# INTRODUCTION TO MODEL THEORY WITH APPLICATIONS

RONNIE NAGLOO

## 4. Proof of the Compactness Theorem and Quantifier elimination

In this lecture we give a proof of the Compactness Theorem using ultraproducts. We begin by introducing filters and ultrafilters and show how they can be used to build new structures from the direct product of a family of structures. This will go through defining the ultraproduct of the family of structures. We then look at quantifier elimination, which is a property that holds for some theories and which greatly helps with the problem of understanding the definable sets.

4.1. **Ultraproduct.** Throughout, $I$ will be a nonempty set and we denote by $\mathscr{P}(I)$ the power set of $I$, that is the set of all subsets of $I$.

**Definition 4.1.** A subset $\mathcal{F} \subseteq \mathscr{P}(I)$ is said to be a (proper) **filter** on $I$ if

- $I \in \mathcal{F}$ and $\emptyset \notin \mathcal{F}$.
- If $A, B \in \mathcal{F}$, then $A \cap B \in \mathcal{F}$.
- If $A \in \mathcal{F}$ and $A \subseteq B \subseteq I$, then $B \in \mathcal{F}$.

Let us look at some examples.

**Example 4.2.**

(1) The set $\mathcal{F} = \{X \subseteq I : I \setminus X \text{ is finite}\}$ is a filter on $I$ and is called the *Fréchet filter*.
(2) More generally, let $\kappa$ be an infinite cardinal such that $|I| \geq \kappa$. Then the set $\mathcal{F} = \{X \subseteq I : |I \setminus X| < \kappa\}$ is a filter on $I$.
(3) In the case $I = \mathbb{R}$, then the set $\mathcal{F} = \{X \subseteq \mathbb{R} : \mathbb{R} \setminus X \text{ has Lebesgue measure } 0\}$ is a filter on $\mathbb{R}$

As we can see, a filter gives a notion of largeness for subset of $I$.

**Definition 4.3.** An **ultrafilter** on $I$ is a maximal filter, i.e., a filter that is not properly contained in any other filter on $I$. Let $x \in I$. A **principal filter** $\mathcal{F}_x$ on $I$ the filter defined as $\mathcal{F}_x = \{X \subseteq I : x \in X\}$.

**Proposition 4.4.** *Let $x \in I$, then $\mathcal{F}_x$ is an ultrafilter.*

*Proof.* Assume for contradiction that $\mathcal{F}_x$ is not maximal. Let $\mathcal{U}$ be a filter on $I$ such that $\mathcal{F}_x \subset \mathcal{U}$. So there is $X \in \mathcal{U} \setminus \mathcal{F}_x$. By the definition of $\mathcal{F}_x$, it must be that $x \notin X$. But then since $\{x\} \in \mathcal{F}_x \subset \mathcal{U}$ and $X \in \mathcal{U}$ it follows that $\emptyset = \{x\} \cap X \in \mathcal{U}$. A contradiction. $\square$

Let $\mathcal{F}$ be any filter on $I$. Using Zorn's Lemma, one can show that the set $S = \{\mathcal{F}' : \mathcal{F} \subset \mathcal{F}' \text{ and } \mathcal{F}' \text{ is a filter on I}\}$ has a maximal element $\mathcal{U}$. So we can always extend $\mathcal{F}$ to an ultrafilter $\mathcal{U}$ on $I$. If we assume for example that $\mathcal{F}$ is the Fréchet filter on $I$, then $\mathcal{U}$ will be non-principal since for any $x \in I$ the cofinite set $I \setminus \{x\} \in \mathcal{F} \subseteq \mathcal{U}$. The following is a useful characterization of ultrafilters.

**Lemma 4.5.** *A filter $\mathcal{U}$ is an ultrafilter if and only if for every $X \subseteq I$, either*

$$X \in \mathcal{U} \quad \textit{or} \quad I \setminus X \in \mathcal{U}.$$

*Proof.* $\Rightarrow$ Assume $\mathcal{U}$ is an ultrafilter and let $X \subseteq I$. Suppose that $X \notin \mathcal{U}$. It is not hard to see that the set

$$\mathcal{F} = \{Y \subseteq I : A \setminus X \subseteq Y \text{ for some } A \in \mathcal{U}\}$$

is a filter on $I$ that contains $\mathcal{U}$. Since $\mathcal{U}$ is an ultrafilter $\mathcal{F} = \mathcal{U}$. By construction $I \setminus X \in \mathcal{F}$.
$\Leftarrow$ Let $\mathcal{U}$ be a filter and assume for all $X \subseteq I$, either $X \in \mathcal{U} \quad$ or $\quad I \setminus X \in \mathcal{U}$. Let $\mathcal{V}$ be a filter on $I$ so that $\mathcal{U} \subset \mathcal{V}$. So there is $Y \in \mathcal{V} \setminus \mathcal{U}$. Since $Y \notin \mathcal{U}$, by assumption $I \setminus Y \in \mathcal{U} \subset \mathcal{F}$. But it follows that $\emptyset = Y \cap I \setminus Y \in \mathcal{F}$ a contradiction. Hence $\mathcal{U}$ is maximal. $\square$

So given a set $X \subseteq I$, an ultrafilter is a decision about which of $X$ or its complement is large. We are now ready to bring back model theory

4.2. **Proof of Compactness.** Let $\mathcal{L}$ be a language. Assume that we have a family $\{\mathcal{M}_i : i \in I\}$ of $\mathcal{L}$-structures. Recall that the direct product $\prod_{i \in I} \mathcal{M}_i$ of the $\mathcal{M}_i$s is the set

$$\prod_{i \in I} \mathcal{M}_i = \{f : I \to \bigcup_{i \in I} \mathcal{M}_i : f(i) \in \mathcal{M}_i \text{ for all } i \in I\}.$$

Let $\mathcal{U}$ be an ultrafilter on $I$. We define a relation $\sim$ on $\prod_{i \in I} \mathcal{M}_i$ as follows:

$$f \sim g \quad \text{if and only if} \quad \{i \in I : f(i) = g(i)\} \in \mathcal{U}.$$

**Proposition 4.6.** *The relation $\sim$ is an equivalence relation on $\prod_{i \in I} \mathcal{M}_i$.*

*Proof.* Let $f, g, h \in \prod_{i \in I} \mathcal{M}_i$. Clearly reflexivity and symmetry holds. Let us prove transitivity: assume $f \sim g$ and $g \sim h$. Then it follows that $X = \{i \in I : f(i) = g(i)\} \cap \{i \in I : g(i) = h(i)\} \in \mathcal{U}$. It is not hard to see that the set $\{i \in I : f(i) = h(i)\}$ contains $X$. Hence since $\mathcal{U}$ is a filter, we have that $\{i \in I : f(i) = h(i)\} \in \mathcal{U}$. So $f \sim h$. $\square$

For $f \in \prod_{i \in I} \mathcal{M}_i$ we will denote by $\underline{f}$ the equivalence class of $f$ under $\sim$. Using this equivalence relation, we can construct a new $\mathcal{L}$-structure.

**Definition 4.7.** Let $\mathcal{L}$ be a language and $\{\mathcal{M}_i : i \in I\}$ a family of $\mathcal{L}$-structures. The **ultraproduct** of $\{\mathcal{M}_i : i \in I\}$ is the $\mathcal{L}$-structure $\mathcal{M}$ defined as follows[1]:

[1] You will show in the problem session that the definition does not depend on the choice of representatives.

- The universe is the set
$$M = \{\underline{f} : f \in \prod_{i\in I} \mathcal{M}_i\}.$$
- If $c \in L_{\mathcal{C}}$ is a constant symbol, then
$$c^{\mathcal{M}} = \underline{f}$$
where $f \in \prod_{i\in I} \mathcal{M}_i$ is the function such that $f(i) = c^{\mathcal{M}_i}$.
- If $f \in L_{\mathcal{F}}$ is a function symbol and $\underline{g_1}, \ldots, \underline{g_{n_f}} \in M$, then we define
$$f^{\mathcal{M}}(\underline{g_1}, \ldots, \underline{g_{n_f}}) = \underline{g}$$
where $g \in \prod_{i\in I} \mathcal{M}_i$ is the function such that $g(i) = f^{\mathcal{M}_i}(g_1(i), \ldots, g_{n_f}(i))$.
- If $R \in L_{\mathcal{R}}$ is a relation symbol and $\underline{g_1}, \ldots, \underline{g_{n_R}} \in M$, then
$$(\underline{g_1}, \ldots, \underline{g_{n_R}}) \in R^{\mathcal{M}} \text{ if and only if } \{i \in I : (g_1(i), \ldots, g_{n_R}(i)) \in R^{\mathcal{M}_i}\} \in \mathcal{U}$$

We will usually write $\prod \mathcal{M}_i/\mathcal{U}$ for the ultraproduct $\mathcal{M}$ above.

**Theorem 4.8** (Loś' Theorem)**.** *Let $\mathcal{U}$ be an ultrafilter. Assume $\{\mathcal{M}_i : i \in I\}$ is a family of $\mathcal{L}$-structures and let us write $\mathcal{M}$ for the ultraproduct $\prod \mathcal{M}_i/\mathcal{U}$. Then*

(1) *Let $t(x_1, \ldots, x_n)$ be an $\mathcal{L}$-term and $f_1, \ldots, f_n \in \prod_{i\in I} \mathcal{M}_i$. Then*
$$t^{\mathcal{M}}(\underline{f_1}, \ldots, \underline{f_n}) = \underline{g}$$
*where $g \in \prod_{i\in I} \mathcal{M}_i$ is the function such that $g(i) = t^{\mathcal{M}_i}(f_1(i), \ldots, f_n(i))$.*

(2) *Let $\phi(x_1, \ldots, x_n)$ be an $\mathcal{L}$-formula and $f_1, \ldots, f_n \in \prod_{i\in I} \mathcal{M}_i$. Then*
$$\mathcal{M} \models \phi(\underline{f_1}, \ldots, \underline{f_n}) \iff \{i \in I : \mathcal{M}_i \models \phi(f_1(i), \ldots, f_n(i))\} \in \mathcal{U}$$

*Proof.* (1) It is not too hard to see that this follows using induction on the complexity of terms. We only give the details for the case $t$ is $f(t_1, \ldots, t_{n_f})$ with $f \in L_{\mathcal{F}}$ a function symbol and $t_1, \ldots, t_{n_f}$ $\mathcal{L}$-terms for which the result already holds. Then
$$t^{\mathcal{M}}(\underline{f_1}, \ldots, \underline{f_n}) = f^{\mathcal{M}}\left(t_1^{\mathcal{M}}(\underline{f_1}, \ldots, \underline{f_n}), \ldots, t_{n_f}^{\mathcal{M}}(\underline{f_1}, \ldots, \underline{f_n})\right).$$
By induction
$$t^{\mathcal{M}}(\underline{f_1}, \ldots, \underline{f_n}) = f^{\mathcal{M}}(\underline{g_1}, \ldots, \underline{g_{n_f}})$$
where $g_k(i) = t_k^{\mathcal{M}_i}(f_1(i), \ldots, f_n(i))$ for $k = 1 \ldots n_f$. By definition
$$f^{\mathcal{M}}(\underline{g_1}, \ldots, \underline{g_{n_f}}) = \underline{g}$$

where $g(i) = f^{\mathcal{M}_i}(g_1(i), \ldots, g_{n_f}(i))$. However, notice that

$$\begin{aligned} g(i) &= f^{\mathcal{M}_i}\left(t_1^{\mathcal{M}_i}(f_1(i), \ldots, f_n(i)), \ldots, t_{n_f}^{\mathcal{M}_i}(f_1(i), \ldots, f_n(i))\right) \\ &= t^{\mathcal{M}_i}(f_1(i), \ldots, f_n(i)). \end{aligned}$$

So the result follows.

(2) Using (1) it is not hard to prove the result using induction on the complexity of $\phi(x_1, \ldots, x_n)$. We leave it to the reader to do the atomic steps, the inductive steps of $\wedge$ and $\vee$ (using the fact that ultrafilters are closed under intersections and unions). We highlight that the assumption that $\mathcal{U}$ is an ultrafilter (as oppose to just a filter) is required in the case when $\phi(\overline{x})$ is of the form $\neg\psi(\overline{x})$ and the result already holds for $\psi(\overline{x})$. Indeed

$$\begin{aligned} \mathcal{M} \models \phi(\underline{f_1}, \ldots, \underline{f_n}) &\iff \mathcal{M} \not\models \psi(\underline{f_1}, \ldots, \underline{f_n}) \\ &\iff \{i \in I : \mathcal{M}_i \models \psi(f_1(i), \ldots, f_n(i))\} \notin \mathcal{U} \\ &\iff \{i \in I : \mathcal{M}_i \models \neg\psi(f_1(i), \ldots, f_n(i))\} \in \mathcal{U}. \end{aligned}$$

The last equivalence follows using Lemma 4.5 since $\mathcal{U}$ is an ultrafilter. We are hence left (using the usual argument about $\forall$) to consider the case when $\phi(\overline{x})$ is of the form $\exists y \psi(\overline{x}, y)$ and the result holds for $\psi(\overline{x}, y)$. First, if $\mathcal{M} \models \phi(\underline{f_1}, \ldots, \underline{f_n})$ then there is $\underline{g} \in M$ such that $\mathcal{M} \models \psi(\underline{f_1}, \ldots, \underline{f_n}, \underline{g})$. Using the induction hypositehsis

$$X := \{i \in I : \mathcal{M}_i \models \psi(f_1(i), \ldots, f_n(i), g(i))\} \in \mathcal{U}.$$

But the set $\{i \in I : \mathcal{M}_i \models \exists y \psi(f_1(i), \ldots, f_n(i), y)\}$ contains $X$. Since $\mathcal{U}$ is an ultrafilter, the result follows. On the other hand if $Y = \{i \in I : \mathcal{M}_i \models \exists y \psi(f_1(i), \ldots, f_n(i), y)\} \in \mathcal{U}$, then we can define $g \in \prod_{i \in I} \mathcal{M}_i$ such that for each $i \in Y$,

$$\mathcal{M}_i \models \psi(f_1(i), \ldots, f_n(i), g(i)).$$

Since $\{i \in I : \mathcal{M}_i \models \psi(f_1(i), \ldots, f_n(i), g(i))\}$ contains $Y$ it must be in $\mathcal{U}$. By induction $\mathcal{M} \models \psi(\underline{f_1}, \ldots, \underline{f_n}, \underline{g})$ and hence $\mathcal{M} \models \phi(\underline{f_1}, \ldots, \underline{f_n})$. □

We are now ready to prove the Compactness Theorem.

**Theorem 3.5** (Compactness Theorem)**.** *Suppose $T$ is an $\mathcal{L}$-theory. $T$ is consistent if and only if every finite subset of $T$ is consistent.*

*Proof.* We only need to prove the right to left assertion. So assume that every finite subset of $T$ is consistent. Let $I$ be the set of all finite subsets of $T$. By assumption, we have a family $\{\mathcal{M}_i : i \in I\}$ of $\mathcal{L}$-structures such that $\mathcal{M}_i \models i$ for each $i \in I$. We build a filter on $I$ as follows (keep in mind that elements of $I$ are subsets of T): first we build a set $F$ of subsets of $I$ that does not contain $\emptyset$, contains $I$ and is closed under intersection. Indeed we can take $F = \{X_i : i \in I\}$, where for each $i \in I$, $X_i \subseteq I$ is the set of all finite subsets of $T$ containing all the elements of $i$. Notice that $I = X_\emptyset$ and $X_i \cap X_j = X_{i \cup j}$. Next, all we need to do is closed under containment by letting $\mathcal{F} = \{X \subseteq I : X_i \subseteq X \text{ for some } i \in I\}$. It follows that $\mathcal{F}$ is a filter on $I$.

Let $\mathcal{U}$ be an ultrafilter on $I$ extending $\mathcal{F}$. We show that the ultraproduct $\mathcal{M} := \prod \mathcal{M}_i/\mathcal{U}$ is a model of $T$. Let $\phi \in T$ and consider the set $X_{\{\phi\}}$ of finite subsets containing $\phi$. If $i \in X_{\{\phi\}}$, then since $\mathcal{M}_i \models i$ and $\phi \in i$, we get that $\mathcal{M}_i \models \phi$. So $X_{\{\phi\}} \subseteq \{i \in I : \mathcal{M}_i \models \phi\}$. Since $\mathcal{U}$ is an ultrafilter and $X_{\{\phi\}} \in \mathcal{U}$, we get that $\{i \in I : \mathcal{M}_i \models \phi\} \in \mathcal{U}$. By Theorem 4.8 (Łoś' Theorem), $\mathcal{M} \models \phi$ and so $T$ is consistent. □

4.3. **Quantifier Elimination.** As already mentioned in the second lecture, for a fixed theory $T$ a major goal of model theory is to study all definable sets in some/any model of $T$. This of course would be hopeless unless one could identify classes of structures where there are some control over the definable sets. In model theory, this leads the distinction between "tame" and "wild" structures or theories. We now discuss one notion of tameness, namely quantifier elimination. Throughout we take $\mathcal{L}$ to be a language and $T$ be a consistent $\mathcal{L}$-theory.

**Definition 4.9.** We say $T$ has **quantifier elimination** (QE) if every $\mathcal{L}$-formula is equivalent modulo $T$ to a quantifier free $\mathcal{L}$-formula. More precisely, for any $\mathcal{L}$-formula $\phi(\overline{x})$ there is a *quantifier free* $\mathcal{L}$-formula $\psi(\overline{x})$ such that

$$T \models \forall \overline{x}(\phi(\overline{x}) \leftrightarrow \psi(\overline{x}))$$

**Example 4.10.** Consider $\mathcal{R} = (\mathbb{R}, +, -, \times, 0, 1, <)$ in $\mathcal{L}_r$. Let

$$\phi(a,b,c) := a \neq 0 \wedge \exists x(a^2x + bx + c = 0)$$

and

$$\psi(a,b,c) := b^2 - 4ac \geq 0.$$

Then $Th(\mathcal{R}) \models \forall a \forall b \forall c(\phi(a,b,c) \leftrightarrow \psi(a,b,c))$.

For theories with quantifier elimination, the definable sets are defined using "simple" formulas. Here are some other consequences of QE.

**Proposition 4.11.** *Suppose $T$ has quantifier elimination and let $\mathcal{M}$ and $\mathcal{N}$ be models of $T$. Then*

(1) *If $\mathcal{A}$ is a substructure of both $\mathcal{M}$ and $\mathcal{N}$, then for any $\mathcal{L}$-formula $\phi(\overline{x})$ and every $\overline{a} \in A^n$ we have that*

$$\mathcal{M} \models \phi(\overline{a}) \iff \mathcal{N} \models \phi(\overline{a}).$$

(2) *$T$ is model complete. That is if $\mathcal{M} \subseteq \mathcal{N}$ then $\mathcal{M} \preceq \mathcal{N}$.*

*Proof.* (1) This follows since we already showed in Proposition 2.13 that if $\mathcal{A} \subseteq \mathcal{M}$ and $\psi(\overline{x})$ is quantifier free, then for $\overline{a} \in A^n$

$$\mathcal{A} \models \psi(\overline{a}) \iff \mathcal{M} \models \psi(\overline{a}).$$

(2) This simply follows from definition after applying (1) to the case $\mathcal{A} = \mathcal{M}$. □

We end this lecture by giving a criterion to check whether a given theory has QE. We will not give a proof (cf. [1, Corollary 3.16]). By a *literal* we mean an atomic or negated atomic formula.

**Proposition 4.12** (Test for QE). *Suppose the $\mathcal{L}$-theory $T$ satisfies the following condition:*

> $(*)$ *For any $\mathcal{M}, \mathcal{N} \models T$ with a common substructure $\mathcal{A}$, and any formula $\phi(x)$ which is a conjunction of $\mathcal{L}_A$-literals in one free variable, if there is $m \in M$ such that $\mathcal{M} \models \phi(m)$ then there is $n \in N$ such that $\mathcal{N} \models \phi(n)$.*

*Then $T$ has quantifier elimination.*

*Steps in the proof.* The proof goes by first showing that a converse of Proposition 4.11 (1) holds: if $\mathcal{A}$ is a substructure of both $\mathcal{M}$ and $\mathcal{N}$ and $\phi(\overline{y})$ is an $\mathcal{L}$-formula such that for every $\overline{a} \in A^n$ we have that $\mathcal{M} \models \phi(\overline{a})$ if and only if $\mathcal{N} \models \phi(\overline{a})$, then $\phi(\overline{y})$ is equivalent modulo $T$ to a quantifier free formula. This is [1, Theorem 3.1.4]. On the other hand, one can also shows that to prove QE, it is enough to restrict to the case of formulas $\phi(\overline{y})$ of the form $\exists x \psi(\overline{y}, x)$ for a quantifier free $\psi(\overline{y}, x)$. This is [1, Lemma 3.1.5]. So in summary, this gives [1, Corollary 3.16]: whenever $\mathcal{A}$ is a substructure of $\mathcal{M}, \mathcal{N} \models T$ and a quantifier free formula $\psi(\overline{y}, x)$, if $\overline{a} \in A^n$ then $\mathcal{M} \models \exists x \psi(\overline{a}, x)$ if and only if $\mathcal{N} \models \exists x \psi(\overline{a}, x)$, then $T$ has quantifier elimination. Noting that $\psi(\overline{a}, x)$ "is" a quantifier free $\mathcal{L}_A$-formula, we almost get $(*)$. It is not hard to show using De Morgan's laws about how negation interacts with conjunctions and disjunctions that it is enough to assume $\psi(\overline{a}, x)$ is a conjunction of $\mathcal{L}_A$-literals. □

**Example 4.13.** We show that $T_\infty$, the theory of infinite sets has QE. Recall that we work in the language $\mathcal{L}_\emptyset = \{\emptyset\}$. Suppose $\mathcal{M}$ and $\mathcal{N}$ are two infinite set that both contain a nonempty set $\mathcal{A}$. It is not hard to see that the only atomic $\mathcal{L}_A$-formulas in one variable $x$ are of the form

$$x = x \quad \text{or} \quad x = a \quad \text{for some } a \in A.$$

The formula $x = x$ is satisfied by any element in any model of $T_\infty$ and the formula $x \neq x$ is not satisfied by any elements of any models of $T_\infty$. So it is not hard to see that we only have to consider the case of an $\mathcal{L}_A$-formula $\phi(x)$ of the form

$$\bigwedge_{i=1}^{p} (x = a_i) \wedge \bigwedge_{i=1}^{q} (x \neq b_i)$$

with $a_1, \ldots, a_p, b_1, \ldots, b_q \in A$. Assume for some $m \in M$, we have that $\mathcal{M} \models \phi(m)$. Then, if at least one of the $x = a_i$ appear in $\phi(x)$, it follows that $m = a_i \in A$ for all $1 \leq i \leq p$ and also $m \neq b_i$ for all $1 \leq i \leq q$. But then $m \in N$ and we have that $\mathcal{N} \models \phi(m)$. On the other hand, if no $x = a_i$ appears in $\phi(x)$, then

$$\phi(x) := \bigwedge_{i=1}^{q} (x \neq b_i)$$

and since $N$ is infinite we can find $n \in N$ so that $n \neq b_i$ for all $1 \leq i \leq q$. Hence $\mathcal{N} \models \phi(n)$.

In the last two lectures we will see examples of more interesting theories that have QE.

## References

[1] D. Marker, Model theory, Graduate Texts in Mathematics, vol. 217, Springer-Verlag, New York, 2002. An introduction.