INTRODUCTION TO MODEL THEORY WITH APPLICATIONS

RONNIE NAGLOO

1. First-order Logic - an introduction

The first lecture is an introduction to the basics of first order logic geared to model theory. We explain how in mathematical logic, statements about mathematical objects and structures can be studied as mathematical objects in their own right. We will formalize notions such as a structure, a property, an axiom (or statement) and satisfaction. I have used David Marker's book "Model Theory: An Introduction" (Springer 2002) and Rahim Moosa's Lecture notes "Set Theory and Model Theory" as a source for ideas about exposition as well as some examples. I am grateful to both of these authors.

1.1. **Languages and Structures.**

Let us start by looking at two mathematical structures that you have likely come across in your math journey.

Example 1.1.

(1) A *group* $\mathcal{G} = (G, *, e)$ consists of
- a set G,
- a binary operation $* : G \times G \to G$,
- a distinguished element $e \in G$,
- the group axioms.

(2) An *ordered field* $\mathcal{F} = (F, +, -, \times, <, 0, 1)$ consists of
- a set F,
- binary operations $+, -, \times : F \times F \to F$,
- a binary relation $<$ on F,
- distinguished elements $0, 1 \in F$.
- the axioms for an ordered field.

Abstracting the common features of the above two examples (while not worrying about the axioms for a moment) brings us to the notion of a structure:

Definition 1.2. A **structure** $\mathcal{M}$ consists of a nonempty underlying set M called its *universe*, together with

- a set $\mathcal{C}$ of distinguished elements of M, called the *constants*,
- a set $\mathcal{F} = \{f : M^{n_f} \to M\}$ of distinguished functions on M called the *basic functions*, and
- a set $\mathcal{R} = \{R \subseteq M^{n_R}\}$ of distinguished subsets called the *basic relations*

Each of the nonzero natural numbers n_f (resp. n_R) above, is called the *arity* of the corresponding function (resp. relation). As we have seen, it could be that one of the set $\mathcal{C}$, $\mathcal{F}$ and/or $\mathcal{R}$ are empty. The set $\mathcal{C} \cup \mathcal{F} \cup \mathcal{R}$ is called the signature of $\mathcal{M}$.

Remark 1.3. An underlying set could be the universe of many structures. For example $(\mathbb{R}, +, 0)$ has the structure of a group and $(\mathbb{R}+, -, \times, <, 0, 1)$ that of an ordered field.

Going back to the example of groups above, it is not hard to see that groups in general will have different signatures. For example, this is clearly the case for $(\mathbb{Z}, +, 0)$ and $(\mathbb{R} \setminus \{0\}, \times, 1)$. Nevertheless in this example it is understood that $+$ and $\times$ are used to represent the group operation (which all groups must have). The following definition will help capture the idea that the (any) two groups have a common language.

Definition 1.4. A **Language** $\mathcal{L}$ consists of the following:

- a set $L_{\mathcal{C}}$ of constant symbols,
- a set $L_{\mathcal{F}}$ of function symbols, together with a nonzero natural numbers n_f for each $f \in L_{\mathcal{F}}$.
- a set $L_{\mathcal{R}}$ of relation symbols, together with a nonzero natural numbers n_R for each $R \in L_{\mathcal{R}}$.

As above, the numbers n_f and n_R are called the *arities* of the corresponding symbols.

Given a language $\mathcal{L}$, an $\mathcal{L}$**-structure** is a structure $\mathcal{M}$ together with bijective maps $\rho^S_{\mathcal{M}} : L_S \to S$ for $S = \mathcal{C}, \mathcal{F}, \mathcal{R}$. The maps $\rho^{\mathcal{F}}_{\mathcal{M}}$ and $\rho^{\mathcal{R}}_{\mathcal{M}}$ are also assumed to preserve the arities.

Notation 1.5. We write $c^{\mathcal{M}}$, $f^{\mathcal{M}}$ and $R^{\mathcal{M}}$ instead of $\rho^{\mathcal{C}}_{\mathcal{M}}(c)$, $\rho^{\mathcal{F}}_{\mathcal{M}}(f)$ and $\rho^{\mathcal{R}}_{\mathcal{M}}(R)$ respectively.

Hence

- for each $c \in L_{\mathcal{C}}$, we have a constant $c^{\mathcal{M}} \in \mathcal{M}$,
- for each $f \in L_{\mathcal{F}}$, we have a basic function $f^{\mathcal{M}} : M^{n_f} \to M$, and
- for each $R \in L_{\mathcal{R}}$, we have a basic relation $R \subset M^{n_R}$.

We call $c^{\mathcal{M}}$, $f^{\mathcal{M}}$ and $R^{\mathcal{M}}$ the interpretations in $\mathcal{M}$ of the corresponding symbols.

Example 1.6.

(1) $\mathcal{L}_{\emptyset} = \emptyset$ the language of pure sets.
(2) $\mathcal{L}_g = \{*, e\}$ the language of groups.
(3) $\mathcal{L}_r = \{+, -, \times, 0, 1\}$ the language of rings.
(4) $\mathcal{L}_{or} = \mathcal{L}_r \cup \{<\}$ the language of ordered rings.

In each cases, you likely have your favourite structure(s).

Important.

(1) Even though we call $\mathcal{L}_g = \{*, e\}$ the language of groups, one should note that an $\mathcal{L}_g$-structure need not be a group. For example $(\mathbb{N}, +, 0)$ is an $\mathcal{L}_g$-structure although it is not a group. The same applies to the other examples above except of course for the first.

(2) To ease notation, often we will use the same letter for the symbol and its interpretation. For example, in $\mathcal{L}_r = \{+, -, \times, 0, 1\}$, we will write $\mathcal{R} = (\mathbb{R}, +, -, \times, 0, 1)$ instead of $\mathcal{R} = (\mathbb{R}, +^{\mathcal{R}}, -^{\mathcal{R}}, \times^{\mathcal{R}}, 0^{\mathcal{R}}, 1^{\mathcal{R}})$.

Let us define what it means for a map between two structures to be "structure preserving":

Definition 1.7. Let $\mathcal{L}$ be a language and $\mathcal{M}, \mathcal{N}$ be two $\mathcal{L}$-structures.

(1) An $\mathcal{L}$-**embedding** from $\mathcal{M}$ to $\mathcal{N}$ is a one-to-one function $\rho : M \to N$ such that[1]
- for each $c \in L_{\mathcal{C}}$,
$$\rho(c^{\mathcal{M}}) = c^{\mathcal{N}}$$
- for all $f \in L_{\mathcal{F}}$ and all $\overline{a} \in M^{n_f}$,
$$\rho(f^{\mathcal{M}}(\overline{a})) = f^{\mathcal{M}}(\rho(\overline{a}))$$
- for all $R \in L_{\mathcal{R}}$ and all $\overline{a} \in M^{n_R}$,
$$\overline{a} \in R^{\mathcal{M}} \iff \rho(\overline{a}) \in R^{\mathcal{N}}.$$

We write $\rho : \mathcal{M} \to \mathcal{N}$.

(2) An $\mathcal{L}$-embedding from $\mathcal{M}$ to $\mathcal{N}$ is an $\mathcal{L}$-**isomorphism** if it is onto. In this case we write $\mathcal{M} \cong \mathcal{N}$

(3) If $M \subseteq N$ and the inclusion map $i : M \to N$ is an $\mathcal{L}$-embedding, then we say that $\mathcal{M}$ is a **substructure** $\mathcal{N}$. In this case we write $\mathcal{M} \subseteq \mathcal{N}$. Notice that $\mathcal{M} \subseteq \mathcal{N}$ if and only if
- for each $c \in L_{\mathcal{C}}$, we have $c^{\mathcal{M}} = c^{\mathcal{N}}$,
- for each $f \in L_{\mathcal{F}}$, we have $f^{\mathcal{M}} = f^{\mathcal{N}}\restriction_{M^{n_f}}$,
- for any $R \in L_{\mathcal{R}}$, we have $R^{\mathcal{M}} = R^{\mathcal{N}} \cap M^{n_R}$.

Example 1.8.

(1) In $\mathcal{L}_g$, we have that $(\mathbb{N}, +, 0)$ is substructure of $(\mathbb{Z}, +, 0)$.

(2) Let $\rho : \mathbb{Z} \to \mathbb{R}$ be the function $\rho(x) = e^x$. Then ρ is an $\mathcal{L}_g$-embedding from $(\mathbb{Z}, +, 0)$ to $(\mathbb{R}, \times, 1)$.

1.2. **Syntax.** So far we have set aside the fact that we would also like to have available certain axioms to help us characterize some structures. By our experience (for example with groups) these axioms are properties that one is able to write down using the symbol in our language together with other logical symbols as well as variables. For example think of the group axiom $\forall x \exists y (x * y = e)$.

In this section, starting with a language $\mathcal{L}$ we will explain how to form syntactically correct expression using the symbols in $\mathcal{L}$. We will come back to axioms later. So throughout, we let $\mathcal{L}$ be a language. We also assume that we have available to us a fixed set of variables $V_{\mathcal{L}} = \{v_i : i \in \mathbb{N}\}$. We also assume we have available the equality symbol "=" and the following logical symbols

[1]If $\overline{a} = (a_1, \ldots, a_n)$ we use here the notation $\rho(\overline{a}) = (\rho(a_1), \ldots, \rho(a_n))$.

$$\begin{array}{ccccc} \neg & \wedge & \vee & \forall & \exists \\ \text{not} & \text{and} & \text{or} & \text{for all} & \text{there exists} \end{array}$$

We begin by describing which functions can be defined from the constant and function symbols in $\mathcal{L}$.

Definition 1.9. The set of **$\mathcal{L}$-terms** is the smallest set $\mathcal{T}$ such that

- if $c \in L_{\mathcal{C}}$ is a constant symbol, then $c \in \mathcal{T}$;
- if $v \in V_{\mathcal{L}}$ is a variable, then $v \in \mathcal{T}$;
- if $f \in L_{\mathcal{F}}$ is a function symbol and $t_1, \ldots, t_{n_f} \in \mathcal{T}$, then $f(t_1, \ldots, t_{n_f}) \in \mathcal{T}$.

Example 1.10. In the language $\mathcal{L}_g = \{*, e\}$ of groups, we have the $\mathcal{L}_g$-terms

$$e \qquad x * y \qquad x * e * y$$

Notice that we have used the usual math notation to write the above terms. Formally we should have written $*(x, y)$ and $*(*(x, e), y)$ instead of $x * y$ and $x * e * y$ respectively. Clearly, this abuse of notation is justified. Also notice that we will use other letters (such as $x, y, z,...$) to denote the variables. Hopefully it will always be clear from the context what is being meant.

Next, let us define the simplest kind of formulas

Definition 1.11. An **atomic $\mathcal{L}$-formula** is a string of symbols of the form

(1) $t_1 = t_2$, where t_1 and t_2 are $\mathcal{L}$-terms, or

(2) $R(t_1 \ldots, t_{n_R})$, where $R \in L_{\mathcal{R}}$ and $t_1 \ldots, t_{n_R}$ are $\mathcal{L}$-terms.

In the language of $\mathcal{L}_{or} = \mathcal{L}_r \cup \{<\}$ of ordered rings,

$$x < y \qquad x - y < 0 \qquad x \times y < x + y$$

are atomic $\mathcal{L}$-formula. Again, for example, we use $x < y$ instead of $<(x, y)$.

Finally, by using the logical symbols we can recursively define all formulas

Definition 1.12. The set of **$\mathcal{L}$-formulas** is the smallest $\mathcal{F}$ such that

(1) if ϕ is an atomic $\mathcal{L}$-formula, then $\phi \in \mathcal{F}$,

(2) if $\phi \in \mathcal{F}$, then $\neg\phi \in \mathcal{F}$,

(3) if $\phi, \psi \in \mathcal{F}$, then $\phi \wedge \psi \in \mathcal{F}$ and $\phi \vee \psi \in \mathcal{F}$,

(4) if $\phi \in \mathcal{F}$ and $v \in V_{\mathcal{L}}$ is a variable, then $\exists v\, \phi \in \mathcal{F}$ and $\forall v\, \phi \in \mathcal{F}$.

Remark 1.13. It is worth noting that there are several logical abbreviations that one can (and we will) use. For example $\forall v\phi$ can be seen as an abbreviation for $\neg\exists v\, \neg\phi$. Similarly $\phi \vee \psi$ is an abreviation for $\neg(\neg\phi \wedge \neg\psi)$. The notations $(\phi \rightarrow \psi)$ and $(\phi \leftrightarrow \psi)$ are abbreviations for $(\neg\phi \vee \psi)$ and $(\phi \rightarrow \psi) \vee (\phi \rightarrow \psi)$ respectively.

Let us look at the example of group.

Example 1.14. We can "now" write down the axioms of group theory in the language $\mathcal{L}_g = \{*, e\}$ and it will not be hard to see that they are formulas as defined above.

$$\forall x \, (e * x = x * e = x)$$
$$\forall x \forall y \forall z \, ((x * y) * z = x * (y * z))$$
$$\forall x \exists y \, (x * y = y * x = e)$$

We can also look at other formulas such as

$$\forall y \, (x * y = y * x)$$

which "expresses that x is a central element".

In the above example, there is one noticeable difference between the three axioms and the fourth formula: in the latter there is one variable x that is not in the scope of (or not *bound* by) any quantifier $\exists x$ and $\forall x$.

Definition 1.15. Suppose ϕ is an $\mathcal{L}$-formula. We say that a variable x is **free** in ϕ, it does not appear in the scope of a quantifier $\exists x$ and $\forall x$. We also say that ϕ is an $\mathcal{L}$**-sentence** it has *no* free variable.

As we have seen above, the $\mathcal{L}$-sentences are the so-called mathematical "axioms". We will explore these formulas in more details in the second lecture.

Notation 1.16. We will write $\phi(x_1, \dots, x_n)$ to highlight that the free variables of a formula ϕ are *from* the set $\{x_1, \dots, x_n\}$. So in the previous example, we would write $\phi(x) := \forall y \, (x * y = y * x)$. We can also write $\phi(y) := \forall x \, (e * x = x * e = x)$.

So far, formulas are simply string of symbols that have been put together in some coherent way. We end this section by formalizing the idea that the formula $\phi(x) := \forall y \, (x * y = y * x)$ "expresses that x is a central element". We first need to describe how to interpret $\mathcal{L}$-terms in a structure.

Definition 1.17. Let $\mathcal{M}$ be an $\mathcal{L}$-structure and $t(x_1, \dots, x_n)$ be an $\mathcal{L}$-terms. The **interpretation of** $t(\overline{x})$ **in** $\mathcal{M}$ is the function $t^{\mathcal{M}} : M^n \to M$ define inductively as

(1) If t is a constant symbol $c \in L_{\mathcal{C}}$, then for any $\overline{a} \in M^n$ we set
$$t^{\mathcal{M}}(\overline{a}) = c^{\mathcal{M}}.$$
(2) If t is a variable symbol $v_i \in V_{\mathcal{L}}$, then for any $\overline{a} \in M^n$ we set
$$t^{\mathcal{M}}(\overline{a}) = a_i.$$
(3) if t is $f(t_1, \dots, t_{n_f})$, where $f \in L_{\mathcal{F}}$ is a function symbol and $t_1(\overline{x}), \dots, t_{n_f}(\overline{x})$ are $\mathcal{L}$-terms, then
$$t^{\mathcal{M}}(\overline{a}) = f^{\mathcal{M}}(t_1^{\mathcal{M}}(\overline{a}), \dots, t_{n_f}^{\mathcal{M}}(\overline{a}))$$

In Example 1.10 above, if $t(x, y)$ is the $\mathcal{L}_g$-term $t(x, y) = x * y$, then in any $\mathcal{L}_g$-structure t is simply function corresponding to pairwise multiplication. We are ready to give the definition of satisfaction in an $\mathcal{L}$-structure.

Definition 1.18. Let $\mathcal{M}$ be an $\mathcal{L}$-structure and $\phi(x_1, \ldots, x_n)$ an $\mathcal{L}$-formula. For $\bar{a} \in M^n$ we inductively define

$$\mathcal{M} \models \phi(\bar{a}),$$

which reads $\phi(\bar{a})$ **is true in** $\mathcal{M}$ as follows:

(1) If $\phi(\bar{x}) := (t_1 = t_2)$ where $t_1(\bar{x}), t_2(\bar{x})$ are $\mathcal{L}$-terms, then $\mathcal{M} \models \phi(\bar{a})$ means that $t_1^{\mathcal{M}}(\bar{a}) = t_2^{\mathcal{M}}(\bar{a})$.

(2) If $\phi(\bar{x}) := R(t_1, \ldots, t_{n_R})$, where $R \in L_{\mathfrak{R}}$ is a relation symbol and $t_1(\bar{x}), \ldots, t_{n_R}(\bar{x})$ are $\mathcal{L}$-terms, then $\mathcal{M} \models \phi(\bar{a})$ means that $(t_1^{\mathcal{M}}(\bar{a}), \ldots, t_{n_R}^{\mathcal{M}}(\bar{a})) \in R^{\mathcal{M}}$.

(3) If $\phi := \neg\psi$ for some $\mathcal{L}$-formula $\psi(\bar{x})$, then $\mathcal{M} \models \phi(\bar{a})$ means that $\mathcal{M} \not\models \psi(\bar{a})$.

(4) If $\phi := \psi \wedge \theta$ for some $\mathcal{L}$-formulas $\psi(\bar{x})$ and $\theta(\bar{x})$, then $\mathcal{M} \models \phi(\bar{a})$ means that $\mathcal{M} \models \psi(\bar{a})$ and $\mathcal{M} \models \theta(\bar{a})$.

(5) If $\phi := \psi \vee \theta$ for some $\mathcal{L}$-formulas $\psi(\bar{x})$ and $\theta(\bar{x})$, then $\mathcal{M} \models \phi(\bar{a})$ means that $\mathcal{M} \models \psi(\bar{a})$ or $\mathcal{M} \models \theta(\bar{a})$.

(6) If $\phi := \exists v\ \psi$ for some $\mathcal{L}$-formulas $\psi(\bar{x}, v)$, then $\mathcal{M} \models \phi(\bar{a})$ means that there is $b \in M$ such that $\mathcal{M} \models \psi(\bar{a}, b)$.

(7) If $\phi := \forall v\ \psi$ for some $\mathcal{L}$-formulas $\psi(\bar{x}, v)$, then $\mathcal{M} \models \phi(\bar{a})$ means that $\mathcal{M} \models \psi(\bar{a}, b)$ for all $b \in M$.

If $\mathcal{M} \models \phi(\bar{a})$ we also say that $\mathcal{M}$ *satisfies* $\phi(\bar{a})$ or that $\bar{a}$ *satisfies* $\phi(\bar{x})$ *in* $\mathcal{M}$.

Going back to the formula $\phi(x) := \forall y\ (x * y = y * x)$, if we consider the $\mathcal{L}_g$-structure $\mathscr{R} = (\mathbb{R}, +, 0)$, then for any $r \in \mathbb{R}$ we have $\mathscr{R} \models \phi(r)$ (this holds of all Abelian group). So $\mathbb{R} = \{r \in \mathbb{R} : \mathscr{R} \models \phi(r)\}$. More generally, if $\mathcal{G} = (G, *, e)$ is a group (so the three sentences in Example 1.14 is true in $\mathcal{G}$), then we see that $\phi(x)$ indeed define its center

$$Z(\mathcal{G}) = \{g \in G : \mathcal{G} \models \phi(g)\}.$$

It follows that $Z(\mathcal{G})$ is an example of a definable set, which is one of the most important notion in modern Model theory.

Definition 1.19. Let $\mathcal{M}$ be an $\mathcal{L}$-structure. A **definable set** $Y \subseteq M^n$ is a set of the form

$$Y = \{\bar{a} \in M^n : \mathcal{M} \models \phi(\bar{a})\}$$

where $\phi(\bar{x})$ is an $\mathcal{L}$-formula with free variables $\bar{x} = (x_1, \ldots, x_n)$.