The Compactness Theorem

A course of three lectures
one hour each
June 20–1, 2015
5th World Congress and School
on Universal Logic
June 20–30, 2015
Istanbul University

David Pierce

June 24, 2015
Edited September 4, 2016

Mathematics Department
Mimar Sinan Fine Arts University
Istanbul
http://mat.msgsu.edu.tr/~dpierce/

Contents

Preface

This document consists of transcriptions of the three hour lectures of my course on the Compactness Theorem in the 5th World Congress and School on Universal Logic, June, 2015.

I submitted the abstract (also included here) some nine months in advance of the School. I have added the footnotes, and the year (1908) of Zermelo's axioms; otherwise the abstract is unchanged, though the references are now incorporated with others at the end of the present document. Every work mentioned with a year has a citation in the references. The submitted abstract was placed in a frame (`t5-compactness.html`) on the website (`http://www.uni-log.org/`) of the School and Congress: apparently this frame represented a conversion from the `pdf` version of my submitted abstract, since hyphens were retained, though they no longer marked the ends of lines. The use of frames means it is not possible to give a single link to the abstract as it was meant to be seen; one must say rather, "go to `http://www.uni-log.org/start5.html`, click on 'Tutorials,' and then click on 'Compactness Theorem.' "

Mostly in the summer of 2014, I composed several collections of notes relevant to the Compactness Theorem. These notes concerned, respectively,

(1) the Löwenheim–Skolem Theorem,
(2) Gödel's Completeness Theorem,
(3) Tarski's 1950 ICM address (giving the Compactness Theorem its current name),

(4) Lindström's Theorem,
(5) logics in general, and
(6) the Compactness Theorem itself.

In July, I gave a two-week course on nonstandard analysis at the Nesin Mathematics Village, Şirince; and in this course I considered the logical relations between the Compactness Theorem, Łoś's Theorem, and the Prime and Maximal Ideal Theorems. I typed up and edited my lectures in this course. I then gave a half-hour talk on the Compactness Theorem at the Caucasian Mathematics Conference, Tbilisi, September 5 & 6, 2014. I used slides for that talk, and the course that is the subject of the present document is basically an expanded version of those slides. All of these documents are among my webpages.

I started typing notes for the lectures of my course before giving them; but then a fortuitous computer malfunction inhibited me from continuing. Notes written at the computer are easier to edit and save and share, which is why I am creating the present document. However, notes to be used in delivering a *lecture* are better written out by hand. Hand writing provides muscle memory as well as, I think, a better visual memory of one's own words. It also forces one to think about just how much writing one will want to do at the board.

Prepared *after* the course, the notes below are a typeset version of my handwritten notes, with changes and footnotes based on my memory of what actually happened in the lectures. Paragraphs labelled as Remarks were only spoken out loud. Other parts *may* only have been spoken aloud; it is hard to be precise on these matters.

The three hours of the course were distinct, the respective themes being the Compactness Theorem (1) in action, (2) as a topological theorem, and (3) proved. Looking at notes from regular university courses that I had recently taught, I had the idea that I could cover three or four of my handwritten pages (size A4) in a one-hour lecture. This turned out to be nearly correct. I tried to be prepared to curtail or jettison material at the end of my notes for each lecture.

Speakers in the other lectures that I attended used slides, unless technical problems prevented it. I think the use of slides is usually a mistake, especially for a *course* of lectures. Slides allow a speaker to pass too quickly through his or her material, and they inhibit the listener from taking notes. Speakers who use slides may also want to write on the board; but since the room will have been darkened for the slides, the speaker's writing will be hard to read.

For my convenience, I typed up the schedule of the two days on which my course would meet. The School of Universal Logic was to have an afternoon session on Saturday, and morning and afternoon sessions on Sunday. Then the schedule was changed, because of the Turkish national university entrance examinations. The Saturday afternoon session was delayed, and the Sunday morning session was shifted to Monday. When the School actually began on Saturday, there was a further delay. I believe the schedule given on the next page represents what actually happened.

	Room I	Room II	Room III

Saturday, June 20, 2015

14:30	Opening: why studying [*sic*] logic		
15:00	Husserl & Frege	Lindström	Completeness
16:00	Jain	Löwenheim–Skolem	Nonsense
17:00	coffee		
17:30	Music	Politics	Relativity
18:30	COMPACTNESS	Category	Quantum

Sunday, June 21, 2015

14:00	COMPACTNESS	Politics	Relativity
15:00	Music	Category	Quantum
16:00	coffee		
16:30	COMPACTNESS	Kant	Nonsense
17:30	Husserl & Frege	Lindström	Dugundji

Abstract

A logic has *a compactness theorem* if a set of sentences of the logic has a model whenever every finite subset of the set has a model. For present purposes, **the Compactness Theorem** is that first-order logic has a compactness theorem. This theorem is fundamental to model theory. One may however note that Hodges's encyclopedic twelve-chapter 1993 volume *Model Theory* finds no need to prove the theorem until Chapter 6. It is worthwhile to think about what needs Compactness and what does not.

One consequence of the Compactness Theorem is that a set of (first-order) sentences with arbitrarily large finite models must have an infinite model. A more purely mathematical consequence is the **Prime Ideal Theorem:** *every nontrivial commutative ring has a prime ideal.* One can prove this by noting first that every *maximal* ideal is prime. Moreover, every *countable* ring has a maximal ideal; for we can obtain a generating set of such an ideal by considering the elements of the ring one by one. In particular then, every finitely generated subring of a given ring has a maximal ideal, because every finitely generated ring is countable. By the Compactness Theorem then, the original ring must have an ideal that is at least prime, although it might not be maximal. The point here is that primeness is a "local" property, while maximality is not.

It is usually understood that every nontrivial commutative ring has, not just a prime ideal, but a maximal ideal. To make

it easy to prove such results, Zorn stated in 1935 the result now known by his name. However, Zorn's Lemma relies on the Axiom of Choice. The Compactness Theorem is strictly weaker than this, with respect to ZF (Zermelo–Fraenkel set theory without Choice). For, Compactness is also a *consequence* of the Prime Ideal Theorem, even the Boolean Prime Ideal Theorem; and *this* is strictly weaker than the Axiom of Choice (as shown by Halpern and Lévy in 1971).

The Compactness Theorem for *countable* sets of sentences needs nothing beyond ZF. Skolem showed this implicitly in 1922 when he established the paradox that Zermelo's [1908] axioms for set theory must have a countable model, if they have a model at all. In 1930, Gödel proved countable Compactness explicitly, though not by that name. Mal′tsev stated the full Compactness Theorem as the General Local Theorem[1] in 1941, having proved it implicitly in 1936; he used it to prove algebraic results in the way we proved the Prime Ideal Theorem above.

In his 1950 address to the International Congress of Mathematicians, Tarski gave the Compactness Theorem its current name and noted its topological meaning. But this meaning is not generally well expressed in today's textbooks of model theory.

The class of structures having a given signature can be given a topology, although the closed "sets" in this topology are proper classes (except for the empty set): they are the classes of models of sets of sentences. The space of all structures has a Kolmogorov (T_0) quotient that is a set: it is the space of complete theories of structures. If one replaces sentences with

[1] I believe this name is the innovation of the translator, but Mal′tsev used no particular name.

their logical equivalence classes, then the set of sentences becomes a Boolean algebra, called a Lindenbaum algebra; and the complete theories of structures become ultrafilters of the Lindenbaum algebra. By means of the Boolean Prime Ideal Theorem, the *Stone space* consisting of *all* ultrafilters of the Lindenbaum algebra is easily shown to be compact. Or one could look instead at the *spectrum,* consisting of the prime ideals of the corresponding Boolean ring. Spectra are always compact. The Compactness Theorem says more: every ultrafilter of the Lindenbaum algebra is derived from the complete theory of a structure.

The compactness theorem for propositional logic can be seen as a version of the theorem, known by the name of Tychonoff, that the product of two-element discrete spaces (or indeed any compact spaces) is compact. *The* Compactness Theorem, for first-order logic, does not follow so readily, though it can be seen to result from a kind of reduction of first-order logic to propositional logic. Then Lindström's Theorem is roughly that there is no such reduction for certain more expressive logics—but see that tutorial for more. Sometimes the Compactness Theorem is derived from the Completeness Theorem: see *that* tutorial for more. Meanwhile, the present tutorial is intended to fill out the foregoing sketch of the Compactness Theorem as such.[2]

[2] My understanding of the history of the Compactness Theorem depends on John Dawson's 1993 article.

1 Compactness in Action

Remark. I have lived in Istanbul for four years (and Turkey for fifteen). Since I moved here with my spouse, the subway station called Vezneciler near Istanbul University has opened. When you come out of that station, you see a mosque, namely

KALENDERHANE CAMII.

Perhaps many people hardly notice it, assuming it is just another mosque. But it is not just another mosque, and it deserves further notice. It is perhaps 800 years old, and it was a church before the Ottoman Turkish conquest of Istanbul in 1453. Similarly, in some model-theory texts at least, the name of the Compactness Theorem is explained in passing as being due to the compactness of a certain Stone space. This casual reader accepts the explanation, because the Stone space in question *is* compact.[1] But *every* Stone space is compact,

[1]For example, Hodges [9, §6.2, p. 274] defines a Stone space as a nonempty compact Hausdorff space with a basis of clopen sets. He then states the theorem that the space of ultrafilters of a Boolean algebra is a Stone space. In the next section, he defines a **complete type** as the complete type of some tuple over a given set X of parameters, in an elementary extension of a given structure A. Using the Compactness Theorem, he shows that such a complete type is precisely a set of formulas that is maximal among the sets of formulas with parameters from X that are finitely realized in A. If $X = \varnothing$ and $T = \mathrm{Th}(A)$, one then speaks of *complete types of* T. The set of complete types of T in n variables is denoted by $S_n(T)$ and called the nth **Stone space** of T, "for reasons that will appear in a moment," namely, that the space

and every logic gives rise to a Stone space, namely the Stone space consisting of the prime filters of the distributive lattice of logical equivalence-classes of sentences.[2] The Compactness Theorem is that, in first-order logic, each of those prime filters actually has a model.

We shall use the notation

$$\mathbb{N} = \{1, 2, 3, \dots\}, \qquad \omega = \{0\} \cup \mathbb{N}.$$

Consider $(\mathbb{N}, 1, x \mapsto x+1)$ as a *structure* in the *signature* $(1, S)$. It satisfies the **Peano Axioms** (1889):

$$\forall x\ \forall y\ (Sx = Sy \rightarrow x = y),$$
$$\forall x\ Sx \neq 1,$$
$$\forall X\ \bigl(1 \in X \land \forall y\ (y \in X \rightarrow Sy \in X) \rightarrow \forall y\ y \in X\bigr).$$

Theorem (Dedekind, 1888)**.** For all models $\mathfrak{A}$ of the Peano Axioms, for all structures $\mathfrak{B}$ in the signature $\{1, S\}$, there is a unique homomorphism from $\mathfrak{A}$ to $\mathfrak{B}$.

Remark. We know the axioms by Peano's name, apparently because he wrote them out formally; but Dedekind recognized them earlier and understood them better. In particular,

"is in fact the Stone space of a boolean algebra generally known as the nth **Lindenbaum algebra** of T... The name 'compactness theorem' originally came from the fact that $S_n(T)$ is a compact space." Strictly, Hodges has not formally defined the Stone space *of* a Boolean algebra; presumably it is to be understood as the space of ultrafilters of the algebra, which space is a Stone space, as noted. If $S_n(T)$ is the Stone space, in this sense, of the nth Lindenbaum algebra of T, then it is automatically compact, regardless of the Compactness Theorem. But if $S_n(T)$ is a space of complete types in the original sense of complete types *of tuples,* then the Compactness Theorem can be understood as the theorem that $S_n(T)$ is compact.

[2]This is spelled out in the second lecture.

Dedekind understood that all three of the axioms were needed to ensure existence of the homomorphism of the theorem.[3]

Corollary. All models of the Peano Axioms are isomorphic to the structure $(\mathbb{N}, 1, x \mapsto x+1)$. Hence if c is a new constant symbol, then[4]

$$\{\text{Peano Axioms}\} \cup \{c \neq S^{(n)}1 \colon n \in \omega\}$$

has no model, although every finite subset has.

Compactness Theorem (Skolem, 1922; Gödel, 1930; Mal′cev, 1936, 1941)**.** In a *first-order* logic, there is a model of a set of sentences, if there is of every finite subset.

First-order means:
1) variables stand for individuals, not sets;
2) formulas are finite.

By Compactness, the Peano Axioms have no first-order formulation. Neither do the axioms of **torsion groups,** which are the abelian-group axioms, along with[5]

$$\forall x \bigvee_{n \in \mathbb{N}} nx = 0.$$

Theorem (Tarski, 1954; Łoś, 1955)**.** In a signature $\mathscr{S}$, suppose T is a (first-order) theory, and σ is a (first-order) sentence that is **preserved in substructures** in $\mathbf{Mod}(T)$, that is,

$$\text{if } \mathfrak{A}, \mathfrak{B} \models T,\ \mathfrak{A} \subseteq \mathfrak{B}, \text{ and } \mathfrak{B} \models \sigma, \text{ then } \mathfrak{A} \models \sigma.$$

[3] I elaborated at the beginning of the second lecture.

[4] Apparently I wrote $S^{(n)}c \neq 1$ rather than $c \neq S^{(n)}1$; I made the correction at the beginning of the second lecture.

[5] There was a question about what nx means. It means $\underbrace{x + \cdots + x}_{n}$.

Then for some *universal* sentence τ of $\mathscr{S}$,

$$T \vdash \sigma \leftrightarrow \tau.$$

Remark. The converse is easy and was alluded to in Nate Ackerman's talk on the Löwenheim–Skolem Theorem.

Proof. By Compactness,

1) $T \vdash \sigma \leftrightarrow \bigwedge \Gamma$, where

$$\Gamma = \{\tau \in \mathrm{Sen}_{\mathscr{S}} \colon \tau \text{ is universal and } T \vdash \sigma \to \tau\};$$

2) $T \vdash \sigma \leftrightarrow \bigwedge \Gamma_0$ for some finite subset Γ_0 of Γ.

Details:

1. We show $T \cup \Gamma \vdash \sigma$. Say $\mathfrak{A} \models T \cup \Gamma$; we show $\mathfrak{A} \models \sigma$. By hypothesis, it is enough to find $\mathfrak{B}$ such that

$$\mathfrak{B} \models T, \qquad \mathfrak{A} \subseteq \mathfrak{B}, \qquad \mathfrak{B} \models \sigma.$$

This means finding a model of

$$T \cup \mathrm{diag}(\mathfrak{A}) \cup \{\sigma\},$$

where $\mathrm{diag}(\mathfrak{A})$ is the **diagram** of $\mathfrak{A}$, namely the quantifier-free theory of $\mathfrak{A}_A$ (the obvious expansion of $\mathfrak{A}$ to $\mathscr{S}(A)$, which has a new constant symbol for each element of the universe A of $\mathfrak{A}$).[6] If there is no model, then by Compactness, for some[7]

[6]The parenthetical remark is in my notes, but I must not have written it all down; for somebody asked me what the Roman A was. Apparently she was unfamiliar with my convention, which comes from Chang and Keisler. I did not mention that Hodges traces diagrams to Wittgenstein's *Tractatus* 4.26.

[7]Out loud I observed that logic is concerned with correct expression, but it is difficult to explain expressions in detail. We should understand that φ is a quantifier-free formula of $\mathscr{S}$, and $\vec{a}$ is a tuple of elements of A.

$\varphi(\vec{a})$ in $\operatorname{diag}(\mathfrak{A})$,

$$\begin{gathered}
T \vdash \sigma \to \neg\varphi(\vec{a}), \\
T \vdash \sigma \to \forall\vec{x}\,\neg\varphi, \\
\forall\vec{x}\,\neg\varphi \in \Gamma, \\
\mathfrak{A} \models \forall\vec{x}\,\neg\varphi, \\
\mathfrak{A} \models \neg\varphi(\vec{a}), \\
\neg\varphi(\vec{a}) \in \Gamma,
\end{gathered}$$

which is absurd.

2. Now $T \cup \Gamma \cup \{\neg\sigma\}$ has no model, so by Compactness, neither has $T \cup \Gamma_0 \cup \{\neg\sigma\}$.[8] □

Remark. Hodges gives the following as an application of a generalization of the theorem.

Theorem (Mal′cev, 1940)**.** A group has a faithful n-dimensional representation (that is, an embedding in $\mathrm{GL}_n(K)$ for some field K) if every finitely generated subgroup does.[9]

The Łoś–Tarski Preservation Theorem yields:

$$\operatorname{Th}(\{\text{substructures of models of } T\}) \subseteq T_\forall$$

(and the converse is easy). From algebra,

$$\{\text{substructures of fields}\} = \{\text{integral domains}\},$$

an **elementary class** (that is, $\mathbf{Mod}(T)$ for some theory T).

[8] I may not have written out even this much. However, in Peter Arndt's earlier talk on Lindström's Theorem, somebody had been confused about applying Compactness to the characteristic of fields. Arndt observed that a theorem about fields of characteristic 0 would be true in fields of large-enough characteristic; the questioner apparently thought that only the converse was the case.

[9] Somebody asked what $\mathrm{GL}_n(K)$ was.

Theorem. For all theories T, the class of substructures of models of T is an elementary class: If $\mathfrak{A} \vDash T_\forall$, then $\mathfrak{A}$ embeds in an element of $\mathbf{Mod}(T)$.

Proof. $T \cup \operatorname{diag}(\mathfrak{A})$ has a model, by Compactness, as in the proof of the Łoś–Tarski Theorem. □

What are the semigroups that embed in groups? They are precisely the models of some universal theory.[10]

[10]It turns out that it was Mal′cev who first found this theory, at least implicitly; but there are infinitely many axioms. See [10].

2 Compactness as a Topological Theorem

As we saw, the set[1]

$$\begin{gathered}\{\forall x\, \forall y\, (Sx = Sy \to x = y),\\ \forall x\, Sx \neq 1,\\ \forall X\, \big(1 \in X \land \forall y\, (y \in X \to Sy \in X) \to \forall y\, y \in X\big),\\ c \neq 1, c \neq S1, c \neq SS1, c \neq SSS1, \dots\}\end{gathered}$$

has no model, although every finite subset does (just let c be large enough).

On $\mathbb{N}$, addition is the unique operation $+$ such that

$$n + 1 = Sn, \qquad\qquad n + Sx = S(n + x).$$

That is, $x \mapsto n + x$ is the unique homomorphism from $(\mathbb{N}, 1, x \mapsto x + 1)$ to $(\mathbb{N}, n + 1, x \mapsto x + 1)$ guaranteed by Dedekind's theorem. The existence of such homomorphisms usually requires all three of the Peano Axioms. In fact, the definitions of addition and multiplication require only the induction axiom (as Landau shows implicitly in *Foundations of Analysis*). This is why *modular* addition and multiplication

[1] This assumes the correction in note 4 on page 12. In my handwritten notes, and in the original June 24, 2015, version of the present document, instead of "As we saw, the set . . . has no model," I wrote "It is the set . . . that has no model."

are possible: these are operations in $\mathbb{Z}/n\mathbb{Z}$, which allows proofs by induction.

But exponentiation needs more than induction. For example, *modulo* 3, we have

$$2^1 \equiv 2, \quad 2^2 \equiv 2 \cdot 2 \equiv 1, \quad 2^3 \equiv 2, \quad 2^4 \equiv 2 \cdot 2 \equiv 1 \not\equiv 2^1,$$

although $4 \equiv 1$.

Cantor Intersection Theorem. Provided F_0 is also bounded,[2] a decreasing chain

$$F_0 \supseteq F_1 \supseteq F_2 \supseteq \cdots$$

of closed intervals of $\mathbb{R}$ has nonempty intersection: for this intersection contains

$$\inf\{\sup F_k \colon k \in \omega\}.$$

This fails for other intervals:

$$\bigcap_{k\in\omega}[k,\infty) = \varnothing, \qquad \bigcap_{k\in\mathbb{N}}\left(0,\tfrac{1}{k}\right) = \varnothing.$$

But it still holds if each F_k is

1) the union of finitely many closed bounded intervals;
2) the intersection of any number of such unions.

Such intersections are the **closed** subsets of $\mathbb{R}$. Let τ be the set of these. Then[3]

[2]Originally I said each F_k was bounded; but then that might seem to imply that all closed sets are bounded.

[3]Somebody asked me after the lecture why the second condition was true. It is true by the distributivity of taking unions over taking intersections:

$$\bigcap A \cup \bigcap B = \bigcap_{a\in A}\bigcap_{b\in B} a \cup b.$$

1) $\varnothing \in \tau$,
2) $X \in \tau \ \& \ Y \in \tau \implies X \cup Y \in \tau$,
3) $\mathscr{X} \subseteq \tau \implies \bigcap \mathscr{X} \in \tau$.
4) $\mathbb{R} \in \tau$ (here $\mathbb{R}$ can be understood as $\bigcap \varnothing$).

Heine–Borel Theorem. A collection $\mathscr{X}$ of bounded closed subsets of $\mathbb{R}$ has non-empty intersection if each finite subcollection does.

Proof. We may assume $\mathscr{X}$ is countable, since each closed interval is the intersection of the larger closed intervals with rational endpoints:[4]

$$[\alpha, \beta] = \bigcap_{\substack{a \leqslant \alpha < \beta \leqslant b \\ a,b \in \mathbb{Q}}} [a, b].$$

Then $\mathscr{X}$ has the form $\{F_k : k \in \omega\}$, and we can apply the Cantor Intersection Theorem to

$$F_0 \subseteq F_0 \cap F_1 \subseteq F_0 \cap F_1 \cap F_2 \subseteq \cdots \qquad \square$$

The theorem is that bounded closed subsets of $\mathbb{R}$ are **compact.** In his address to the 1950 ICM, Tarski called the Compactness Theorem by this name and noted that it did establish compactness of certain topological spaces. We can understand these spaces as $\mathbf{Str}_{\mathscr{S}}$, the spaces of structures in the signature $\mathscr{S}$ for various $\mathscr{S}$.

Let

$$\boldsymbol{A} = \mathbf{Str}_{\mathscr{S}},$$
$$B = \mathrm{Sen}_{\mathscr{S}} = \{\text{first-order sentences of } \mathscr{S}\}.$$

There are

[4] I just made the explanation out loud, until somebody asked for clarification.

- a relation $\vDash$ from $\boldsymbol{A}$ to B,
- a binary operation $\vee$ on B.

If $\sigma \in B$, we define

$$\mathbf{Mod}(\sigma) = \{a \in \boldsymbol{A} \colon a \vDash \sigma\}.$$

Then[5]

$$\mathbf{Mod}(\sigma) \cup \mathbf{Mod}(\tau) = \mathbf{Mod}(\sigma \vee \tau). \tag{$*$}$$

Thus the classes $\mathbf{Mod}(\sigma)$ are a basis for a topology on $\boldsymbol{A}$ where every closed subclass is, for some subset Γ of B,

$$\mathbf{Mod}(\Gamma), \qquad \text{that is,} \qquad \bigcap_{\sigma\in\Gamma} \mathbf{Mod}(\sigma).$$

Every topology on a set or class $\boldsymbol{A}$ can be seen as arising in this way from a structure $(B, \vee)$ and a relation $\vDash$ from $\boldsymbol{A}$ to B such that $(*)$ holds. This topology is **compact,** provided that

$$\mathbf{Mod}(\Gamma) \neq \varnothing$$

whenever $\Gamma \subseteq B$ and, for every finite subset Γ_0 of Γ,

$$\mathbf{Mod}(\Gamma_0) \neq \varnothing.$$

Thus the Compactness Theorem is that $\mathbf{Str}_{\mathscr{S}}$ is compact.

In the general case, we can produce a structure $(C, \vee, \wedge, \bot, \top)$ or $\mathfrak{C}$ such that

- $(B, \vee) \subseteq (C, \vee)$,
- $\mathbf{Mod}(\bot) = \varnothing$ and $\mathbf{Mod}(\top) = \boldsymbol{A}$,

[5] At some point I was asked what the operation "vee" was. My explanation that it was logical disjunction was immediately satisfactory to the questioner. I believe I did point that, more generally, it would be any operation making $(*)$ true.

- for all σ and τ in C,
    - $(*)$ holds,
    - $\mathbf{Mod}(\sigma \wedge \tau) = \mathbf{Mod}(\{\sigma, \tau\})$,
    - $\mathbf{Mod}(\sigma) = \mathbf{Mod}(\Gamma)$ for some subset[6] Γ of B.

Now let

$$\sigma \sim \tau \iff \mathbf{Mod}(\sigma) = \mathbf{Mod}(\tau).$$

Then $\mathfrak{C}/{\sim}$ is a well-defined distributive lattice. If $a \in \boldsymbol{A}$, define

$$\mathrm{Th}(a) = \{\sigma \in C \colon a \vDash \sigma\}.$$

Then

$$\mathrm{Th}(a) = \mathrm{Th}(b) \iff \text{a and b are topologically indistinguishable.}$$

Thus $\{\mathrm{Th}(a) \colon a \in \boldsymbol{A}\}$ is a **Kolmogorov** (or T_0) quotient of $\boldsymbol{A}$.[7]

$$\begin{array}{ccc} \boldsymbol{X} & \overset{\vDash}{\rightsquigarrow} & C \\ {\scriptstyle x \mapsto \mathrm{Th}(x)} \downarrow & & \downarrow {\scriptstyle \sigma \mapsto \sigma^{\sim}} \\ \{\mathrm{Th}(x) \colon x \in \boldsymbol{X}\} & \rightsquigarrow & C/{\sim} \end{array}$$

Moreover,

$$\bot \notin \mathrm{Th}(a), \qquad \top \in \mathrm{Th}(a),$$

and

$$\sigma \vee \tau \in \mathrm{Th}(a) \iff \sigma \in \mathrm{Th}(a) \text{ OR } \tau \in \mathrm{Th}(a).$$

[6]Normally Γ will be finite; but we may allow it to be infinite if we wish.

[7]I think I did not actually express this this last sentence, since several students in the audience had indicated that they did not know anything about topology.

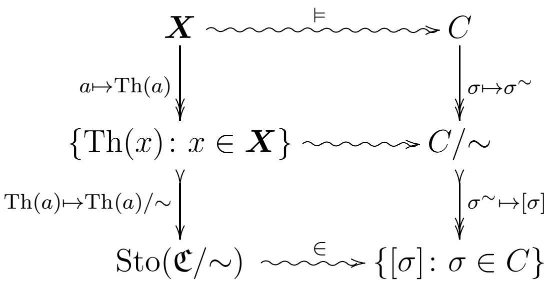
Therefore, by definition, $\mathrm{Th}(a)/{\sim}$ is a **prime filter** of $\mathfrak{C}/{\sim}$. But possibly not every prime filter of $\mathfrak{C}/{\sim}$ is of this form. Let

$$\mathrm{Sto}(\mathfrak{C}/{\sim}) = \{\text{prime filters of } \mathfrak{C}/{\sim}\},$$
$$[\sigma] = \{F \in \mathrm{Sto}(\mathfrak{C}/{\sim})\colon \sigma^{\sim} \in F\}.$$

Then

$$[\sigma] \cup [\tau] = [\sigma \vee \tau],$$

so the $[\sigma]$ induce a topology on $\mathrm{Sto}(\mathfrak{C}/{\sim})$ by the relation $\in$ (that is, the $[\sigma]$ themselves are basic closed sets).



The map $\mathrm{Th}(a) \mapsto \mathrm{Th}(a)/{\sim}$ is a dense topological embedding of the space $\{\mathrm{Th}(a)\colon a \in \boldsymbol{A}\}$ in $\mathrm{Sto}(\mathfrak{C}/{\sim})$ since if $[\sigma] \neq \varnothing$, then $\sigma \nsim \bot$, so $\mathbf{Mod}(\sigma) \neq \varnothing$.

By the Prime Ideal Theorem, $\mathrm{Sto}(\mathfrak{C}/{\sim})$ is compact, since if

$$\bigcap_{\sigma\in\Gamma_0} [\sigma] \neq \varnothing$$

for all finite subsets Γ_0 of Γ, that is, all Γ_0 in $\mathscr{P}_\omega(\Gamma)$, then $\Gamma/{\sim}$ generates a proper filter of $\mathfrak{C}/{\sim}$, and so $\Gamma/{\sim}$ is included in a prime filter F, and then

$$F \in \bigcap_{\sigma\in\Gamma}[\sigma].$$

Thus (except it may not be Hausdorff) $\mathrm{Sto}(\mathfrak{C}/{\sim})$ is a compactification of $\{\mathrm{Th}(a) \colon a \in \boldsymbol{A}\}$. But the latter space can be any Kolmogorov space. In particular, it need not be compact.

3 Compactness Proved

Suppose $\Gamma \subseteq \mathrm{Sen}_{\mathscr{S}}$ and is **consistent,** that is, every finite subset has a model. Why has Γ a model?

Skolem's approach (1922), developed by Gödel (1930), and then by Mal′cev (1936, 1941) for the uncountable case, is first to show that, by adjusting $\mathscr{S}$, we may assume

- the sentences of Γ are in **Skolem normal form,** that is, prenex $\forall\exists$ form;
- $\mathscr{S}$ contains no operation symbols.

If a sentence

$$\forall x_0 \cdots \forall x_{m-1} \exists y_0 \cdots \exists y_{n-1} \varphi$$

has a model, then

- some structures with universe $n+1$, that is, $\{0, \ldots, n\}$, are models of
$$\exists \vec{y}\, \varphi(0, \ldots, 0, \vec{y});$$
- some of these embed in structures with universe $2n+1$ that are also models of
$$\exists \vec{y}\, \varphi(1, 0, \ldots, 0, \vec{y});$$
- some of these embed in structures with universe $3n+1$ that are also models of
$$\exists \vec{y}\, \varphi(1, 1, 0, \ldots, 0, \vec{y});$$
- and so on.

Thus we obtain an infinite, finitely branching tree of structures. By König's Lemma (1926), this tree has an infinite branch, whose union has universe ω and is a model of each sentence

$$\exists\vec{y}\ \varphi(a_0, \dots, a_{m-1}, \vec{y});$$

therefore it is a model of $\forall\vec{x}\ \exists\vec{y}\ \varphi$. Similarly we obtain a model of any consistent set of sentences, though the uncountable case will use the Axiom of Choice.

But suppose we do not want to change $\mathscr{S}$. We say that Γ

- is **complete** if it is consistent and, for all σ in $\mathrm{Sen}_{\mathscr{S}}$, contains σ or $\neg\sigma$;
- **has witnesses** if for every φ in $\mathrm{Fm}_{\mathscr{S}}(x)$ (the set of formulas of $\mathscr{S}$ whose only free variable is x), for some constant symbol c of $\mathscr{S}$,

$$\Gamma \vdash \exists x\ \varphi \to \varphi(c).$$

If it exists, a **canonical model** of Γ has universe consisting of the interpretations of the constant symbols of $\mathscr{S}$.

Theorem (Henkin 1949)**.** A complete subset of $\mathrm{Sen}_{\mathscr{S}}$ with witnesses has a canonical model.

Corollary (Mal′cev 1941)**.** The Compactness Theorem follows is a consequence of the Prime Ideal Theorem.[1]

Proof (Henkin). Given the consistent subset Γ of $\mathrm{Sen}_{\mathscr{S}}$, we can obtain a set C of new constant symbols and a bijection

$$\varphi \mapsto c_\varphi$$

[1] I cite Mal′cev as being the first to state the Compactness Theorem in full generality. I don't know that he referred to the Prime Ideal specifically; probably he did not.

from $\mathrm{Fm}_{\mathscr{S}(C)}(x)$ to C.[2] Now let

$$\Gamma^* = \Gamma \cup \{\exists x\, \varphi \to \varphi(c_\varphi) \colon \varphi \in \mathrm{Fm}_{\mathscr{S}(C)}(x)\}.$$

This is consistent and has witnesses, and the same is true of any complete subset of $\mathrm{Sen}_{\mathscr{S}(C)}$ that includes Γ^*. Such complete sets exist, by Lindenbaum's Theorem (on which there will be a tutorial); this theorem follows from the Prime Ideal Theorem. □

The converse is true by Henkin (1954).

Another corollary of the Canonical Model Theorem is

Łoś's Theorem (1955). Given a signature $\mathscr{S}$ and an index set I, suppose

- $(\mathfrak{A}_i \colon i \in I) \in \mathbf{Str}_{\mathscr{S}}{}^I$;
- $A = \prod_{i \in I} A_i$;
- for each i in I, $\mathfrak{A}_i{}^*$ is the expansion of $\mathfrak{A}_i$ to $\mathscr{S}(A)$ such that
$$a = (a_j \colon j \in I) \implies a^{\mathfrak{A}_i{}^*} = a_i;$$
- $\mathscr{U}$ is a prime filter, or **ultrafilter,** of $\mathscr{P}(I)$, that is,
$$\varnothing \notin \mathscr{U}, \quad I \in \mathscr{U}, \quad X \cup Y \in \mathscr{U} \iff X \in \mathscr{U} \text{ OR } Y \in \mathscr{U};$$
- for each σ in $\mathrm{Sen}_{\mathscr{S}(A)}$,
$$\|\sigma\| = \{i \in I \colon \mathfrak{A}_i{}^* \vDash \sigma\};$$
- $T = \{\sigma \in \mathrm{Sen}_{\mathscr{S}(A)} \colon \|\sigma\| \in \mathscr{U}\}$.

[2] I originally had the bijection from $\mathrm{Fm}_{\mathscr{S}}(x)$ to C, but then recognized out loud that this was not quite right. One will obtain C as a disjoint union $\bigcup_{n \in \omega} C_n$, with bijections from $\mathrm{Fm}_{\mathscr{S}}(x)$ to C_0, and $\mathrm{Fm}_{\mathscr{S}(C_0)}(x) \smallsetminus \mathrm{Fm}_{\mathscr{S}}(x)$ to C_1, and $\mathrm{Fm}_{\mathscr{S}(C_{n+1})}(x) \smallsetminus \mathrm{Fm}_{\mathscr{S}(C_n)}(x)$ to C_{n+2}.

By the Axiom of Choice, T has a canonical model, which is called the **ultraproduct** of $(\mathfrak{A}_i \colon i \in I)$ with respect to $\mathscr{U}$.

Proof. T is consistent because[3]

$$\begin{gathered}\sigma \in T \implies \|\sigma\| \in \mathscr{U} \implies \|\sigma\| \neq \varnothing,\\ X \in \mathscr{U} \;\&\; Y \in \mathscr{U} \implies X \cap Y \in \mathscr{U},\\ \|\sigma\| \cap \|\tau\| = \|\sigma \wedge \tau\|.\end{gathered}$$

Then T is complete because

$$\begin{gathered}X \notin \mathscr{U} \iff I \smallsetminus X \in \mathscr{U},\\ I \smallsetminus \|\sigma\| = \|\neg\sigma\|.\end{gathered}$$

Finally, T has witnesses because if $\varphi \in \mathrm{Fm}_{\mathscr{S}(A)}(x)$, then by the Axiom of Choice there is a, namely $(a_i \colon i \in I)$, in A such that

$$\mathfrak{A}_i{}^* \vDash \exists x\ \varphi \iff \mathfrak{A}_i{}^* \vDash \varphi(a_i) \iff \mathfrak{A}_i{}^* \vDash \varphi(a),$$

and thus

$$\begin{gathered}\|\exists x\ \varphi\| = \|\varphi(a)\|,\\ \|\exists x\ \varphi \to \varphi(a)\| = I,\\ T \vDash \exists x\ \varphi \to \varphi(a).\end{gathered}$$

□

Theorem (Halpern & Levy, 1971)**.** The Prime Ideal Theorem does not imply the Maximal Ideal Theorem.

Theorem (Hodges, 1979)**.** The Maximal Ideal Theorem implies the Axiom of Choice.

[3]Running out of time, I did not write down arguments for consistency and completeness.

Thus Łoś's Theorem is stronger than the Compactness Theorem. To prove the latter from the former, suppose every finite subset Δ of Γ has a model $\mathfrak{A}_\Delta$ (here we use the Axiom of Choice). Then there is an ultraproduct of $(\mathfrak{A}_\Delta \colon \Delta \in \mathscr{P}_\omega(\Gamma))$ that is a model of Γ. Indeed,[4] if we let

$$[\Delta] = \{\Theta \in \mathscr{P}_\omega(\Gamma) \colon \Delta \subseteq \Theta\},$$

then

$$[\Delta] \cap [\Theta] = [\Delta \cup \Theta],$$

so the $[\Delta]$ generate a proper filter of $\mathscr{P}(\mathscr{P}_\omega(\Gamma))$. This filter is therefore included in an ultrafilter, which yields the desired ultraproduct.

In case $\Gamma = \{\sigma_k \colon k \in \omega\}$, we get

$$\mathbf{Mod}(\sigma_0) \supseteq \mathbf{Mod}(\sigma_0 \wedge \sigma_1) \supseteq \mathbf{Mod}(\sigma_0 \wedge \sigma_1 \wedge \sigma_2) \supseteq \cdots$$

as in the Cantor Intersection Theorem.

[4] I did not have time to write down the details.

Bibliography

[1] C. C. Chang and H. J. Keisler. *Model theory.* North-Holland Publishing Co., Amsterdam, third edition, 1990. First edition 1973.

[2] John W. Dawson, Jr. The compactness of first-order logic: from Gödel to Lindström. *Hist. Philos. Logic*, 14(1):15–37, 1993.

[3] Richard Dedekind. *Essays on the Theory of Numbers. I: Continuity and Irrational Numbers. II: The Nature and Meaning of Numbers.* authorized translation by Wooster Woodruff Beman. Dover Publications Inc., New York, 1963.

[4] Kurt Gödel. The completeness of the axioms of the functional calculus of logic. In van Heijenoort [19], pages 582–91. First published 1930.

[5] J. D. Halpern and A. Lévy. The Boolean prime ideal theorem does not imply the axiom of choice. In *Axiomatic Set Theory (Proc. Sympos. Pure Math., Vol. XIII, Part I, Univ. California, Los Angeles, Calif., 1967)*, pages 83–134. Amer. Math. Soc., Providence, R.I., 1971.

[6] L. Henkin. Boolean representation through propositional calculus. *Fund. Math.*, 41:89–96, 1954.

[7] Leon Henkin. The completeness of the first-order functional calculus. *J. Symbolic Logic*, 14:159–166, 1949.

[8] Wilfrid Hodges. Krull implies Zorn. *J. London Math. Soc. (2)*, 19(2):285–287, 1979.

[9] Wilfrid Hodges. *Model theory*, volume 42 of *Encyclopedia of Mathematics and its Applications.* Cambridge University Press, Cambridge, 1993.

[10] Christopher Hollings. Embedding semigroups in groups: not as simple as it might seem. *Arch. Hist. Exact Sci.*, 68(5):641–692, 2014.

[11] Edmund Landau. *Foundations of Analysis. The Arithmetic of Whole, Rational, Irrational and Complex Numbers.* Chelsea Publishing Company, New York, N.Y., third edition, 1966. Translated by F. Steinhardt; first edition 1951; first German publication, 1929.

[12] Jerzy Łoś. Quelques remarques, théorèmes et problèmes sur les classes définissables d'algèbres. In *Mathematical interpretation of formal systems*, pages 98–113. North-Holland Publishing Co., Amsterdam, 1955.

[13] Anatoliĭ Ivanovič Mal′cev. A general method for obtaining local theorems in group theory. In *The metamathematics of algebraic systems. Collected papers: 1936–1967* [15], chapter 2, pages 15–21. First published in Russian in 1941.

[14] Anatoliĭ Ivanovič Mal′cev. Investigations in the realm of mathematical logic. In *The metamathematics of algebraic systems. Collected papers: 1936–1967* [15], chapter 1, pages 1–14. First published in German in 1936.

[15] Anatoliĭ Ivanovič Mal′cev. *The metamathematics of algebraic systems. Collected papers: 1936–1967.* North-Holland Publishing Co., Amsterdam-London, 1971. Translated, edited, and provided with supplementary notes by Benjamin Franklin Wells, III, Studies in Logic and the Foundations of Mathematics, Vol. 66.

[16] Giuseppe Peano. The principles of arithmetic, presented by a new method. In van Heijenoort [19], pages 83–97. First published 1889.

[17] Thoralf Skolem. Some remarks on axiomatized set theory. In van Heijenoort [19], pages 290–301. First published 1923.

[18] Alfred Tarski. Some notions and methods on the borderline of algebra and metamathematics. In Lawrence M. Graves et al., editors, *Proceedings of the International Congress of Mathematicians, Cambridge, Mass., 1950, vol. 1*, pages 705–720. Amer. Math. Soc., Providence, R. I., 1952. Held at Harvard University, Cambridge, Massachusetts.

[19] Jean van Heijenoort, editor. *From Frege to Gödel: A source book in mathematical logic, 1879–1931.* Harvard University Press, Cambridge, MA, 2002.

[20] Ludwig Wittgenstein. *Tractatus Logico-Philosophicus.* Routledge, London, 1922.

[21] Ernst Zermelo. Investigations in the foundations of set theory I. In van Heijenoort [19], pages 199–215. First published 1908.

[22] Max Zorn. A remark on method in transfinite algebra. *Bull. Amer. Math. Soc.*, 41(10):667–670, 1935.