INTRODUCTORY NOTES ON MODEL THEORY

by

R. KOPPERMAN

1.- INTRODUCTION

We all realize that the properties of a set of axioms are related to those of the class of structures which satisfy those axioms. Model Theory is a study of the relationships which may hold. Before we can study the relationships between axioms and structures which satisfy - them, we must know what we mean by a structure, a set of axioms, and how we make sense out of a set of axioms to see wether or not the - structure satisfies them. While we all agree that an «axiom» such - as the following makes sense:

$$(\forall x)(\forall y)(\forall z)(x < y, \wedge, y < z \Rightarrow x < z),$$

we don't know what to say about

$$(\forall x)(x <, \wedge, < y \Rightarrow < x * y)$$

How can even decide wether or not a proposed «axiom» even makes - sense? This is clearly the first question one must answer.

2.- TYPES, STRUCTURES AND LANGUAGES

A tipe t is simply a map from an ordinal number ρ (called the do main of t, $\mathcal{D}(t)=\rho$), into the set of non-negative intergers ω. (For those who prefer, a type is simply an ordered set of non-negative whole numbers).

We say that R is a relation on A iff for some positive interger N, $R \subseteq A^N$.

A first-order structure $\mathfrak{A} = \langle A, R_\lambda \rangle_{\lambda<\rho}$ is something that satisfies the following conditions:

(a) ρ is an ordinal and $A \neq \emptyset$ a set.

(b) for each $\lambda < \rho$, $R_\lambda \subseteq A^{N\lambda}$ is a relation, or $R_\lambda = e_\lambda \in A$ is a <<distinguished>> element of A.

We define the type of the structure $\mathfrak{A}$ by $\mathcal{D}(t)=\rho$, and for $\lambda < \rho$, $t(\lambda) = N\lambda$ (if $R_\lambda \subseteq A^{N\lambda}$), or $t(\lambda) = 0$ (if $R_\lambda = e_\lambda \in A$).

A class of structures K is called a class of similar structures - iff any two structures in K have the same type.

EXAMPLE:

Groups may be considered as structures of type <3,2,0> as follows: let $G = \langle A, \circ, {}^{-1}, e \rangle$, where $\circ = R_0$ is the "binary operation for the group G (i.e., $\langle a,b,c \rangle \in R_0 \Leftrightarrow a \circ b = c$), R_1 is the invertíve relation (i.e., $\langle a,b \rangle \in R_1 \Leftrightarrow a = b^{-1}$), and e is the neutral element, i.e., $e \circ a = a \circ e = a$, for each $a \in A$ (where A is the set of - elements of G).

NOTE:

From now on, any class of structures mentioned will be implicity assumed to be a class of similar structures (unless otherwise indicated), and similarly any pair, etc., of structures will be assumed similar.

We still don't know (officially). How to talk about our structures

Let t be a type. The (lower predicate calculus) language L_t is defined as follows: it will have a set $V = \{v_o, v_1, \dots\}$ of variables. For each $\lambda < \rho = \mathcal{D}(t)$, if $t(\lambda) = 0$, we will have a constant c_λ; if $t(\lambda) > 0$, we will have a $t(\lambda)$- ary predicated P_λ (i.e., P_λ is simply a place-holder for $t(\lambda)$ variables or constants). Now, letting x_i be variables or constants, we set:

(A) Σ_{o_t} is the set of atomic formulas of type t, where an atomic formula of type t is of the form,

(i) $x_i = x_j$, or

(ii) $P_\lambda (x_{i_1}, \dots, x_{i_{t(\lambda)}})$ (where we note once more that the x_i are variables or constants)

(B) Assume we have defined Σ_{r_t}. Them

$$\Sigma_{(r+1)_t} = \Sigma_{r_t} \cup \{\neg F ; F \in \Sigma_{r_t}\} \cup \{F \vee G ; F, G \in \Sigma_{r_t}\} \cup \{(\exists v_i) F ; i \in \omega, F \in \Sigma_{r_t}\}$$

(C) $L_t = \bigcup_{r=o}^{\infty} \Sigma_{r_t}$

NOTE:

$\forall, \wedge, \Rightarrow, \Leftrightarrow$, etc., are defined as appropriate abbreviations. For exemple, $F \wedge G$ stands for $\neg(\neg F \vee \neg G)$. Note also that our defini-- tion is inductive, so other definitions and proofs based on it will be inductive.

We now come to the final task of assigning meanings (and truth values) to elements of L_t. First let us consider an exemple: $+(v_1, v_2, v_3)$ is an atomic formula of $L_{<3>}$. It has no definite value of its own, i;e. without any knowledge of elements a_1, a_2, a_3 which we shall correspond to

v_1, v_2, v_3, we don't know whether the formula $+(v_1, v_2, v_3)$ is true about a_1, a_2, a_3 (that is, whether $a_1 + a_2 = a_3$).

Thus, the truth or falseness of a formula, in a given structure, may also depend on a sequence of elements in that structure.

Let $v^{\mathfrak{A}} : L_t \times A^{\omega} \to \{0,1\}$ the unique funtion satisfying the following:

Let $a = \langle a_0, a_1, \ldots \rangle \in A^{\omega}$, $x_1 \ldots, x_n$ be variables or constants, and set $b_j = a_i$ if $x_j = v_i$, $b_j = e_\lambda$ if $x_j = c_\lambda$.

Then :

(D) $v^{\mathfrak{A}}(x_k = x_j, a) = 1$ iff $b_k = b_j$

$v^{\mathfrak{A}}(P_\lambda(x_1, \ldots, x_{t(\lambda)}, a) = 1$ iff $\langle b_1, \ldots, b_{t(\lambda)} \rangle \in R_\lambda$

(E) $v^{\mathfrak{A}}(\neg F, a) = 1 - v^{\mathfrak{A}}(F, a)$

(F) $v^{\mathfrak{A}}(F \vee G, a) = v^{\mathfrak{A}}(F, a) + v^{\mathfrak{A}}(G, a) - v^{\mathfrak{A}}(F, a) . v^{\mathfrak{A}}(G, a)$

(G) $v^{\mathfrak{A}}((\exists v_i) F, a) = \max \{ v^{\mathfrak{A}}(F, a(i/b)); b \in A \}$,

where $a(i/b) = \langle a_0, a_1, \ldots, a_{i-1}, b, a_{i+1}, \ldots \rangle$

As for the existence and uniqueness of function $v^{\mathfrak{A}}$, we don't prove them here. However, the assertions contained in (D)-(G) can be shown by induction.

We say that F in the language L_t <u>is satisfied by</u> $a \in A^{\omega}$ in the -

structure $\mathfrak{A}$, and write $\mathfrak{A} \models F[a]$ iff $v^{\mathfrak{A}}(F,a)=1$. We have the following properties:

(i) $\mathfrak{A} \models P_\lambda(v_1,\ldots, v_{t(\lambda)})[a]$ iff $\langle a_1\ldots, a_{t(\lambda)}\rangle \in R_\lambda$

(ii) $\mathfrak{A} \models \neg F[a]$ iff not $\mathfrak{A} \models F[a]$

(iii) $\mathfrak{A} \models F \vee G[a]$ iff $\mathfrak{A} \models F[a]$ or $\mathfrak{A} \models G[a]$

(iv) $\mathfrak{A} \models (\exists v_i) F[a]$ iff there exists a $b \in A$ such that $\mathfrak{A} \models F[a(i/b)]$

We have now acomplished our first major goal. We know what is meant by a strcture and given a structure, we have a language in which we may talk about it (naturelly, if our structure is of type t, we may use L_t), and a way to discover whether or not what we say is true on a sequence in the given structure.

EXAMPLE:

Let us consider the real number system $\mathbb{R} = \langle R, +, \cdot, 0, 1, < \rangle$ of type $t = \langle 3,3,0,0,2 \rangle$. The least upper bound axiom states that each subset of R with an upper bound has at least one least upper bound. That axiom in the presence of certain others determines $\mathbb{R}$ up to an isomorphism. It can not obviously be written in L_t (but the others can for example conmutativity:

$$(\forall v_0)(\forall v_1)(\forall v_2)(+(v_0,v_1,v_2) \Rightarrow +(v_1,v_0,v_2))$$

and later we shall prove that it cannot be written in any L_s.

Thus, not every fact can be expressed in an lower predicated calculus. In general, <<algebraic>> facts can, but facts about analysis cannot. However theorems in both analysis and algebra have shown using model theory.

One minor problem remains: our formulas may be true on some sequences and false on others, in the same structure, whereas axioms should either hold or not hold for a structure. For this, we say that a variable is free in F iff it is not after a quantifier (i.e., iff $(\exists v_i)$, $(\forall v_i)$ never appears in F). This concept can be formally defined by induction, and the following can be shown:

LEMMA 1: Let a, $a' \in A^\omega$ such that for any free variable v_i of F, $a_i = a'_i$. Them $\mathfrak{A} \models F[a]$ iff $\mathfrak{A} \models F[a']$. ▌

A predicated $\sigma \in L_t$ is called a structure iff σ has no free variables. Let us put Λ_t for the set of sentences in L_t

Then we have the following corollary of the lemma above:

COROLLARY.

Let $\mathfrak{A}$ be of type t, $\sigma \in \Lambda_t$, a, $a' \in A^\omega$. Then $\mathfrak{A} \models \sigma[a]$ iff $\mathfrak{A} \models \sigma[a']$. ▌

Let $\sigma \in \Lambda_t$, then we say that the structure $\mathfrak{A}$ of type t is a model of σ (written $\mathfrak{A} \in M(\sigma)$ or $\mathfrak{A} \models \sigma$) iff for each $a \in A^\omega$, $\mathfrak{A} \models \sigma[a]$. If $S \subseteq \Lambda_t$, we write

$$M(S) = \left\{ \mathfrak{A};\ (\forall \sigma \in S)\ (\mathfrak{A} \models \sigma) \right\} = \bigcap_{\sigma \in S} M(\sigma)$$

We may also write $\mathfrak{A} \models S$ for $\mathfrak{A} \in M(S)$.

Thus, for us, a set of axioms is simply a set of sentences.

3. ULTRAPRODUCTS.

Let I be a set; a filter D on I is a non-empty collection of subsets of I, such that:

(H) $\emptyset \notin D$

(I) $a, b \in D \Rightarrow a \cap b \in D$

(J) $a \in D,\ a \subseteq b \subseteq I \Rightarrow b \in D$

Now let $\{\mathfrak{A}_i ; i \in I\}$ be a set of structures, D a filter on I. Then we define the <u>reduced product</u>.

$$\prod_{i \in I} \mathfrak{A}i/D$$

as follows;assume $\mathfrak{A}_i = \langle A_i, R^i_\lambda \rangle_{\lambda<\rho}$. Then $\prod_{i\in I} \mathfrak{A}i/D = \langle A, R_\lambda \rangle_{\lambda<\rho}$, where the terms are defined as follows

(K) If $B = \prod_{i\in I} A_i = \{f : I \to \bigcup_{i\in I} A_i ; f(i) \in A_i\}$, we say that

$f \equiv_D g$, $f, g \in B$, iff $\{i ; f(i) = g(i)\} \in D$. It is easy to show that $\equiv_D$ is an equivalence relation (transitivity using (I) and (J) - from the definition of a filter, and reflexivity requiring the fact - that $I \in D$). Then we set $A = B/\equiv_D$.

(L) Let $f \in B$; if f/D denotes the class of f modulus $\equiv_D$, we say that

$$\langle f_1/D, \ldots, f_{t(\lambda)}/D \rangle \in R_\lambda \text{ iff}$$

$$\{i ; \langle f_1(i), \ldots, f_{t(\lambda)}(i) \rangle \in R^i_\lambda\} \in D$$

Let us consider two filters D and D' on I. We say that D is <u>finer than</u> D' ($D \leq D$) iff $D \subseteq D'$. An <u>ultrafilter</u> D on I is a filter on I such that if $D \leq D'$, then $D = D'$ (i.e., D is <u>maximal</u> with respect to the <u>order</u> $\leq$). If the filter in the definition of a reduced product is an ultrafilter on I, then $\prod_{i \in I} \mathfrak{A}_i/D$ is called an <u>ultraproduct</u>.

Example:

Let $a \subseteq I$. Then $D_a = \{b \subseteq I;\ a \subseteq b\}$ is a filter on I, and is called <u>principal</u> (<u>ortrivial</u>). If $a = \{x\}$, then D_a is an ultrafilter.

<u>THEOREM 1</u>:

<u>Each filter can be imbedded in an ultrafilter</u>

Proof: By Zorn s Lemma ▮

<u>COROLLARY</u>:

<u>Let E be a set of subsets of I such that for</u> $e_1, e_2, \ldots, e_k \in E$, $e_1 \cap e_2 \cap \ldots \cap e_k \neq \emptyset$. <u>Then E can be imbedded in an ultrafilter</u>.

Proof: The set $D = \{a \subseteq I;\ (\exists e_1, \ldots, e_k)(e_1 \cap \ldots \cap e_k \subseteq a\}$ can be shown to be a filter. Thus D can be imbedded in an ultrafilter D' But $E \subsetneq D \subsetneq D'$. ▮

<u>THEOREM 2</u>:

<u>If D is an ultrafilter on I, then for each</u> $a \subseteq I$, $a \in D$ <u>or</u> $(I \sim a) \in D$ (<u>but not both</u>)

Proof: see [3]. ▮

<u>COROLLARY 1</u>:

<u>Let</u> D <u>be an ultrafilter on</u> I; <u>if</u> $a \cup b \in D$, <u>then</u> $a \in D$ <u>or</u> $b \in D$.

Proof: If $a \notin D$, then $(I \sim a) \in D$, so $(a \cup b) \cap (I \sim a) = b \in D$ ▮

<u>COROLLARY 2</u>:

If $a_1 \cup \ldots \cup a_k \in D$ D an ultrafilter then some $a_i \in D$

Proof: By induction using last corollary ▮

THEOREM 3:

If I infinite, then there are non-principal ultrafilters on I

Proof: Extend $E = \{ I \sim \{x\} \; ; \; x \in I \}$ to an ultrafilter. It must be non-principal, since $(\forall x)(I \sim \{x\} \in E)$

Notation:

We recall that in general, truth is a function of sequences (etc.) Our ultraproduct also, uses sequences (or functions, at least). To avoid confusion, we use the following conventions:

Suppose that $f = \langle f_1/D, f_2/D, \ldots \rangle \in (\prod_{i \in I} A_i/D)^{\omega}$ By $f(i)$ we denote $\langle f_1(i), f_2(i), \ldots \rangle \in A_i^{\omega}$, and f_k we shall mean a representative of f_k/D in $\prod_{i \in I} A_i$.

We state now our basic result:

THEOREM 4:

Let $\mathfrak{A} = \prod_{i \in I} \mathfrak{A}i/D$ be the product of the $\mathfrak{A}_i$, $f = \langle f_1/D, f_2/D, \ldots \rangle \in (\prod_{i \in I} Ai/D)$. Then

$$\mathfrak{A} \models F[f] \iff \{ i \; ; \; \mathfrak{A}_i \models F[f(i)] \} \in D \quad (\forall F \in L_t)$$

Proof: By induction: For $F \in \Sigma_{o_t}$, our theorem is true by definition of $\equiv_D$ and R_λ. Assume now that our theorem is true for all f, - all $F \in \Sigma_{r_t}$. Let $G \in \Sigma_{(r+1)_t} \sim \Sigma_{r_t}$. If $G = \neg F$, then $\mathfrak{A} \models G[f]$ iff not $\mathfrak{A} \models F[f]$ iff (induction step):

$$\{ i \; ; \; \mathfrak{A}_i \models \neg F[f(i)] \} \in D \text{ iff } \{ i \; ; \; \mathfrak{A}_i \models G[f(i)] \} \in D.$$

The case of $G = F \vee E$ is done similarly. For the case $G = (\exists \sim_j) F$,

$\mathfrak{A} \models G[f]$ iff for some $g/D \in \prod_{i\in I} A_i/D$, $\mathfrak{A} \models F[f(j/(g/D))]$

iff by induction $\{i ; \mathfrak{A}_i \models F[f(j/g)(i)]\} \in D$ for some g iff

$\{ i ; \mathfrak{A}_i \models (\exists \sim_j) F[f(i)]\} \in D$ iff $\{i ; \mathfrak{A}_i \models G[f(i)]\} \in D$.

We have now shown the theorem for all $G \in \Sigma_{(r+1)_t}$. By induction, it is true for all $G \in \bigcup_{r=0}^{\infty} \Sigma_{r_t} = L_t$

COROLLARY 1:

Let $\sigma \in \Lambda_t$. Then $\mathfrak{A} \models \sigma$ iff $\{i ; \mathfrak{A}_i \models \sigma\} \in D$ ▮

Let $\mathfrak{A}$ and $\mathfrak{B}$ be estructures of type t. Then $\mathfrak{A} \equiv \mathfrak{B}$ iff for all $\sigma \in \Lambda_t$, $\mathfrak{A} \models \sigma \Leftrightarrow \mathfrak{B} \models \sigma$. In this case, we also say that $\mathfrak{A}$ is elementarily equivalent to $\mathfrak{B}$ We can then state the following corollary.

COROLLARY 2:

If for all i, j $\in$ I, $\mathfrak{A}_i \equiv \mathfrak{A}_j$, then for each j $\in$ I, $\mathfrak{A}_j \equiv \prod_{i\in I} \mathfrak{A}_i/D$. In particular, if for all $i \in I$, $\mathfrak{A}_i = \mathfrak{A}$, then: $\prod_{i\in I} \mathfrak{A}_i/D$ is called an ultrapower of $\mathfrak{A}$ and denoted by $\mathfrak{A}^I/D$.

In this case, $\mathfrak{A} \equiv \mathfrak{A}^I/D$ ▮

THEOREM 5:

Let $S \subseteq \Lambda_t$ be such that for any finite subset $S' \subseteq S$, $M(S') \neq \emptyset$. Then $M(S) \neq \emptyset$.

Proof: For each finite set $S' \subseteq S$ let $\mathfrak{A}_{S'} \in M(S')$. Now let

I be the set of finite subsets of S. For each $\sigma \in S$, let $S_\sigma = \{ S_i \in I;$ $\sigma \in S_i \}$ and let $D' = \{ S_\sigma ; \sigma \in S \}$. Then D' is a set of subsets of I. Let $S_{\sigma_1}, \ldots\ldots\ldots, S_{\sigma_k} \in D$; then $\{\sigma_1, \ldots, \sigma_k\}$ $\in S_{\sigma_1} \cap \ldots \cap S_{\sigma_k}$, so $S_{\sigma_1} \cap \ldots \cap S_{\sigma_k} \neq \emptyset$. Thus $D' \subseteq D$ for some ultrafilter D on I. Now let $\mathfrak{A} = \prod_{i \in I} \mathfrak{A}i/D$. We shall show $\mathfrak{A} \in M(S)$. For any $\sigma \in S$, the set $\{ S' \in I; \mathfrak{A}_S \sqsubset \sigma \}$ contains $\{ S_i \in I ; \sigma \in S_i \} = S_\sigma \in D' \subseteq D$. Thus $\{ S' \in I; \mathfrak{A}_{S'} \sqsubset \sigma \} \in D$, so $\mathfrak{A} \sqsubset \sigma$; thus $\mathfrak{A} \sqsubset S$ ∎

<u>THEOREM</u> <u>6</u>:

<u>Let $S \subseteq \Lambda_t$ be such that each finite N there is an $\mathfrak{A} \in M(S)$ such that $N \leq \mathcal{K}(\mathfrak{A})$ (i. e., A, The set of elements of $\mathfrak{A}$, has more than N members) then for each cardinal M (including infinite ones) there is an $\mathfrak{A} \in M(S)$ such that $M \leq \mathcal{K}(\mathfrak{A})$.</u>

Proof: (A <u>trick</u> is involved - we change languages). Let $\rho = \mathcal{D}(t)$, thus if $\mathfrak{A} \in M(S)$, $\mathfrak{A} = \langle A, R_\lambda \rangle_{\lambda < \rho}$. Now choose ρ' such that $\mathcal{K}(\rho' \sim \rho) \geq M$, and t' such that if $\lambda < \rho$, $t'(\lambda) = t(\lambda)$, and if $\rho \leq \lambda < \rho'$; $t'(\lambda) = 0$ (and $\rho' = \mathcal{D}(t)$).
We have adjoined constants, $c_\rho, c_{\rho+1}, \ldots, c_\lambda, \lambda < \rho'$). We assert that

$$T = S \cup \{ \neg c_\lambda = c_{\lambda'} ; \rho \leq \lambda < \lambda' < \rho' \} \subseteq \Lambda_{t'}$$

is consistent (i. e., $M(T) \neq \emptyset$). Let $T_1 \subseteq T$ be finite. Then $T_1 \subseteq S \cup \{ \neg c_{\lambda_i} = c_{\lambda_j} ; i<j, \rho \leq \lambda_1 < \ldots < \lambda_k < \rho' \} = T_k$.

Now take $\mathfrak{A} \in M(S)$, $\mathcal{K}(\mathfrak{A}) \geq K$, and let $\mathfrak{A}' = \langle A, R_\lambda, e_\lambda \rangle \lambda < \rho \leq \lambda' < \rho'$ (where $\mathfrak{A} = \langle A, R_\lambda \rangle_{\lambda < \rho}$). Let $a_1, \ldots, a_k \subseteq A$, be unequal, and set $e_{\lambda_i} = a_i$, $e_\lambda = a_1$ si $\lambda \neq \lambda_1, \ldots, \lambda_k$.

We claim $\mathfrak{A} \in M(T_k)$. Since $\mathfrak{A} \sqsubset S$ (its relations are the same, as far as S is concerned). But since $e_{\lambda_i} \neq e_{\lambda_j}$, we have $\mathfrak{A} \sqsubset \{\neg c_{\lambda_i} ; i < j, \rho \leq \lambda_1 < \ldots < \lambda_k < \rho'\}$, thus $\mathfrak{A}' \in M^j(T_1)$. Now by theo. 5, we can find $\mathfrak{B} \in M(T), \mathfrak{B} = \langle B, S_\lambda, d_\lambda \rangle \; \lambda < \rho \leq \lambda' < \rho'$. Since

$$\mathfrak{B} = \{\neg c_\lambda = c_\lambda \; ; \; \rho \leq \lambda < \lambda' < \rho'\}$$

$d_{\lambda'} \neq d_{\lambda''}$ for $\rho \leq \lambda' < \lambda'' < \rho'$, so $\mathcal{K}(B) \geq \mathcal{K}(\rho' - \rho) \geq M$.

Now consider $\mathfrak{B}_0 = \langle B, S_\lambda \rangle_{\lambda<\rho}$. $\mathfrak{B}_0$ is of type t.

$\mathcal{K}(\mathfrak{B}_0) = \mathcal{K}(B) \geq M$, and since $\mathfrak{B} \sqsubset S \subseteq \Lambda_t$, so does $\mathfrak{B}$. ∎

<u>COROLLARY</u>:

The least upper bound axiom is not in Λ_t <u>for any</u> t (<u>we say that it is not a first-order axiom</u>)

Proof: Assume the contrary. Thus all the axioms for $\mathbb{R}$ can be written in some Λ_t. Let $S \subseteq \Lambda_t$, be this set of axioms.
Since $\mathbb{R} \in M(S)$, for each finite N, we have $\mathbb{R} \in M(S)$ with $N \leq \mathcal{K}(\mathbb{R})$
Thus by theo.6, for each cardinal M, including $2^{\mathcal{K}(\mathbb{R})}$ ( ), we have an $\mathfrak{A} \in M(S), \mathcal{K}(\mathbb{R}) \geq M$.
But we now that all models of S are isomorphic to $(\mathbb{R})$, thus if $\mathfrak{A} \in M(S)$
$\mathcal{K}(\mathbb{R}) = \mathcal{K}(\mathbb{R}) \geq 2^{\mathcal{K}(\mathbb{R})}$, a contradiction, for $\mathcal{K}(\mathbb{R}) < 2^{\mathcal{K}(\mathbb{R})}$

This shows that we need <<higher>> predicated calculus to express in this way all we want to do in mathematics.

<u>BIBLIOGRAFIA</u>

[1] ROBINSON, A.: Meta-mathematics of Algebra, North-Holland Pub. Co., Amsterdam, 1951.

[2] SUPPES, P.: An introduction to Logic, Van Nostrand, N. York, 1957

[3] KLEENE, Introduction to Meta-mathematics, Van Nostrand, New York, 1954.

[4] BOURBAKI, N.: Thèorie des Ensembles, Hermann, Paris, 1957.

SUMARIO

Se trata de una introducción a la teoría de modelos.

Department of Mathematics
University of Rhode Island

(Received August 1.966)