds:

Every model M' of T' can be expanded in one and only one way to a model M of T

It is an interesting fact that the converse of this implication - that implicit definability entails explicit definability - also holds. This result

---

[26] Recursive definitions are found almost exclusively within mathematics, something that has to do with the circumstance that they are suitable for domains that have the special "recursive" structure that such definitions presuppose.

differs from the statement that explicit definability entails implicit definability in that it depends on specific properties of first order predicate logic and is not generalisable to other logical formalisms (such as, for instance, higher order predicate logic). The result is known as *Beth's Definability Theorem*, after the Dutch logician E.W. Beth (1908-1964) who formulated and proved the theorem. To do justice to its importance we state Beth' Theorem once more, as a separate theorem with its own number.

Theorem 15 (Beth's Definability Theorem)

> Let L be a language of first order logic, α a non-logical constant of L and T a theory of L. If α is implicitly definable in T, then α is explicitly definable in T.

The proof of Beth's Theorem that we will present here is not the proof which Beth gave himself. But it is, I believe, the most popular proof of the theorem today. It makes use of another important theorem abolut first order logic, the so-called "Craig Interpolation Lemma". Craig proved this theorem on the way towards some other result in proof theory in which he was interested ar that point, hence the name "Interpolation *Lemma*". But it states a proposition which has come to be recognised as a salient fact about first order predicate logic in its own right. As in the case of Beth's Definability Theorem, there are other logical formalisms than first order logic for which the Interpolation Lemma does not hold, and in fact, validity of the Lemma has become (like the validity of Beth's Theorem) an important property in terms of which logical formalisms are classified. (Satisfying Craig's Lemma can be seen as a certain kind of well-behavedness for formal systems.)

The Interpolation Lemma says that if A and B are sentences of first order logic and $A \vdash B$, then there is a sentence C in the comon vocabulary of A and B such that $A \vdash C$ and $C \vdash B$. We can roughly paraphrase this as: That which is responsible for the fact that A is logically at least as strong as B can be articulated in just the terminology that is common to them both. A formal statement of the Interpolation Lemma is given as Theorem 16.

Theorem 16 (Craig's Interpolation Lemma).

Suppose that A is a sentence belonging to some first order language L_1, B a sentence belonging to some first order language L_2 and that L_1 and L_2 are compatible in that L_1 and L_2 assign the same signature to the

symbols they have in common. We denote the language whose non-logical constants are those common to L_1 and L_2 as L. Suppose that $A \vdash B$. Then there is a sentence C belonging to L such that $A \vdash C$ and $C \vdash B$.

The Interpolation Lemma can be proved quite easily on the basis of the completeness proof for first order logic that is given in the Appendix to Ch. 1. A proof of the Interpolation Lemma along those lines is given at the end of that Appendix. Here we will, as last item of this section, present a proof in which the same construction is used that is central to the completeness proof given in the main body of the text of Ch. 1 (see Section 1.2). This proof has an interest in its own right as a further application of the method used to prove completeness there, but it is more complicated than the one from the Appendix. (The central idea of this latter proof can be grasped immediately, although its technical details take up a certain amount of space.)

Proof of Beth's Theorem.

Beth's Theorem holds for arbitrary non-logical constants α. However, we will first give the proof for the case where α is an n-place predicate P. After completion of that proof we will then show how the case where α is a function constant can be reduced to the case where α is a predicate.

Suppose that L, T and α are as in the statement of the Theorem and that α is implicitly definable in T. Further assume that α is an n-place predicate P, that $L' = L\backslash\{P\}$ and that T' is the theory of L' defined by: $T' = T \cap \{A: A \text{ is a sentence of } L\}$. Let P_1 and P_2 be symbols not occurring in L and let L_1 and L_2 be the languages which result when we add, respectively, P_1 and P_2 as n-place predicates to L'. Let T_1 be the theory of L_1 which we get by replacing P in all theorems of T everywhere by P_1, and let, analogously, T_2. be the theory of L_2 which we get by replacing P in T everywhere by P_2. Let T_3 be the theory $CN_{L_3}(T_1 \cup T_2)$ in the language $L_3 = L_1 \cup L_2$. Then the following sentence (1) is a theorem of T_3:

$$(\forall x_1) \ldots (\forall x_n)\ (P_1(x_1, \ldots, x_n) \leftrightarrow P_2(x_1, \ldots, x_n)) \qquad (1)$$

That (1) is a theorem of T_3 can be seen as follows. Suppose that M_3 is any model of T_3. Let M_1 be the reduction of M_3 to L_1, M_2 the reduction of M_3 to L_2 and M' the reduction of M_3 to L'. Then M_1 is a model of T_1, M_2 is a model of T_2 and M' is a model of T'. Since by assumption P is

implicitly defined in T, the same is evidently true of P_1 in relation to T_1 and of P_2 in relation to T_2. Since P_1 is implicitly defined in T_1, there is exactly one expansion M_1' of M' which is a model of T_1. So $M_1' = M_1$, which means that the extension of P_1 in M_1' is the same as it is in M_1. Since T_2 is just like T_1 except for renaming of the predicate P_1 as P_2, the unique expansion M_2' of M to L_2 that is a model of T_2 will assign to P_2 exactly the same extension as M_1' assigns to P_1. And, as before, the extension of P_2 in M_2' is the same as the extension of P_2 in M_2. So all these extensions are the same and in particular the extension of P_1 in M_1 is the same as the extension of P_2 in M_2. As these are also the respective extensions of P_1 and P_2 in M_3, P_1 and P_2 have the same extension in M_3. So it follows that (1) holds in M_3. Since this is true for arbitrary models M_3 of T_3, (1) is a logical consequence of T_3.

Since $T_3 \vdash$ (1), we also have $T_3 \vdash$ (2), where (2) is the result of dropping the universal quantifiers of (1) and replacing the variables $x_1,.., x_n$ by fresh individual constants $c_1,.., c_n$, which do not belong to L':

$$P_1(c_1,.,c_n) \leftrightarrow P_2(c_1,.,c_n) \qquad (2)$$

Since $T_3 = CN_{L_3}(T_1 \cup T_2)$ and $T_3 \vdash$ (2), there are finitely many sentences $D_{11},.., D_{1m}$ from T_1 and there are finitely many sentences $D_{21}, .., D_{2n}$ from T_2 such that

$$\{D_{11},.,D_{1n}, D_{21},.,D_{2m}\} \vdash P_1(c_1,.,c_n) \leftrightarrow P_2(c_1,.,c_n). \qquad (3)$$

We can choose the sentences $D_{11},..,D_{1n}, D_{21},..,D_{2m}$ in such a way that n = m and that D_{2i} is the result of replacing P_1 in D_{1i} by P_2. Forming the conjunction D_1 of the D_{1i} and the conjunction D_2 of the D_{1i} we get

$$D_1 \,\&\, D_2 \vdash P_1(c_1,.,c_n) \leftrightarrow P_2(c_1,.,c_n) \qquad (4)$$

and

$$D_2 = D_1\,[P_2/\,P_1]. \qquad (5)$$

(4) entails (6):

$$D_1 \,\&\, P_1(c_1,.,c_n) \vdash D_2 \rightarrow P_2(c_1,.,c_n) \qquad (6)$$

Note that in (6) the formula to the left of $\vdash$ belongs to L'_1 and the formula to its right belongs to L'_2, where $L'_1 = L_1 \cup \{c_1,.,c_n\}$, and

similarly for L'_2. So the Craig Interpolation Lemma applies: There is a sentence C from the common language $L' = L \cup \{c_1,., c_n\}$ such that

$$D_1 \,\&\, P_1(c_1,.,c_n) \vdash C \qquad (7)$$

and

$$C \vdash D_2 \rightarrow P_2(c_1,.,c_n). \qquad (8)$$

Since C does not contain any occurrences of P_2, the proof of $D_2 \rightarrow P_2(c_1,.,c_n)$ from C will turn into a proof of $D_1 \rightarrow P_1(c_1,.,c_n)$ from C when we replace all occurrences of P_2 by P_1. So we have

$$C \vdash D_1 \rightarrow P_1(c_1,.,c_n), \text{ or, equivalently:} \qquad (9)$$

$$D_1 \vdash C \rightarrow P_1(c_1,.,c_n), \qquad (10)$$

Also, (7) can be turned into

$$D_1 \vdash P_1(c_1,.,c_n) \rightarrow C, \qquad (11)$$

and (10) and (11) give us

$$D_1 \vdash P_1(c_1,..,c_n) \leftrightarrow C. \qquad (12)$$

Since D_1 is a sentence from L_1, it does not contain any of the constants $c_1,..,c_n$. So (12) entails:

$$D_1 \vdash (\forall x_1)..(\forall x_n)\ (P_1(x_1,.,x_n) \leftrightarrow C'), \qquad (13)$$

where C' is the formula of L which we get by replacing the occurrences of $c_1,.,c_n$ in C by the variables $x_1,.,x_n$. Replacing P_1 in (13) throughout by P gives us

$$D \vdash (\forall x_1) .. (\forall x_n)\ (P(x_1,.,x_n) \leftrightarrow C'), \qquad (14)$$

where D is a sentence from T' and C'' is a formula from L. So

$$T' \vdash (\forall x_1) .. (\forall x_n)\ (P(x_1,..,x_n) \leftrightarrow C') \qquad (15)$$

which shows that P is explicitly definable in T'. q.e.d.

This concludes the proof for the case where α is a predicate. Suppose now that α is an n-place function constant f. We can reduce this case to the case where α is a predicate by replacing f by an n+1-place predicate P, where "$P(x_1,.,x_n,x_{n+1})$" expresses that $f(x_1,.,x_n) = x_{n+1}$. Let P be a symbol not occurring in L and let 'L be the language $(L\backslash\{f\}) \cap \{P\}$. Corresponding to each model $M = \langle U,F\rangle$ for L there is a model $'M = \langle U,'F\rangle$ for 'L, where for any n+1-tuple $\langle u_1,.,u_n,u_{n+1}\rangle$ of elements of U, $('F(P))(\langle u_1,.,u_n,u_{n+1}\rangle) = 1$ iff $(F(f))(\langle u_1,.,u_n\rangle) = u_{n+1}$. Conversely, for any model 'M for 'L there is a model M for L such that 'M corresponds to M in the manner indicated.

Let $^+$ be the translation function from L to 'L defined in Exercise EA2 of the Appendix to Ch. 1. $^+$ translates terms τ into formulas $\tau^+(y)$ and formulas A of L into formulas A^+ of 'L. As shown in EA2, $^+$ has the property that for any model M for L, corresponding model 'M for 'L and assignment **a** in M, $[[\tau^+(y)]]^{'M,\mathbf{a}} = 1$ iff $[[\tau]]^{M,\mathbf{a}} = \mathbf{a}(y)$ and $[[A^+]]^{'M,\mathbf{a}} = [[A]]^{M,\mathbf{a}}$.

Let 'T be the deductive closure of the set of $^+$-translations of the sentences in T: $'T = Cl_{'L}(\{A^+: A \in T\})$. Then it follows from the above remarks about $^+$ that for any model M for L we have $M \models T$ iff $'M \models 'T$, where 'M is the 'L-model corresponding to M. Moreover, the "reduction" of T to L' = L\{f} - i.e. the theory $T' = T \cap \{A: A$ is a sentence of L'} - is the same as the "reduction" of 'T to L'. (Note that the language L' can also be written as 'L\{P}.) From these observations we can infer that P is implicitly definable in 'T. For suppose that M' is a model of T'. Then there is by assumption a unique way to expand M' to a model M of T. It follows from what we have said that the model 'M for 'L coresponding to M is a model of 'T. So there exists an expansion of M' to a model of 'T. Moreover, if there were two different expansions $'M_1$ and $'M_2$ of M' that were both models of 'T, then the corresponding models M_1 and M_2 for L would be also different and they would be expansions of M' that would be both models of T, which would contradict the assumption that f is implicitly definable in T.

Since P is implicitly definable in 'T we can apply Beth's Theorem for the case of predicates and obtain as theorem of 'T an explicit definition for P of the form given in (16).

$$(\forall x_1) .. (\forall x_n)(\forall x_{n+1})(P(x_1,.,x_n,x_{n+1}) \leftrightarrow A) \qquad (16)$$

where A is a formula of the language L'.

At this point we must refer once more to the properties of the translation function $^+$. One further property of $^+$ is that the formula $(f(x_1,.,x_n) = x_{n+1})^+$ is logically equivalent to the formula $P(x_1,.,x_n,x_{n+1})$ and that this equivalence is preserved by logical operations which combine these atomic formulas with each other and with formulas from L' (which are not affected by $^+$). This entails that (16) is logically equivalent to the $^+$-translation of (19).

$$(\forall x_1) .. (\forall x_n)(\forall x_{n+1})(f(x_1,.,x_n) = x_{n+1} \leftrightarrow A) \qquad (17)$$

So since (16) is a theorem of 'T, (17) is a theorem of T.

This concludes the proof of Beth's Theorem.

q.e.d.

There is a striking similarity between Beth's Definability Theorem and the Correctness-and-Completeness Theorem for first order predicate logic. Each theorem states an equivalence between (i) a syntactic and (ii) a semantic condition, and in each case the one condition is existential and the other universal. In our original formulation of the (Correctness and) Completeness Theorem the syntactic condition is existential - **there exists** a proof of B from the premises A_1, .., A_n - and the semantic condition universal - **every** model which verifies A_1, .., A_n also verifies B. Similarly, in the case of Beth's Theorem the syntactic condition - explicit definability, i.e. the existence of an explicit definition of α which is a theorem of T - is existential and the semantic condition - implicit definability, the unique expandability of every model of T' to a model of T - is universal. But we can also turn things around by taking contrapositives. The two conditions connected by the Completeness Theorem are then an existential semantic condition - **there exists** a model which verifies A_1, .., A_n but fails to verify B and a universal syntactic condition - **no** formally correct proof is a proof of B from A_1, .., A_n. Similarly, taking contrapositives in the case of Beth's Theorem turns it into an equivalence statement between an existential semantic condition - **there is** a model of T' that either cannot be expanded to a model of T at all or else can be expanded to a model of T in more than one way - and a universal syntactic condition - **no** explicit definition of α is a theorem of T.

When the Correctness-and-Completeness Theorem is stated as the equivalence between the negated conditions mentioned above - there

exists a "countermodel", in which A_1, .., A_n are true and B is false iff there is no derivation of B from A_1, ..,A_n -, then the hard part (completeness) is to prove that non-existence of a proof of B from A_1, .., A_n entails the existence of a countermodel. The converse - that the existence of a countermodel entails that there is no proof of B from A_1, .., A_n; in other words, the correctness of the given proof procedure - is generally easier (although how easy will depend somewhat on the proof procedure for which correctness and completeness are being proved). In the case of Beth's Theorem the difference between the two directions is even more striking. When there is a model of T' which either has no expansion or else more than one expansion to a model of T, then obviously it cannot be the case that T contains an explicit definition of α as a theorem. It was Beth's striking accomplishment to succeed in proving the converse of that.

In fact, the easy direction of the equivalence between implicit and explicit definition had been known for many years before Beth proved his Theorem. And it was one half of that easy direction - that the non-existence of an explicit definition of α in T can be established by finding a model of T' that can be expanded in more than one way to a model of T - which had gained currency under its own name, viz. as the "Method of Padoa", after the Italian mathematician Alessandro Padoa (1868-1937). It was by pursuing the question whether Padoa's Method was a necessary as well as a sufficient condition for the non-existence of an explicit edfinition of α in T that Beth was led to the proof of his definability theorem.

Internal definability questions - Is, for given T and $\alpha \varepsilon L_T$, α definable in T? - are sometimes easy to answer, but they can also be very hard. Examples of fairly easy questions of this kind we have observed earlier in this Chapter in connection with the Theory of Boolean Lattices and the Theory of Algebras. In the theory T_{bla} of Boolean Algebras given in Section 2.1.3 the operation $\cap$ is definable in terms of $\cup$ and - and, conversely, $\cup$ is definable in terms of $\cap$ and -. To show this is straightforward since in this case explicit definitions are easy to find: $\cap$ is definable in T_{bla} in terms of $\cup$ and - by the definition $(\forall x)(\forall y)(\forall z)(x \cap y = z \leftrightarrow z = -(-x \cap -y))$ and $\cup$ is similarly definable in terms of $\cap$ and - by a definition that is the "dual" of the one just given (i.e. one whose definiens is obtained by replacing in that of the given definition $\cup$ everywhere by $\cap$ and $\cap$ everywhere by $\cup$). We also saw that - is definable in T_{bla} in terms of $\cup$, $\cap$, 0 and 1, viz. by the definition $(\forall x)(\forall y)(-x = y \leftrightarrow (x \cup y = 1 \ \& \ x \cap y = 0))$.

In fact, there are even stronger definability results in this case: (i) the complement operation - is definable just in terms of $\cap$, for instance by the definition

(20) $(\forall x)(\forall y)(-x = y \leftrightarrow (x \cap y = 0 \ \& \ (\forall z)(x \cap z = 0 \rightarrow y \cap z = z)))$

where "$a \cap b = 0$" is short for: "$(\exists u)((\forall v)(u \cap v = u) \ \& \ a \cap b = u)$"

and (ii) - is definable just in terms of $\cup$, for instance by the definition

(21) $(\forall x)(\forall y)(-x = y \leftrightarrow (x \cup y = 1 \ \& \ (\forall z)(x \cup z = 1 \rightarrow y \cup z = z)))$.

(where "$a \cup b = 1$" is a similar abrreeviation as "$a \cap b = 0$")

The reason why (20) is a proper definition of - in T_{bla} is that it is one of the theorems of T_{bla} that for each x there is among the elements y such that $x \cap y = 0$ a unique largest one. Likewise, (21) is a proper definition of - in T_{bla} because T_{bla} has the theorem that for each x there is a unique smallest element y such that $x \cup y = 1$.

For the same reason the pseudo-complement - of pseudo-complemented lattices is definable in terms of $\cup$, $\cap$, 0 and 1. (See Section 2.2.1) For recall that one of the axioms of the theory of pseudo-complemented lattices says that for each x there is a unique largest y such that $x \cap y = 0$. But when the uniqueness requirement is dropped, the possibility of defining "-" in terms of these operations also disappears. More precisely, let T be the theory of the language $\{\cup,\cap,0,1,-\}$ which we get by adding to the axioms of T_{lata} the following sentence, which says that the meet of x and -x is always equal to the minimal element 0:

(21) $(\forall x)\ x \cap -x = 0$

In this theory there is no longer any guarantee that -x is unique and so there is no hope of defining - in terms of $\{\cup,\cap,0,1\}$.
That - is no longer definable can be seen as follows. Let $V = \langle U,\cup,\cap,\mathbf{0},\mathbf{1}\rangle$ be the lattice whose universe U consists of the elements $\{\mathbf{0},\mathbf{1},\mathbf{a}\}$ and the infinite set of elements $\{\mathbf{b}_n : n \in N\}$, where **1** is as always the largest and **0** the smallest element of the lattice and where the operations $\cup$ and $\cap$ are fixed by: (i) for all n, $\mathbf{a} \cup \mathbf{b}_n = \mathbf{1}$ and $\mathbf{a} \cap \mathbf{b}_n =$

0, and (ii) for all n, m such that n ≤ m, $\mathbf{b}_n \cup \mathbf{b}_m = \mathbf{b}_m$ and $\mathbf{b}_n \cap \mathbf{b}_m = \mathbf{b}_n$. Evidently V is a model of the theory T' consisting of those theorems of T that are expressible in the language {∪,∩,0,1}. We can extend V to a model of T in several ways by adding an extension for -. That is, we can choose $F_V(-)$ to be any of the following functions $-_n$ on U. The functions $-_n$ all coincide insofar as (i) $-_n(0) = 1$, (ii) $-_n(1) = 0$ and (iii) for all m, $-_n(\mathbf{b}_m) = \mathbf{a}$. But they differ from each other in the values they return for the argument **a**: for each n, $-_n(\mathbf{a}) = \mathbf{b}_n$. It is easily seen that each function $-_n$ yields a model of T when added to the model V of T'. So there is more than one way to expand V to a model of T.

Note that this argument is an application of Padoa's Method. In fact, to reach the conclusion that - is not definable in T it suffices to consider just two of the functions $-_n$, e.g. $-_0$ and $-_1$.

In the discussion above we have repeatedly used the phrase "α is definable in T in terms of ...", where the ... mention some of the other non-logical constants of L_T, but not necessarily all of them. We have so far only used this turn of phrase in connection with explicit definitions, and there it is immediately clear what is meant: a definition in which the definiens A contains only those non-logical constants that are mentioned in the dot part ...); thus α isn't merely claimed to be definable in the language L \ {α}, but in the sublanguage L' of L \ {α} which consists just of the symbols mentioned in the dor part. It is straightforward to also extend the characterisation of implicit definability to this more general case. All we need to do is to restrict the earlier characterisation of implicit definability to the sub-theories T' and T'' of T in the sublanguages L' and L'', where L' is the sublanguage just mentioned and L'' = L' ∪ {α}. To be precise, the characterisation of implicit definability of α in T in terms of the non-logical constants of L' now takes the following form:

(22) Let T be a theory of the language L. Let α be a non-logical constant of T, let L' ⊆ L\{α} and let L'' = L' ∪ {α}. Let T' = T ∩ {A: A is a sentence of L'} and T'' = T ∩ {A: A is a sentence of L''}.
Then α is said to be *implicitly definable in* T *in terms of* L' iff for each model M' of T' there is a unique expansion M'' of M' that is a model of T''.

It is left as an exercise to the reader to show that the corresponding version of Beth's Theorem holds:

(23) Let T, L and L' as in (22). If α is implicitly definable in T in terms of L', then there exists an explicit definition of α in terms of L' which is a theorem of T.

These generalised characterisations of implicit and explicit definability are convenient in particular in connection with a kind of application which we haven't yet mentioned, but of which there are many instances of the greatest importance. In such applications the focus is on particular structures - or, more precisely, on the descriptions of those structures in particular logical languages. Relevant examples that we have already entcountered are the structure of the rational numbers as described in the language {<}, and the Tarski Lattices for particular first order languages L as described in the language $\{\cup,\cap,0,1,-\}$.

Given a particular structure and a particular language in which it is described we can ask questions about the definability "within the given structure" of some of the notions represented in the describing language in terms of one or more of the others. Such questions can be phrased as definability questions of the kind we have been asking so far, i.e. as questions about the definability in a first order theory T of one non-logical constant α from the language of T, L_T, in terms of certain others. More specifically, they are questions of the form given in (24), where $\mathbb{S}$ is the structure in question, Th($\mathbb{S}$) is the set of all sentences of L_T that are true in $\mathbb{S}$ and L' is some sublanguage of $L_T \setminus \{\alpha\}$.

(24) Is α definable in the theory Th($\mathbb{S}$) in terms of the non-logical constants of L'?

In the next section we will study two structures that are at the very centre of mathematics. The first of these is "natural number arithmetic", i.e. the structure consisting of the natural numbers with the number null, the successor function S (where S(n) = n+1) and the operations of addition and multiplication; more explicitly, we will study the theory of natural number arithmetic as a theory of first order predicate logic formulated in the "language of Peano Arithmetic" - the first order language L_{PA} = {0,S,+,.}, where 0 is an individual constant, S a 1-place function constant and + and . are 2-place function constants. The second structure is that of real number arithmetic, i.e. the structure of the real numbers described in the first order language {+,.,<,0,1}, where + and . are 2-place function constants, < is a 2-place predicate constant and 0 and 1 are individual constants. About these and some other, related structures a range of questions of the general

form (24) can be asked - some easy, some hard and some with answers that have important further consequences.

The "Non-Circularity Requirement"

In the opening paragraphs of this section we promised a few words on the notion of definitional circularity. Many philosophical discussions of definitions make a big thing out of circularity, as something that is bad and should be avoided at all cost. Informally speaking, the basic cocern is something like this: Suppose you define a concept C in terms of certain other concepts C_1, ..., C_n. Suppose moreover that at the same time you define one of the C_i in terms of some further concepts one of which is C. That wouldn't be right, as the second definition would in all likelihood defeat the purpose of the first definition. For suppose you want to use the first definition to determine whether some given entities fall under C; then there is good chance that that will lead you consider whether certain entities, and quite possibly the same ones, fall under C_i. But to determine that you will, in all likelihood, be led to apply the second definition and that may get you involved in turn in questions about what falls under C; in particular, it may lead you back to the very same question that you started with.

We noted that circularity isn't really a topic that can be properly dealt with within the setting we have adopted - that of fully articulated theories formalised within first order logic. The difficulty can be illustrated at the hand of a very simple example. Consider the theory T_{lin} of arbitrary non-trivial linear orderings in the language $\{<,\leq\}$ according to which $<$ and $\leq$ stand in the familiar relation of a strict linear ordering and the corresponding weak ordering. We can axiomatise this theory by means of the axioms L1-L3 of Section 1.2.1 together with the sentences (25.i) and (25.ii).

(25) i. $(\exists x)(\exists y)\ x \neq y$

ii. $(\forall x)(\forall y)(x \leq y \leftrightarrow (x = y \vee x < y))$

Among the theorems of T_{lin} we find on the one hand the definition (25.ii) of $\leq$ in terms of $<$ and on the other - this is just as trivial to show - the definition (26) of $<$ in terms of $\leq$.

(26) $(\forall x)(\forall y)(x < y \leftrightarrow (x \leq y \ \&\ x \neq y))$

An obvious implication of this result is that for any given structure $\mathbb{S}$ which involves some linear ordering of its universe the weak ordering $\leq$ of the universe of $\mathbb{S}$ can be *defined in terms of* the strict ordering $<$ in the sense that (25.i) will be a theorem of the theory Th($\mathbb{S}$) for any first order language which includes the predicate symbols $<$ and $\leq$ and where these symbols are interpreted in $\mathbb{S}$ as $<$ and $\leq$. Conversely, in the same sense of 'define' $<$ can be defined in $\mathbb{S}$ in terms of $\leq$.

To repeat: $\leq$ can be defined for such structures in terms of $<$ and $<$ in terms of $\leq$. Does this mean that there is any circularity involved, of a sort that should be cause for worry? The answer would seem to be an obvious "no". You can define $\leq$ in terms of $<$ or you can define $<$ in terms of $\leq$; either is fine. What you *cannot* do, of course, is at the same time "define $\leq$ in terms of $<$ and $<$ in terms of $\leq$" - not at least if that were to mean that on the one hand you formulate the theory of linear orderings as one which uses $<$ as "primitive" - i.e. as a theory in the language $\{<\}$ - and then add $\leq$ as a defined concept (by extending the language $\{<\}$ to $\{<,\leq\}$ and adding, say, definition (25.ii) as a new axiom) - and also formulate the theory as a theory in the language $\{\leq\}$ and then extend that theory with a definition of $<$ (such as (26)). You have to make a choice: either formulate your theory in the language $\{<\}$ and then, if you wish, add $\leq$ by definition, or else formulate it in the language $\{\leq\}$ and then, if you wish, add a definition for $<$.

Surely the warning to avoid circularity can't be a warning against anything as obviously impossible as constructing a formal theory T whose language L_T is different from what it is. But then, what are the dangers of which we are being warned? To answer this it is important to realise that theory development is in general a very complex and protracted process, which typically runs through a number of successive stages. First, a body of data whose internal connections will often be quite poorly understood at hte outset must be structured into an organic, explanatory whole - into a "theory", in other words - and an essential part of that is to design the concepts in terms of which the central principles of the theory are to be stated. Exactly what these concepts stand for need not be fully clear from the start; often their true meaning will reveal itself only gradually, as the principles which make use of them become more firmly entrenched and their implications better understood (in particular those which link them to the data). Among the means of concept clarification that can be helpful

during this stage of theory development are definitions of one concept in terms of some of the others.[27] And such definitions may be useful even if some of the other concepts occurring in the definiens are still in need of further clarification in their turn. If, however, one then attempts to clarify one of those other concepts by means of a definition that employs the original concept C in its definiens, then that is a sign that something has gone awry. Trying to back a given definition of C with a further definition that makes use of C is a bit like putting up one piece of real estate as collateral when acquiring another, and then offering the second one as a collateral in an attempt to refinance the first. In business this is regarded as a form of fraud. Circular definitions won't land you in jail, but they too are violations of sound general principles and ought to be be avoided.

---

27 It is a remarkable fact that progress can be made in this way at all. Philosophers call this the "Paradox of Analysis": If we understand a conecpt C well enough to be able to judge a proposed definition as a correct definition of C, then how can that definition tell us anything about C that we didn't know already? There are, it would seem, just two possibilities: either we didn't know everything that the definition tells us, but then we are not in a position to recognise the definition as correct; or else we did already know all that it tells us, but then the definition cannot tell us anything about C that is really new to us; the best that it could do would be to give us something that we knew already in a different form. And yet it is undeniable that "explanatory" definitions - definitions of concepts we already have that seem right to us and that nonetheless reveal something new about the concepts they define - do play a sginificant part in theory development, and in concept formation generally.

How can a definition ever be explanatory in this sense? There are no easy answers to this question. But I think it is intuitively clear that any satisfactory answer must have to do in some essential way with the nature of human cognition. A person's thoughts form a complex web of propositional representations in which concepts are the principal building blocks. At the same time some of these concepts are linked to the external world by complex application criteria - criteria that determine for at least some real world entities that they belong to the extension of the concept, and for certain others that they do not, and which also enable us to recognise when this to the case. However, much of this - propositional representations as well as linking criteria - can be *implicit knowledge*: we can apply the criteria without being able to articulate them and we can draw inferences from the network of representations without necessarily being able to name or state all those parts of the network that serve as premises to the inference. Definitions which purport to reduce one concept to a number of others are among the most effective prompts for dragging to the surface of our awareness connections between two or more concepts that up to then were just implicit knowledge. In this way something that was known to us already in some hidden and nebulous way can acquire a new quality - become a "clear and distinct idea", to use Descartes' phrase. This may give us on the one hand the sense that we are learning something new while at the same time we can perceive that "new" piece of knowledge as agreeing with the implicit knowledge we already had As I said, this isn't much of an answer. But I think it indicates the direction in which we should look for one.

It doesn't follow from what has just been said that definitional circularity is a trap that it is easy to fall into. But it doesn't follow either that it is harmless altogether. There are at least two concomitant factors that contribute to the danger of being caught in it. First, definitional circles can be more concealed than they are in the simple case I have mentioned - they may involve not just two, but three definitions (D1 defines C with the help of C', D2 C' with the help of C'' and D3 C'' with the help of C) or even more than three. At a stage where one is still struggling for a better grasp of each of these concepts it is perfectly possible - and legitimiate - for all three definitions to be on the drawing board, each indicating a possible avenue of conceptual clarification. In this context the non-circularity principle can be seen as urging that a choice between those definitions will have to be made eventually: At least one of the definitions will have to be abandoned.

A second contributing factor is that theory development, and the conceptual analysis that is almost always an indispensable part of it in its earlier stages, is usually not a one-person enterprise but one that involves a group of investigators or even a whole scientific community. Different members of the group or community may come up with different definitions for different concepts. Taken together these definitions may well contain loops that no one member of the group or community is aware of; or else, individual members may not even be much concerned by such loops even if they see them, since they feel no commitment to one or more of the definitions involved. Once again, as a temporary state of affairs during the exploratory stage of theory development this situation need not be particularly objectionable. But, of course, by the time the theory has reached its definitive form all loops will have had to be eliminated.

When conceptual clarification has progressed to the point where logical formalisation becomes a meaningful option the explorations and debates that can lead to definitional circularity will normally have come to an end. At that point the hardest conceptual work that goes into developing the given theory will have been done as well. But this does not mean that logical formalisation should be seen as little more than a logician's pass time, from which nothing of substance can be learned that could not have been gathered just as easily from the theory before it is formalised. Within mathematics formalisation has led to numerous results that are not just of interest to formal logicians but are considered important by the community of mathematicians who deal with the branch of mathematics to which the given theory belongs, and who may have no particular interest in formal logic as such. Within the

empirical sciences formalisation has led to many important new insights too. Perhaps the single most important advance that has been achieved in this way within the general domain of empirical sicence is the formalisation of the concept of probability by Kolmogorov (1903-1987) (*About the Analytical Methods of Probability Theory*, 1831). Probability has become a central concept in all the empirical sciences, since it enters almost invariably in evaluating the truth or tenability of scientific hypotheses in the light of relevant data. Kolmogorov's axiomatisation has given us an understanding of the essentials of probability that, it seems fair to say, could not have been reached in any other way.

Proof of the Craig Interpolation Lemma.

Our last act in this section is the promised proof of the Craig Interpolation Lemma. (We remind the reader: an alternative proof can be found in the Appendix to Ch. 1.)

Proof of the Craig Interpolation Lemma.

Let A and B are as in the statement of the Interpolation Lemma and suppose that there is no C of L such that $A \vdash C$ and $C \vdash B$. We extend L to a language L' by adding an infinite sequence $\{c_i\}_{i \varepsilon N}$ of new constants. Similarly we extend, by adding this same set of constants, L_1 to L'_1 and L_2 to L'_2. Let $\{D_{i+1}\}_{i \varepsilon N}$ be an infinite sequence of sentences such that (i) the even-numbered sentences D_{2i} constitute a complete enumeration of the set of all sentences of L'_1 and the odd-numbered sentences D_{2i+1} a complete enumeration of the set of all sentences of L'_1. We proceed in a way reminiscent of the completeness theorem, extending once more given consistent sets in an infinite number of steps to maximal consistent sets. However this time we extend two sets in tandem and it is not just the consistency of the individual sets that we are interested in, but a kind of mutual consistency between them. More precisely, we generate two infinite sequences, a sequence $\{\Delta_{1i}\}_{i \varepsilon N}$ of finite but growing sets of sentences from L'_1 and a sequence $\{\Delta_{2i}\}_{i \varepsilon N}$ of finite but growing sets of sentences from L'_2. At each stage the pair $< \Delta_{1i}, \Delta_{1i}>$ is "compatible modulo L' " in the following sense:

(1) there is no sentence C of L' such that (i) $\Delta_{1i} \vdash C$ and (ii) $\Delta_{2i} \vdash \neg C$.

Note that if (1) holds, then both Δ_{1i} and Δ_{2i} are consistent. For suppose e.g. that Δ_{1i} were inconsistent. Then $\Delta_{1i} \vdash \perp_{L'}$, where $\perp_{L'}$ is some logical contradiction of L'; but then we would also have $\Delta_{2i} \vdash \neg\neg \perp_{L'}$, which would contradict (1). Consistency of Δ_{2i} is entailed for the same reason.

Our initial sets are singletons: $\Delta_{1o} = \{A\}$ and $\Delta_{2o} = \{\neg\neg B\}$, and our first task is to verify that these two satisfy (1). This, however, follows directly from the reductio assumption we have made about A and B.

The construction of the sequences $\{\Delta_{ji}\}$ proceeds as follows: At the even steps 2.i we operate on the set $\Delta_{1,2.i}$ and at the odd steps 2.i + 1 we operate on the set $\Delta_{2,2.i+1}$. We will state the rules according to which the sets are modified only for the even steps. The case for the odd steps is entirely symmetric.

Step 2.i:

Consider $D_{2.i}$. (i) When (1) holds for $\Delta_{1,2.i} \cup \{D_{2.i}\}$ and $\Delta_{2,2.i}$, then we add $D_{2.i}$ to $\Delta_{1,2.i}$, and, as in the Completeness Proof, we add, in case $D_{2.i}$ is an existential sentence $(\exists v_i)E$, then we also add a "witness sentence" $E[c_k/ v_i]$, where c_k is a constant which does not occur in either $\Delta_{1,2.i}$ or $\Delta_{2,2.i}$. Much as in the case of the Completeness Proof we can show that (1) is preserved also in the case where $\Delta_{1,2.i+1} = \Delta_{1,2.i} \cup \{D_{2.i}, E[c_k/ v_i]\}$, given that it holds for the pair $\langle\Delta_{1,2.i}, \Delta_{2,2.i}\rangle$.

(ii) When (1) does not hold for $\Delta_{1,2.i} \cup \{D_{2.i}\}$ and $\Delta_{2,2.i}$, then we add $\neg\neg D_{2.i}$ to $\Delta_{1,2.i}$: $\Delta_{1,2.i+1} = \Delta_{1,2.i} \cup \{\neg\neg D_{2.i},\}$.

We need to show that in each of the three cases condition (1) is preserved. Case (i) is automatic in case $D_{2.i}$ is not an existential sentence. Suppose instead that $D_{2.i}$ is the sentence $(\exists v_i)E$. In that case is $\Delta_{1,2.i+1} = \Delta_{1,2.i} \cup \{(\exists v_i)E, E[c_k/ v_i]\}$, with c_k a constant not previously used. Suppose that (1) fails for $\Delta_{1,2.i+1}$ and $\Delta_{2,2.i+1} = \Delta_{2,2.i}$. Then there is a sentence C of L' such that

(2) (i) $\Delta_{1,2.i} \cup \{(\exists v_i)E, E[c_k/ v_i]\} \vdash C$, and
(ii) $\Delta_{2,2.i} \vdash \neg\neg C$

From (2.i) we get that there is a sentence G ε $\Delta_{1,2.i}$ such that

(3) $\{G, (\exists v_i)E\} \vdash E[c_k/ v_i] \rightarrow C(c_k)$

(Here we have made explicit that C may contain the constant c_k.)

Since c_k does not occur in $\{G, (\exists v_i)E\}$, (3) entails

(4) $\{G, (\exists v_i)E\} \vdash (\forall v_i)(E \rightarrow C(v_i)/c_k))$, and from this

(5) $\{G, (\exists v_i)E\} \vdash (\exists v_i)E \rightarrow (\exists v_i)C(v_i)/c_k)$,

(5) evidently entails

(6) $\{G, (\exists v_i)E\} \vdash (\exists v_i)C(v_i)/c_k)$.

On the other hand, since does not contain c_k, (2,ii) entails

(7) $\Delta_{2,2.i} \vdash (\forall v_i)\neg C(v_i)/c_k)$, and thus

(8) $\Delta_{2,2.i} \vdash \neg(\exists v_i)\, C(v_i)/c_k)$.

Thus $(\exists v_i)\, C(v_i)/c_k)$ is a sentence of L' which is provable from $\Delta_{1,2.i}$, while its negation is provable from $\Delta_{2,2.i}$. This contradicts the assumption that (1) holds for $\Delta_{1,2.i}$ and $\Delta_{2,2.i}$.

Case (ii) is also somewhat different from the corresponding argument in the completeness proof. The argument now takes the following form.

Suppose that (1) fails for $\Delta_{1,2.i} \cup \{\neg D_{2.i}\}$ and $\Delta_{2,2.i}$. Then

(9) there is a sentence C of L' such that $\Delta_{1,2.i} \cup \{\neg D_{2.i}\} \vdash C$ and $\Delta_{2,2.i} \vdash \neg C$.

Recall, however, that in this case we also have a failure of (1) for the pair $<\Delta_{1,2.i} \cup \{D_{2.i},\}, \Delta_{2,2.i}>$. So

(10) there is a C' of L' such that $\Delta_{1,2.i} \cup \{D_{2.i}\} \vdash C'$ and $\Delta_{2,2.i} \vdash \neg C'$.

(9) and (10) entail on the one hand that $\Delta_{2,2.i} \vdash \neg\neg C$ and $\Delta_{2,2.i} \vdash \neg\neg C'$ and thus that $\Delta_{2,2.i} \vdash \neg\neg (C \vee C')$. On the other hand $\Delta_{1,2.i} \cup \{\neg\neg D_{2.i}\} \vdash C$ entails $\Delta_{1,2.i} \vdash \neg\neg D_{2.i} \rightarrow (C \vee C')$ and $\Delta_{1,2.i} \cup \{D_{2.i}\} \vdash C'$ entails $\Delta_{1,2.i} \vdash D_{2.i} \rightarrow (C \vee C')$.

These last two consequence relations jointly entail $\Delta_{1,2.i} \vdash (D_{2.i} \vee \neg\neg D_{2.i}) \rightarrow (C \vee C')$ and thus also $\Delta_{1,2.i} \vdash (C \vee C')$. So there is a sentence C'' of L' (viz. C v C') such that $\Delta_{1,2.i} \vdash C''$ and $\Delta_{2,2.i} \vdash \neg\neg C''$. So (1) fails for the pair $<\Delta_{1,2.i}, \Delta_{2,2.i}>$, contrary to assumption.

We now form $\Delta_1 = \cup \{\Delta_{1i}\}_{i \in N}$ and $\Delta_2 = \cup \{\Delta_{2i}\}_{i \in N}$. Much as in the proof of the Completeness Theorem, we can show that (1) holds for the pair $<\Delta_1, \Delta_2>$. This entails, as we have seen, that Δ_1 is a maximal consistent theory of L'_1 and that Δ_2 is a maximal consistent theory of L'_2. So, again as in the Completeness Proof, we can convert Δ_1 into a model M_1 for the language L'_1 which verifies precisely the sentences of Δ_1, and Δ_2 into a model M_2 for the language L'_2 which verifies precisely the sentences of Δ_2. We note the following: M_1 and M_2 have the same universe. For recall that if we proceed as in the Completeness Proof, then the universe U_1 of M_1 consists of equivalence classes $[c_i]_{\sim_1}$, where c_i is one of the new constants of L' and $\sim_1$ is the relation which holds between constants c_i and c_j iff the sentence $c_i = c_j$ belongs to Δ_1. Similarly, the universe U_2 of M_2 consists of equivalence classes $[c_i]_{\sim_2}$, where c_i is one of the new constants of L' and $\sim_2$ is the relation which holds between constants c_i and c_j iff the sentence $c_i = c_j$ belongs to Δ_2. It is to be stressed that in the present construction we take the elements of the universes U_1 and U_2 to consist *solely* of the new constants $c_i \in L' \setminus L$. (This means that we need a separate clause to deterine the denotations of the individual constants $c \in L_1 \cup L_2$. But this is unproblematic. For instance, assume that $c \in L_1$. Then there is a constant $c_i \in L' \setminus L$ such that $c_i = c \in \Delta_1$, i.e. $c \sim_1 c_i$. In this case we can unambiguously stipulate that $c_{M_1} = [c_i]_{\sim_1}$.) This entails that the two universes are in fact identical, since the relations $\sim_1$ and $\sim_2$ coincide. To see this, suppose for instance that $c_i \sim_1 c_j$. Then $c_i = c_j \in \Delta_1$. But then also $c_i = c_j \in \Delta_2$. For if not, then, by maximality of Δ_2, $c_i \neq c_j \in \Delta_2$. But then there would be a sentence C of L' (viz. $c_i = c_j$), such that $\Delta_1 \vdash C$ and $\Delta_2 \vdash \neg C$. So (1) would fail for $<\Delta_1, \Delta_2>$, which we

know already that (1) holds for these sets. Since $c_i = c_j \varepsilon \Delta_2$, $c_i \sim_2 c_j$. In the same way we show that if $c_i \sim_2 c_j$, then $c_i \sim_1 c_j$.

Not only do M_1 and M_2 have the same universes, they also assign the same interpretations to each of the non-logical constants of L'. For the new constants c_i this is immediate: $[c_i]_{M1}$ is the equivalence class $[c_i]_{\sim 1}$ and $[c_i]_{M2}$ is the equivalence class $[c_i]_{\sim 2}$, but these equivalence classes are the same. Now consider any non-logical constant α of L. Let us for simplicity assume that α is a 1-place predicate P. From the construction of M_1 we know that the extension of P in M_1, $[P]_{M1}$, consists of those equivalence classes $[c_i]_{\sim 1}$ such that the sentence $P(c_i) \varepsilon \Delta_1$. And by the same token, $[c_i]_{\sim 2} \varepsilon [P]_{M1}$ iff the sentence $P(c_i) \varepsilon \Delta_2$. But again we can infer from the fact that (1) holds for $<\Delta_1, \Delta_2>$ that $P(c_i) \varepsilon \Delta_1$ iff $P(c_i) \varepsilon \Delta_2$. For if not then we would have, say, $P(c_i) \varepsilon \Delta_1$ and $\neg P(c_i) \varepsilon \Delta_2$, so $P(c_i)$ would be a sentence C of L' contradicting (1). For non-logical constants of other types the argument is analogous.

We thus conclude that the reduction of M_1 to L' is identical with the reduction of M_2 to L'. This means that we can form the common expansion M_3 of M_1 and M_2 in that we add to their common reduction (i) the interpretations in M_1 of the non-logical constants of $L_1 \backslash L$ and (ii) the interpretations in M_2 of the non-logical constants of $L_2 \backslash L$. Since M_1 is the reduction of M_3 to L'_1, A, which is a sentence of L'_1, will have the same truth value in M_3 as in M_1. So A is true in M_3. An analogous argument shows that $\neg B$ is true in M_3. But this contradicts the assumption that $A \vdash B$. q.e.d.

2.6. Formalisations of Arithmetic

The first theory we looked at in this chapter aimed at giving as accurate a description as possible of one particular structure, viz. the ordering of the rationals. In that case our effort was as successful as a first order description of an infinite structure can be: the theory T_{rat} we formulated proved to be not only complete - in the sense that it captured as theorems all that can be said truly about that structure in the given first order language {<} in which T_{rat} was formulated - it even proved to be categorical in the cardinality of the target structure; every countably infinite model of T_{rat}, we found, is isomorphic to the ordering structure of the rationals.

The theories we have been looking at since then - lattices, distributive lattices, boolean algebras, groups - have for the most part been incomplete, and they were meant to be that. The aim of those theories was to capture what is common to a whole range of similar but non-identical structures, many of which differ from each other in ways that can actually be expressed in the language of the theory. In such cases the common core - the theory which consists of all sentences of the given language that are true in all the structures - is necessarily incomplete. It was only in a few cases - when we considered the theories of some particular orderings such as the ordering of the integers and that of the natural numbers or the theories of the Tarski Lattices of particular first order languages - that we were confronted once again with questions of the form: "What is the theory of *this* particular structure?

In this section we will focus once again on axiomatisation tasks connected with particular structures. We will be concerned with axiomatisations of two structures that occupy a central place in both pure and applied mathematics: (i) the structure of 'natural number arithmetic', i.e. the structure consisting of the natural numbers with the arithmetical operations + and . ; and (ii) the structure of 'real number arithmetic', i.e. the structure of the real numbers, also with the arithmetical operations + and . . The main results about axiomatisability of these two strutures are strikingly different, and at first sight they seem to contradict each other. The axiomatisation we will give for arithnmetic on the natural numbers will be, like any other axiomatisation for natural number arithmetic, incomplete and undecidable; these are the famous incompleteness and undecidability results for natural number arithmetic that we owe to Gödel. (Gödel's resuts will not be proved in this chapter). On the other hand, the axioms that we will give for arithmetic on the real numbers provide us with a complete axiomatisation of this kind of artihmetic. (For this result an explicit proof will be given here.)

How can this be, one might be tempted to ask? For it would seem obvious that airhtmetic on the real numbers is much richer than arithmetic on the natural numbers? and that the first includes the second as a part (and as a comparatively small and simple part at that). To put this intuition into a more concrete form: Couldn't one determine whether any arbitrary statement of natural number arithmetic is true by interpreting it as a statement of real number arithmetic (which speaks only of a small part of the real numbers, viz. the natural numbers) and then either derive or else refute this statement (as a statemrent about the reals) from our complete axiom

system for real number arithmetic? That would give as a decision method for number-theoretic truth; but that, Gödel proved, cannot be. The just mentioned results about natural and real number arithmetic thus entail that tstatements about antural number arithmetic cannot be interpreted as statements about the arithmetic of the real numbers. But why not? One of our tasks in this chapter will be to elucidate this apparent contradition.

2.6.1 The Natural Numbers and Peano Arithmetic.

The arithmetical structure $\mathbb{N}$ of the natural numbers consists of the numbers 0,1,2, ... ad infinitum, with the familiar operations of addition and multiplication. Our task in this subsection is to describe this structure by means of a first order theory.

Our first decision is to choose a suitable language. As we have seen repeatedly in this chapter, there usually is a certain freedom regarding this choice: We can choose one set of 'primitives' and then define the missing members of some other set in terms of them, or we can choose the other set and use those to define the missing members of the first set. Also it is not always desirable to keep the set of primitives as small as possible; it can be more perspicuous to choose a larger set, some members of which could also be defined in terms of the others and thus could have been dispensed with in principle.

This is the case for the language we will adopt for the description of $\mathbb{N}$. With the help of the operations + and . we can, given the right axioms, define a number of other notions, such as that of the number 0 (the unique number x with the property that for any number y, $y + x = y$); the number 1 (the unique number y such that $y \times y = y$); the successor function S, which assigns to each number the next one after it (this is the function which maps each number x onto $x + 1$), and the relation $<$ (which holds between x and y iff there is number $z \neq 0$ such that $y = x + z$). So none of these are absolutely indispensable. However, it has become standard practice to include both the constant 0 and the successor function S among the non-logical constants of the language of natural number arithmetic. Quite often the relation $<$ is included as well, but we won't do that here. So the language L_{PA} in which we will describe $\mathbb{N}$ has besides the 2-place function constants + and . the individual constant 0 and the 1-place function constant S, and that is it: $L_{PA} = \{0,S,+,.\}$.

In the literature on formal natural number arithmetic the following axiom set has gained wide currency. It is known as '(First Order) Peano Arithmetic', after the Italian mathematician Giuseppe Peano (1958-1932), who first formulated a set of axioms much like these. We refer to the theory axiomatised by PA1-PA7 simply as 'PA'.

PA1. $(\forall x)\ (x \neq 0 \leftrightarrow (\exists y)\ x = Sy)$[28]

PA2. $(\forall x)(\forall y)\ (Sx = Sy \rightarrow x = y)$

PA3. $(\forall x)\ x + 0 = x$

PA4. $(\forall x)(\forall y)\ x + Sy = S(x + y)\ )$

PA5. $(\forall x)\ \ x \ . \ 0 = 0$

PA6. $(\forall x)(\forall y)\ (x \ . \ Sy = (x \ . \ y) + x)$

PA7. $(\forall y_1)..(\forall y_n)((A[0/x] \ \& \ (\forall x)(A \rightarrow A[S(x)/x])) \rightarrow (\forall x)A)$, where $y_1, .., y_n$ are all the variables other than x which have free occurrences in A.

The rationale behind these axioms is as follows. The first two concern only the constant 0 and the function S and say that 0 is the only element that is not in the range of S and that S is 1-1. These axioms guarantee that 0 is the first of an infinite series of elements 0, S0, SS0, .. all of which are different from each other, and thus that all models of the axioms will be infinite. The next four axioms 'recursively define' the operations of addition and multiplication - PA3 and PA4 do this for +, PA5 and PA6 build on this definition in the recursive definition for . . The specifications of these axioms can be regarded as recursive definitions in that they specify an algorithm for computing the results of these operations, reducing all instances ultimately of cases involving 0. Thus PA3 and PA4 define n + m for any two numbers n and m, by reducing the result via n + (m-1), n + (m-2), ... eventually to n + 0. Likewise for PA5, PA6 and the terms 'n.m'.

This leaves PA7. Here, for the first time, we are dealing not with a single axiom, but with an *axiom schema,* which can be instantiated to an infinite number of different axioms by substituting different formulas of L_{PA} for the schematic letter A. The idea behind this schema is the following. The structure of the natural numbers makes it possible to prove that all natural numbers have a certain property P by mathematical induction: Show (i) that 0 has P and (ii) that for any

---

28 Where there is no danger of confusion we will wríte 'St' instead of 'S(t)'. Note that in the literature one often uses a prime ' instead of S. Thus one writes ' t' ' instead of 'S(t)'. Thus, in partocular the term " 0' " will be a term denoting the number 1.

number x that has P Sx also has P. That it follows from (i) and (ii) that every natural number must have P can be argued in a number of different (if fairly closely similar) ways. One informal argument goes like this: (i) tells us that 0 has P. From this and one application of (ii), taking x to be 0, we get that 1 has P. From this and a second application of (ii), now taking x to be 1, we get that 2 has P, and so forth. In this way we eventually reach every number n and establish that n has P.

Peano recognised that the Principle of Induction - that (i) and (ii) suffice to show that all numbers have P irrespective of what P may be - is one of the central characteristics of the natural number system. And he made it into the corner stone of his axiomatisation of $\mathbb{N}$. PA7' states this principle with the force he intended it to have, but in the notation of formal logic as we know it today.

PA7' $(\forall P)(P(0) \ \& \ (\forall x)(P(x) \rightarrow P(Sx)) \rightarrow (\forall x)P(x))$

The problem with (1) is that it is not a formula of first order logic. It isn't because it quantifies over the predicate symbol P. This means that P is a predicate variable and predicate variables are not part of first order logic. They are part of what is called 'Second Order Logic', a very powerful extension of First Order Logic in which we can quantify not only over individuals but also over sets of individuals. Second Order Logic has formal properties that are very different from those of First Order Logic.

We can use PA7' to obtain an axiomatisation of $\mathbb{N}$ within second order logic in which the other axioms are PA1-PA6. In one sense this axiom system is the perfect answer to our desire for an exhaustive description of the properties of $\mathbb{N}$. For it has the property that any model of it is isomorphic to $\mathbb{N}$. To see that this is so, we first need to make explicit what is meant by a predicate quantification like that in PA7'. The standard semantics of quantifications over predicate variables is that for any set X of individuals of the model M in which the formula containing the quantification is evaluated there is a predicate that can be a value for the variable and which has X as its extension in M. This means that predicate quantification comes to the same thing as quantification over sets, more precisely: over arbitrary subsets of the universe of the model. In particular, PA7' can be stated equivalently in the form PA7''.

PA7'' $(\forall X)(0 \in X \ \& \ (\forall x)(x \in X \rightarrow Sx \in X) \rightarrow (\forall x)\ x \in X)$

Given this interpretation of the quantification in PA7', we can argue as follows. Let M = <U,F> and M' = <U',F'> be two models of {PA1-PA6, PA7'}. Consider the universe U of M. It contains denotations in M of all the terms 0, S0, SS0, .. of L_{PA}. (We will refer to these terms as the *numerals* of L_{PA}. Thus a numeral is a term in which the constant 0 is preceded by some number n of occurrences of the function constant S, where $n \geq 0$.) Let us denote the element of U that is denoted in M by the term 'S...S0', in which '0' is preceded by n occurrences of 'S', as n_M. Let N_M be the set of all $u \in U$ that are denotations of numerals:

$$N_M = \{u \in U: u = n_M \text{ for some natural number } n\}$$
$$(= \{u \in U: \text{there is numeral } \nu \text{ of such that } u = [[\nu]]^M\})$$

Then N_M is the extension in M of a possible value for the predicate variable P in PA7'. It is clear that when P ia assigned this extension in M, then the formulas P(0) and $(\forall x)(P(x) \rightarrow P(Sx))$, which form the antecedent of the conditional in PA7', are satisfied in M. So it follows that the consequent of the conditional is satisfied as well, i.e. $(\forall x)P(x)$. But that means that every element of the model belongs to N_M and thus is the denotation of some numeral.

This argument is just as applicable to M', so its universe too consists of all and only the elements that are denotations of numerals. Given this it is easy to define an isomorphism h from M onto M': for every numeral ν, $h([[\nu]]^M) = [[\nu]]^{M'}$. (It follows from the argument above that h is well-defined and onto, from PA1 and PA2 that h is 1-1, from the definition of 'numeral' that h preserves S and from PA3-PA6 that h preserves + and . .)

This means that the theory PA^2 axiomatised by {PA1-PA6, PA7'} is semantically complete: For any sentence A of L_{PA} we have either $PA^2 \models A$ or $PA^2 \models \neg A$. (For either $\mathbb{N} \models A$, but then, since all models of PA^2 are isomorphic to $\mathbb{N}$, for all M such that $M \models PA^2$, $M \models A$; or else $\mathbb{N} \models \neg A$, and so for all M such that $M \models PA^2$, $M \models A$. But unfortunately this is not much help in deciding which sentences are true in $\mathbb{N}$ and which are false. For second order Logic has no complete proof procedure - there is no completeness result for Second Order Logic comparable to the completeness of First Order Logic we proved in Ch. 1. In fact, it follows from Gödel's Incompleteness Theorems that there can be no sound and complete proof procedure for second Order Logic, for then we would have a decision procedure for natural number arithmetic: to decide

whether a sentence A is true or false in $\mathbb{N}$, launch a simultaneous search for a proof of A from PA^2 and a proof of $\neg A$ from PA^2 and go on until one or the other is found; this must happen at some point in a systematic proof search, since one of A and $\neg A$ must be derivable from PA^2. But what Gödel proved is that there cannot be such a decision procedure.

PA7' is thus too much of a good thing. If we want to stay within First Order Logic, which does have completeness, the best we can do is to save from PA7' as much as can be expressed in first order terms. Presumably PA7 is the best one can do towards this end (though it seems hard to turn this intuition into a well-defined statement that we might be able to demonstrate formally). PA7 saves from PA7' all those cases in which the value of the predicate variable P is a property that is defined by some formula A(x) of the language L_{PA}. We write 'A(x)' to indicate that we think of x as the 'predicate bearer': A(x) is to be understood as the predicate that is true of an individual d in a model M iff $M \models A(x)[d]$. This means that the interesting cases are those where A has free occurrences of x. (If x does not occur free in A, then the 'predicate' A is either true of all individuals in the model or else of none.) On the other hand we allow A to have other free variables besides x. This form of PA7 is more comprehensive and thus gives a stronger axiom system. In some inductive proofs this extra strength is actually needed and in many others, where it could strictly speaking be avoided, it can be quite convenient. We will see an example of this in our sample derivation below.

To prove general properties of the natural numbers from the Peano axioms almost always involves induction, and thus an appeal to one or more instances of PA7. As an example we derive the 'commutative law for +', i.e. the sentence $(\forall x)(\forall y)\ y + x = x + y$. This is a very simple statement, which most people - and in particular non-mathematicians - would be inclined to think hardly worth attention. But even the derivation of this intuitively simple law takes some doing.

The strategy we will follow is the following. We will apply induction to the property that is expressed by the formula $A(x) \equiv (\forall y)\ x + y = y + x$. That is, we use the following instance of PA7:

$$((\forall y)\ 0 + y = y + 0\ \&\ (\forall x)((\forall y)\ x + y = y + x \rightarrow (\forall y)\ Sx + y = y + Sx)) \rightarrow (\forall x)(\forall y)\ x + y = y + x \qquad (1)$$

To derive the consequent $(\forall x)(\forall y)\ x + y = y + x$ of (1) we must prove the two conjuncts that make up its antecedent. We begin with the first conjunct:

$$(\forall y)\ 0 + y = y + 0 \qquad (2)$$

According to PA3 $y + 0 = y$. But how do we prove that $0 + y = y$? This requires another induction, this time wrt. y. To this end we use the following instance of PA7. (Of course this involves renaming variables, but we know we can always do that in the sense that every sentence logically entails all of its alphabetic variants. See Section 1.2.? of Ch. 1)

$$0 + 0 = 0 + 0\ \&\ (\forall y)(0 + y = y + 0 \rightarrow 0 + Sy = Sy + 0) \rightarrow (\forall y)\ 0 + y = y + 0 \qquad (3)$$

To prove the antecedent of (3) first observe that its first conjunct - $0 + 0 = 0 + 0$ - is a logical truth. To prove the second conjunct,

$$(\forall y)(0 + y = y + 0 \rightarrow 0 + Sy = Sy + 0), \qquad (4)$$

assume that $0 + y = y + 0$. We must show that $0 + Sy = Sy + 0$. We argue as follows. $0 + Sy =_{(PA4)} S(0 + y) =_{(Ass)} S(y + 0) =_{(PA3)} Sy =_{(PA3)} Sy + 0$. This shows that $0 + y = y + 0 \rightarrow 0 + Sy = Sy + 0$ and so by Universal Generalisation we get (4). From (3) and (4) we get (2) by M.P.

We now turn to the second conjunct of (1):

$$(\forall x)((\forall y)\ x + y = y + x \rightarrow (\forall y)\ Sx + y = y + Sx) \qquad (5)$$

Suppose that

$$(\forall x)((\forall y)\ x + y = y + x \qquad (6)$$

We must derive from this

$$(\forall y)\ Sx + y = y + Sx \qquad (7)$$

Take any y. By PA4 we have $y + Sx = S(y + x)$, which by assumption (6) equals $S(x + y)$, which by another application of PA4 equals $x + Sy$. But unfortunately $y + Sx = x + Sy$ is not what we want; what we want is

y + Sx = Sx + y. There is nothing for it but to prove the missing equality, x + Sy = Sx + y, separately, and that requires yet another induction.

In other words we must prove

$$(\forall x)(\forall y)\ x + Sy = Sx + y \qquad (8)$$

It is now convenient to take some arbitrary x and prove by induction that

$$(\forall y)\ x + Sy = Sx + y \qquad (9)$$

This requires another induction and thus another instance of PA7, to wit

$$(\forall x)((x + S0 = Sx + 0\ \&\ (\forall y)(x + Sy = Sx + y \rightarrow x + SSy = Sx + Sy)) \rightarrow (\forall y)\ x + Sy = Sx + y)^{29} \qquad (10)$$

(10) entails the free variable formula (11)

$$x + S0 = Sx + 0\ \&\ (\forall y)(x + Sy = Sx + y \rightarrow x + SSy = Sx + Sy) \rightarrow (\forall y)\ x + Sy = Sx + y \qquad (11)$$

To prove (9) from (11) we have to prove the antecedent of (11). Its first conjunct is straightforward:

$$x + S0 =_{(PA4)} S(x + 0) =_{(PA3)} Sx =_{(PA3)} Sx + 0$$

To prove the second conjunct,

$$(\forall y)(x + Sy = Sx + y \rightarrow x + SSy = Sx + Sy), \qquad (12)$$

assume that x + Sy = Sx + y in order to show that x + SSy = Sx + Sy. We have:

$$x + SSy =_{(PA4)} S(x + Sy) =_{(Ass)} S(Sy + x) =_{(PA4)} Sx + Sy$$

This shows (12). From (12) together with the first conjunct of (11) we get (9) and from this by Universal Generalisation (8). We already saw

---

29 Here we make use of the strong form of PA7, according to which A(x) may have free variables other than the "induction variable" x,

that with the help of (8) we can complete our derivation of (7) from (6). This completes the proof of (5) and thus of the second conjunct of the antecedent of (1). (5) and (2) give us the desired conclusion $(\forall x)(\forall y)\ y + x = x + y$.

q.e.d.

Exercise: Give a complete formal derivation of this result from the axioms of PA, using the rules of MP and UG, together with the axioms of predicate logic and previously proved logical theorems.

Exercise: Prove from PA1-7 the following theorems:

(i) $(\forall x)(\forall y)\ (x + y = y + x)$

(ii) $(\forall x)(\forall y)(\forall z)\ ((x + y) + z = (x + (y + z))$

(iii) $(\forall x)(\forall y)\ (x \cdot y = y \cdot x)$

(iv) $(\forall x)(\forall y)(\forall z)\ ((x \cdot y) \cdot z = (x \cdot (y \cdot z))$

(v) $(\forall x)(\forall y)(\forall z)\ ((x + y) \cdot z = (x \cdot z) + (y \cdot z))$

(vi) $(\forall x)(\forall y)\ (x = y \vee (\exists z)(z \neq 0 \ \&\ x = y + z) \vee (\exists z)(z \neq 0 \ \&\ y = x + z))$

Exercise: In PA we can define the order relation between the natural numbers by: $(\forall x)(\forall y)(x < y \leftrightarrow (\exists z)\ x + Sz = y)$.

(a) Show that for any numbers n and m, n is less than m (in the standard sense) iff $\mathbb{N} \models ((\exists z)\ x + Sz = y)[n,m]$.

(b) Interpreting 'x < y' as an abbreviation for '$(\exists z)\ x + Sz = y$' prove that the following are theorems of PA:

(i) $(\forall x)(\forall y)(x < y \rightarrow \neg(y < x))$

(ii) $(\forall x)(\forall y)(\forall z)(x < y \ \&\ y < z \rightarrow x < z)$

(iii) $(\forall x)(\forall y)(x < y \ \vee\ x = y \ \vee\ y < x)$

(iv) $(\forall x)(\forall y)(x < Sy \leftrightarrow (x = y \ \vee\ x < y))$

(v) $(\forall x)(\forall y)(\forall z)(x < y \leftrightarrow Sx < Sy)$

(vi) $(\forall x)(\forall y)(\forall z)(x < y \leftrightarrow x + z < y + z)$

Induction and Well-Foundedness

The validity of the method of mathematical induction rests on the fact that the "less than" relation between natural numbers is *well-founded*. This means that every non-empty set of natural numbers has a smallest member, a number such that no other number in the set is less than it. WF, in which X is assumed to range over subsets of N, expresses this fact formally.

(WF) $(\forall X)\ (X \neq \emptyset \rightarrow (\exists z)\ (z \in X\ \&\ (\forall u)(u \in X\ \&\ u \neq z \rightarrow \neg\ (u < z))))$[30]

WF entails the Principle of Induction. Consider for instance the 'subsets of N' version of the principle PA7''. That PA7'' follows from WF is easy to show. Suppose that X is a subset of N such that (i) $0 \in X$ and (ii) $\forall x)(x \in X \rightarrow Sx \in X)$. Suppose for the sake of arriving at a contradiction that it is not the case that X = N. Then the set Y = N\X is non-empty. So according to WF Y has a smallest member y_o. Since by assumption $0 \in X$, $y_o \neq 0$. So y_o must be a successor, i.e. there must be a number z such that $y_o = Sz$; obviously this entaikls that $z < y_o$. Since y_o is the smallest number of Y, z is not a member of Y and therefore a member of X. But then by property (ii) of X Sz - i.e. y_o - must also be in X and thus not in Y; and with thia we have our contradiction.

The relation between well-foundedness and the validity of the method of proof by induction holds more generally. First, well-foundednes is a property that can be defined for arbitrary strict partial orderings.

Def. 15 Let $<U,<>$ be a strict partial ordering.
$<U,<>$ is *well-founded* iff every non-empty subset of U has a minimal element. Formally:

$$(\forall X \subseteq U)(X \neq \emptyset \rightarrow (\exists z)\ (z \in X\ \&\ (\forall u)(u \in X\ \&\ u \neq z \rightarrow \neg\ u < z)))$$

To this general notion of well-foundedness corresponds a more general induction principle on partial orderings, As in Def, 15 let $<U,<>$ be a strict partial ordering.

(GIP) $(\forall X \subseteq U)((\forall x \in U)((\forall y \in U)(y < x \rightarrow y \in X) \rightarrow x \in X) \rightarrow U \subseteq X)$

Prop. 6 GIP holds for all well-founded strict partial orderings.

---

30 For the 'definition' of "<" see the Exercise at the end of the last section.

Prop. 6 can be proved in the same way as the special case we considered above where <U, < > was the ordering of the natural numbers.

The converse of Prop. 6 also holds: If GIP holds for <U,< >, then <U, < > is well-founded. The proof is left to the reader.

Well-foundedness is equivalent to the non-existence of infinite descending chains. An i*nfinite descending chain in* a strict partial ordering <U,< > is a function f from the set of the natural numbers N into U such that for all n f(n+1) < f(n). Clearly, well-foundedness of <U,< > entails the non-existence of such chains. For if there were such a chain, then Ran(f) would be a non-empty set without a first element. Conversely, if <U,< > is without infinite descending chains, then <U,< > must be well-founded. For suppose <U,< > were not well-founded. Then there would be a non-empty subset X of U without a minimal element. Let x_o be any element of X. We put $f(0) = x_o$. Since x_o is not a minimal element of X, there is an element x_1 in X such that $x_1 < x_o$. Put $f(1) = x_1$. Since x_1 is not minimal, there must be an element x_2 in X such that $x_2 < x_1$. So we can put $f(2) = x_2$; and so on ad infinitum. In this way we obtain an infinite descending chain f. (Warning: this second argument involves the Axiom of Choice. See Ch. 3 for discussion.)

Inductive proofs on well-founded partial orderings are very common in formal logic. We already encountered many examples of this, in particular in all those cases where we found it necessary or convenient to prove results by induction on the complexity of formulas. The partial order invoked in those proofs is that which holds between two grammatical expressions whenever the first is a constituent of the second. That such relations are always well-founded is plain: The easiest way to see this is to consider a well-formed expression together with all its syntactic constituents, Obviously there are no infinite descending chains of expressions, no infinite sequences of expressions in which each next element is a constituent of the last one. For each expression is built from basic expressions in a finite number of steps; so when we decompose an expression into its constituents, then we will get again to the bottom also in a finite number of steps.

As an example consider a language L of propositional logic with propositional constants p_o, p_1, .. and connectives $\neg$, &, v, $\rightarrow$, $\leftrightarrow$. The constituent relation between formulas of this language is of course

well-known by now, but for present purposes we will define it once again explicitly. We do this by first defining the relation of *immediate constituency*. The immediate constituency relation $<_{ic}$ for formulas of the given language consists of all pairs of the following forms:

<A, ¬A>, <A, (A & B)>, <B, (A & B)>, <A, (A v B)>, <B, (A v B)>,
<A, (A → B)>, <B, (A → B)>, <A, (A ↔ B)>, <A, (A ↔ B)>,

The general relation of constituency $<_{co}$, which holds also between A and B when A is not an immediate constituent of B, but, for instance, an immediate constituent of an immediate constituent of B, is defined as the *transitive closure* of $<_{ic}$. That is: $<_{co}$ holds between two formulas A and B iff there is a finite chain of formulas C_o = A, C_2, ... C_n = B, with n ≥ 1, so that for all i, $C_i <_{ic} C_{i+1}$.

Let U be the set of all formulas of L. Since <U,$<_{co}$>is well-founded, we can use GOP to prove that all formulas in U have a certain property P. Here is an example of such a property. Let NPC(A) be the number of occurrences of propositional constants in A and NBC(A) the number of occurrences of binary connectives in A. Then P is the property defined in (1)

$$NPC(A) = NBC(A) + 1 \qquad (1)$$

To prove that all formulas of L have P, suppose that X is the set of all formulas in U that have P. We show that

$$(\forall A \varepsilon U)((\forall B \varepsilon U)(B <_{co} A \rightarrow B \varepsilon X) \rightarrow A \varepsilon X) \qquad (2)$$

Suppose that A is any formula and that $(\forall B \varepsilon U)(B <_{co} A \rightarrow B \varepsilon X)$. We must show that A ε X. First suppose that A is a propositional constant. Then NPC(A) = 1 and NBC(A) = 0, so (12) is satisfied and A ε X.

Second suppose that A is of the form ¬C. Then C $<_{co}$ A. So C ε X and thus (12) holds for C. But then clearly (12) also holds for A, since adding a negation sign changes neither NPC nor NBC. So A ε X.

Finally suppose that A is built from two immediate constituents C and D, combined via a binary connective. For instance let A = (C & D). Then C $<_{co}$A and D $<_{co}$ A, so by assumption C, D ε X and therefore (12) holds for C and for D. Furthermore

$$NPC(A) = NPC(C) + NPC(D) \quad (3)$$

and

$$NBC(A) = NBC(C) + NBC(D) + 1 \quad (4)$$

So $NPC(A) = NPC(C) + NPC(D) =_{(Ind.Hyp.)} (NBC(C) + 1) + (NBC(D) + 1) = (NBC(C) + NBC(D) + 1) + 1 = NBS(A) + 1.$

So once more $A \varepsilon X$.

This concludes the proof of (2). With GIP we conclude that $X = U$, i.e. that all formulas have the property P and thus satisfy (1).

q.e.d.

Another way to justify the method of proof by induction on well-founded partial orderings is to reduce it to induction on the natural numbers via the notion of *rank*.

Def. 16 (of *rank*)

Suppose that $<U, <>$ is a well-founded strict partial ordering. Then we can assign elements x of U a rank by the following condition::

(i) If for no $y \in U$, $y < x$, then $rank(x) = 0$.

(ii) Otherwise $rank(x) = max(\{rank(y): y < x\}) + 1$

In general it is not clear that this will assign a rank to every element of U. For it is in principle possible that certain elements have 'infinite rank;. (For details see Ch. 3.). But in the case considered above, and similarly for all other cases where we have proved results by induction on partial orders so far in the Notes, every element of the ordering has 'finite rank', and in that case the interpretation of Def. 16 is unproblematic, and each element of U is assigned a finite number.

Given that all members of U have finite rank we can prove that all members of U have a certain property P by using induction over $\mathbb{N}$ to prove the following related property P' of natural numbers n, defined by

$$P'(n) \text{ iff } (\forall x \varepsilon U)(rank(x) \leq n \rightarrow P(x))$$

It is straightforwardly verified that the following two statements are equivalent:

(i) The instantiaton of GIP to the set X of all members of U that have P;

(ii) The instantiaton of PA7'' to the set X of all n that have P'.

Exercise: Check this for the example discussed above in which P is the property given by (1).

Extensions of PA and Non-Standard models of Arithemtic

It follows from Gödel's Incompleteness Theorems that PA is essentially undecidable: every consistent axiomatisable extension of PA is undecidable. Exactly what is meant by 'axiomatisable' here is something that we cannot properly account for with the means available to us. (Any account will presuppose a certain amount of Recursion Theory and as things stand, Recursion Theory is entirely missing from these Notes.) But for what we want to say here it suffices to note that finitely axiomatisable extensions of PA - extensions obtained by adding a finite number of axioms to those of PA - are axiomatisable extensions in the relevant sense. So it is true in particular that all finitely axiomatisable extensions of PA are undecidable.

This entails that every consistent finitely axiomatisable extension of PA must be incomplete. For suppose $T = CL_{L_{PA}}(PA \cup \{A_1,..., A_n\})$ were consistent and complete. Then we would have the following decision procedure for T: for any sentence B of L_{PA} start simultaneously a search for a derivation of B from T and a search for a derivation of $\neg$ B from T. A search for such a derivation can be set up in such a way that if there exists a derivation, then it will eventually be found: just enumerate all finite lists of sentences of L_{PA} and check whether they are correct derivations from T and whether they yield the target sentence as a theorem. When no finite list is left out by the search, the proof must be turned up at some point. Since by assumption T is complete, there must either exist a derivation from it of B or a derivation from it of $\neg$ B. So if both searches are carried out in tandem, then a proof of one of the two formulas will eventually turn up and that then tells us

whether B is a theorem of T: It is if the derivation that has been found is of B itself; it is not if the derivation is of its negation.

The fact that no finitely axiomatisable extension of PA is consistent and complete, means that the Tarski lattice $\mathbb{T}_{L_{PA},PA}$ is very rich. On the other hand, part of it admits of a comparatively simple characterisation. Let A_1, A_2,... be a complete enumeration of all sentences of L_{PA}. Pick the first sentence A from this list that is neither provable nor refutable in PA and form the to extensions $PA_{<0>} = CL_{L_{PA}}(PA \cup \{A\})$ and $PA_{<1>} = CL_{L_{PA}}(PA \cup \{\neg A\})$. (From now on we refer to a sentence that is neither provable nor refutable form a given theory as *independent from* T.) Both extensions will be consistent and incomplete. Consider $PA_{<0>}$. Since it is incomplete, there will be sentences that are neither provable nor refutable from it. Let $A_{<0>}$ be the first of these in our list. We form the extensions $PA_{<0,0>} = CL_{L_{PA}}(PA_{<0>} \cup \{A_{<0>}\})$ and $PA_{<0,1>} = CL_{L_{PA}}(PA_{<0>} \cup \{\neg A_{<0>}\})$ of $PA_{<0>}$. Similarly we can form consistent, but necessarily incomplete extensions $PA_{<1,0>} = CL_{L_{PA}}(PA_{<1>} \cup \{A_{<1>}\})$ and $PA_{<1,1>} = CL_{L_{PA}}(PA_{<1>} \cup \{\neg A_{<1>}\})$ of $PA_{<1>}$. Each of these four theories can then be extended in its turn into a pair of consistent, incomplete and mutually incompatible theories, and so on. In this way we obtain an infinitely branching binary tree all of whose branches are infinite.

Each branch B determines a theory T_B consisting of all sentences that belong to some node of the tree. (Exercise: prove that T_B is a theory.) Let us denote the successive nodes of B as $T_{B,1}$, $T_{B,2}$, .. . It is obvious that T_B is consistent. For its successive nodes are increasing in strength - for all n $T_{B,n} \subseteq T_{B,n+1}$. So if a contradiction were derivable from T_B, it would be derivable from some $T_{B,n}$, which is impossible since $T_{B,n}$ is consistent. Second, T_B is complete. For let C be any sentence of L_{PA}. Then C occurs somewhere in our list, say $C = A_k$. Then during the construction of the first k nodes $T_{B,1}$,.., $T_{B,k}$ of B C must have been considered at least once as a possible candidate for extending the theory $T_{B,i}$ that was up for extension. At that point there were two possibilities (a) C was independent from $T_{B,i}$. Then either $T_{B,i+1} = CL_{L_{PA}}(T_{B,i} \cup \{C\})$ or $T_{B,i+1} = CL_{L_{PA}}(T_{B,i} \cup \{\neg C\})$, so either C or $\neg C$ belongs to T_B. (b) C was not independent from $T_{B,i}$. That means that either C or $\neg C$ belongs to $T_{B,i}$; so again one of C and $\neg C$ belongs to T_B.

Furthermore, it is easy to show (i) that no T_B is finitely axiomatisable over PA - this follows from the fact that the theories are strictly

increasing in strength - and (ii) that if B and B' are different branches, then $T_B \neq T_{B'}$ - there must be some node T in the tree where B and B' part company and the two daughters R' and T'' of T that belong to B and B', respectively, will then differ in that for some sentence C, T' contains C and T'' $\neg$C. We conclude that there is a 1-1 correspondence between the complete consistent extensions of PA and the branches of our tree. From this we can infer that the cardinality of the set of all complete extensions of PA is that of the power set $P(\mathrm{N})$ of the set of natural numbers N.

So our tree gives a complete description of the complete extensions of PA. But it is not by any means an exhaustive representation of $\mathbb{T}_{L_{PA},PA}$. For one thing the extensions it represents, by its nodes and by its branches, are either finitely axiomatisable over PA (the nodes) or else complete (the branches). However, there are also many incomplete extensions of PA that are not finitely axiomatisable. Also, which finitely axiomatisable extensions turn up as nodes of the tree depends on the enumeration $A_1, A_2,..$ of the sentences of L_{PA}. And each enumeration will leave some of them out.

Exercise: Let the enumeration $A_1, A_2,..$ and the tree *T* of extensions of PA constructed on the basis of that enumeration be as described above.

(a) Show that the cardinality of the set of branches of *T* is that of the power set $P(\mathrm{N})$. (Hint: Show that there is a 1-1 correspondence between the set of branches and the set of all denumerably infinite sequences of 0's and 1's. Note that there is a 1-1 correspondence between the countable sequences of 0' and 1' on the one hand and the subsets of N on the other.)

(b) Show that for every complete and consistent extension T of PA there is a branch B of *T* such that $T = T_B$.

(c) Show that there are incomplete extensions of PA that are not finitely axiomatisable over PA.

(d) Show (i) that there are finitely axiomatisable extensions of PA that do not occur as nodes of *T*.

The second topic of this section concerns the models of PA. Models of PA that are not isomorphic to the standard model $\mathbb{N}$ are usually referred to as *non-standard* models. Even when we stay within the realm of the

denumerably infinite, the variety of models is very great. First, since PA is incomplete, many models differ from $\mathbb{N}$ in that they do not even verify the same sentences. Such models will not be considered here. Instead we concentrate on non-standard models of the theory Th($\mathbb{N}$), consisting of all sentences of L_{PA} that are true in $\mathbb{N}$. Even of such models there exists a great variety. (The cardinality of the set of isomorphism types of denumerable models of Th($\mathbb{N}$) is again that of $P(N)$.) Here we will only show how certain non-standard models of Th($\mathbb{N}$) can be constructed with the comparatively simple methods that are available to us.

The geneRAL method we will use consists in adding new individual constants to the first order language of the theory of departure and adding new sentences involving those constants to the theory. In the case at hand the language is L_{PA} and the theory is Th($\mathbb{N}$).

First a matter of terminology. The *numerals of* L_{PA} are the terms 0, S0, SS0, ... - in other words, all terms of the form S...S0 consisting of the constant '0' preceded by zero or more occurrences of the symbol 'S'. Note that in the standard model $\mathbb{N}$ every individual is the denotation of some numeral: If $n \varepsilon N$, then $n = [[\nu_n]]^{\mathbb{N}}$, where ν_n is the term consisting of one occurrence of 0 preceded by n occurrences of S.

We begin by adding just a single constant c to L_{PA}, thus obtaining the language $L_{PA} \cup \{c\}$, which we will denote for simplicity as L(c). Let S be the set Th($\mathbb{N}$) together with all sentences of the form $c \neq \nu$, where ν is a numeral of L_{PA}: $S = Th(\mathbb{N}) \cup \{c \neq \nu: \nu \text{ a numeral of } L_{PA}\}$. It is easy to show that S is consistent. For this it suffices to show that Th($\mathbb{N}$) together with any finite subset of $\{c \neq \nu: \nu \text{ a numeral of } L_{PA}\}$ is consistent. Sol et S' be such a finite subset. Let k be the largest number n such that the numeral ν_n occurs in the sentences of S'. Expand $\mathbb{N}$ to a model $\mathbb{N}'$ for L(c) by adding that interpretation of c which assigns it as its denotation the number k+1. Then the sentences of Th($\mathbb{N}$) are true in $\mathbb{N}'$ for the same reason that they are true in $\mathbb{N}$ and the sentences in S' are true in $\mathbb{N}'$ since the numerals they contain all denote numbers that are distinct from the denotation of c. So S' has a model and thus is consistent.

Since S is consistent, S has a model. And since any model of S will be infinite - this is because all models of Th($\mathbb{N}$) are necessarily infinite - it has a denumerably infinite model. Let M be such a model. Then M is

not isomorphic to $\mathbb{N}$. To see this, let us consider what an isomorphism h from $\mathbb{N}$ into M would have to be like. We begin by observing that every numeral will denote a unique element of M. We refer to the denotations of 0, S0, SS0, .., S^n0,.. in M as 0_M, 1_M, 2_M,., n_M,.. . It is clear that the number 0, which is the element of $\mathbb{N}$ that is the denotation in of the constant '0', can only be mapped by h onto the element 0_M of M. For if h is to be an isomorphism from $\mathbb{N}$ to M, then it must preserve in particular the denotation of '0'. So we have: h(0) = 0_M. By the same token, the number 1 can only be mapped onto 1_M, since 1 and 1_M are the denotations in $\mathbb{N}$ and M, respectively of the term 'S0'; the same applies for the numbers 2,3,.. and the elements 2_M, 3_M of M; and so on ad infinitum. So we have for every natural number n in N that h(n) = n_M.

This specifies h for all of N. But the range of h will not consist of all of M. For the truth in M of all the sentences in S entails that the denotation of c is different from all denotations of numerals in M and thus from all elements in the range of h. It is not hard to verify that h is indeed an isomorphism. (The axioms PA3-PA6 fix the extensions of + and . in both $\mathbb{N}$ and M in terms of the extensions of 0 and S. So if the latter are preserved by h, then so are the former.) But h cannot be onto M. Since there can be no isomorphism from *onto* M, $\mathbb{N}$ and M are not isomorphic.

It would be natural to try and push this method further to show that there are more isomorphism types of denumerable models of Th($\mathbb{N}$) than the two we have so far identified. But that is not easy. Additional or alternative techniques are needed to make further progress on this particular question, and many others like that. We do not pursue this issue any further here.

2.6.2. Arithmetic on the Reals.

We now turn to arithmetic of the real numbers. We mentioned in the introduction that this arithmetic admits a complete axiomatisation. Again the choice of a first order language in which the axiomatisation is to be formulated leaves some latitude. We follow the standard in adopting as language the language L_{Rea} = {0,1,+,.,<} (where 0 and 1 are individual constants, + and . are 2-place function constants and < 2-place predicate). Let $\mathbb{R}$ be the structure of the real numbers cast in hte form of a model for the language L_{Rea}. That is, $\mathbb{R}$ = <R,**0**,**1**,**+**, **.** ,**<** >, where R is the set of real numbers and **0**,**1**,**+**, **.** and **<** are the number

zero, the number one, the operations of addition and multiplication on the reals and the standard ordering of the reals, respectively.

The theory T_{Rea} in the language L_{Rea} that we will consider is also standard. Its axioms are REA1-REA18.

REA1. $(\forall x)(x + 0 = x)$

REA2. $(\forall x)(\forall y)\ (x + y = y + x)$

REA3. $(\forall x)(\forall y)(\forall z)\ ((x + y) + z = (x + (y + z))$

REA4. $(\forall x)(\forall y)(\exists z)(\ x = y + z)$

REA5. $(\forall x)\ \ (x \,.\, 1 = x)$

REA6. $(\forall x)(\forall y)\ (x \,.\, y = y \,.\, x)$

REA7. $(\forall x)(\forall y)(\forall z)\ ((x \,.\, y) \,.\, z = (x \,.\, (y \,.\, z))$

REA8. $(\forall x)(\forall y)((y \neq 0 \rightarrow (\exists z)(\ x = y \,.\, z))$

REA9. $(\forall x)(\forall y)(\forall z)\ ((x + y) \,.\, z = (x \,.\, z) + (y \,.\, z))$

REA10. $(\forall x)(\forall y)(x < y \rightarrow \neg (y < x))$

REA11. $(\forall x)(\forall y)(\forall z)\ (x < y \ \&\ y < z \rightarrow x < z)$

REA12 $(\forall x)(\forall y)(x = y \ \vee\ x < y \ \vee\ y < x)$

REA13. $0 < 1$

REA14 $(\forall x)(\forall y)(\forall z)(y < z \rightarrow x + y < x + z)$

REA15 $(\forall x)(\forall y)(\forall z)(x < 0 \ \&\ y < z \rightarrow x \,.\, z < x \,.\, y)$

REA16. $\forall x)(0 < x \rightarrow (\exists z)(\ x = z \,.\, z))$

REA17. $(\forall a_o) \,..\, (\forall a_{2n+1})(\ a_{2n+1} \neq 0 \rightarrow$
$(\exists x)(a_{2n+1}.x^{2n+1} + a_{2n}.x^{2n} + .. + a_o = 0),$

for all n, where " x^n " is short for x.x. x
(multiplication of x with itself n times)

REA18. $(\forall x_1) \,..\, (\forall x_n)(x_1{}^2 + ... + x_n{}^2 \neq -1)$, for all n

The models of T_{Rea} are known among algebraists as *real-closed fields*.

Note that the last two axioms are, like PA7 in our formalisation of the arithmetic of the natural numbers, schemata: They are not single sentences of the language but infinite collections thereof. Once again this is essential; there are no finite axiomatisations of $\mathbb{R}$ in L_{Rea} that are equivalent to the axiomatisation presented.

As in the case of PA, it is not too difficult to see that all axioms REA1-REA18 are true in the model $\mathbb{R}$ that T_{Rea} is meant to describe. No more than standard high school knowledge is needed to verify all but REA17. REA17 expresses the fact, the proof of which requires a certain amount of algebra, that every polynomial in and 'unknown' x in which the highest occurring power of x is odd takes on the value 0. (This has to do with the fact that such polynomials always become negative for sufficiently large negative values of x and positive for sufficiently high positive values of x, together with the Mean Value Theorem for continuous functions on the reals. We do not go into the details here.)

We noted in Chapter 1 that the set of the real numbers is non-denumerable: There are as many real numbers as there are sets of natural numbers. Since T_{Rea} is a first order theory, it will also have denumerable models. In neither cardinality - that of the reals or that of the denumerable sets - is T_{Rea} categorical. For the case of the reals themselves this can be easily shown by the same trick which we used to show the existence of non-standard models of arithmetic. The standard model $\mathbb{R}$ of T_{Rea} has a property reminiscent of the property of $\mathbb{N}$ we used to show the existence of a non-standard model of Th($\mathbb{N}$) and which is known by the name *archemedean* (after the great Greek mathematician Archimedes.) $\mathbb{R}$ is archemedean in that for every real number r there is a natural number n such that $r < +n$, where '+n' is short for '1 + ... + 1 n times'. i. e. for the term of L_{Rea} in which '1' is followed by n-1 occurrences of '+ 1', and -n is the unique number such that $(-n) + (+n) = 0$. The existence of a non-standard model of T_{Rea}, which is not isomorphic to $\mathbb{R}$, follows from the fact that the following set S of sentences of the language $T_{Rea} \cup \{c\}$ is consistent:
$S = Th(\mathbb{R}) \cup \{+n < c: n \varepsilon N\}$. Clearly no model of S is archemedean. So, since S has models of any infinite cardinality, it will have a non-archemedean model M of the same cardinality as $\mathbb{R}$. M cannot be isomorphic to $\mathbb{R}$, since the denotation in $\mathbb{R}$ of any term +n must be mapped by any isomorphism h onto the denotation of +n in M. But that will mean that no matching element to c_M can be found in $\mathbb{R}$. Dor suppose that $h(r) = c_M$. Since is archemedean, there is an n such that r satisfies the formula "x < +n" in $\mathbb{R}$. But on the other hand the sentence "+n < c" is true in M. so r stands in the relation **<** to the denotation of +n in $\mathbb{R}$ whereas c_M does not stand in the relation $<_M$ to $(+n)_M$ in M. Thus h would not preserve <.

A similar argument is also possible for the denumerable case, provided we can show that T_{Rea} has denumerable models that are archemedean.

This can be done. But it requires some techniques we haven't developed. So we will let this matter rest.

The remainder of this section will be devoted to a proof of the completeness of T_{Rea}. This proof rests in part on deep properties of real-closed fields and otherwise on general results and arguments in general Model Theory. We closely follow the proof presented in (Hodges, 1993), which has the merit of separating the algebraic and model-theoretic components of the argument very clearly.

As in (Hodges,1993) we take the following two facts about real-closed fields for granted. Fact 2 is a 'deep' fact about real closed fields, and an algebraist would properly argue that that is really the crux of the entire argument. We also follow Hodges in using sometimes capital letters A, B, C, .. to denote models. Given the need that will arise more than once to talk about three models at once, this is somewhat more perspicuous than using M, M', M'', .. , as we have done so far.

Fact 1.

Let M be a model of T_{Rea} and let $p(x, y_1,..,y_k)$ be polynomial in x and the parameters $y_1,..,y_k$ - i. e. a term of L_{Rea} which has occurrences of x and $y_1,..,y_k$ (where $k \geq 0$, so the case without parameters in included) and which is of the form "$a_n.x^n + a_{n-1}.x^{n-1} + .. + a_o$", where the a_i are terms of L_{Rea} not containing x - and two elements u_1 and u_2 of M such that $u_1 <_M u_2$ and $M \models p(u_1).p(u_2) < 0$. Then there is an element u in M sich that $u_1 <_M u <_M u_2$ and
$M \models p(u) = 0$.

Fact 2.

Let A be real-closed field, i.e. A is a model for L_{Rea} such that $A \models T_{Rea}$ and C an ordered subfield of A, i.e. a submodel of A which satisfies the axioms of an ordered field, that is REA1-REA15. Then there exists an extension of C to a real closed field A' within A that is 'minimal' in the following sense:

(0) $C \subseteq A' \subseteq A$, $A' \models T_{Rea}$ and if B is a model of T_{Rea} such that $C \subseteq B$, then there is an isomorphic embedding f of A' into B which is the identity on C.

The strategy of the proof is as follows. We prove that T_{Rea} has Quantifier Elimination (QE) and from this that the theory is complete.

Def. 17 A theory T of a language L *has Quantifier Elimination* iff for every formula $A(y_1,..,y_k)$ of L there is a quantifier-free formula $B(y_1,..,y_k)$ such that $T \models (\forall y_1)..(\forall y_k)\,(A \leftrightarrow B)$.

To prove that T has QE it suffices to show

(1) For every quantifier-free formula $B(x,y_1,..,y_k)$of L there is a quantifier-free formula $C(y_1,..,y_k)$ such that

$$T \models (\forall y_1)..(\forall y_k)(\exists x)(B(x,y_1,..,y_k) \leftrightarrow C(y_1,..,y_k)).$$

That (1) entails that T has QE is easily verified. Let $A(y_1,..,y_k)$ be as in Def. 17, and let $(Q_1x_1)..(Q_mx_m)\,D(x_1,..,x_m, y_1,..,y_k)$ be a formula in prenex normal form that is logically equivalent to A, where D is quantifier free and for each i = 1,.., k, Q_i is either $\exists$ or $\forall$. We can of course always arrange for this to be so, by renaming.) Suppose first that Q_k is $\exists$. Then because of (1) there is a quantifier-free formula $D_m(x_1,..,x_{m-1}, y_1,..,y_k)$ so that

(2) $$T \models (\forall y_1)..(\forall y_k)((\exists x_m)D(x_1,.,x_m,y_1,..,y_k) \leftrightarrow D_m(x_1,.,x_{m-1},y_1,.,y_k)).$$

The equivalence (2) entails that in (3), where we have replaced $(\exists x_m)\,D(x_1,..,x_m, y_1,..,y_k)$ by $D_m(x_1,.,x_{m-1},y_1,.,y_k)$ in the normal form for A:

(3) $$T \models (Q_1x_1)..(Q_mx_m)\,D(x_1,..,x_m, y_1,..,y_k) \leftrightarrow (Q_1x_1)..(Q_mx_{m-1})\,D_m(x_1,.,x_{m-1},y_1,.,y_k)$$

In case Q_m is $\forall$, we proceed analogously, but making use of the equivalence between $\forall$ and $\neg\exists\neg$: According to (1) there is a quantifier-free $D'_m(x_1,.,x_{m-1},y_1,.,y_k)$ such that

(4) $$T \models (\forall y_1)..(\forall y_k)((\exists x_m)\neg D(x_1,.,x_m,y_1,..,y_k) \leftrightarrow D'_m(x_1,.,x_{m-1},y_1,.,y_k)).$$

So defining $D_m(x_1,.,x_{m-1},y_1,.,y_k)$ as $\neg D'_m(x_1,.,x_{m-1},y_1,.,y_k))$, we get

(5) $$T \models (\forall y_1)..(\forall y_k)((\forall x_m)D(x_1,.,x_m,y_1,..,y_k) \leftrightarrow \neg D_m(x_1,.,x_{m-1},y_1,.,y_k)).$$

Again (5) enables us to eliminate the innermost quantifier $(Q_m x_m)$ from the normal form. In this way we continue until all quantifiers have been eliminated and we have found a quantifier-free formula $D_1(y_1,..,y_k)$ that is provably equivalent in T to the normal form of A and thus also to A itself. So T has QE.

Before we go on, here is a brief comment on the two Facts we have stated. It is important to realise that these are facts about the theory T_{Rea}: What is claimed here is that the facts hold in any model of T_{Rea}, not just in its standard model, or perhaps one or two other models familiar from Field Theory as a branch of Analysis or Algebra. Thus there is an important difference in particular between Fact 1 and the appeal to the Mean Value Theorem that we made when discussing the truth of the axioms of T_{Rea} in $\mathbb{R}$. The proof we appealed to there can make use of any acknowledged form of argumentation that mathematicians as legitimate for proving results about the reals. Ín contrast, the claim made by Fact 1 is that there exists a formal derivation of the fact claimed from T_{Rea} - i.e. an axiomatic derivation in the sense of Ch. 1 or the construction of closed semantic tableau in the sense of the Appendix to Ch.1. So establishing these facts requires careful checking that all steps can be justified by the axioms REA1-REA18.

To prove that T_{Rea} has QE we need two intermediate steps. First we derive the following two properties of T_{Rea}:

(6) Let A, B be models of T_{Rea} and that $A \subseteq B$. Suppose that $D(x, y_1,..,y_k)$ is a quantifier-free formula of and that $a_1,..,a_k$ are elements of A. Then if $B \models (\exists x)D(x, y_1,..,y_k)[a_1,..,a_k]$, also $A \models (\exists x)D(x, y_1,..,y_k)[a_1,..,a_k]$.

(7) Suppose that $A \models T_{Rea}$ and C a submodel of A. Then the condition of Fact 2 is fulfilled: There exists an A' such that

(0) $C \subseteq A' \subseteq A$, $A' \models T_{Rea}$ and if B is a model of T_{Rea} such that $C \subseteq B$, then there is an isomorphic embedding of A' into B which is the identity on C.

Proof of (6) Suppose that A,B, D are described in (6) and that $B \models (\exists x)D(x, y_1,..,y_k)[a_1,..,a_k]$. We make use of the fact that because we are dealing with the language and theory of rela-closed fields,

$(\exists x)D(x, y_1,..,y_k)$ can be written in a special form. First, we note that, quite generally, $D(x, y_1,..,y_k)$ can be written in disjunctive normal form and the existential quantifier then distributed over the disjuncts. So it suffices to show that if B verifies on of the disjuncts, then A does too. Each disjunct will be of the form

(8) $(\exists x)(\alpha_1 \& .. \& \alpha_r)$

where the α_j are literals of L_{Rea} - atomic formulas or negations of atomic formulas. Note that atmoic formulas of are either equations $\sigma = \tau$ or inequalities $\sigma < \tau$, where σ and τ are terms of L_{Rea}.

Our next observation is that T_{Rea} allows us to replace the negations of atomic formulas by disjunctions of atomic formulas: $\neg(\sigma = \tau)$ is provably equivalent to $\sigma < \tau \vee \tau < \sigma$ and $\neg(\sigma < \tau)$ to $\sigma = \tau \vee \tau < \sigma$. When we substitute these disjunctions for the negative literals in (8), we get a conjunction of disjunctions following the quantifier $(\exists x)$. This conjunction can be transformed once more into a disjunctive normal form and the quantifier distributed once more over the disjuncts so that we end up with a disjunction of formulas of the form (8) where now the α_j are all positive literals.

It now helps to think of the terms σ and τ that occur in these atomic formulas as polynomials in x and to think of the elements $a_1,..,a_k$ from A as 'parameters' of these polynomials. (If we want to be very formal, we can extend the language L_{Rea} to a language $L' = L_{Rea} \cup \{\underline{a}_1,..,\underline{a}_k\}$, where $\underline{a}_1,..,\underline{a}_k$ are new individual constants and expand A and B to models of L' by adding the specification that $\underline{a}_i$ denotes a_i.) This means that the conjuncts of (8) are either of the form $p(x) = q(x)$ or of the form $p(x) < q(x)$, where p and q are polynomials in x with coefficients built from the constants 0,1, $\underline{a}_1,..,\underline{a}_k$. As a next step we can, by familiar algebraic manipulations of which it can easily be seen that they can be justified in T_{Rea}, Transform atomic formulas of the form $p(x) = q(x)$ into atomic formulas of the form $r(x) = 0$ (essentially by 'subtracting q from p or vice versa, though the matter is a little more involved, since we haven't introduced - as a separate operation into our language, and similarly reduce formulas of the form $p(x) < q(x)$ to formulas of the form $r(x) > 0$. This turns (8) into a formula of the form:

(9) $(\exists x)(p_1(x) = 0 \& .. \& p_m(x) = 0 \& p_{m+1}(x) > 0 \& .. \& p_r(x) > 0)$

We now distinguish two cases. (i) First suppose that there is at least one non-trivial equation among the conjuncts. Here 'non-trivial' means that not every element of A is a solution ot the equation, i.e. every possible assignment to x verifies the equation in A. It is a well-known fact of real-closed fields that if a polynomial equation is non-trivial inm this sense, then it has only finitely many solutions; moreover, any solution that exists in a real-closed field that extends A already exists in A. (This is really what 'real-closed' means, and it follows directly from the definitions of the notion that are found in mathematics. if real-closed fields are defined as the models of T_{Rea}, then more work is necessary here. of course that work needs to be done one way or another, for as we remarked above, models of is what we are concerned with, whatever we choose to call them. It too belongs to the results that we are taking for granted here, but that an exhaustive proof would have to supply. (As should be intuitively clear, the crucial part in demonstrating this fact is played by the solution axioms REA17.)

Suppose then that the equation $p_i(x) = 0$ ($1 \leq i \leq m$) is non-trivial. Since by assumption $B \models$ (9) there is an element b in B such that

(10) $B \models (p_1(x) = 0 \ \& \ .. \ \& p_m(x) = 0 \ \& \ p_{m+1}(x) > 0 \ \& \ .. \ \& p_r(x) > 0)[b]$

Since b is a solution of $p_i(x) = 0$ in B b must by the remark above belong to A. Furthermore, all the other equations and inequalities of (9) are also satisfied by b in B and thus, since they are all quantifier-free, will be equally satisfied by b in A. So we have

(11) $A \models (p_1(x) = 0 \ \& \ .. \ \& p_m(x) = 0 \ \& \ p_{m+1}(x) > 0 \ \& \ .. \ \& p_r(x) > 0)[b]$

From this we can conclude

(12) $A \models (\exists x)(p_1(x) = 0 \ \& \ .. \ \& p_m(x) = 0 \ \& \ p_{m+1}(x) > 0 \ \& \ .. \ \& p_r(x) > 0)$,

which concluds the first case.

The second case is that where there are no non-trivial equations in (9). In this case, (9) reduces to

(13) $\exists x)(p_{m+1}(x) > 0 \ \& \ .. \ \& p_r(x) > 0)$

We now make use of Fact 1. Let b_1, ... b_s be all the solutions of the equations $p_{m+1}(x) = 0$, .., $p_r(x) = 0$ in B, given in order of magnitude in B. (I.e. er have $b_1 <_B b_2$,. .) For the same reason that was mentioned in

the argument for case (i) all of b_1, ... b_s belong to A. Moreover, since these are all the solutions to these equations, there will bee no other switches from positive(negative to negative/positive values of any of the polynomials $p_{m+1}(x)$,.., $p_r(x)$. Since by assumption B verifies (13), there is a b in B such that

(14) $B \models (p_{m+1}(x) > 0 \ \& \ .. \ \& p_r(x) > 0)[b]$

this element b will be situated somewhere with regard to the sequence of elements b_1, ... b_s, e.g. between b_j and b_{j+1}; that is, $b_j <_B b <_B b_{j+1}$. This means that in B b verifies all the inequalities occurring as conjuncts in (14). Since b_1, ... b_s are all the places where any of the polynomials $p_{m+1}(x)$,.., $p_r(x)$ changes sign, the formula (14) will be satisfied by any element in the interval (b_j,b_{j+1}), whether in A or in B. There must be elements in (b_j,b_{j+1}) in A, since the order relation in any real-closed filed is dense. (This is yet another thing that must be derived from T_{Rea}, but this is quite straightforward.) Any such element a will satisfy the formula in (9) in A. So we get:

(15) $A \models (\exists x)(p_{m+1}(x) > 0 \ \& \ .. \ \& p_r(x) > 0)$

and so, in the light of the assumptions of case (ii), we have once more (12) and we are done.

We now proceed to the proof of (7)

Let A and C be as stated in (7). We must show that there exists A' such that

(0) $C \subseteq A' \subseteq A$, $A' \models T_{Rea}$ and if B is a model of T_{Rea} such that $C \subseteq B$, then there is an isomorphic embedding of A' into B which is the identity on C.

This is almost what Fact 2 tells us. The only difference is that our assumption is that in the assumption of Fact 2 C is an ordered subfield of A, whereas what we are given in (7) is only that C is a submodel of A. To bridge this gap we argue aas follows. Assume that A,B are models of T_{Rea} and that C is a submodel of both A and B. Here we must appeal to another general fact of real-closed fields: There is unique way of extending C to an ordered subfield C' of A. (C' can be obtained as the quotient field of C, a familiar construction which among other things leads to the arithmetical structure of the rationals starting from the integers.) This minimal field extension of C can be embedded also into

B, and in fact we may as well assume that C' is within the intersection of A and B. replacing elements in B by their originals from C' under the given embedding. This gives as the situation described in the assumptions of Fact 2. So there is a real-closed field A' such that $C \subseteq C' \subseteq A' \subseteq A$, such that A' can be embedded into B by a map which preserves C' and therefore also C.

Our next step is to prove from (6) and (7) the following condition (16):

(16) If A and B are models of T_{Rea}, and $<a_1, \ldots a_k>$, $<b_1, \ldots b_k>$ are k-tuples from A and B respectively such that

(i) $(A,a_1, \ldots a_k) \equiv_o (B,b_1, \ldots b_k)$,

then

(ii) $(A,a_1, \ldots a_k) \Rightarrow_1 (B,b_1, \ldots b_k)$

First we must explain the notation. For any models A, B for some language L and tuples $<a_1, \ldots a_k>$, $<b_1, \ldots b_k>$ from these models $(A,a_1, \ldots a_k) \equiv_o (B, b_1,\ldots b_k)$ means that the tuples satisfy the same quantifier-free formulas in A and B, respectively; and $(A,a_1, \ldots a_k) \Rightarrow_1 (B, b_1,\ldots b_k)$ means that every purely existential formula $(\exists x_1)..(\exists x_m)D(x_1,.,x_m,y_1,..,y_k)$ that is verified by $a_1, \ldots a_k$ in A is verified by $b_1, \ldots b_k$ in B.

Proof of (16) from (6) and (7).

Assume that $A, B \models T_{Rea}$, $(A,<a_1, .., a_k>) \equiv_o (B,<b_1, .., b_k>)$, and that $D(x_1,.,x_m,y_1,.,y_k))$ is a quantifier-free formula of L_{Rea} such that

(17) $A \models (\exists x_1)..(\exists x_m)D(x_1,.,x_m,y_1,.,y_k))[a_1, .., a_k]$.

We have to show that

(18) $B \models (\exists x_1)..(\exists x_m)D(x_1,.,x_m,y_1,.,y_k))[b_1, .., b_k]$.

Because of (17) there are elements $c_1, \ldots c_m$ of A such that

(19) $A \models D(x_1,.,x_m,y_1,.,y_k))[c_1, \ldots c_m,a_1, .., a_k]$.

We first show that there exists an elementary extension B_1 of B and an element d_1 in B_1 such that

(20) $(A,c_1,a_1, .., a_k) \equiv_o (B_1,d_1,b_1, .., b_k)$

Let $\Psi(x_1,y_1,.,y_k)$ be the set of all quantifier-free formulas satisfied by $<c_1,a_1, .., a_k>$ in A:

(21) $\Psi(x_1,y_1,.,y_k) = \{\psi(x_1,y_1,.,y_k): A \models \psi(x_1,y_1,.,y_k)[c_1,a_1, .., a_k]\}$

We infer that

(22) For each $\psi \varepsilon \Psi$, $A \models (\exists x_1)(x_1,y_1,.,y_k)[a_1, .., a_k]$.

Consider the subset $\{a_1, .., a_k\}$ of U_A. Since L_{Rea} contains function constants the restriction of A to $\{a_1, .., a_k\}$ will not be a submodel of A. But we can close $\{a_1, .., a_k\}$ under the operations of A and obtain in this way a (uniquely determined) extension A_o of this restriction which is a submodel of A. Since by assumption $(A,a_1, .., a_k) \equiv_o (B_1,b_1, .., b_k)$, the b's satisfy the same relations in B as the a's in A. This remains the case when we close $\{b_1, .., b_k\}$ to a submodel B_o of B. That is, we can extend the map $(a_1, .., a_k) \mapsto (b_1, .., b_k)$to an isomorphism f from A_o to B_o. We can rearrange things so that f becomes the identity function by taking an isomorphic copy B' of B which contains A_o as a submodel in lieu of $f(A_o)$. In other words we may assume that A_o is both a submodel of A and of B'.

We are now in a position to apply (7): There is a model A' of T_{Rea} such that $A_o \subseteq A' \subseteq A$ and such that A' has an embedding h in B' which is the identity on A_o. Since $A' \subseteq A$ and A' and A are both models of T_{Rea}, we can apply (6) and infer from (22) that,

(23) for each $\psi \varepsilon \Psi$, $A' \models (\exists x_1)(x_1,y_1,.,y_k)[a_1, .., a_k]$.

Since h is an embedding of A' in B' which preserves $a_1, .., a_k$ we can conclude that for each $\psi \varepsilon \Psi$, $B' \models (\exists x_1)(x_1,y_1,.,y_k)[a_1, .., a_k]$, and since B is an isomorphic copy of B' under an isomorphism which maps $a_1, .., a_k$ onto $b_1, .., b_k$, we get (24).

(24) for each $\psi \varepsilon \Psi$, $B \models (\exists x_1)(x_1,y_1,.,y_k)[b_1, .., b_k]$.

This entails that there is an elementary extension B_1 of B and an element d_1 in B_1 such that

(25) $B_1 \models \Psi(x_1,y_1,..,y_k)[d_1,b_1, .., b_k]$

(The argument is the same as in the proof of Thm. 8 of Ch. 1: We extend the language with names for all elements of B and form the theory Th'(B) of B in this language. Let $\Psi(\underline{d}_1,\underline{b}_1,..,\underline{b}_k)$ be the set of all sentences $\psi(\underline{d}_1,\underline{b}_1,..,\underline{b}_k)$, where $\psi \in \Psi$, $\underline{b}_1,..,\underline{b}_k$ are the names in the extended language for $b_1, .., b_k$ and $\underline{d}_1$ is an additional new constant (in yet a further extension of the language). It is then easily shown using (22) that Th'(B) $\cup$ $\Psi(\underline{d}_1,\underline{b}_1,..,\underline{b}_k)$ is consistent. Any model of this set will be an elementary extension of B in which the sentences of $\Psi(\underline{d}_1,\underline{b}_1,..,\underline{b}_k)$ are true.) Let d_1 be the denotation of $\underline{d}_1$ in B_1. Then for all $\psi \in \Psi$ $B_1 \models \psi(x_1,y_1,..,y_k)[d_1,b_1, .., b_k]$. So we have (23).)

Since contains all quantifier-free formulas ψ such that $A \models \psi(x_1,y_1,..,y_k)[c_1,a_1, .., a_k]$ we have (20).

We can now reiterate the argument above for A and B_1. in this way we obtain an elementary extension B_2 of B_1 and an element d_2 in B_2 such that $(A,c_1,c_2,a_1, .., a_k) \equiv_o (B_2,d_1,d_2,b_1, .., b_k)$; and, continuing, we eventually get an elementary extension B_m of B and $d_1, .., d_m$ in B_m such that

(26) $(A,c_1,..,c_m,a_1, .., a_k) \equiv_o (B_m,d_1,..,d_m,b_1, .., b_k)$

From (26) and (19) we infer that

(27) $B_m \models D(x_1,..,x_m,y_1,..,y_k)[d_1,..,d_m,b_1, .., b_k]$

So

(28) $B_m \models (\exists x_1)..(\exists x_m)D(x_1,..,x_m,y_1,..,y_k)[b_1, .., b_k]$

Since B_m is an elemrntary extension of B and $b_1, .., b_k$ belong to B we reach the desired conclusion:

(29) $B \models (\exists x_1)..(\exists x_m)D(x_1,..,x_m,y_1,..,y_k)[b_1, .., b_k]$.

q.e.d.

We now come to the last step in our proof that T_{Rea} has QE. We have seen that it suffices to show that T_{Rea} has the property (1).

Let $D(x,y_1,..,y_k)$ be a quantifier-free formula of L_{Rea}. We must show that there is a quantifier-free formula $E(y_1,..,y_k)$ such that

(30) $T_{Rea} \models (\forall y_1)..(\forall y_k)((\exists x)D(x,y_1,..,y_k) \leftrightarrow E(y_1,..,y_k))$

Let $\Psi(y_1,..,y_k)$ be the set of all quantifier-free formulas $\psi(y_1,..,y_k)$ of L_{Rea} such that $T_{Rea} \models (\exists x)D(x,y_1,..,y_k) \rightarrow \psi(y_1,..,y_k))$. We show that the set $T_{Rea} \cup \Psi \cup \{\neg (\exists x)D(x,y_1,..,y_k)\}$ is inconsistent. Suppose the set was consistent. Then there would be a model B of T_{Rea} and elements $b_1,..,b_k$ of B such that

(31) $B \models \Psi(y_1,..,y_k)[b_1,..,b_k]$ and $B \models \neg (\exists x)D(x,y_1,..,y_k)[b_1,..,b_k]$.

Let $Dia(B,b_1,..,b_k)$ be the set of all quantifier-free formulas $\chi(y_1,..,y_k)$ such that $B \models \chi(y_1,..,y_k)[b_1,..,b_k]$. Then $T_{Rea} \cup Dia(B,b_1,..,b_k) \cup \{ (\exists x)D(x,y_1,..,y_k)\}$ is inconsistent. For if the set were consistent, then there would be a model A of T_{Rea} with elements $a_1,..,a_k$ such that

(32) $A \models (\exists x)D(x,y_1,..,y_k)[a_1,..,a_k]$

(33) $A \models Dia(B,b_1,..,b_k)[a_1,..,a_k]$

But (33) entails that $(A,a_1,..,a_k) \equiv_o (B,b_1,..,b_k)$. So by (16) and (32), it follows that $B \models (\exists x)D(x,y_1,..,y_k)[b_1,..,b_k]$, which contradicts the assumptions about B.

The inconsistency of $T_{Rea} \cup Dia(B,b_1,..,b_k) \cup \{ (\exists x)D(x,y_1,..,y_k)\}$ entails that there is a finite conjunction $\chi(y_1,..,y_k)$ $(= \chi_1(y_1,..,y_k)$ & .. & $\chi_r(y_1,..,y_k))$ of formulas $\chi_i(y_1,..,y_k)$ from $Dia(B,b_1,..,b_k)$ such that

(34) $T_{Rea} \models (\exists x)D(x,y_1,..,y_k) \rightarrow \neg \chi(y_1,..,y_k)$

This means that $\neg \chi(y_1,..,y_k)$ belongs to the set $\Psi(y_1,..,y_k)$. But according to (33) $B \models \Psi(y_1,..,y_k)[b_1,..,b_k]$, so $B \models \neg \chi (y_1,..,y_k)[b_1,..,b_k]$. But this is impossible since on the other hand χ is a conjunction of members of $Dia(B,b_1,..,b_k)$.

This concludes the argument that $T_{Rea} \cup \Psi \cup \{\neg (\exists x)D(x,y_1,..,y_k)\}$ is inconsistent. From the inconsistency of $T_{Rea} \cup \Psi \cup \{\neg (\exists x)D(x,y_1,..,y_k)\}$

we infer that there is a finite conjunction $\psi(y_1,..,y_k)$ (= $\psi_1(y_1,..,y_k)$ & .. & $\psi_s(y_1,..,y_k)$) of formulas $\psi_i(y_1,..,y_k)$ from Ψ such that

(35) $T_{Rea} \models \psi(y_1,..,y_k) \rightarrow (\exists x)D(x,y_1,..,y_k)$

Since $\psi(y_1,..,y_k) \varepsilon \Psi$, we also have $T_{Rea} \models (\exists x)D(x,y_1,..,y_k) \rightarrow \psi(y_1,..,y_k)$.

So we get

(36) $T_{Rea} \models (\forall y_1)..(\forall y_k)(\psi(y_1,..,y_k) \leftrightarrow (\exists x)D(x,y_1,..,y_k))$

which concludes the proof of (1) and thus of the fact that T_{Rea} has QE.

q.e.d.

Our only remaining task is to derive the completeness of T_{Rea} from the fact that it has QE. This is easy. Let a be any sentence of L_{Rea}. Then there is a quantifier-free sentence B of L_{Rea} such that $T_{Rea} \models A \leftrightarrow B$. We already saw in the proof of (6) that any atomic formula of L_{Rea} can be transformed into a formulas of a very special form that is provably equivalent to it in T_{Rea}: every such formula is equivalent to a disjunction $V_j \gamma_j$ of conjunctions γ_j of atomic formulas. In the present case, moreover we are dealing with sentences. That is, our quantifier-free sentence B can be rewritten as an equivalent disjunction $V_j \gamma_j$ in which each atomic conjunct of each γ_j is a sentence that is either of the form $\sigma = \tau$ or of the form $\sigma < \tau$. The terms σ and τ occurring in these atomic sentences are all built up form the individual constants 0 and 1 with hte help of the functions constants + and . . It is not hard to verify that each such term τ can be transformed into a 'canonical' term τ' which is either 0 or 1 or a sum of the form 1 + .. + 1 involving two or more 1's. ('Transformed' in the sense that the equation "$\tau = \tau'$ " can be proved from T_{Rea}.) it is also straightforward to verify that T_{Rea} enables us to either prove or disprove any equation $\sigma' = \tau'$ and inequality $\sigma' < \tau'$, when σ' and τ' are both canonical.

This means that T_{Rea} will either prove or refute B. T_{Rea} will prove B iff there is at least one disjunct γ_j of its rewritten form $V_j \gamma_j$ such that T_{Rea} proves every conjunct of γ_j. Otherwise T_{Rea} refutes $V_j \gamma_j$, and with it B. The same is true for the sentence A we started with. So T_{Rea} either proves or refutes every sentence from L_{Rea}.

q.e.d.

In the introduction to Section 2.6 we remarked on the intuitively paradoxical result that the arithmetic of the reals admit formalisation as a complete and decidable explicitly axiomatised theory, whereas the arithmetic of the natural numbers does not. Now that we have shown the first of these two facts at the hand of the the theory T_{Rea} is, the paradox may seem even more striking. it is true that the axioms of T_{Rea} is that capture the behaviour of the operations + and . are quite different from those of PA. The latter cannot be used here, since - obviously- there is no way of reducing what happens when these operations are applied to numbers other than the natural numbers recursively to what happens when one of the arguments is 0. But on the other hand, it is not hard to see that the axioms of T_{Rea} force the behaviour of + and . on the natural numbers to be the way that PA describes them. Specifically, let M be any model of T_{Rea} and let N_M be the set of all elements of U_M that are the denotations of some closed canonical term τ' of L_{Rea}. Then the submodel $\mathbb{N}_M$ of M with universe N_M will be isomorphic to the standard model $\mathbb{N}$ of PA. This might suggest that it should be possible to translate every sentence A from the language of PA into a sentence A' of L_{Rea} which talks only about the submodels $\mathbb{N}_M$ of models M of T_{Rea}. However, that would give us a method to check for any A whether or not it is true in $\mathbb{N}$ and that is precisely what Gödel proved to be impossible.

What then is wrong with the suggestion? The answer is - and must be - that we cannot translate sentences from Peano Arithmetic intoi sentences of L_{Rea} that 'speak only about the submodels $\mathbb{N}_M$. And that in turn implies that there can be no formula N(x) of L_{Rea} that defines the set of natural numbers in T_{Rea}, in the sense that

(37) For all models M of T_{Rea}, $N_M = \{d \varepsilon U_M: M \models N(x)[d]\}$.

Exercise: Show that if there were a formula N(x) satisfying (39), then it would be possible to define a translation function tr from L_{PA} to L_{Rea} such that for every sentence A of L_{Rea} $\mathbb{N} \models A$ iff $T_{Rea} \models tr(A)$.

That the set of natural numbers cannot be defined in T_{Rea} in spite of the fact that in every model M of the theory it consists (modulo isomorphism) of precisely the denotations in M of canonical closed terms of L_{Rea}, is itself a surprising result, which has to do with deep properties of real-closed fields. (It is a result that is entailed by Gödel's Incompleteness Theorems and the completeness of but it does not in

any direct and obvious way entail one of those two results given the other.)

That the undefinability of the natural numbers within T_{Rea} is connected with special properties of real-closed fields is indicated by the fact that arithmetic on the rational numbers is crucially different in this respect. It is possible to give a (necessarily incomplete) axiomatisation $T_{\mathbb{Q}}$ of the arithmetic of the rational numbers - for instance in the language L_{Rea} - and to define a formula N(x) such that (39) holds for models of $T_{\mathbb{Q}}$:

(38) For all models M of $T_{\mathbb{Q}}$, $N_M = \{d \varepsilon U_M: M \models N(x)[d]\}$.

(This quite difficult result is due to (Robinson, 1949).)

(38) entails that $T_{\mathbb{Q}}$ must be undecidable and incomplete, just as PA and all its axiomatisable extensions.

2.2.6. Rooted Feature Structures.

Let A be a set. An n-*ary feature structure relative to* A is an algebra S = <U, $\mathbf{f}_1$,.., $\mathbf{f}_n$>, consisting of a *universe* U and n partial unary functions $\mathbf{f}_1$,.., $\mathbf{f}_n$ over U such that

(i) no feature $\mathbf{f}_i$ is defined on any element of U that belongs to A

The elements of U ∩ A are called *the atoms of* S. We refer to the members of U \ A as the *variables* of S. S is said to be *finite* whenever U is finite. Sometimes we will refer to the elements of U also as *nodes*.

We will be especially interested in *rooted* n-*ary feature structures*. Suppose that <U, $\mathbf{f}_1$,.., $\mathbf{f}_n$> is an n-ary feature structure relative to A, u ε U and u has the property:

(*) for each v ε U, v ≠ u, there is a composition $\mathbf{f}^1{}_o$ ....o$\mathbf{f}^j$ of features such that u = $\mathbf{f}^1{}_o$ ....o$\mathbf{f}^j$(u_o) and for each r = 1,...,k there is an i ≤ n such that $\mathbf{f}^r = \mathbf{f}_i$ (In other words, each element v of U can be reached from u via a "feature path").

Then u is called a*root of* <U, $\mathbf{f}_1$,.., $\mathbf{f}_n$>. By a *rooted n-ary feature structure relative to* A we understand an n+2-tuple <U, u, $\mathbf{f}_1$,.., $\mathbf{f}_n$> such that <U, $\mathbf{f}_1$,.., $\mathbf{f}_n$> is an n-ary feature structure relative to A and u is a root of <U, $\mathbf{f}_1$,.., $\mathbf{f}_n$>.

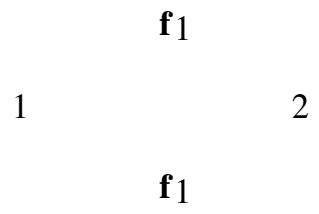
The relation "v can be reached from u' via some feature path" where u, v are elements of the universe U of a feature structure, is clearly a transitive relation. We denote this relation as $\leqslant_S$. S is called *well-founded* if $\leqslant_S$ is irreflexive (or "has no loops", as it is also put). Well-founded feature structures are also called *unfolded*. A well-founded feature structure S is called a *feature tree* if for no u, u' ε U there are distinct paths $\mathbf{f}^1{}_o$ ....o$\mathbf{f}^j$ and $\mathbf{g}^1{}_o$ ....o$\mathbf{g}^k$ such that u = $\mathbf{f}^1{}_o$ ....o$\mathbf{f}^j$(u') = $\mathbf{g}^1{}_o$ ....o$\mathbf{g}^k$(u').

For any rooted feature structure S = <U, u, $\mathbf{f}_1$,.., $\mathbf{f}_n$> and any v ε U, let S ⌈ v (*the restriction of* S *to* v) be the rooted structure <U', v, $\mathbf{f'}_1$,.., $\mathbf{f'}_n$> where U' consists of all w ε U such that there is a path from v to w and for i = 1, ..., n $\mathbf{f'}_i$ is the restriction of $\mathbf{f}_i$ to U' - i.e. for any v ε U' $\mathbf{f'}_i$(v) = $\mathbf{f}_i$(v), provided $\mathbf{f}_i$(v) is defined, and $\mathbf{f'}_i$ is undefined otherwise. (It is

easily verified that <U', $\mathbf{f}'_1$,.., $\mathbf{f}'_n$> is an n-ary feature structure relative to A and that v is a root of the structure <U', $\mathbf{f}'_1$,.., $\mathbf{f}'_n$>.)

Notation: It is common in the feature structure literature to write "u'$\mathbf{f}^1$....$\mathbf{f}^j$ in stead of $\mathbf{f}^1{}_o$ ....$_o\mathbf{f}^j$(u).

Often the root u_o of a feature structure S is the unique element of S which satisfies condition (iv). But this need not be so. It is not so, for instance, for the 1-ary structure S_o = < {1,2}, 1, $\mathbf{f}_1$>, where $\mathbf{f}_1$ is the function {<1,2>, <2,1>}. S_o can be graphically represented as follows:



Here not only 1 but also 2 satisfies condition (*) of definition of rooted feature structures; so < {1,2}, 2, $\mathbf{f}_1$> is a rooted feature structure as well.

Note however that if <U, u, $\mathbf{f}_1$,.., $\mathbf{f}_n$> is well-founded, then u will always be the unique element satisfying (*). (Show this!). So every well-founded rooted feature structure has a unique root.

The first language we choose to describe n-ary feature structures is L_n = {F_1, .. , F_n, At}, where the F_i (i = 1, ... ,n) are "partial one-place functors" and At (for "Atom"!) is a one-place predicate. (Partial one-place functors are really two-place predicate constants that are consistently interpreted as partial one-place functions; given this interpretative convention, it is possible to adopt a functor-like notation for them; see below.) An n-ary feature structure S = <U, $\mathbf{f}_1$,.., $\mathbf{f}_n$> relative to A can be regarded as a model <U,F> for L_n, where F(At) = U $\cap$ A and $F(F_i) = \mathbf{f}_i$.

Exercise: Formulate sentences of L_n which describe the following feature structures up to isomorphism. (Letters in the first half of the alphabet denote atoms - i.e. elements of A - letters in the second half denote variables.)

Often feature structures are described with the help of languages $L'_{n,B}$ that are minor variants of the languages L_n. The languages $L'_{n,B}$ differ

from their counterparts L_n in that they have, in lieu of the 1-place predicate At, a set B of individual constants. We will assume that these sets B are subsets of some given set of "canonical names" of the members of A. That is, each constant in B is taken to denote that atom in A of which it is the canonical name. It will be harmless, and simplify matters, to assume that the elements of A act as their own canonical names, so that B is simply a subset of A.

Exercise. For each of the feature structures of the last exercise, give a uniquely identifying description of it by a sentence belonging to some appropriate language $L'_{n,B}$.

As is always the case for finite structures, every finite n-ary feature structure can be uniquely described in L_n up to isomorphism. The same is true for models for L_n which consist of finite sets of disjoint finite feature structures. Unique characterization up to isomorphism is not possible, however, for "universal models" of finite feature structures, models for L_n in which all and only the finite feature structures are represented. Let us have a closer look at such models.

To define such a universal model we have to confine the universes of the feature structures it contains to some given set V. We assume that V is denumerably infinite. An easy set-theoretical argument shows that the set $\mathbb{S}(V)$ of all finite n-ary feature structures whose universes are included in V is also denumerable. To build a model in which all the finite n-ary feature structures are represented we have to proceed carefully. We cannot simply form the union of the structures in $\mathbb{S}(V)$, for then the elements in V would have to do multiple duty and that would lead to conflicts; for instance, an element u would have to act in one complex structure as a node on which the feature $\mathbf{f}_1$, say, is defined and in some other feature structure as a node on which $\mathbf{f}_1$ is not defined. Clearly we cannot have it both ways.

To avoid this difficulty we can proceed in one of two ways. The first way is to make the variable parts of the universes of all the represented finite structures disjoint. Note that $\mathbb{S}(V)$ contains many copies of what is intuitively just one feature structure. We can get rid of such spurious duplication by forming equivalence classes of ismorphic structures. Since the set of equivalence classes is again denumerable, it can be enumerated. Using this enumeration we can then replace each equivalence type in turn by an instance $S_i = \langle U_i, F_i \rangle$ of it such that the "non-atomic part" $U_i \backslash A$ of U_i consists of elements of $V \backslash A$ that do not

occur in any of the instances chosen for the equivalence types which, in the enumeration, occur before it. In this way we obtain representatives of all the equivalence types no two of which share any variables (i.e. elements that do not belong to A). We can now form the model $M_1(V)$ = <U,*F*> as the union of all the S_i: $U = \bigcup_i U_i$ and $F(F_j) = \bigcup_i F_i(F_j)$.

Exercise: Check whether the sentences you formulated in the two preceding exercises are true in $M_1(V)$. If not, then formulate other sentences which also describe the given graphs up to isomorphism and which are true in $M_1(V)$.

Just as one can construct a universal model of all finite n-ary feature structures we can also construct, by the same method, a universal model for all finite n-ary rooted feature structures.

The second way of constructing a universal model works smoothly only for rooted structures <U, u_o, $\mathbf{f}_1$,.., $\mathbf{f}_n$> with distinguished root u_o. This time we let the finite n-ary rooted feature structures themselves be the *elements* of the model. On this universe we must define interpretations of At and of the features F_i. We take as interpretation of At the set of all feature structures that consist of single atoms, i.e. all structures <{a}, *F*> such that a ε A and $F(At) = F(F_1) = F(F_2) = ... = \emptyset$. (Note that there is an obvious bijection between this interpretation of At and A.) Furthermore, we define $F(F_j)$ as follows. We put $F(F_j)$ = the set of all pairs < S, S'> such that S = <U, u_o, $\mathbf{f}_1$,.., $\mathbf{f}_n$>, S' = <U', u'_o, $\mathbf{f}'_1$,.., $\mathbf{f}'_n$>, u'_o ε U and S' = S⌈u'_o. We refer to the model thus constructed as $M_2(V)$.

It is easy to see that neither the model $M_1(V)$ nor the model $M_2(V)$ is identified up to isomorphism by the set of sentences true in it. The reason is a quite general one: If a first order theory has models of arbitrarily large finite size, it also has infinite models. By the same token, the sentences that are true in a model in which there are objects of any finite size (no matter how large), will also have models in which there are besides these finite objects also infinite objects which can be regarded as "limits" of chains of ever larger finite ones.

The proof of this fact rests on the compactness of first order logic. We consider the model $M_1(V)$. We extend L_n to a new language L_n' by adding a new individual constant c and $Th(M_1(V))$ to a new theory Th' by adding an infinite collection of sentences which together express that c is the root of an infinite feature structure. There are many different ways in which we can do this. Perhaps the simplest way is to

state that c is the root of an infinite path consisting exclusively of applications of the feature f1. That is, Th' = $Th(M_1(V)) \cup \{A_n\}_{n \in \omega}$, where A_n is the sentence

$(\exists x)$ (f1 o f1 o ... o f1(c) = x). Clearly Th' is consistent. For let G be a finite subset of Th'. Then G is consistent. For among the sentences A_n that it contains there is one with highest index, say A_{n_o}. This sentence will entail all other sentences A_n in the set. It is clear, however that A_{n_o} is true in the model M for L_n' which we get by adding to $M_1(V)$ an interpretation for c which makes c denote a feature structure that has an f1-path of length at least n_o. In this model all sentences from G which belong to $h(M_1(V))$ will be true as well. So G is satisfiable. By compactness Th' is satisfiable. So it has a model M'. In this model the denotation of c will be the root of an infinite f1-path.

Exercises Ch. 2.

1. a. Let {} be the language {} which has no non-logical constants, and let T be any complete consistent theory of {} .
Show that T is κ-categorical for every cardinality κ (both finite and infinite)

b. Give a complete description of the complete theories of {}. Which of these are finitely axiomatisable?

c. Show that for any first order language L there are complete theories of L that have infinite models and that are κ-categorical for all infintie cardinals κ.

Moreover, how that there are such theories that finitely axiomatisable whenever L is finite.

2. Let L = {P}, where P is a 1-place predicate.

a. Define countably many complete theories of L that only have infinite models and that are categorical for all infinite cardinalities.

b. Specify a complete theory of L that is ω-categorical but not κ-categorical for uncountable cardinals κ.

Hint: One can express, by means of an infinite number of axioms of L, that (i) the extension of P is infinite, and (ii) that the complement of P's extension (the set of individuals that do not satisfy P) is also infinite. It is easy to show (i) that this theory has infinite models; (ii) that any model of it is infinite; (iii) that any two countable models of the theory are isomorphic; and (ii) that for any uncountable cardinality κ there are models of the theory of cardinality κ which are not isomorphic. (N.B. follows from the fact that M is a model of the theory and $|U_M|$ is uncountable, then the extension of P in M could be either countable or uncountable.)

c. Let L' = {R}, where R is a 2-place predicate.

Define countably many complete theories of L' that only have infinite models and that are ω-categorical but not κ-categorical for uncountable cardinals κ.

3. Show that the theory T_{rat} is not categorical for uncountable cardinalities.

Hint: In view of Morley's Theorem it suffices to show this for just one uncounbtable carfinality. Choose the cardinality 2^{ω} of the set $\mathbb{R}$ of real numbers.
Compare the following two models for the language {<}:

$M_1 = <\mathbb{R},<_{\mathbb{R}}>$, where $<_{\mathbb{R}}$ is the standard ordering of $\mathbb{R}$.
$M_2 = <\mathbb{Q} \otimes \mathbb{R},<'>$, where $\mathbb{Q}$ is the set of rational numbers and $<'$ is the "alphabetic ordering of $\mathbb{Q} \otimes \mathbb{R}$ induced by the standard orderings of $\mathbb{Q}$ and $\mathbb{R}$" - that is, for $q,q' \varepsilon \mathbb{Q}$ and $r,r' \varepsilon \mathbb{R}$ $<q,r> <' <q',r'>$ iff (i) $q <_{\mathbb{Q}} q'$ or (ii) $q = q'$ and $r <_{\mathbb{R}} r'$.
It follows from general facts of set theory that M_2 has cardinality 2^{ω} and thus that M_1 and M_1 are of the same uncountable cardinality.
Show that M_1 and M_1 are not isomorphic.

4. Let DS(T,L) be the lattice of all extensions of a given theory T of a some 1-st order language L.

Show: If DS(T,L) is a boolean algebra, then DS(T,L) is finite.

5. (Stone Representation Theorem for Boolean Lattices.)

Let $BL = <U, \leq>$ be any boolean lattice. For each $b \varepsilon U$, let $I_b = \{d \varepsilon U: d \leq b\}$. ($I_b$ is called the *prime ideal determined by* b.)

Show that BL is isomorphic to the structure $<U', \subseteq>$, where $U' = \{I_b: b \varepsilon U\}$ and $\subseteq$ is set-theoretical inclusion.

N.B. The intuitive significance of Stone's Representation Theorem is that all different types of boolean lattices (and thus also all types of all boolean algebras) are realised by set-theoretical structures, whose universe consists of subsets of some given set and in which lattice relation is set-theoretic inclusion.

(This is the purport of all representation theorems in mathematics: Every structure satisfying some general requirements (such as that of being a model of a given set of axioms) is isomorphic to - and thus can be "represented" as - a structure of some special form.)

6. Show that 0 and S are definable within PA in terms of +.

7. Show that S is not definable within PA in terms of . (multiplic.).

(N.B. intuitively this means: the successor operation on the natural numbers is not definable within PA just with the help of multiplication.)

Hint: Let $T_{PA,\{.\}}$ be the set of all sentences from the sublanguage {.} of L_{PA} that are theorems of PA.

i. Show that any denumerable model M of $T_{PA,\{.\}}$ is isomorphic to the model $\mathbb{N}_{\{.\}} = \langle N, .^N \rangle$, where $.^N$ is the multiplication operator from the standard model $\mathbb{N}$ of arithmetic.

To show this, first observe that "the number zero" and "the number one" are definable in PA from . alone (i.e. we can define in terms of . the predicate "is equal to the number zero" and the predicate "is equal to the number one"); and further that with . we can also define the predicate "is a prime". Once this has been established it is easily seen that among the things that $T_{PA,\{.\}}$ asserts is that there are infinitely many primes and that these are all different from both zero and one. This means that denumerable model M of $T_{PA,\{.\}}$ has a unique zero, a unique one and infinitely distinct primes. it is then easy to show that any bijection between the primes of M and the primes of the standard model of arithmetic N is an isomorphism between M and the model $\mathbb{N}_{\{.\}}$.

ii. Show that (i) entails the non-definability of S in $T_{PA,\{.\}}$.

8. Let L be the language (0, S}, with 0 a 0-place function constant and S a 1-place dunction constant. Let T be the L-theory that is axiomatised by the set { A_1,A_2}, where:

A_1 := $(\forall x)(x \neq 0 \leftrightarrow (\exists y)(x = Sy))$

A_2 := $(\forall x)(\forall y)(S(x) = S(y) \rightarrow x = y)$

Let *N* be the L-model <N, I>, where:

(i) N is the set of natural numbers;
(ii) I(0) = 0 (i.e. I(0) is the number zero); and
(iii) for every natural number nm I(S)(n) = n+1.

Evidently N is a model of T.

Show: There exist countably infinite models of T that are not isomorphic to N .

9. Let L be the language {0, 1, S, P}, where 0 and 1 are individual constants and S and P are 2-place predicate constants. LetT be the theory of L that is axiomatised by the following set of axioms:

A1. $(\forall x)(\forall y)\forall z)(S(x,y) \ \& \ S(x, z) \rightarrow y = z)$
A2. $(\forall x)(\forall y)\forall z)(S(y,x) \ \& \ S(z, x) \rightarrow y = z)$
A3. $(\forall x)((\exists y)(S(x,y)) \leftrightarrow x \neq 1)$
A4. $(\forall x)((\exists y)(S(y,x)) \leftrightarrow x \neq 0)$

A5. $(\forall x)(\forall y)\forall z)(P(x,y) \ \& \ P(x, z) \rightarrow y = z)$
A6. $(\forall x)(\forall y)\forall z)(P(y,x) \ \& \ P(z, x) \rightarrow y = z)$
A7. $(\forall x)((\exists y)(P(x,y)) \leftrightarrow x \neq 0)$
A8. $(\forall x)((\exists y)(P(y,x)) \leftrightarrow x \neq 1)$

A9. $(\forall x)(x \neq 1 \rightarrow (\exists y)(S(x,y) \ \& \ P(y,x))) \ \&$
$(\forall x)(x \neq 0 \rightarrow (\exists y)(P(x,y) \ \& \ S(y,x)))$

(Intuitively the conent of T as as follows:

(i) (A1 -A4) say that S sdenotes a partial 1-1 function, such that all elements of the universe U except for 1 belong to its domain and all elements of U except for 0 belong to its range;

(ii) (A5-A8) say that the same applies to P, except that in this case it is 0 that is missing from the domain and 1 that is missing from the range.

(iii) (A9) says that the function from U\{0} onto U\{1} that is denoted by P is the inverse of the function from U\{1} onto U\{0} that is denoted by S.)

1. Show that T has an infirnite model and that all models of T are infinite.

2. The constants 0, 1, P are all definable in T using just the constant S. (That is, for each of these three constants there is an explicit definition in which the only non-logical constant appearing on the right hand side is S.)

   Give explicit definitions for 0, 1 and P in terms of S in T.

10. Let L be the language {=, <, I, 0,S}, in which < is a 2-place predicate, I a 1-place predicate, 0 an individual constant and S a 2-place predicate. Let T be the theory of L that is axiomatised as follows:

A1 $\forall x \forall y\ (x < y \rightarrow \neg\ y < x)$
A2 $\forall x \forall y \forall z\ (x < y\ \&\ y < z \rightarrow x < z)$
A3 $\forall x \forall y\ (x < y\ v\ x = y\ v\ y < x)$
A4 $\forall x \forall y\ (x < y \rightarrow \exists z\ (x < z\ \&\ z < y)$
A5 $I(0)$
A6 $\forall x \forall y\ (S(x,y) \rightarrow I(x)\ \&\ I(y))$
A7 $\forall x \forall y \forall z\ (\ S(x,y)\ \&\ S(x,z) \rightarrow y = z)$
A8 $\forall x \forall y \forall z\ (\ S(x,z)\ \&\ S(y,z) \rightarrow x = y)$
A9 $\forall x \forall y\ (S(x,y) \rightarrow x < y)$

It is easily verified that T holds in the following model M_o:

(i) the universe of M_o is the set Q of rational numbers;
(ii) $<_{M_o}$ is the "less than"-relation between rational numbers;
(iii) I_{M_o} is the set of integers;
(iv) 0_{M_o} is the number zero; and
(v) S_{M_o} is the successor relation between integers.

Show that there is apart from M_o at least one ohter countable infinite model of T which is not isomorphic to M_o .

11. Let L_{PA} (= {0, s, +, .}) be the language of Peano Arithmetic. Let L_1 be the extension $L_{PA} \cup \{c_1,P\}$ of L_{PA} where c_1 is an individual constant and P a 2-place predicatem and let L_2 be the extension $L_1 \cup \{c_2\}$ of L_1 where c_2 is an individual constant . (So L_1 = {0, s, +, ., <, c_1} and L_2 = {{0, s, +, ., <, c_1, c_2}.)
Let T_1 = Cl(PA $\cup$ {$(\forall x)(\forall y)(x < y \leftrightarrow (\exists z)(\ z \neq 0\ \&\ x + z = y)\} \cup$
$\{0 < c_1, S0 < c_1, SS0 < c_1, ...\}$).

Let $T_2 = T_1 \cup \{c_1 < c_2, Sc_1 < c_2, SSc_1 < c_2, ...\}$. and let M_1 be any model of T_1.

Show that M_1 dcan be expanded to a model M_2 of T_2 by adding a suitable interpretation of the constant c_2.

12. Let L be the language {<}, where< is a 2-place predicate constant and öet L' = L ∪ {S}, with S a 1-place function constant. Let T' be the theory $Cn_{L'}(\{A.1, ..., A.4\})$, where:

A.1 $(\forall x)(\forall y)\ (x < y \rightarrow \neg\ (y < x))$

A.2 $(\forall x)(\forall y)(\forall z)\ ((x < y\ \&\ y < z) \rightarrow x < z)$

A.3 $(\forall x)(\forall y)\ (x < y\ \vee\ x = y\ \vee\ y < x)$

A.4 $(\forall x)(x < S(x)\ \&\ (\forall z)(x < z \rightarrow (S(x) < z\ \vee\ S(x) = z)))$

Show that S is definable in T' (i.e. in terms of <).

13. Deduce the fllowing statement from the axioms PA1-PA7:

$(\forall x)(\ (\exists y)(x = 2.y) \leftrightarrow \neg\ (\exists y)(x = 2.y + 1)\ )$

14. a. We extend the language of arithmetic L_{PA} with a new 1-place predicate G to the language $L' = L_{PA} \cup \{G\}$ and extend the theory PA to a theory T' of L' by adding as a new axiom the following definition D of G in terms of +:

(D) $(\forall x)(G(x) \leftrightarrow (\exists y)(\ x = y + y))$

(Intuitively D says that G denotes the property "is an even number".)

Show that the sentence $(\forall x)(G(x)\ \vee\ G(S(x)))$ is derivable from T'.

b. This time we extend L_{PA} with a new 2-place predicate < to the language $L'' = L_{PA} \cup \{<\}$ and extend PA to a theory T'' of L'' by adding as a new axiom the following definition D' of < in terms of + and 0:

(D') $(\forall x)(\forall y)(\ x < y \leftrightarrow (\exists z)(z \neq 0\ \&\ z + x = y))$

Show that the sentence (1) is deducible from T''.

(1) $(\forall x)(\forall y)(x < y \rightarrow x \neq y)$

(Hint: One way to show this is to prove first that (1) is equivalent to (2)

(2) $(\forall x)(\forall v)(x + Sv \neq 0)$

(Intuitively D says that G denotes the property "is an even number".)

Show that the sentence $(\forall x)(G(x) \vee G(S(x)))$ is derivable from T'.

b. This time we extend L_{PA} with a new 2-place predicate < to the language $L'' = L_{PA} \cup \{<\}$ and extend PA to a theory T'' of L'' by adding as a new axiom the following definition D' of < in terms of + and 0:

(D') $(\forall x)(\forall y)(x < y \leftrightarrow (\exists z)(z \neq 0 \ \& \ z + x = y))$

Show that the sentence (1) is deducible from T''.

(1) $(\forall x)(\forall y)(x < y \rightarrow x \neq y)$

(Hint: One way to show this is to prove first that (1) is equivalent to (2)

(2) $(\forall x)(\forall v)(x + Sv \neq 0)$

and then to prove (2) be mathematical induction.)

15. Let L and L' be languages of first order logic and let T and T' be theories of L and L', repsectively. Let I be a function from the sentences of L to the sentences of L'. (We may call such a function I a "translation" from L to L'.) We say that I *interprets* T i*n* T' iff for every sentence A of L such that $T \models A$, $T' \models I(A)$. Second, let $\mathbb{I}$ be a set of translation finctions from l to L'. Then we say that T is *interpretable in* T' *relative to* $\mathbb{I}$ iff there is an I in $\mathbb{I}$ which interprets T in T'. i

Let now T be the theory od stron partial orderings in the language L = {<}, whose axioms are

$$(\forall x)(\forall y)(\ x < y \ \rightarrow \neg\ y < x\ )$$
$$(\forall x)(\forall y)(\forall z)(\ x < y\ \&\ \ y < z\ \rightarrow\ x <\ z)$$

We can interpret T in the theory PA formulated in the language L_{PA} of Section 2.6.1 by means of the function I which is "based on" the following definition of "<" in L_{PA}:

$(D_<)$ $\qquad (\forall x)(\forall y)(x < y \ \leftrightarrow\ (\exists z)(z \neq 0\ \&\ x + z = y))$

Here, when we say that I is "based on" $(D_<)$ what we mean is that for any sentence A of L, I(A) is the sentence which we get by replacing each subformula "u < w" of A by the right hand side of $(D_<)$, replacing x by u and v by w (and if necessary renaming z in order to avoid variable clashes).

Show that I interprets T in PA (and therewith that T is interpretable in PA relative to the set of all translations from L into L_{PA} that are based on possible definitions of "<" in L_{PA}).

16. Let T be a theory of some first order language L and let α be a non-logical constant of L. Let D be the set of all possible explicit definitions of α in terms of the remaining vocabulary of L. (That is, if α is an n-place predicate P, then D will be the set of all sentences of the form $(\forall v_1)..(\forall v_n)(P(v_1,.,v_n) \leftrightarrow A)$, where A is a formula of L \{α} in which only $v_1,.,, v_n$ may have free occurrences; and if α is an n-place function constant f, then D is the set of all formulas $(\forall v_1)..(\forall v_n)(f(v_1,.,v_n) = v_{n+1} \leftrightarrow A)$, where A is a formula of L \{α} in which the only free occurrences are of the variables $v_1,.,, v_{n+1}$.)

Let $\mathbb{I}$ be the set of all translations of L into L \{α} that are based on definitions in D, where for a definjition d ε D with right hand side A_d the translation I_d based on d is the one which replaces in any formula B of L all occurrences of atomic formulas involving α by the corresponding instantiations of A_d. (See also the previous exercise.)

Let T' be the theory T ∩ {C: C is a sentence of L \{α}}

Show: T is interpretable in T relative to $\mathbb{I}$ iff α is definable in T.

Lösungen von einigen Aufgaben.

9. Wir bezeichnen das zu beweisende Theorem
$(\forall x)((\exists y)(x = 2.y) \leftrightarrow \neg (\exists y)(x = 2.y + 1))$ als (*).

Wir verfahren nach Induktion und zeigen (*) indem wir zeigen:

(**) $(\exists y)(0 = 2.y) \leftrightarrow \neg (\exists y)(0 = 2.y + 1)$

(***) $((\exists y)(x = 2.y) \leftrightarrow \neg (\exists y)(x = 2.y + 1)) \rightarrow$
$((\exists y)(Sx = 2.y) \leftrightarrow \neg (\exists y)(Sx = 2.y + 1))$

(**): Einerseits haben wir PA $\vdash 0 = 2.0$. Also auch PA $\vdash (\exists y)(0 = 2.y)$. Andererseits gilt: PA $\vdash \neg (\exists y)(0 = 2.y + 1)$. Denn nehmen wir an, dass $(\exists y)(0 = 2.y + 1)$, dann gibt es ein y, so daß $0 = S(2.y)$, was dem PA-Axiom widerspricht, dass 0 nicht von der Form Sx ist.

(***): Nehmen wir an: $((\exists y)(x = 2.y) \leftrightarrow \neg (\exists y)(x = 2.y + 1))$.
Dann gilt also entweder

(i) $(\exists y)(x = 2.y) \ \& \ \neg (\exists y)(x = 2.y + 1)$ oder
(ii) $\neg (\exists y)(x = 2.y) \ \& \ (\exists y)(x = 2.y + 1)$

Im ersten Fall gibt es ein n, so daß $x = 2.n$. Also gilt $Sx = S(2.n) = 2.n + 1$ und deshalb auch $(\exists y)(Sx = 2.y + 1)$. Wäre es der Fall, daß $(\exists y)(Sx = 2.y)$, so gäbe es ein n, so daß $Sx = 2.n$. Offenbar kann n nicht gleich 0 sein. Also ist $n = Sm$ für irgendein m. Dann aber $Sx = 2. Sm = Sm.2 = Sm + Sm = S(Sm + m)$. Also ist $x = Sm + m = m + m + 1 = m.2 + 1 = 2.m + 1$. Also $(\exists y)(x = 2.y + 1)$, was dem zweiten Konjunkt in (i) widerspricht. Also führt die Annahme, daß
$(\exists y)(Sx = 2.y)$ zu einem Widerspruch. Somit haben wir
$(\exists y)(Sx = 2.y + 1) \ \& \ (\exists y)(Sx = 2.y)$ und damit
$(\exists y)(Sx = 2.y + 1) \leftrightarrow \neg (\exists y)(Sx = 2.y)$.

Der zweite Fall, (ii), erledigt sich ähnlich.

ii. Zu zeigen:

$$(\forall x)(\forall y)(Sx \cdot Sx = (x \cdot x) + y \rightarrow \neg (\exists u)(y = 2.u))$$

(Intuitiv besagt diese Formel, daß die Differenz zwischen zwei aufeinanderfolgenden Quadraten immer eine ungrade zahl ist.) Wir agumentieren wie folgt:

$Sx \,.\, Sx = (Sx \,.\, x) + Sx = (x \,.\, Sx) + Sx = ((x \,.\, .x\,) + x) + x + 1$
$= (x \,.\, .x\,) + (x + x + 1) = (x \,.\, .x\,) + (2..x + 1)$. Also, wenn
$Sx \,.\, Sx = (x \,.\, x) + y$, dann ist $y = 2..x + 1$. (Siehe unten!).
Wenn aber $y = 2..x + 1$, dann gilt auch $(\exists\, u)(y = 2.u + 1)$.
Dann gilt aber nach (i), daß $\neg\, (\exists\, u)(y = 2.u)$.

(Wir haben hier von dem Prinzip Gebrauch gemacht, nach dem aus $x + y = x + z$ folgt, daß $y = z$. Dieses Prinzip läßt sich leicht nach Induktion beweisen:

(i) Wenn $o + y = o + z$, dann natürlich $y = z$.
(ii) Wenn gilt, dass (a) wenn $x + y = x + z$, dann $y = z$, dann gilt auch, dass (b) wenn $Sx + y = Sx + z$, dann $y = z$. Denn sei $Sx + y = Sx + z$. Dann $y + Sx = S(y + x) = z + Sx = S(z + x)$. Dann aber $y + x = z + x$. Also $x + y = x + z$ und nach Induktionshypothese $y = z$.)