LECTURE 25: ULTRAPRODUCTS

CALEB STANFORD

First, recall that given a collection of sets and an ultrafilter on the index set, we formed an ultraproduct of those sets. It is important to think of the ultraproduct as a *set-theoretic* construction rather than a *model-theoretic* construction, in the sense that it is a product of sets rather than a product of structures. I.e., if X_i are sets for $i = 1, 2, 3, \dots$, then $\prod X_i/\mathcal{U}$ is another set. The set we use does not depend on what constant, function, and relation symbols may exist and have interpretations in X_i. (There are of course profound model-theoretic consequences of this, but the underlying construction is a way of turning a collection of sets into a new set, and doesn't make use of any notions from model theory!)

We are interested in the particular case where the index set is $\mathbb{N}$ and where there is a set X such that $X_i = X$ for all i. Then $\prod X_i/\mathcal{U}$ is written $X^{\mathbb{N}}/\mathcal{U}$, and is called the **ultrapower of $\boldsymbol{X}$ by $\boldsymbol{\mathcal{U}}$**. From now on, we will consider the ultrafilter to be a fixed nonprincipal ultrafilter, and will just consider the **ultrapower of $\boldsymbol{X}$** to be the ultrapower by this fixed ultrafilter. It doesn't matter which one we pick, in the sense that none of our results will require anything from $\mathcal{U}$ beyond its nonprincipality.

The ultrapower has two important properties. The first of these is the Transfer Principle. The second is $\aleph_0$-saturation.

1. The Transfer Principle

Let $\mathcal{L}$ be a language, X a set, and $X_{\mathcal{L}}$ an $\mathcal{L}$-structure on X. Let $\mathcal{U}$ be a nonprincipal ultrafilter on $X^{\mathbb{N}}$. Let $Y = X^{\mathbb{N}}/\mathcal{U}$. Łoś's theorem tells us that we can interpret Y as an $\mathcal{L}$-structure $Y_{\mathcal{L}}$ in a natural way, and that for any $\mathcal{L}$-sentence φ:

$$\begin{aligned} Y_{\mathcal{L}} \vDash \varphi &\iff \{i \;:\; X_i \vDash \varphi\} \in \mathcal{U} \\ &\iff \{i \;:\; X_{\mathcal{L}} \vDash \varphi\} \in \mathcal{U} \\ &\iff X_{\mathcal{L}} \vdash \varphi. \end{aligned}$$

In other words, in the Ultrapower case, Łoś's theorem collapses down to the simple statement that $Y_{\mathcal{L}}$ is elementarily equivalent to $X_{\mathcal{L}}$.

But wait! Remember that Y itself is simply a set that is a function of the set X. The above will be true for ANY $\mathcal{L}$-structure that we put on X; we will always get a corresponding $\mathcal{L}$-structure on Y that satisfies the elementary equivalence.

So, let $\mathcal{L}$ be the set of ALL functions, relations, and constants on X. This includes a symbol for every function and relation and constant you may be interested in, but it also includes symbols for everything else: many functions and relations and constants which you never thought to consider, and many which can't be written down explicitly. Then $Y_{\mathcal{L}}$ and $X_{\mathcal{L}}$ are elemantarily equivalent. This is known as the Transfer Principle. We restate it below, introducing some new notation that is often used.

Theorem 1. (Transfer Principle) Let X be a set. Let ${}^*X$ (what we have been calling Y) be an ultrapower of X. Let $c_1, c_2, c_3, \dots$ be constant elements of X, let $R_1, R_2, R_3, \dots$ be relations on X, and let $F_1, F_2, F_3, \dots$ be functions on X. Let ${}^*c_1, {}^*c_2, \dots$, ${}^*R_1, {}^*R_2, \dots$, ${}^*F_1, {}^*F_2, \dots$ be the corresponding constants, relations, and functions on ${}^*X$. Let φ be a first-order sentence over the language $\{c_i, R_i, F_i\}$, and let ${}^*\varphi$ be the corresponding first-order sentence over the language $\{{}^*c_i, {}^*R_i, {}^*F_i\}$. Then $X \vDash \varphi$ if and only if ${}^*X \vDash {}^*\varphi$.

Proof. From Łoś's theorem, as described above. □

Here are some examples.

Example 1. Let $X = \mathbb{R}$. In $\mathbb{R}$ it is true that $\forall x : |x| \geq 0$. Therefore, in ${}^*\mathbb{R}$, $\forall x : {}^*|x| \geq 0$.

Date: 2015-04-10.

Example 2. Let X be the set of finite binary strings. Let $\circ$ denote concatenation. In X it is the case that $\forall x \forall y \forall z : (x \circ y) \circ z = x \circ (y \circ z)$. Therefore, in ${}^*X$, $\forall x \forall y \forall z : (x \,{}^*{\circ} y) \,{}^*{\circ} z = x \,{}^*{\circ} (y \,{}^*{\circ} z)$.

First-order logic generally forces us to quantify over the entire set X and not over a subset of X. But a subset of X is actually an element of our language now (a unary relation), so we can formalize e.g. the statement "x is in the Cantor set" as Cx, where C denotes this unary relation. Thus the Transfer Principle also gives us things like:

Example 3. (open set) A set A is open in $\mathbb{R}$ if and only if

$$\forall x \in A \; \exists \epsilon \in (0, \infty) \; \forall y \in \mathbb{R} \;:\; (|x - y| < \epsilon \rightarrow y \in A)$$

Therefore, for any open set A of $\mathbb{R}$ we have

$$\forall x \in {}^*A \; \exists \epsilon \in {}^*(0, \infty) \; \forall y \in {}^*\mathbb{R} \;:\; ({}^*|x - y| \,{}^*{<} \epsilon \rightarrow y \in {}^*A).$$

Since X is embedded in ${}^*X$, in the future, we will generally think of X as being a subset of ${}^*X$. We will then generally omit the * before constants, functions, and relations, with the exception that we distinguish between a set $A \subseteq X$ and its corresponding set ${}^*A$. Just to illustrate why this is not ambiguous:

- For constants, if we say "let $c \in X$", this also implies that $c \in {}^*X$, as we are thinking of X as a subset of ${}^*X$. We are also allowed to use c in any transfer principle arguments.
- If on the other hand we say "let $c \in {}^*X$", it is perfectly clear what we mean, but we are not then allowed to apply the transfer principle to sentences involving c.
- For relations, if we say "let R be a binary relation on X", it is clear what we mean. Then R extends naturally to a binary relation on ${}^*X$, so we can compare both things in X and things in ${}^*X$ using R. Note that it is important here that ${}^*R$ agrees with R on X.
- If we say "let R be a binary relation on ${}^*X$", then R is also well-defined on all of ${}^*X$ as well as X, but we can't apply the transfer principle to a sentence involving R.
- Similarly, functions f defined on X are assumed to be extended automatically in the natural way to ${}^*X$, but functions defined on ${}^*X$ originally cannot be dealt with by the transfer principle.
- Finally, when we fix a subset A of X, this is to be thought of as different than a unary relation on X. A unary relation would extend naturally to ${}^*X$ via the transfer principle, but A is considered to be a fixed subset of X which is in turn a subset of ${}^*X$; A is the same subset of ${}^*X$ as it is of X. If we want to consider the corresponding (different) subset of ${}^*X$, we will write ${}^*A$.

In summary, we are able to keep things straight if we just remember whether the function or relation was defined originally on X, or on ${}^*X$.

2. $\aleph_0$-Saturation

Up to this point, it has seemed that ${}^*X$ is just a bigger version of X; elementarily equivalent, in fact. So what use is there in defining it? If it is so similar to X, why not just use X?

The answer is that we also have a lot more than just was in X, and we can exploit that. For instance, the archimedian property (which is not first-order) holds in $\mathbb{R}$ but not in ${}^*\mathbb{R}$, and this turns out to allow us to define calculus of $\mathbb{R}$ using infinitesimal elements of ${}^*\mathbb{R}$.

Theorem 2. ($\aleph_0$-saturation) Let $\mathcal{T} = (\varphi_1, \varphi_2, \ldots)$ be a countable set of formulas in free variables $u_1, u_2, \ldots, u_k$. Suppose that every finite subset Σ of $\mathcal{T}$ has a solution in X^k. Then $\mathcal{T}$ has a solution in $({}^*X)^k$.

Proof. Define infinite sequences $u_1, u_2, \ldots, u_k \in {}^*X$ by

$$\begin{aligned} u_1 &:= ({}_1u_1, {}_2u_1, {}_3u_1, \ldots) \\ u_2 &:= ({}_1u_2, {}_2u_2, {}_3u_2, \ldots) \\ &\cdots \\ u_k &:= ({}_1u_3, {}_2u_3, {}_3u_3, \ldots) \end{aligned}$$

such that ${}_iu_1, {}_iu_2, \ldots, {}_iu_k$ is a solution in X^k for the finite set of formulas $\Sigma_i := (\varphi_1, \varphi_2, \ldots, \varphi_i)$. Then observe that for each φ_j, there are only finitely many i such that φ_j is not in Σ_i, and hence there are only finitely many i for which ${}_iu_1, {}_iu_2, \ldots, {}_iu_k$ is not a solution to φ_j. Therefore, the set

$$\{i \;:\; X_i \vDash \varphi_j({}_iu_1, {}_iu_2, \ldots, {}_iu_k)\}$$

is cofinite, therefore being a member of our ultrafilter. This implies by the definition of satisfaction in ${}^*X$ and by Łoś's theorem that

$$ {}^*X \vDash \varphi_j(u_1, u_2, \dots, u_k) $$

This is true for any φ_j, so ${}^*X \vDash \mathcal{T}(u_1, u_2, \dots, u_k)$. □

Example 4. In $\mathbb{R}$, let $\mathcal{T}$ be the set of formulas

$$ \{x < 1, x < 1/2, x < 1/3, \dots\} $$

Then there is an element $x \in {}^*\mathbb{R}$ satisfying the above. This is called an **infinitesimal**.

Example 5. In $\mathbb{N}$, let $\mathcal{T}$ be the set of formulas

$$ \{1 \mid x, 2 \mid x, 3 \mid x, \dots\} $$

Then there is an element $x \in {}^*\mathbb{N}$ satisfying all the above, i.e. there is a hypterinteger divisible by every integer.

Example 6. In $\mathbb{R}$, take $\mathcal{T}$ to be

$$ \{x > 1, x > 2, x > 3, \dots\} \cup \{y > x, y > x^2, y > x^3, \dots\} $$

The result is two hyperreal numbers, x and y, both infinite, but such that y is much bigger than x.

3. Applications

3.1. **Infinitely many primes.** In class, we proved that there are infinitely many primes. The idea is to form a hyperinteger divisible by every standard prime number, then to add one. The resulting hyperinteger must be divisible by a hyperprime, but it isn't divisible by any standard primes. As a lemma, a subset $S \subseteq X$ is finite if and only if ${}^*S = S$. So the fact that there are hyperprimes which are not prime means that the set of primes is infinite.

3.2. **ZFC.** (Background: ZFC is the first-order theory of set theory. The language of ZFC consists only of the binary relation $\in$.)

Assume ZFC is consistent. Then there is a model of ZFC, call it M. Let ω be the element of M corresponding to the natural numbers, the first uncountable ordinal. Consider the set of formulas

$$ \mathcal{T} = \{x \in \omega, x \neq 1, x \neq 2, x \neq 3, \dots\} $$

Every finite subset of $\mathcal{T}$ is satisfiable in M. Therefore, by $\aleph_0$-saturation, $\mathcal{T}$ is satisfiable in ${}^*M$. That is, the set of natural numbers ω in ${}^*M$ actually contains something which is not a natural number.

Why is this a problem? Well, it is true in ZFC that every nonzero natural number has a predecessor, so x has a predecessor x_1, which has a predecessor x_2, and so on. None of these are equal to any of $1, 2, 3, \dots$, else x would be equal to one of $1, 2, 3, \dots$ just by applying taking a few successors of x_i.

So $x_1, x_2, x_3, \dots$ is an infinite decreasing chain of "natural numbers". Worse, it is an infinite decreasing chain of ordinals; each is contained in the next! And taking the set

$$ S = \{x_1, x_2, x_3, \dots\} $$

we find that S contains no element disjoint from itself, which violates one of the axioms of set theory (axiom of regularity).

What gives? Well, we can externally write down S in our meta-theory, and claim it is a set, but the bizarre model of ZFC ${}^*M$ does not know about S. Nor does this model of ZFC have any way to form the infinite decreasing chain $x_1, x_2, x_3, \dots$ and compile it into a single list. Just to illustrate this, notice that the usual definition of a countable list is a function from ω to a set; yet ω is much larger in ${}^*M$ than it is in our standard understanding of ZFC.