Part III Logic in Lent Term 2017
Chapter 2: Model Theory

Thomas Forster

February 13, 2017

Contents

I will try to adhere to the habit of using $\mathfrak{FRAKTUR}$ font[1] for letters denoting structures and the corresponding upper-case Roman letter for the carrier set.

A reminder of two bits of jargon: an **expansion** of a structure $\mathcal{B}$ is a structure with the same carrier set and more gadgets. e.g. the rationals as a field are an expansion of the rationals as an additive group. The converse relation is a **reduction**: the rationals as an additive group are a reduction of the rationals as a field.

Definition 1 *A sentence is "universal" iff it is in PNF and its quantifier prefix consists entirely of universal quantifiers. By a natural extension we say a theory is "universal" iff, once you put its axioms into PNF, their quantifier prefixes consist entirely of universal quantifiers. We define "universal-existential" sentences and theories[2] similarly as theories all of whose axioms, when in PNF have a block of universal quantifiers followed by a block of existential quantifiers, and so on.*

Definition 2 *The* **diagram** $D_{\mathfrak{M}}$ *of a structure* $\mathfrak{M}$ *is the theory obtained by expanding* $\mathfrak{M}$ *by giving names to every* $m \in M$*, and collecting all true atomic assertions about them.*

Lemma 1 *For any consistent theory* T *and any model* $\mathfrak{M}$ *of* $T_\forall$*, the set of universal consequences of* T*, the theory* $T \cup D_{\mathfrak{M}}$ *is consistent.*

Proof:

Let $\mathfrak{M}$ be a model of $T_\forall$, with carrier set M. Add to $\mathcal{L}(T)$ names for every member of M. Add to T all the (quantifier-free) assertions about the new constants that $\mathfrak{M}$ believes to be true. This theory is $T \cup D_{\mathfrak{M}}$. We want this theory to be consistent. How might it not be? Well, if it isn't, there must be an inconsistency to be deduced from a conjunction ψ of finitely many of the new axioms. This rogue ψ mentions finitely many of the new constants. We have a proof of $\neg\psi$ from T. T knows nothing about these new constants, so clearly we must have a UG proof of $(\forall\vec{x})\neg\psi$. But this would contradict the fact that $\mathfrak{M}$ satisfies every universal consequence of T. ■

Theorem 1

T is universal iff every substructure of a model of T is a model of T.

Proof:

L → R is easy. We prove only the hard direction.

Suppose that T is a theory such that every substructure of a model of T is also a model of T. Let $\mathfrak{M}$ be an arbitrary model of $T_\forall$. We will show that it must be a model of T. We know already from the foregoing that the theory $T \cup D_{\mathfrak{M}}$ is consistent, and so it must have a model—$\mathfrak{M}^*$, say. $\mathfrak{M}^*$ is a model of

[1] Often called 'Gothic' by the ignorant. The Goths had a different alphabet (and a different language!) not just a different font.

[2] PTJ calls such theories "inductive" in his lectures.

T, and $\mathfrak{M}$ is a submodel of $\mathfrak{M}^*$ and therefore (by assumption on T) a model of T—as desired.

But all we knew about $\mathfrak{M}$ was that it was a model of the universal consequences of T. So any old $\mathfrak{M}$ that was a model of the universal consequences of T is a model of T. So T is axiomatised by its universal consequences. ■

There are lots of theorems with this flavour: "The class of models of T is closed under operation burble iff T has an axiomatisation satisfying syntactic condition blah"

Definition 3 *The* **Skolem Hull** *of a structure $\mathfrak{M}$ is what one obtains as follows. For each sentence $\exists x \phi(x)$ true in $\mathfrak{M}$ pick the first such x. For each sentence $\forall x \exists y \psi(x, y)$ true in $\mathfrak{M}$ let f_ψ send each x to the first y such that $\psi(x, y)$. Close under these operations. The result is the Skolem Hull.*

Of course we can generalise this by requiring that the Skolem hull should contain some specified things to start with. It's another recursive datatype.

Definition 4 *An embedding $i : \mathfrak{M} \to \mathfrak{N}$ is* Γ-**elementary** *iff, for all $\phi \in \Gamma$, $\mathfrak{M} \models \phi(x_1 \ldots x_n)$ iff $\mathfrak{N} \models \phi(i(x_1) \ldots i(x_n))$*

If i is Γ-elementary where Γ is the set of all formulæ we say i is just plain elementary.

Remember elementary equivalence?!

Some examples:

End-extensions are elementary for formulæ in which all quantifiers are restricted!

Inclusion embedding from the rationals-as-an-ordered set into the reals (ditto) is elementary. Not as an ordered field.

The simplest application of the idea of elementary embeddings known to me is the usual proof that classical monadic predicate logic is decidable.

Remark 1 *Classical monadic predicate logic is decidable.*

Proof: Suppose we have a monadic formula Φ, and let $\mathfrak{M}$ be a model. Φ contains only finitely many monadic predicate letters, say $\psi_1 \ldots \psi_i$. Let $\mathcal{L}_\Psi$ be the language with these monadic predicates and no other predicates or function letters. The various ψ divide the carrier set M into 2^i classes in the obvious way: a typical class looks like $\{x : \psi_1(x) \wedge \neg\psi_2(x) \wedge \ldots\}$. Any selection set for this partition gives a submodel of $\mathfrak{M}$ for which (we prove by induction on the recursive datatype of $\mathcal{L}_\Psi$-formulæ) the inclusion embedding is $\mathcal{L}_\Psi$-elementary. The submodel is finite and it will only take a finite time to check the truth value in it of any formula. ■

One more definition before we get stuck in.

Definition 5 *$\mathcal{I} = \langle I, \leq_\mathcal{I} \rangle$ is a* **set of indiscernibles** *for a model $\mathfrak{M}$ for a language $\mathcal{L}$ iff for all $\phi \in \mathcal{L}$, if ϕ is a formula with n free variables in it then for all distinct n-tuples $\vec{x}$ and $\vec{y}$ from $\mathcal{I}$* **taken in increasing order** *we have $\mathfrak{M} \models \phi(\vec{x}) \longleftrightarrow \phi(\vec{y})$.*

The idea of a set of indiscernibles is due to Ramsey, in *that paper*. He doesn't identify the idea or give it a name, but it's there.

1 Products

1.1 Direct Products and Reduced Products

If $\{\mathcal{A}_i : i \in I\}$ is a family of structures, we define the product

$$\prod_{i \in I} \mathcal{A}_i$$

to be the structure whose carrier set is the set of all functions f defined on the index set I such that $(\forall i \in I)(f(i) \in A_i)$ and the relations of the language are interpreted *"pointwise"*: the product believes $f\ R\ g$ iff $(\forall i \in I)(f(i)\ R\ g(i))$.

The $\{\mathcal{A}_i : i \in I\}$ are said to be the *factors* of the product $\prod\limits_{i \in I} \mathcal{A}_i$.

For this operation to make sense it is of course necessary that all the $\mathcal{A}_i$ should have the same signature!

Products are nice in various ways.

Definition 6 *Let Γ be a class of formulæ. Products* **preserve** Γ *if whenever $\prod\limits_{i \in I} A_i$ is a product of a family $\{\mathcal{A}_i : i \in I\}$ and $\phi \in \Gamma$ then*

$$\prod_{i \in I} \mathcal{A}_i \models \phi \text{ as long as } (\forall i \in I)(\mathcal{A}_i \models \phi).$$

By definition of product, products preserve atomic formulæ. Clearly they also preserve conjunctions of anything they preserve, and similarly universal quantifications over things they preserve.

What about more complex formulæ? You know that products preserve equational theories (a product of rings is a ring, after all). They also preserve Horn formulæ

Definition 7 .

A Horn clause	*is a disjunction of atomics and negatomics of which at most one is atomic.*
A Horn property	*is a property captured by a [closure of a] Horn expression;*
A Horn theory	*is a theory all of whose axioms are universal closures of (conjunctions of) Horn clauses.*

Remark 2 *Products preserve Universal Horn formulæ*

Proof:

Suppose every factor $\mathcal{A}_n$ believes $(\forall \vec{x})((\bigwedge_{i<j} \phi_i(\vec{x})) \to \phi_j(\vec{x}))$, where all the ϕ are atomic. We want to show that the product believes it too. So let

$\vec{f} = f_1 \ldots f_k$ be a tuple of things in the product satisfying the antecedent. That is to say, for each factor $\mathcal{A}_n$, we have $\mathcal{A}_n \models \phi_i(f_1(n), f_2(n) \ldots f_k(n))$ for each $i < j$. But then every $\mathcal{A}_n$ believes $\phi_j(f_1(n), f_2(n) \ldots f_k(n))$ so the product believes $\phi_j(f_1, f_2 \ldots f_k)$ as desired. ■

In particular an arbitrary product of transitive relations is a transitive relation. [This is a good point of departure]. An arbitrary product of posets is a poset (being a poset is horn) but an arbitrary product of tosets is not reliably a toset because the totality condition (trichotomy, connexity) is not Horn.

This illustrates how products do not always preserve formulæ containing $\vee$ or $\neg$. This suggests that remark 2 is best possible. (We won't prove it) How so? If ϕ is preserved, then the product will fail to satisfy it if even *one* of the factors does not satisfy it (c.f. Genesis [19:23-33] where not even one righteous man is enough to save the city) but all the rest do. (The product is not righteous unless *all* its factors are). In these circumstances the product $\models \neg\phi$ but it is not the case that all the factors $\models \neg\phi$. As for $\vee$, if ϕ and ψ are preserved, it can happen that $\phi \vee \psi$ is not, as follows. If half the factors satisfy ϕ and half satisfy ψ, then they all satisfy $\psi \vee \phi$. Now the product will satisfy $\psi \vee \psi$ iff it satisfies one of them. But in order to satisfy one of them, that one must be true at *all* the factors, and by hypothesis it is not. Something similar happens with the existential quantifier.

1.1.1 Reduced products

Quick revision [not written out here]. We assume ultrafilters, why they exist etc. This is all in [2], or—if you begrudge me the royalties—is covered in my lecture notes for Part II Set Theory and Logic in Michaelmas 2016, available on `www.dpmms.cam.ac.uk/~tf/partiilectures2016.pdf`

Worth making the point that the collection of filters on X is a complete poset and the collection of proper filters is merely a chain-complete poset.

Given a filter F over the index set, we can define $f \sim_F g$ on elements of the product if $\{i \in I : f(i) = g(i)\} \in F$. Then we *either* take this $\sim_F$ to be the interpretation of '=' in the new product we are defining, keeping the elements of the carrier set of the new product the same as the elements of the old *or* we take the elements of the new structure to be equivalence classes of functions under $\sim$. These we will write $[g]_{\sim_F}$ or $[g]_F$ or even $[g]$ if there is no ambiguity.

This new object is denoted by the following expression:

$$(\prod_{i \in I} \mathcal{A}_i)/F$$

Similarly we have to revise our interpretation of atomic formulæ so that

$$(\prod_{i \in I} \mathcal{A}_i)/F \models \phi(f_1, \ldots f_n) \ \text{ iff } \ \{i : \phi(f_1(i), \ldots f_n(i)\} \in F.$$

REMARK 3 *$\sim_F$ is a congruence relation for all the operations that the product inherits from the factors.*

Can't do any harm to write out a proof. [Not lectured but supplied for the notes]

Let H be an operation, of arity h, and let $\vec{f}$ and $\vec{g}$ be two h-tuples in the product, with $f_i \sim_F g_i$ for each $i \leq h$. That is to say: for each $i \leq h$, $\{n : f_i(n) = g_i(n)\} \in F$. Since h is finite, we can conclude that $\{n : \bigwedge_{i \leq h} f_i(n) = g_i(n)\} \in F$.

We want $H(\vec{f}) \sim_F H(\vec{g})$. That is to say we desire that $\{n : H(f_1(n) \cdots f_h(n)) = H(g_1(n) \cdots g_h(n))\} \in F$. But we know (by our assumption that $f_i \sim_F g_i$ for each $i \leq h$) that $\bigwedge_{i \leq k}(f_i(n) = g_i(n))$ holds for an F-large set of n, so if H is given the same tuple of arguments it can hardly help but give back the same value.

■

It may be worth bearing in mind that to a certain extent the choice between thinking of elements of the carrier set of the reduced product as the set of $\sim_F$-equivalence classes or thinking of them as the functions is a real one and might matter. I have proceeded here on the basis that the carrier set is the set of $\sim_F$-equivalence classes because that seems more natural. However, in principle there are set-existence issues involved in thinking of a product this way—how do we know that the $\sim_F$-equivalence classes are sets?—so we want to keep alive in our minds the possibility of doing things the second way. This will matter when we come to consider reduced products where the factor structures are proper classes (= have carrier sets that are proper classes). This happens in the extensions of ZF(C) with large cardinal axioms (specifically measurable cardinals). In practice these issues are usually swept under the carpet; this is a safe strategy only because it is in fact possible to sort things out properly! There is of course also the possibility of picking representatives from the equivalence classes, possibly by means of AC.

(For those of a philosophical turn of mind, there is an interesting contrast here with the case of quotient structures like, say, integers mod p. I have the impression that, on the whole, mathematicians do not think of integers-mod-p as sets of integers, nor as integers equipped with a nonstandard equality relation, but rather think of them as objects of a new kind. These reflections may have significance despite not really belonging to the study of *mathematics*: the study of *how we think about mathematics* is important too.

The reason for proceeding from products to reduced products was to complicate the construction and hope to get more things preserved. In fact nothing exciting happens (we still have the same trouble with $\vee$ and $\neg$—think: tosets) unless the filter we use is ultra. Then everything comes right.

1.2 Ultraproducts and Łoś's theorem

Theorem 2 *(Łoś's theorem)*

Let $\mathcal{U}$ be an ultrafilter on I. For all expressions $\phi(f, g, h \ldots)$,

$$(\prod_{i \in I} \mathcal{A}_i)/\mathcal{U} \models \phi(f, g, h \ldots) \text{ iff } \{i : \mathcal{A}_i \models \phi(f(i), g(i), h(i) \ldots)\} \in \mathcal{U}.$$

Proof: We do this by structural induction on the rectype of formulæ. For atomic formulæ it is immediate from the definitions.

[wouldn't hurt to write out the details for the fainthearted!]

As we would expect, the only hard work comes with $\neg$ and $\vee$, though $\exists$ merits comment as well.

Disjunction

Suppose we know that $(\prod_{i\in I}\mathcal{A}_i)/\mathcal{U} \models \phi$ iff $\{i : \mathcal{A}_i \models \phi\} \in \mathcal{U}$ and $(\prod_{i\in I}\mathcal{A}_i)/\mathcal{U} \models \psi$ iff $\{i : \mathcal{A}_i \models \psi\} \in \mathcal{U}$. We want to show $(\prod_{i\in I}\mathcal{A}_i)/\mathcal{U} \models (\phi \vee \psi)$ iff $\{i : \mathcal{A}_i \models \phi \vee \psi\} \in \mathcal{U}$.

The steps in the following manipulation will be reversible. Suppose

$$(\prod_{i\in I}\mathcal{A}_i)/\mathcal{U} \models \phi \vee \psi.$$

Then

$$(\prod_{i\in I}\mathcal{A}_i)/\mathcal{U} \models \phi \text{ or } (\prod_{i\in I}\mathcal{A}_i)/\mathcal{U} \models \psi.$$

By induction hypothesis, this is equivalent to

$$\{i : \mathcal{A}_i \models \phi\} \in \mathcal{U} \text{ or } \{i : \mathcal{A}_i \models \psi\} \in \mathcal{U},$$

both of which imply

$$\{i : \mathcal{A}_i \models \phi \vee \psi\} \in \mathcal{U}.$$

$\{i : \mathcal{A}_i \models \phi \vee \psi\}$ is $\{i : \mathcal{A}_i \models \phi\} \cup \{i : \mathcal{A}_i \models \psi\}$. Now we exploit the fact that $\mathcal{U}$ is ultra: for all A and B it contains $A \cup B$ iff it contains at least one of A and B, which enables us to reverse the last implication.

Negation

We assume $(\prod_{i\in I}\mathcal{A}_i)/\mathcal{U} \models \phi$ iff $\{i : \mathcal{A}_i \models \phi\} \in \mathcal{U}$ and wish to infer $(\prod_{i\in I}\mathcal{A}_i)/\mathcal{U} \models \neg\phi$ iff $\{i : \mathcal{A}_i \models \neg\phi\} \in \mathcal{U}$.

Suppose $(\prod_{i\in I}\mathcal{A}_i)/\mathcal{U} \models \neg\phi$. That is to say,

$$(\prod_{i\in I}\mathcal{A}_i)/\mathcal{U} \not\models \phi.$$

By induction hypothesis this is equivalent to

$$\{i : \mathcal{A}_i \models \phi\} \notin \mathcal{U}.$$

But, since $\mathcal{U}$ is ultra, it must contain I' or $I \setminus I'$ for any $I' \subseteq I$, so this last line is equivalent to

$$\{i : \mathcal{A}_i \models \neg\phi\} \in \mathcal{U},$$

as desired.

Existential quantifier

The step for $\exists$ is also nontrivial:

$$(\prod_{i\in I}\mathcal{A}_i)/\mathcal{U} \models \exists x\phi$$

$$\exists f(\prod_{i\in I}\mathcal{A}_i)/\mathcal{U} \models \phi(f)$$

$$\exists f\{i \in I : \mathcal{A}_i \models \phi(f(i))\} \in \mathcal{U},$$

and here we use the axiom of choice to pick a witness at each factor

$$\{i \in I : \mathcal{A}_i \models \exists x\phi(x)\} \in \mathcal{U}.$$

■

You will notice that in the induction step for the existential quantifier you use the axiom of choice to pick a witness from each factor, and this use of AC seems unavoidable. This might lead you to suppose that Łoś's theorem is actually equivalent to AC, but this seems not to be the case. Try it! I am endebted to Phil Freeman for drawing my attention to Paul Howard, Proc Am Math Soc Vol. 49, No. 2, Jun., 1975.

This has the incredibly useful corollary (which we shall not prove) that

Corollary 1 *A formula is equivalent to a first-order formula iff the class of its models is closed under elementary equivalence and taking ultraproducts.*

Theorem 2 enables us to show that a lot of things are not expressible in any first order language. Since, for example, an ultraproduct of finite p-groups (which are all simple) is not simple, it follows that the property of being a simple group is not capturable by a language in which you are allowed to quantify only over elements of the object in question.

Miniexercise: If the ultrafilter is principal ($\{J \subseteq I : i \in J\}$), then the ultraproduct is isomorphic to the ith factor. So principal ultrafilters are no use.

In contrast if the ultrafilter is nonprincipal you can make good use of the construction even if all the models you feed into it are the *same*.

Definition 8 *If all the factors are the same, the ultraproduct is called an* **ultrapower**, *and we write '$A^K/\mathcal{U}$' for the ultraproduct where there are $|K|$-many copies of A, where K is a set and $\mathcal{U}$ an ultrafilter on K.*

Not only are $\mathfrak{M}$ and the ultrapower $\mathfrak{M}^\kappa/\mathcal{U}$ elementarily equivalent by Łoś's theorem , we also have the following, of which we will make frequent use.

Lemma 2 *The embedding $i : \mathfrak{M} \hookrightarrow \mathfrak{M}^\kappa/\mathcal{U}$ defined by $\lambda m_{\mathfrak{M}}\lambda f_{\mathfrak{M}^\kappa/\mathcal{U}}.m$ is elementary.*

(This embedding i is just a typed version of the K combinator!)
Proof:

It will be sufficient to show that, for any $m \in \mathfrak{M}$, if there is an $x \in \mathfrak{M}^\kappa/\mathcal{U}$ such that $\mathfrak{M}^\kappa/\mathcal{U} \models \phi(x, i(m))$ then there is $x \in \mathfrak{M}$ s.t. $\mathfrak{M} \models \phi(x, m)$. Consider such an $x \in \mathfrak{M}^\kappa/\mathcal{U}$. It is the equivalence class of a family of functions which almost everywhere (in the sense of $\mathcal{U}$) are related to m by ϕ so—by Łoś's theorem —there must be something x in $\mathfrak{M}$ such that $\mathfrak{M} \models \phi(x, m)$. Then $i \mapsto x$ will do.

■

If you are doing Set Theory you will see the utility of this later in connection with measurable cardinals.

Ultraproducts enable us to give a particularly slick proof of the compactness theorem for predicate calculus.

Theorem 3 *(Compactness theorem for predicate logic)*
Every finitely satisfiable set of sentences of predicate calculus has a model.

Proof: Let Δ be a set of wffs that is finitely satisfiable. Let $\mathcal{S}$ be the set of finite subsets of Δ (elsewhere in these notes notated $\mathcal{P}_{<\aleph_0}(\Delta)$), and let $X_s = \{t \in \mathcal{S} : s \subseteq t\}$. Pick $\mathfrak{M}_s \models s$ for each $s \in \mathcal{S}$. Notice that $\{X_s : s \in \mathcal{S}\}$ generates a proper filter. Extend this to an ultrafilter $\mathcal{U}$ on $\mathcal{S}$. Then

$$(\prod_{s\in\mathcal{S}} \mathfrak{M}_s)/\mathcal{U} \models \Delta.$$

This is because, for any $\phi \in \Delta$, $X_{\{\phi\}}$ is one of the sets that generated the filter that was extended to $\mathcal{U}$. For any $s \in X_{\{\phi\}}$, $\mathfrak{M}_s \models \phi$, so $\{s : \mathfrak{M}_s \models \phi\} \in \mathcal{U}$.
■

Notice we are not making any assumption that the language is countable.

Notice the relation between Arrow's paradox and the nonexistence of non-principal ultrafilters on finite sets. Consider an ultraproduct of finitely many linear orders: it must be isomorphic to one of the quotients. This is Arrow's "dictatorship" condition.

Exercise 1 *Let $\{A_i : i \in \mathbb{N}\}$ be a family of finite structures, and $\mathcal{U}$ a non-principal ultrafilter on $\mathbb{N}$. Show that the ultraproduct is finite if there is a finite bound on the size of A_i and is of size $2^{\aleph_0}$ if every infinite subset of $\{A_i : i \in \mathbb{N}\}$ contains arbitrarily large elements.*

2 Infinitesimals

The effect of the ultraproduct construction is to add lots of things whose presence cannot be detected by finitistic first-order methods. Thus we can add infinitesimals to the reals. Hence Nonstandard Analysis: an ultraproduct of $\mathbb{R}$ (modulo a countably incomplete ultrafilter at least) is saturated (theorem 5) and therefore contains infinitesimals. This means we can reconstruct the seventeenth century theory of differentiation and integration!

3 Saturated Models

Review countable categoricity. See a property S so that any two models of a complete theory that are both S are not only elementarily equiv but iso.

There is a very beautiful theorem of Ryll-Nardzewski concerning countably categorical structures which i shall not prove, though i shall throw out a couple of hints.

Theorem 4 *Let $\mathfrak{M}$ be a countable structure. Then the following are equivalent*

- *For all $n \in \mathbb{N}$ $Aut(\mathfrak{M})$ has only finitely many orbits on n-tuples from $\mathfrak{M}$;*
- *$Th(\mathfrak{M})$ is countably categorical.*

Proof:

Sketch: One way we use a back-and-forth construction. The other way we use the omitting types theorem, theorem 9 below. ■

It was once an exercise on a PTJ example sheet. You might like to try to prove it.

Here is as good a place as any to introduce the idea of a saturated model. Informally a structure is saturated iff it realises as many types as possible.

Definition 9

A **type** *is a set of formulæ, typically all wioth the same number of frfee variables.*

If Σ is a type with free variables $\vec{x}$ we say that a tuple $\vec{a}$ (in a structure $\mathfrak{M}$) **realizes** *Σ if $\mathfrak{M} \models \sigma(\vec{a})$ for every $\sigma \in \Sigma$.*[3]

A type is finitely satisfiable if every finite subset of it can be realized.

*A model is $\aleph_1$-***saturated** *iff every finitely satisfiable countable type.*

We can use ultraproducts to prove the existence of saturated models.

We need one more definition:

Definition 10 *α-complete filter, α a cardinal. Means "an intersection ofg fewer than α things in the filter is also in the filter". Filters are automatically $\aleph_0$-complete. "Countably complete" always means "$\aleph_1$-complete".*

Theorem 5 *(The existence of saturated models)*

Let $\mathcal{L}$ be a countable language, and let $\mathcal{U}$ be an ultrafilter over an index set I, where $\mathcal{U}$ is not countably complete. Then for every family $\{A_i : i \in I\}$ the ultraproduct $(\prod_{i \in I} \mathcal{A}_i)/\mathcal{U}$ is $\aleph_1$-saturated.

What do we mean by $\aleph_1 -$ *saturated*

[3] Model theorists tend to use capital Greek letters for types (in this sense of 'type') and corresponding lower-case Greek letters for formulæ in them.

Proof:

We must show that for every countable set $\{f_i : i \in \mathbb{N}\}$ of elements of $(\prod_{i \in I} \mathcal{A}_i)/\mathcal{U}$ and every set $\Sigma(x)$ of formulæ from $\mathcal{L}$ (with countably many new constants $c_i \ldots$), if each finite subset of $\Sigma(x)$ is satisfiable in $(\prod_{i \in I} \mathcal{A}_i)/\mathcal{U}$ (with names for the f_i) then so is $\Sigma(x)$ itself.

Since $\mathcal{L}$ with the new constants is also a countable language it will be sufficient to prove it without the constants.

Suppose every finite subset of $\Sigma(x)$ is satisfiable in $(\prod_{i \in I} \mathcal{A}_i)/\mathcal{U}$. $\Sigma(x)$ is countable, so we can think of it as $\{\sigma_i : i \in \mathbb{N}\}$. Since $\mathcal{U}$ is countably incomplete, we find a $\subseteq$-descending ω sequence $\langle I_i : i \in \mathbb{N}\rangle$ of $\mathcal{U}$-large subsets of I whose intersection is empty.

Define a new sequence $\langle X_i : i \in \mathbb{N}\rangle$ by

$$X_0 =: I$$

and thereafter

$$X_n =: I_n \cap \{i \in I : \mathcal{A}_i \models (\exists x)(\bigwedge_{j<n} \sigma_j(x))\}.$$

$(\prod_{i \in I} \mathcal{A}_i)/\mathcal{U}$ satisfies every finite subset of Σ so, by Łoś's theorem ,

$$\{i \in I : \mathcal{A}_i \models (\exists x)(\bigwedge_{j<n} \sigma_j(x))\} \in \mathcal{U}.$$

This ensures that (i) each X_n is in $\mathcal{U}$, (ii) the X_i are nested and (iii) their intersection is empty. From (iii) it follows that for each $i \in I$ there is a last $n \in \mathbb{N}$ s.t. $i \in X_n$. Let this last n be $n(i)$. We are now going to construct an $f \in \prod_{i \in I} \mathcal{A}_i$ such that $[f]_{\mathcal{U}}$ realizes Σ. If $n(i) = 0$ then $f(i)$ can be anything. Otherwise set $f(i)$ to be any x such that $\mathcal{A}_i \models \bigwedge_{j<n(i)} \sigma_j(x)$.

■

Exercise 2 *Show that any two countably saturated countable elementarily equivalent structures are isomorphic.*

3.1 The Ehrenfeucht-Mostowski theorem

Ultraproducts contain lots of nonstandard funny stuff, but they don't obviously admit automorphisms. However we can use them to create models that do. The theorem of this section was proved in the 1950's by Ehrenfeucht and Mostowski, using methods of Ramsey theory and compactness, and we shall give that proof. We also give another proof due to Gaifman that uses ultrapowers.

We can create a nonstandard model of the reals by adding to our language a constant symbol c_α for each countable ordinal α, and—whenever $\alpha < \beta$—an axiom $c_\alpha < c_\beta$. By compactness this gives a consistent theory, so there is definitely going to be a nonstandard model of the reals containing a copy of the countable ordinals—or any other total order we want, come to think of it! However nothing in this construction will ensure that constants $\{c_\alpha : \alpha < \omega_1\}$ are embedded as a set of indiscernibles. The fact that this apparently much more difficult feat can be achieved is the content of

Theorem 6 *(Ehrenfeucht-Mostowski theorem)*
Let I be a total order, T a theory with infinite models and a formula $P()$ with one free variable s.t. T thinks that the extension of P is an infinite total order. Then T has a model $\mathfrak{M}$ in which I is embedded in (the interpretation of) P and in which every automorphism of I extends to an automorphism of $\mathfrak{M}$. Finally the copy of I in $\mathfrak{M}$ is a set of indiscernibles.

Notice that there is no suggestion that the copy of I in the model we build is a set of that model, or is in any way definable.

We give first an outline of the original proof, due to Ehrenfeucht and Mostowski.

Proof:
Add to the language of T names c_i for every element of I, and axioms to say the c_i are all distinct. Next we add axioms providing correct order information about the c_i. Let this theory be T^*. By compactness we know that this theory is consistent, since T has an infinite model, and believes that the domain of $<$ is infinite. That is to say, if T believes that the domain of $<$ is infinite, we can find a model of T in which I is embedded in the domain of $<$. This much is a straightforward application of completeness and compactness.

Now we add axioms saying that these constants are a set of indiscernibles; these axioms will say things like

$$\phi(c_i, c_j) \longleftrightarrow \phi(c_k, c_l)$$

whenever $i \leq_I j$ and $k \leq_I l$ and ϕ is a two-place formula in $\mathcal{L}(T)$ [analogously for higher arities] and there will be infinitely many of them. This gives us a theory T^I. We want to prove T^I consistent. The obvious thing to try is to show that every finite fragment of T^I is consistent. Let T' be one such finite fragment. It mentions only finitely many constants—c_1, c_3, c_4 and c_5, say—and it says that they form a set of indiscernibles for finitely many predicates—ϕ_1, ϕ_2, ϕ_3 and ϕ_4, for the sake of argument.

Now the task of proving this theory consistent is precisely the same task as proving consistent the theory T'' obtained from T' by replacing c_1, c_3, c_4 and c_5 by any other sequence of constants of length 4. So if we liked we could drop the names c_1, c_3, c_4 and c_5 and call them something noncommittal like a, b, c and

d. Once we have the model for the noncommittal version of T'' we can restore the c labels.

Now let $\mathfrak{M}$ be a model of T^*. A bit of notation: Let $[X]^m$ be $\{Y \subseteq X : |Y| = m\}$. These predicates—$\phi_1$, ϕ_2, ϕ_3 and ϕ_4, together with the order relation $<$ that T^* imposes on the c_i—divide up $[\{c_i : i \in I\}^{\mathfrak{M}}]^m$ (where m is the supremum of the arities of the ϕ's) into finitely many pieces. How do they do this? If ϕ_4 is of arity m then it obviously splits $[\{c_i : i \in I\}^{\mathfrak{M}}]^m$ into two bits. But what if $m = 3$ and ϕ_4 is of arity two? What is it to do with a triple from $\{c_i : i \in I\}^{\mathfrak{M}}$. Well, any such triple gives rise to three pairs, and we feed each pair into ϕ_4 in increasing order. So ϕ_3 splits $[\{c_i : i \in I\}^{\mathfrak{M}}]^m$ into eight pieces. So the number of pieces into which we split $[\{c_i : i \in I\}^{\mathfrak{M}}]^m$ is the product of the numbers of pieces mandated by each ϕ_i mentioned in T'. Call this partition Π.

Now $\{c_i : i \in I\}^{\mathfrak{M}}$ is infinite, so there must be a monochromatic set for Π of size 4, and we take its elements, read in increasing order, to be a, b, c and d.

Thus T' has a model.

Now we invoke compactness to conclude that T^I has a model. Any model of T^I has I embedded as a set of indiscernibles.

Then the model we desire is the Skolem hull of the indiscernibles. ∎

3.2 Gaifman's proof

We now give a second proof, which uses ultrapowers but no Ramsey Theory.

We start with some standard observations about direct limits. Given a directed family of structures with embeddings (satisfying commutation conditions) there is a well-defined notion of **direct limit** which you should know, but (in case you don't) …

Limits and Colimits: revision

A poset is **directed** if any two elements have a common upper bound.

What follows is a Part II Set Theory exercise.

Colimits

Let $\langle I, \leq_I \rangle$ be a directed poset and, for each $i \in I$, let A_i be a set and, for all $i \leq_I j$, let $\sigma_{i,j} : A_i \hookrightarrow A_j$ be an injection, and let the injections commute.

Show that there is a set A_I with, for each $i \in I$, an injection $\sigma_i : A_i \hookrightarrow A_I$ and the $\sigma_{i,j}$ commute with the σ_i.

Show also that A_I is minimal in the sense that if B is any set such that for each $i \in I$ there is an injection $\tau_i : A_i \hookrightarrow B$ and the τ_i commute with the $\sigma_{i,j}$, then there is a map $A_I \hookrightarrow B$.

Limits

Let $\langle I, \leq_I \rangle$ be a directed poset and, for each $i \in I$, let A_i be a set and, for all $i \leq_I j$, let $\sigma_{j,i} : A_j \twoheadrightarrow A_i$ be a surjection, and let the surjections commute.

Show that there is a set A_I with, for each $i \in I$, a surjection $\pi_i : A_I \twoheadrightarrow A_i$.

Show also that A_I is minimal in the sense that, if B is any set such that for each $i \in I$ there is a surjection $\tau_i : B \twoheadrightarrow A_i$ and the τ_i commute with the $\sigma_{i,j}$, then there is a map $B \twoheadrightarrow A_I$.

We will need the following important banality:

Remark 4 *Every structure for a first-order language is the colimit of its finitely generated substructures.*

We need this only in the [particularly simple] case of total orders, where it is obvious. After all the finitely generated substructures are just the finite suborderings. Every total ordering is the direct limit of the family of its finite suborderings equipped with the inclusion embedding!

Remark 5

1. *A direct limit of structures preserves* Π_2 *sentences;*
2. *A direct limit of an elementary family (one where the embeddings are elementary) preserves everything.*

The idea underlying the proof of Gaifman's is that one can recover any structure from the embedding relations between its finitely generated substructures: it's a direct limit of them (co-limit if you're a categorist).

Specifically if $\langle I, \leq_{\mathcal{I}} \rangle$ is an ordered set then it is the direct limit of its finite substructures where the embedding relations are the obvious inclusion embeddings. Remarkably, this banal fact is almost all we need!

3.2.1 The Construction

We start with an infinite model $\mathfrak{M}$ of T. We are going to create a directed family of elementary embeddings and iterated ultrapowers of $\mathfrak{M}$ indexed by the set of finite substructures of $\langle I, \leq_{\mathcal{I}} \rangle$, and the desired model will be a substructure of the direct limit, $\mathfrak{M}_\infty$. We will use the letters 's' and 't' to range over these finite substructures and we will notate the corresponding models $\mathfrak{M}_s$.

Let P be $\{x : P(x)\}^{\mathfrak{M}}$. We will assume that P has no last element in the sense of the ordering of P according to $\mathfrak{M}$. $\mathcal{U}$ will be an ultrafilter on P that contains all terminal segments of P. (So P had better not have a last element!)

Now to define the models in the family. $\mathfrak{M}_s$ will simply be the result of doing the ultrapower construction $|s|$ times to $\mathfrak{M}$, so that what $\mathfrak{M}_s$ actually *is* depends only on the *length* of s and in no way on what the members of s are. $\mathfrak{M}_\emptyset$ is just the $\mathfrak{M}$ we started with.

Now we have to define a family of embeddings and establish that they commute. We need to recall some notation:

> `last`(s) is the last member of s (remember s is thought of as an increasing sequence) and
>
> `butlast`(s) is s minus its last element.

We now define by recursion a family $\{I(s,t) : s \subseteq t \in I^{<\omega}\}$ of embeddings: $I(s,t)$ will be an elementary embedding from $\mathfrak{M}_s$ into M_t. The recursion needs two constructions:

1. K is the standard elementary embeddding by constant functions from a structure into its ultrapower as in lemma 2.

2. If i is an embedding from $\mathfrak{M}$ to $\mathfrak{N}$ then there is an injection from $\mathfrak{M}^\kappa/\mathcal{U}$ into $\mathfrak{N}^\kappa/\mathcal{U}$ "compose with i on the right". Perhaps a picture will help.

$$\begin{tikzcd}
\mathfrak{M}_s \arrow[r, "i"] \arrow[d, "K"'] & \mathfrak{M}_t \arrow[d, "K"] \\
(\mathfrak{M}_s)^P/\mathcal{U} \arrow[r, "{\lambda[f].[\lambda\alpha.i(f(\alpha))]}"'] & (\mathfrak{M}_t)^P/\mathcal{U}
\end{tikzcd}$$

Let us call this operation L, so that $L =: \lambda i.\lambda f. i \circ f$.[4]

Now we can give the recursive definition of $I(s,t)$ when $s \subseteq t$.

`If` $s = t$ `then` $I(s,t)$ is the identity `else`
`If` `last`(s) = `last`(t) `then` $I(s,t) =: L(I($`butlast`(s),`butlast`$(t)))$
`else` $I(s,t) =: K \circ I(s,$ `butlast`$(t))$.

Notice that $\lambda[f].[\lambda\alpha.i(f(\alpha))]$ sends "new stuff" in $\mathfrak{M}_s^P/\mathcal{U}$ (by which I mean $(\mathfrak{M}_s^P/\mathcal{U}) \setminus K``\mathfrak{M}_s$) to to "new stuff" in $\mathfrak{M}_t^P/\mathcal{U}$ (by which I mean $(\mathfrak{M}_t^P/\mathcal{U}) \setminus K``\mathfrak{M}_t$). This will be essential later.

To check that this system of models and embeddings is genuinely a directed system it remains only to show that the embeddings are elementary and that they commute.

K is elementary by lemma 2. L of an elementary embedding is elementary as follows. Suppose $i : \mathfrak{M}_s \hookrightarrow \mathfrak{M}_t$ is elementary, and that $\mathfrak{M}_s^P/\mathcal{U} \models \phi(f_1 \ldots f_n)$. That is to say, $\{p : \mathfrak{M}_s \models \phi(f_1(p) \ldots f_n(p)\} \in \mathcal{U}$. Now $i : \mathfrak{M}_s \hookrightarrow \mathfrak{M}_t$ is elementary so this is equivalent to $\{p : \mathfrak{M}_t \models \phi(i(f_1(p)) \ldots i(f_n(p))\} \in \mathcal{U}$ which is equivalent to $\mathfrak{M}_s^P/\mathcal{U} \models \phi((i \circ f_1) \ldots (i \circ f_n))$ as desired.

To check that the family is commutative it is sufficient to check that the representative diagram below is commutative.

[4] A Curry-Howard point. The constructor L explains why '$M \to N. \to (K \to M) \to (K \to N)$' is intuitionistically correct. It is also the embedding underlying the cardinal arithmétic banality that $\alpha \leq \beta. \to \alpha^\zeta \leq \beta^\zeta$.

$$\begin{tikzcd}
\mathfrak{M}_1 \arrow[r, "i"] \arrow[d, "K"'] & \mathfrak{M}_{1,2} \arrow[d, "K"] \\
\mathfrak{M}_{1,3} \arrow[r, "L(i)"] & \mathfrak{M}_{1,2,3}
\end{tikzcd}$$

That is to say that, for any i, $K \circ i = L(i) \circ K$.

- $K \circ i$ gives: $x \mapsto i(x) \mapsto \lambda p.i(x)$;
- $L(i) \circ K$ gives: $x \mapsto K(x) \mapsto L(i)(K(x)) = \lambda p.(i \circ p)(K(x))$
$$\begin{aligned} &= i \circ K(x) \\ &= \lambda p.(i \circ K(x))p \\ &= \lambda p.(i(K(x)p)) \\ &= \lambda p.i(x). \end{aligned}$$

This will show that all paths from $\mathfrak{M}_s$ to $\mathfrak{M}_t$ (and the number of such paths is presumably $(|t| - |s|)!$) correspond to the same injection.

3.2.2 Embedding I in the direct limit...

The point of this direct limit construction was to obtain a structure $\mathfrak{M}_\infty$ in which I was embedded. To achieve this we ensure that each finite subset $s \subset I$ is embedded in $\mathfrak{M}_s$ in such a way that the manifestations of the elements of I in the $\mathfrak{M}_s$ get stitched together properly. That means that inside $\mathfrak{M}_s$ we must be able to point to $|s|$ distinct **things**. We will find these **things**[5] by a recursive construction, and we will prove by induction on n that the construction works for s of length n. We can think of $\mathfrak{M}_s$ as a segmented structure: it has $|s|$ segments, and each new segment consists of the junk added by the ultraproduct construction applied to the object so far, and each segment contains a **thing**—each time we zap the model with our ultraproduct wand we add a new **thing**.

What is the $|t|$th **thing** in $\mathfrak{M}_t$ to be? The following train of tho'rt gives us a fix on what it must be, and tells us how we might find the $|t|$th thing by recursion on $|t|$. Suppose we know how to find the n **thing**s in $\mathfrak{M}_s$ when $|s| = n$, and let $|t| = n+1$. Now let t' be t with its *penultimate* element deleted (so we are assuming that $n \geq 2$). By induction hypothesis we know already what the last **thing** in $\mathfrak{M}_{t'}$ is. But we also know the embedding $I(t', t)$ from $\mathfrak{M}_{t'} \hookrightarrow \mathfrak{M}_t$. This tells us that the last **thing** in $\mathfrak{M}_t$ is that object to which $I(t', t)$ sends the last **thing** in $\mathfrak{M}_{t'}$. (We flagged this earlier.)

It's worth checking that we have got the base case right. This construction giving us the $n+1$th **thing** is guaranteed to work only as long as $\mathfrak{M}_s$ (where $|s| = n$) is already an ultrapower, since any embedding created by the recursion

[5] 'Things'?? I've got to call them *something*!!

expects its domain to be an ultrapower. But this is all right; $\mathfrak{M}_\emptyset$ is not an ultrapower (it is $\mathfrak{M}$) but then it isn't expected to have a **thing** in it.

So how do we decide what the first **thing** is? We know that as the various **thing**s get added they must form an increasing sequence according to the (extension of the) order (in $\mathfrak{M}$) to which I will belong. It would be helpful to ensure that the first **thing**—which we see for the first time in $\mathfrak{M}^I/\mathcal{U}$—is later than everything in $\mathfrak{M}$. This will be the case if (i) the total order in $\mathfrak{M}$ has a last element and (ii) the index set I is precisely the extension of the order in $\mathfrak{M}$ and the ultrafilter $\mathcal{U}$ contains all terminal segments of I.

So we let the first **thing** be an arbitrary object in $\mathfrak{M}^P\mathcal{U} \setminus \mathfrak{M}$.

A last thought: how do we identify the nth **thing** in $\mathfrak{M}_s$ with the nth **thing** in $\mathfrak{M}_t$ where $|t| > |s| \geq n$? Well, K is an embedding from $\mathfrak{M}_s$ into $\mathfrak{M}_{s\cup\{i\}}$ and obviously we want the nth **thing** in $\mathfrak{M}_{s\cup\{i\}}$ to be K of the nth **thing** in $\mathfrak{M}_s$. Repeat as necessary.

3.2.3 . . . as a set of indiscernibles

The **things** end up in $\mathfrak{M}_\infty$ as a subclass $\{c_i : i \in I\}$. We want to show that the c_i form a set of indiscernibles in $\mathfrak{M}_\infty$. Let $\vec{s}$ and $\vec{t}$ be two finite subsets of I (tho'rt of as increasing sequences). We want to show that $\phi(\vec{s})$ iff $\phi(\vec{t})$ (identifying each c_i with i for the moment). Now the embedding from $\mathfrak{M}_s$ into $\mathfrak{M}_\infty$ is elementary, so $\mathfrak{M}_\infty \models \phi(\vec{s})$ iff $\mathfrak{M}_s \models \phi(\vec{s})$; similarly the embedding from $\mathfrak{M}_t$ into $\mathfrak{M}_\infty$ is elementary, so $\mathfrak{M}_\infty \models \phi(\vec{t})$ iff $\mathfrak{M}_t \models \phi(\vec{t})$. Now comes the step at which we exploit the fact that $\mathfrak{M}_s = \mathfrak{M}_t$ as long as $|s| = |t|$. This fact tells us that $\mathfrak{M}_\infty \models \phi(\vec{s})$ iff $\mathfrak{M}_\infty \models \phi(\vec{t})$.

■

We can now do various other clever things. We can consider the Skolem hull of the indiscernibles. We then find that any order-automorphism of $\mathcal{I} = \langle I, \leq_\mathcal{I}\rangle$ extends to an automorphism of the Skolem hull.

4 Unsaturated models

"*Any fool can realize a type: it takes a model theorist to omit one.*"

Gerald Sacks.

Sacks is right—omitting types is hard!

We start by proving a theorem about propositional logic, with the intention of proving a version for predicate logic later.

A *type* in a propositional language $\mathcal{L}$ is a set of formulæ (a *countably infinite* set unless otherwise specified).

For T an $\mathcal{L}$-theory a T*-valuation* is an $\mathcal{L}$-valuation that satisfies T. A valuation v *realises* a type Σ if $v(\sigma) = \texttt{true}$ for every $\sigma \in \Sigma$. Otherwise v *omits* Σ. We say a theory T *locally omits* a type Σ if, whenever ϕ is a formula such that T proves $\phi \to \sigma$ for every $\sigma \in \Sigma$, then $T \vdash \neg\phi$.

Theorem 7 *The Omitting Types Theorem for Propositional Logic.*

Let T be a propositional theory, and $\Sigma \subseteq \mathcal{L}(T)$ a type. If T locally omits Σ then there is a T-valuation omitting Σ.

Proof:

By contraposition. Suppose there is no T-valuation omitting Σ. Then every formula in Σ is a theorem of T so there is an expression ϕ (namely '$\top$') such that $T \vdash \phi \to \sigma$ for every $\sigma \in \Sigma$ but $T \not\vdash \neg\phi$. Contraposing, we infer that if $T \vdash \neg\phi$ for every ϕ such that $T \vdash \phi \to \sigma$ for every $\sigma \in \Sigma$ then there is a T-valuation omitting Σ. ■

However, we can prove something stronger.

Theorem 8 *The Extended Omitting Types Theorem for Propositional Logic*

Let T be a propositional theory and, for each $i \in \mathbb{N}$, let $\Sigma_i \subseteq \mathcal{L}(T)$ be a type. If T locally omits every Σ_i then there is a T-valuation omitting all of the Σ_i.

Proof:

We will show that whenever $T \cup \{\neg A_1, \ldots \neg A_i\}$ is consistent, where $A_n \in \Sigma_n$ for each $n \leq i$, then we can find $A_{i+1} \in \Sigma_{i+1}$ such that $T \cup \{\neg A_1, \ldots \neg A_i, \neg A_{i+1}\}$ is consistent.

Suppose not, then $T \vdash (\bigwedge\limits_{1 \leq j \leq i} \neg A_j) \to A_{i+1}$ for every $A_{i+1} \in \Sigma_{i+1}$. But, by assumption, T locally omits Σ_{i+1}, so we would have $T \vdash \neg \bigwedge\limits_{1 \leq j \leq i} \neg A_j$ contradicting the assumption that $T \cup \{\neg A_1, \ldots \neg A_i\}$ is consistent.

Now, as long as there is an enumeration of the formulæ in $\mathcal{L}(T)$, we can run an iterative process where at each stage we pick for A_{i+1} the first formula in Σ_{i+1} such that $T \cup \{\neg A_1, \ldots \neg A_i, \neg A_{i+1}\}$ is consistent. This gives us a theory $T \cup \{\neg A_i : i \in \mathbb{N}\}$ which is consistent by compactness. Any model of $T \cup \{\neg A_i : i \in \mathbb{N}\}$ is a model of T that omits each Σ_i.

■

4.1 Omitting Types for First-Order Logic

First some some definitions

Definition 11

1. *An n-**type** is a set of formulæ all with at most n free variables*

2. *A model $\mathfrak{M}$ **realises** an n-type Σ if there is a tuple $\vec{x}$ s.t $\mathfrak{M} \models \phi(\vec{x})$ for every $\phi \in \Sigma$;*

3. *T **locally omits** an n-type Σ if, whenever ϕ is a formula s.t. $T \vdash (\forall x)(\phi(\vec{x}) \to \sigma(\vec{x}))$ for all σ in Σ, then $T \vdash (\forall \vec{x})(\neg\phi(\vec{x}))$.*

The property of theories of locally omitting a particular type is universal Horn:

$$(\forall \phi)(\bigwedge_{\sigma \in \Sigma} (T \vdash (\forall x)(\phi(x) \to \sigma(x))) \ \to \ T \vdash (\forall x)(\neg\phi(x)))$$

(It looks like $(\forall x)((\bigwedge_{i \in I} p_i) \to q)$.) The following is a consequence of this observation:

Remark 6 *An intersection of an arbitrary family of theories each locally omitting a type Σ also locally omits Σ.*

Proof: Obvious. . . ■

. . . but worth noting, since it means that we have a good notion of closure: keep adding axioms to T until you obtain something that locally omits Σ.

Definition 12 *Let T^0 be T. Obtain $T^{\alpha+1}$ from T^α as follows. Whenever ϕ is a formula s.t. $T^\alpha \vdash (\forall x)(\phi(x) \to \sigma(x))$ for all σ in Σ, then add to T^α the new axiom $(\forall x)(\neg\phi(x))$. The result of doing this for all $\sigma \in \Sigma$ is $T^{\alpha+1}$.*

I can't think of any reason why this process should close up at ω so we iterate transfinitely until it closes or becomes inconsistent. Let the result be T_∞ where ∞ is the closure ordinal (countable if $\mathcal{L}_T$ is countable). I suppose the 'Σ' should appear in this notation somewhere!

Theorem 9 *If T locally omits Σ then it has a model omitting Σ.*

Proof:

Let T be a theory locally omitting a type Σ and let $C = \langle c_i : i \in \mathbb{N}\rangle$ be a countable set of new constant letters. Let $\langle \phi_i : i \in \mathbb{N}\rangle$ be an enumeration of the sentences of $\mathcal{L}_T$.

We will construct recursively a $\subseteq$-increasing sequence $\langle T_i : i \in \mathbb{N}\rangle$ of finite extensions of T with the property that, for each $m \in \mathbb{N}$,

1. T_{m+1} decides ϕ_n for all $n \leq m$;
2. If ϕ_m is $(\exists x)\psi(x)$ and $\phi_m \in T_{m+1}$, then $\psi(c_p) \in T_{m+1}$ where c_p is the first constant not occurring in T_m or ϕ_m;
3. There is a formula $\sigma(x) \in \Sigma$ such that $(\neg\sigma(c_m)) \in T_{m+1}$.

Given T_m, we construct T_{m+1} as follows. Think of T_m as $T \cup \{\theta_1 \ldots \theta_r\}$, and the conjunction of the theta's as Θ. Let $\{c_1 \ldots c_n\}$ be the constants from C that have appeared in Θ, and let $\Theta(\vec{x})$ be the result of replacing 'c_i' by 'x_i' in Θ. Then (obviously!) $\Theta(\vec{x})$ is consistent with T. Therefore, for some $\sigma(x) \in \Sigma$, $\Theta \wedge \neg\sigma(x_m)$ is consistent with T. Put '$\neg\sigma(c_m)$' into T_{m+1}. This makes (3) hold.

If ϕ_m is consistent with $T_m \cup \{\neg\sigma(c_m)\}$, put it into T_{m+1}. Otherwise put in $\neg\phi_m$. This takes care of (1). If ϕ_m is $(\exists x)\psi(x)$ and is consistent with

$T_m \cup \{\neg\sigma(c_m)\}$, put $\psi(c_p)$ into T_{m+1}. This takes care of (2). This ensures that (1-3) hold for T_{m+1}.

Now consider $T^* = \bigcup_{i \in \mathbb{N}} T_i$. T^* is complete by construction. Consider an arbitrary countable model of T^* and the submodel of that model generated by the constants in C. This will be a model of T^*, and condition 3 ensures that it omits Σ. ■

Do we need this next bit?

Lemma 3 *If T locally omits a type Σ, then so does any finite extension of T.*

Proof: Suppose T locally omits Σ; we will show that $T \cup \{p\}$ locally omits Σ too. Suppose for each $\sigma \in \Sigma$, $T \cup \{p\} \vdash (\forall x)(\phi(x) \to \sigma(x))$. Then, for each $\sigma \in \Sigma$,

$$T \vdash p \to (\forall x)(\phi(x) \to \sigma(x))$$

so

$$T \vdash (\forall x)(p \to (\phi(x) \to \sigma(x)))$$

and

$$T \vdash (\forall x)((p \wedge \phi(x)) \to \sigma(x))$$

so, since T locally omits Σ,

$$T \vdash (\forall x)(\neg(p \wedge \phi(x))$$

whence

$$T \cup \{p\} \vdash (\forall x)(\neg\phi(x))$$

as desired. ■

This seems to work for finite extensions only. One might think that one can do it for arbitrary extensions by using compactness but the proof has a hole. (Check it!) In particular i can see no reason why the union of a $\subseteq$-chain of theories each locally omitting Σ should omit Σ—unless of course the chain has uncountable cofinality. However it does enable us to prove the following

Corollary 2 *Let T be a theory, Σ a type $\subseteq \mathcal{L}_T$ and T^∞ the least theory $\supseteq T$ that locally omits Σ. Let T^* be the theory of all models of T that omit Σ. (That is to say $T^* = \bigcap\{Th(\mathfrak{M}) : \mathfrak{M} \models T$ and $\mathfrak{M}$ omits $\Sigma\}$)*

Then $T^\infty = T^*$.

Proof:

Clearly T^* locally omits Σ, so $T^\infty \subseteq T^*$. Suppose the inclusion is proper, so that there is $p \in T^* \setminus T^\infty$. But then, by lemma 3, $T^\infty \cup \{\neg p\}$ locally omits Σ. Therefore, by theorem 9, there will be a model of $T^\infty \cup \{\neg p\}$ that omits Σ. But then p cannot be in T^* which—after all—is $\bigcap\{Th(\mathfrak{M}) : \mathfrak{M} \models T$ and $\mathfrak{M}$ omits $\Sigma\})$. ■

Omitting Types matters because the standard model of PA omits the type that says of a constant that it denotes a nonstandard natural.

5 Preservation Theorems

(This lemma is probably not going to be lectured. It's here beco's, well ... it's the key lemma one uses for proving preservation theorems)

Lemma 4 *Let T be a consistent theory in $\mathcal{L}$ and let Δ be a set of sentences of $\mathcal{L}$ which is closed under $\vee$. Then the following are equivalent*

1. *T has a set Γ of axioms where $\Gamma \subseteq \Delta$;*
2. *If $\mathcal{A}$ is a model of T and $(\forall \delta \in \Delta)(A \models \delta \to \mathcal{B} \models \delta)$ then $\mathcal{B} \models T$.*

Proof:
It is obvious that 1 implies 2. For the converse, assume (2), and suppose Δ and T given. Let $\Gamma = \{\phi \in \Delta : T \vdash \phi\}$. Then $T \vdash \Gamma$. We will show that Γ entails the whole of T. Let $\mathcal{B}$ be a model of T. Let

$$\Sigma = \{\neg\delta : \delta \in \Gamma \wedge \mathcal{B} \models \neg\delta\}$$

We show that $T \cup \Sigma$ is consistent. T is consistent by hypothesis; Suppose $T \cup \Sigma$ is inconsistent. Then there are $\neg\delta_1 \ldots \neg\delta_n$ all in Σ such that $T \vdash \neg(\neg\delta_1 \wedge \ldots \wedge \neg\delta_n)$ which is to say $T \vdash \delta_1 \vee \ldots \vee \delta_n$. Since Δ is closed under $\vee$ this theorem belongs to Δ, and therefore to Γ and therefore holds in $\mathcal{B}$. But this contradicts the fact that these δ_i are *false* in $\mathcal{B}$. So $\Sigma \cup T$ must have been consistent, and has a model $\mathcal{A}$. Then every sentence $\delta \in \Delta$ which holds in $\mathcal{A}$ holds also in $\mathcal{B}$ (by (2)). So Γ is an axiomatisation of T as desired.

Definition 13

The triple $\langle \mathcal{A}, \mathcal{B}, \mathcal{C} \rangle$ form a **sandwich** *if $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{C}$ and $\mathcal{A} \prec \mathcal{C}$;*

$\mathcal{A}$ is **sandwiched by** *$\mathcal{B}$ if there are elementary extensions $\mathcal{A}'$ of $\mathcal{A}$ and $\mathcal{B}'$ of $\mathcal{B}$ such that $\mathcal{B} \subseteq \mathcal{A}' \subseteq \mathcal{B}'$.*

Theorem 10 *(Chang-łoś-Suszko*
the following are equivalent

1. *T has a set of universal-existential axioms*
2. *T is preserved under unions of chains of models*
3. *if whenever $\mathcal{A} \models T$ and $\mathcal{A}$ is sandwiched by $\mathcal{B}$ then $\mathcal{B} \models T$.*

Proof:
We prove $1 \to 2 \to 3 \to 1$.
$1 \to 2$ is easy; $3 \to 1$ follows from lemma 4; we now prove $2 \to 3$.
Suppose $\mathcal{A}$ is sandwiched by $\mathcal{B}$. We shall construct a chain of models

$$\mathcal{B}_0 \subseteq \mathcal{A}_0 \subseteq \mathcal{B}_1 \subseteq \mathcal{A}_1 \ldots \mathcal{B}_n \subseteq A_n \ldots$$

where $\mathcal{B}_0 = \mathcal{B}$; each triple $\mathcal{B}_n, \mathcal{A}_n, \mathcal{B}_{n+1}$ forms a sandwich; $\mathcal{A} \prec \mathcal{A}_0$ and each $\mathcal{A}_n$ is elementarily equivalent to $\mathcal{A}$. We will attempt to construct this sequence by recursion, and to do this we will need to be able—on being presented with $\mathcal{B}_n, \mathcal{A}_n$ and $\mathcal{B}_{n+1}$ forming a sandwich—to find $\mathcal{A}_{n+1}$ elementarily equivalent to $\mathcal{A}_n$ and an elementary extension $\mathcal{B}_{n+2}$ of $\mathcal{B}_{n+1}$ so that $\mathcal{B}_{n+1}, \mathcal{A}_{n+1}$ and $\mathcal{B}_{n+2}$ form a sandwich.

How do we do this? We extend the language-in-hand by adding a new one-place predicate U and a constant name c_b for every element $b \in \mathcal{B}_{n+1}$. Let us call this new language Ł$'$. Note let T' be the theory

(elementary diagram of $\mathcal{B}_{n+1}$) $\cup \{\phi^U : \mathcal{A}_n \models \phi\} \cup \{U(c_b) : b \in B_{n+1}\}$
where ϕ^U is the relativisation of ϕ.

Thus any model of T' will be an elementary extension $\mathcal{B}_{n+2}$ of $\mathcal{B}_{n+1}$ which contains a subset U that includes at least all the elements of $\mathcal{B}_{n+1}$. Also the submodel determined by the extension of U (in $\mathcal{B}_{n+2}$) is elementarily equivalent to $\mathcal{A}_n$. That T' is consistent can be shown as follows:

- A sentence $F(c_{b_1} \ldots c_{b_n})$ such that $\mathcal{B}_{n+1} \models S(b_1 \ldots b_n)$
- A sentence ϕ^U such that $\mathcal{A}_n \models \phi$
- The sentence $U(c_{b_1}) \wedge U(c_{b_2}) \wedge \ldots U(c_{b_n})$.

Since $\mathcal{B}_n \prec \mathcal{B}_{n+1}$ there will be $d_1 \ldots d_n \in B_n$ such that

$$\mathcal{B}_n \models S(d_1 \ldots d_n), \quad \mathcal{B}_{n+1} \models S(d_1 \ldots d_n)$$

Now we find that B_{n+1} is a model of T' if we interpret U as membership in A_n, and the constant c_{d_i} as d_i.

Now consider the sequence of models in this chain we are building. Clearly all the $\mathcal{A}_n$ are models of T and so (since we are assuming 2) the union is also a model of T. But this union is also the union of the $\mathcal{B}_n$, which are all elementarily equivalent, and is therefore elementarily equivalent to them too, and is a model of T, so $\mathcal{B}$ was a model of T too.

We have used the fact that a direct limit of a family of elementary embeddings is an elementary extension. This was theorem 5.

This lemma is the crucial lemma in the proof of lots of completeness theorems: a formula is equivalent to a [syntactic property] formula iff the class of its models is closed under [some operations].

5.1 Ultralimits and Frayne's Lemma

Lemma 5 *Suppose $\mathcal{A}$ and $\mathcal{B}$ are elementarily equivalent. Then there is an ultrapower $\mathcal{A}^I/\mathcal{U}$ of A and an elementary embedding from $\mathcal{B}$ into it.*

Recall here the terminology from model theory: expansion, reduction, extension

Proof:

Supply names $\mathtt{b}$ for every member b of B. Let $\mathcal{L}$ be the language with the new constants. There is an obvious way of expanding $\mathcal{B}$ to a structure for this new language, namely to let each constant $\mathtt{b}$ denote that element b of B which gave rise to it. (Of course this is not the only way of doing it: *any* map $B \to B$ will give rise to an expansion of $\mathcal{B}$ of this kind—and later we will have to consider some of those ways). Let us write '$\mathcal{B}'$' to denote this obvious expansion of $\mathcal{B}$, and let I be the set of sentences of $\mathcal{L}$ true in $\mathcal{B}'$. (Use of the letter 'I' for this is a bit of a give-away!)

Consider ϕ a formula in I. It will mention finitely many constants—let us say two, for the sake of argument. Replace these two constants by new variables 'v_1' and 'v_2' (*not* mentioned in ϕ!) to obtain ϕ'' and bind them both with '$\exists$' to obtain $(\exists v_1)(\exists v_2)\phi''$ which we will call 'ϕ'' for short. This new formula is a formula of the original language which is true in $\mathcal{B}$ and is therefore also true in $\mathcal{A}$ (since $\mathcal{A}$ and $\mathcal{B}$ are elementarily equivalent).

The next step is to expand $\mathcal{A}$ to a structure for the language $\mathcal{L}$ by decorating it with the with the extra constants $\mathtt{b}$ etc that we used to denote members of B. Of course any function $B \to A$ gives us a way of decorating $\mathcal{A}$ but with ϕ in mind we are interested only in those decorations which give us a structure that satisfies ϕ. If ϕ contained the constants $\mathtt{b}$ and $\mathtt{b'''}$ for example then the obvious way to expand A involves using those two constants to denote the witnesses in $\mathcal{A}$ for the two existential quantifiers in ϕ'. Since ϕ contains only finitely many constants this nails down denotations for only finitely many of the constant-names-for-members-of-B. However any finite map from B to A can be extended to a total function $B \to A$ so we can extend this to a way of labelling members of A with these constants in such a way that the decorated version of $\mathcal{A}$ satisfies the original formula ϕ.

Pick one such labelling and call it $a(\phi)$. (Thus $a(\phi)$ is merely an element of $B \to A$ satisfying an extra condition parametrised by ϕ. We can think of a as a function $\mathcal{L} \to (B \to A)$ or as a function $(\mathcal{L} \times B) \to A$ *ad libitum*). $\mathcal{A}$ expanded with this decoration we call $\langle \mathcal{A}, a(\phi) \rangle$. Now consider the set

$$J(\phi) =: \{\psi \in I : \langle \mathcal{A}, a(\psi) \rangle \models \phi\}$$

It is easy to check that the family $\{J(\phi) : \phi \in I\}$ of subsets of I has the finite intersection property and so gives rise to a ultrafilter[6] $\mathcal{U}$ on I and thence to an ultrapower $\mathcal{A}^I/\mathcal{U}$. Evidently if $\phi \in I$ then $J(\phi) \in \mathcal{U}$ and the ultrapower will believe ϕ.

We have to find an elementary embedding from $\mathcal{B}$ into this ultrapower. Given $b \in B$ whither do we send it? The obvious destination for b is the equivalence class of the function $\lambda\phi.a(\phi, b)$ that sends ϕ to $a(\phi, b)$. The function that sends b to $[\lambda\phi.a(\phi, b)]$ is $\lambda b.[\lambda\phi.a(\phi, b)]$—which we will write '$h$' for short. We must show that h is elementary.

[6] There doesn't seem to be any reason to conclude that this ultrafilter will be nonprincipal, but then nor does it seem to matter if it *isn't*. Bell and Slomson don't say that it will be nonprincipal. Thanks to Phil Ellison for drawing my attention to this point.

The best way to understand what h does and why it is elementary is to think of the ultrapower as a reduction of the ultraproduct

$$\prod_{\psi \in I} \langle \mathcal{A}, a(\psi) \rangle.$$

("**expand** the factors; take an ultraproduct; **reduce** the ultraproduct—to obtain a ultrapower of the factors ...")

Each of the factors $\langle \mathcal{A}, a(\psi) \rangle$ is a structure for $\mathcal{L}$ and therefore the ultraproduct is too. By the same token, for each $b \in B$, each of the factors has an element which is pointed to by $\mathtt{b}$-the-constant-name-of-b, and therefore the ultraproduct will too. The key fact is that h is the function that sends each $b \in B$ to the thing in the ultraproduct that is pointed to by $\mathtt{b}$ the constant-name-of-b.

As for the elementarity of h, suppose $\mathcal{B} \models \phi(\vec{v})$. Then, for some choice of constants $\vec{\mathtt{b}}$, $\mathcal{B} \models \phi(\vec{\mathtt{b}})$, and $\mathcal{B}' \models \phi'$. But now $J(\phi)$ is $\mathcal{U}$-large, so the ultrapower believes ϕ. ■

I lifted this proof from [1].

But what we really need is Scott's lemma:

Lemma 6 *Suppose $g : \mathcal{A} \hookrightarrow \mathcal{B}$ is an elementary embedding. Then there is an ultrapower $\mathcal{A}^I/\mathcal{U}$ of A and an elementary embedding from $\mathcal{B}$ into it making the triangle commute.*

$$\begin{tikzcd}
 & \mathcal{A}^I/\mathcal{U} \\
\mathcal{A} \arrow[ur, "K"] \arrow[r, "g"'] & \mathcal{B} \arrow[u]
\end{tikzcd}$$

Proof:

The ideas are the same, but we need to be slightly more careful in the definition of $a(\phi)$. Fix once for all a member a of A. As before, we extend the language by adding names for every member of B, thus obtaining the language $\mathcal{L}$ as before. Now we expand $\mathcal{B}$ by decorating B with these names, but not in the obvious way. If b is in the range of g we allow $\mathtt{b}$ the constant-name-of-b to denote b; if b is not in the range of g, then $\mathtt{b}$ will denote $g(a)$. Let's call this expanded structure $\mathcal{B}'$.

If we are to expand $\mathcal{A}$ to obtain a structure for $\mathcal{L}$ then we must ensure that, for each $b \in B$, the constant-name-$\mathtt{b}$-of-b points to something in A. The obvious way to do this is to ordain that $\mathtt{b}$ point to g^{-1} of the thing that that $\mathtt{b}$ points to in the expansion $\mathcal{B}'$ of $\mathcal{B}$. This decorated version of $\mathcal{A}$ and the decorated version

$\mathcal{B}'$ of $\mathcal{B}$ are elementarily equivalent (with respect to the extended language with the names) (*)

As before, let I be the set of sentences of $\mathcal{L}$ true in $\mathcal{B}'$. Consider a formula $\phi \in I$. Recall what we did at the same stage in the proof of Frayne's Lemma. This time we replace with existentially-quantified variables only those constants denoting elements of B not in the range of g. Let's call this formula ϕ' like last time. Evidently $\mathcal{B}' \models \phi'$ and so, by the remark (*) at the end of the last paragraph, the decorated version of $\mathcal{A}$ also satisfies ϕ'. So, as before, there is another decoration of $\mathcal{A}$ which actually satisfies the original ϕ. Pick one such decoration and call it $a(\phi)$, and call the structure thus decorated $\langle \mathcal{A}, a(\phi) \rangle$. We define

$J(\phi) =: \{\psi \in I : \langle \mathcal{A}, a(\phi) \rangle \models \phi\}$

as before, and it has the finite intersection property as before and gives us an ultrafilter $\mathcal{U}$ as before, and we have the same elementary embedding h from $\mathcal{B}$ into the ultrapower as before. It remains only to check that the diagram is commutative. I think this can safely be left as an exercise to the reader. ■

$$\begin{tikzcd}
A \arrow[r] & A_1 \arrow[r] & A_2 \arrow[r] & \cdots & A_\infty \\
B \arrow[r] \arrow[ur] & B_1 \arrow[r] \arrow[u] \arrow[ur] & B_2 \arrow[r] \arrow[u] \arrow[ur] & \cdots & B_\infty
\end{tikzcd}$$

References

[1] J.L Bell and A Slomson: Models and Ultraproducts.

[2] Forster: Logic, Induction and sets. CUP