Math 225A – Model Theory

Speirs, Martin

Autumn 2013

General Information

These notes are based on a course in *Metamathematics* taught by Professor Thomas Scanlon at UC Berkeley in the Autumn of 2013. The course will focus on Model Theory and the course book is Hodges' *a shorter model theory.*

As with any such notes, these may contain errors and typos. I take full responsibility for such occurences. If you find any errors or typos (no matter how trivial!) please let me know at mps@berkeley.edu.

Lecture 3

We continue studying the connection between language and signature. Last time we discussed how one might interpret a language in a structure. Today we will go the other way and associate to each structure a certain set of formulas which describe the structure.

We let τ be a signature and $\mathfrak{A}$ a τ-structure with domain A. Recall that a *sentence* in $\mathscr{L}(\tau)$ is a formula with no free variables.

Definition. The **theory of** $\mathfrak{A}$, written $\mathrm{Th}(A)$ (or $\mathrm{Th}_{\mathscr{L}(\tau)}(\mathfrak{A})$ to emphasize the signature), is the set of sentences ψ in $\mathscr{L}(\tau)$ that such that $\mathfrak{A} \models \psi$. I.e.

$$\mathrm{Th}(\mathfrak{A}) = \{\psi \in \mathscr{L}(\tau) : \psi \text{ is a sentence and } \mathfrak{A} \models \psi\}.$$

The theory of $\mathfrak{A}$ contains all that can be said about the structure $\mathfrak{A}$ using the language $\mathscr{L}(\tau)$. There are some subclasses of $\mathrm{Th}(\mathfrak{A})$ which are also of interest, for instance we might look at all *quantifier-free* sentences true in $\mathfrak{A}$ or all *existential* sentences true in $\mathfrak{A}$.

If we want to describe the basic structure $\mathfrak{A}$ itself, (i.e. answer questions such as; What are the relations? What are the functions? What are the constants inside of $\mathfrak{A}$?) then we look at the *diagram* of $\mathfrak{A}$. To define this object we first need to say what a *literal* sentence is.

Definition. A sentence of the form φ or $\neg(\varphi)$ where φ is atomic is called a **literal**.

Definition. The **diagram of** $\mathfrak{A}$, written $\mathrm{diag}(\mathfrak{A})$, is the set of literals that are true in $\mathfrak{A}_A$, i.e.

$$\mathrm{diag}(\mathfrak{A}) = \{\varphi \in \mathscr{L}(\tau_A) : \varphi \text{ is a literal and } \mathfrak{A}_A \models \varphi\}.$$

The **positive diagram of** $\mathfrak{A}$, written $\mathrm{diag}^+(\mathfrak{A})$ is the set of atomic formulas true in $\mathfrak{A}_A$, i.e.

$$\mathrm{diag}^+(\mathfrak{A}) = \{\varphi \in \mathscr{L}(\tau_A) : \varphi \text{ is a atomic and } \mathfrak{A}_A \models \varphi\}.$$

The diagram should be thought of as the "multiplication table" of the structure in analogy with the multiplication table of a group – even though the diagram is not *strictly* a generalization of the multiplication table in the case where $\mathfrak{A}$ is a group. But it contains the same information. For example, suppose G is a structure in the language of groups and we have $a \cdot b = c$ in G. Then the sentence $a \cdot^G b = c$ will be in the diagram. Or suppose $\mathfrak{A}$ is in the structure of partial orders. Then the diagram will include information like $a < b$ and $a \not< b$ in the partial order. In this case the *positive* diagram will contain different information than the diagram. To see this, consider two elements of the partial order which are not related, i.e. $a \not< b$. So neither $a < b$ nor $b < a$ will be in the positive diagram, but we could have some other structure $\mathfrak{B}$ such that $\mathfrak{B}$ satisfies all sentences of $\text{diag}^+(\mathfrak{A})$ *and* $a < b$. In which case the positive diagram of $\mathfrak{B}$ will strictly contain $\text{diag}^+(\mathfrak{A})$.

Of course we would like to say that $\mathfrak{B}$ looks more or less the same as $\mathfrak{A}$ if $\mathfrak{B} \models \text{diag}(\mathfrak{A})$. Indeed there is a relation. To demonstrate it we first need a lemma, which states that homomorphisms commute with arbitrary terms.

Lemma. *Let $\rho : \mathfrak{A} \to \mathfrak{B}$ be a homomorphism of τ-structures, and $t(x_0, \ldots, x_{n-1}) \in \mathscr{L}(\tau_{\{x_1,\ldots,x_n\}})$ and $a_0, \ldots, a_{n-1} \in A$. Then $t(\rho(a_0), \ldots, \rho(a_{n-1})) = \rho(t(a_0, \ldots, a_{n-1}))$.*

Proof. We proceed by induction on the complexity of t.

- if t is $c \in \mathcal{C}_\tau$ then $\rho(c^{\mathfrak{A}}) = c^{\mathfrak{B}}$ since ρ is a homomorphism.
- if t is x_i $(0 \leq i \leq n-1)$ then $\rho(t(a_0, \ldots, a_{n-1})) = \rho(a_i) = t(\rho(a_0), \ldots, \rho(a_{n-1}))$.
- if t is $f(t_0, \ldots, t_{n-1})$ and the statement is true for $t_0, \ldots, t_{n-1}$ then

$$\begin{aligned} \rho(t^{\mathfrak{A}}(\bar{a})) &= \rho(f^{\mathfrak{A}}(t_0^{\mathfrak{A}}(\bar{a}), \ldots, t_{n-1}^{\mathfrak{A}}(\bar{a}))) \\ &= f^{\mathfrak{B}}(\rho(t_0^{\mathfrak{A}}(\bar{a})), \ldots, \rho(t_{n-1}^{\mathfrak{A}}(\bar{a}))) \\ &= f^{\mathfrak{B}}(t_0^{\mathfrak{B}}(\rho\bar{a})), \ldots, t_{n-1}^{\mathfrak{B}}(\rho\bar{a})) \\ &= t^{\mathfrak{B}}(\rho\bar{a}) \end{aligned}$$

□

With the lemma at hand we can now state the relationship between structures and satisfying the (positive) diagram.

Proposition. *Let $\mathfrak{A}$ be a τ-structure. The following are equivalent for a τ-structure $\mathfrak{B}$:*

- *The exists an expansion of $\mathfrak{B}$ to $\mathfrak{B}'$ in τ_A such that $\mathfrak{B}' \models \text{diag}^+(\mathfrak{A})$.*
- *There exists a homomorphism $\rho : \mathfrak{A} \to \mathfrak{B}$.*

Proof. "⇒" Let $\mathfrak{B}'$ be an expansion to τ_A such that $\mathfrak{B}' \models \text{diag}^+(\mathfrak{A})$. Define $\rho : A \to B$ by $\rho(a) = a^{\mathfrak{B}'}$. We check that ρ is a homomorphism.

- For $c \in \mathcal{C}_\tau$. We must show that $\rho(c^{\mathfrak{A}}) = c^{\mathfrak{B}}$. By definition $\rho(c^{\mathfrak{A}}) = (c^{\mathfrak{A}})^{\mathfrak{B}'}$. Now consider the sentence φ which is $c^{\mathfrak{A}} = c$ (this is an atomic formula in $\mathscr{L}(\tau_A)$. Now $\mathfrak{A}_A \models \varphi$ since $(c^{\mathfrak{A}})^{\mathfrak{A}} = c^{\mathfrak{A}}$. So $\varphi \in \mathrm{diag}^+(\mathfrak{A})$ which, by assumption, implies $\mathfrak{B}' \models \varphi$.
- For $f \in \mathcal{F}_\tau$ with arity$(f) = n$ and $a_1, \ldots, a_n \in A$ then the formula ψ : $f(a_1, \ldots, a_n) = b$ with $b := f^{\mathfrak{A}}(a_1, \ldots, a_n)$. Now $\mathfrak{A}_A \models \psi$ i.e. $\psi \in \mathrm{diag}^+(\mathfrak{A})$. But then by assumption $\mathfrak{B}' \models \psi$ i.e. $f^{\mathfrak{B}'}(a_1^{\mathfrak{B}'}, \ldots, a_n^{\mathfrak{B}'}) = b^{\mathfrak{B}'}$, thus

$$f^{\mathfrak{B}}(\rho(a_1), \ldots, \rho(a_n)) = \rho(b) = \rho(f^{\mathfrak{A}}(a_1, \ldots, a_n)).$$

- For $R \in \mathcal{R}_\tau$ with arity$(R) = n$ and $a_1, \ldots, a_n \in A$ such that $R^{\mathfrak{A}}(a_1, \ldots, a_n)$. Now let $\theta : R(a_1, \ldots, a_n) \in \mathscr{L}(\tau_A)$. Then $\theta \in \mathrm{diag}^+(\mathfrak{A})$ and so $\mathfrak{B}' \models \theta$ so $R^{\mathfrak{B}'}(a_1^{\mathfrak{B}'}, \ldots, a_n^{\mathfrak{B}'})$ which by definition is $R^{\mathfrak{B}'}(\rho(a_1), \ldots, \rho(a_n))$ so $R^{\mathfrak{B}}(\rho(a_1), \ldots, \rho(a_n))$.

Thus ρ is indeed a homomorphism.

"$\Leftarrow$" Let $\rho : \mathfrak{A} \to \mathfrak{B}$ be a homomorphism. We expand $\mathfrak{B}$ to $\mathfrak{B}'$ in τ_A by setting $a^{\mathfrak{B}'} = \rho(a)$. This is clearly an expansion of $\mathfrak{B}$ to τ_A. We must show that $\mathfrak{B}' \models \mathrm{diag}^+(\mathfrak{A})$. We do this by induction on atomic formulas.

- Suppose $t, s \in \mathscr{T}(\tau_A)$ and $\mathfrak{A}_A \models s{=}t$. Now both s and t are closed terms in τ_A so there are $\tilde{s}$ and $\tilde{t}$ in $\mathscr{T}(\tau_{\{x_i : i \in \omega\}})$ such that $s^{\mathfrak{A}_A} = \tilde{s}^{\mathfrak{A}}(\bar{a})$ and $t^{\mathfrak{A}_A} = \tilde{t}^{\mathfrak{A}}(\bar{a})$ for some $\bar{a} \in A^n$. By assumption $\tilde{s}^{\mathfrak{A}}(\bar{a}) = \tilde{t}^{\mathfrak{A}}(\bar{a})$. Now by the lemma preceding this proposition, ρ commutes with terms and so

$$\rho(\tilde{s}^{\mathfrak{A}}(\bar{a})) = \tilde{s}^{\mathfrak{B}}(\rho\bar{a}) = \tilde{s}^{\mathfrak{B}}(\bar{a}^{\mathfrak{B}'}).$$

but this is the same as $\tilde{s}^{\mathfrak{B}'}(\bar{a}^{\mathfrak{B}'})$ since the interpretation of function symbols in $\mathfrak{B}$ doesn't change under the extension to $\mathfrak{B}'$. Likewise $\rho(\tilde{t}^{\mathfrak{A}}(\bar{a})) = \tilde{t}^{\mathfrak{B}'}(\bar{a}^{\mathfrak{B}'})$. Now since ρ is a function and $\tilde{s}^{\mathfrak{A}}(\bar{a}) = \tilde{t}^{\mathfrak{A}}(\bar{a})$ we have $\tilde{s}^{\mathfrak{B}'}(\bar{a}^{\mathfrak{B}'}) = \tilde{t}^{\mathfrak{B}'}(\bar{a}^{\mathfrak{B}'})$, i.e. $s^{\mathfrak{B}'} = t^{\mathfrak{B}'}$, so $\mathfrak{B}' \models s{=}t$ as well.
- Similar reasoning applies to atomic formulas given by relation symbols and so by induction $\mathfrak{B}' \models \mathrm{diag}^+(\mathfrak{A})$.

□

We have a similar but stronger relationship for the diagram. First we define the notion of *embedding* to be a morphism which respects negations of relations.

Definition. A homomorphism $\iota : \mathfrak{A} \to \mathfrak{B}$ is an **embedding** if it is injective and if $(a_1, \ldots, a_n) \in R^{\mathfrak{A}}$ if and only if $(\iota(a_1), \ldots, \iota(a_n)) \in R^{\mathfrak{B}}$.

Proposition. *Let $\mathfrak{A}$ be a τ-structure. The following are equivalent for a τ-structure $\mathfrak{B}$.*

1. *There exists an expansion* $\mathfrak{B}'$ *of* $\mathfrak{B}$ *to* τ_A *such that* $\mathfrak{B}' \models \mathrm{diag}(\mathfrak{A})$.
2. *There exists an embedding* $\iota : \mathfrak{A} \to \mathfrak{B}$.
3. *There exists a substructure* $\mathfrak{A}' \subseteq \mathfrak{B}$ *such that* $\mathfrak{A} \cong \mathfrak{A}'$.

Proof. First note that *2* $\Leftrightarrow$ *3* by definition. Most of the rest of the proof is done exactly as before when looking at the positive diagram. In going from *1* to *2* we take $\mathfrak{B}' \models \mathrm{diag}(\mathfrak{A})$ and define $\rho : \mathfrak{A} \to \mathfrak{B}$ by $a \mapsto a^{\mathfrak{B}'}$ and check (like above) that this is a homomorphism. Now ρ will be injective since, if $a \neq b$ in A then it follows that $\mathfrak{A}_A \models \neg(a = b)$, i.e. $\neg(a = b) \in \mathrm{diag}(\mathfrak{A})$. Thus, $\mathfrak{B}' \models \neg(a = b)$ and so $\rho(a) \neq \rho(b)$. The rest of the proof is much the same as before. □

These propositions show the first steps of how model theory works by going between syntax and semantics. We can convert properties which are purely structural into statements about satisfying certain sorts of formulas and sentences.

What we have called the diagram might also be called the **quantifier-free diagram**, since we only include sentences without quantifiers. If we want more information about the structure we can also look at the **elementary diagram**, el$-$diag($\mathfrak{A}$) which by definition is $\mathrm{Th}(\mathfrak{A}_A)$.

We now prove a result about the existence of a model of a theory in much the same way as with the term algebra. We take a theory where we would like to find a model and basically just letting the language serve this goal.

Definition. A set T of $\mathscr{L}(\tau)$-sentences is **=-closed** if

- for all closed terms t, s in $\mathscr{T}(\tau)$ and for all formulas φ with one free variable x, if $\varphi(t) \in T$ and if $t = s \in T$ then $\varphi(s) \in T$,
- for all closed terms t we have $t = t \in T$.

Remark. If S is any set of $\mathscr{L}(\tau)$-sentences then there is a smallest =-closed set $\tilde{S}$ containing S.

So the following proposition could be applied to any set of $\mathscr{L}(\tau)$-sentences by passing to the =-closure of the given set first.

Proposition. *If* T *is an =-closed set of atomic sentences then there exists a structure* $\mathfrak{A}$ *such that* $\mathfrak{A} \models T$ *and such that for any* $\mathfrak{B}$ *with* $\mathfrak{B} \models T$ *there is a unique homomorphism* $\mathfrak{A} \longrightarrow \mathfrak{B}$.

Remark. If $T = \emptyset$ then $\mathfrak{A}$ will be the term algebra.

Proof. The domain of $\mathfrak{A}$ will be $\mathscr{T}(\tau)$ modulo the equivalence relation given by $s \sim t$ if and only if $s = t \in T$. Let us show that this is indeed an equivalence relation.

- *Reflexivity*: For all t we have $t \sim t$ since by assumption $t = t \in T$.

- *Symmetry*: Suppose $s \sim t$ so that $s = t \in T$. Consider the formula $\varphi(x)$ given by $x = s$. φ is an atomic formula with one free variable, x. Now $\varphi(s)$ is in T and so by $=$-closedness of T we have $\varphi(t)$ in T, i.e. $t = s \in T$ and so $t \sim s$.
- *Transitivity*: Suppose $s \sim t$ and $t \sim r$. Let $\varphi(x)$ be $x = r$. Then $\varphi(t) \in T$ and since $t \sim s$ by symmetry we have that $\varphi(s) \in T$ so $s \sim r$.

Thus, $\sim$ is an equivalence relation. We now let the domain of $\mathfrak{A}$ be $A := \mathscr{T}(\tau)/\sim$, and denote the equivalence class containing t by $[t]_\sim$. To define the τ-structure on $\mathfrak{A}$ we set

- for $c \in \mathcal{C}_\tau$, $c^{\mathfrak{A}} = [c]_\sim$
- for $f \in \mathcal{F}_\tau$ of arity n we define $f^{\mathfrak{A}}([t_0]_\sim, \ldots, [t_{n-1}]_\sim) = [f(t_0, \ldots, t_{n-1}]_\sim$
- for $R \in \mathcal{R}_\tau$ of arity n then $([t_0]_\sim, \ldots, [t_{n-1}]_\sim) \in R^{\mathfrak{A}}$ if and only if $R(t_0, \ldots, t_{n-1}) \in T$.

We must show that these definitions are well-defined and that $\mathfrak{A}$ has the desired property. We shall do this next time. $\square$