İstanbul Model Theory Seminar Notes, Fall 2012

David Pierce

October 14, 2012

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

We are reading

Matthias Aschenbrenner, Alf Dolich, Deirdre Haskell, Dugald McPherson, and Sergei Starchenko, *Vapnik–Chervonenkis Density in Some Theories Without the Independence Property, I*, arXiv:1109.5438v1 [math.LO], [1].

These notes are based on:

1. Ayhan Günaydın's talk on October 4, 2012, in which all theorems in but Theorem 4 in the main text—§§1 and 2—were stated;
2. my own preparations to speak on October 11;
3. my experience in speaking then (I discussed §A first, and then garbled the proof of Theorem 1 in §1—the notes were OK, my use of them not);
4. a discussion of the role of the Axiom of Choice in mathematics: §B is meant to suggest that paying attention to the use of this Axiom may be worthwhile.

Contents

1. Combinatorics

Definition. Let $\mathscr{S}$ be a **set system,** that is, a set of sets. Its **Vapnik–Chervonenkis dimension,** denoted by

$$\mathrm{VC}(\mathscr{S}),$$

is the size of the largest finite set A such that

$$\{A \cap S \colon S \in \mathscr{S}\} = \mathscr{P}(A).$$

Here $\mathscr{S}$ is said to **shatter** the set A. If there is no bound on the size of A, then $\mathrm{VC}(\mathscr{S}) = \infty$.

There is no need to give a name to a set X such that $\mathscr{S} \subseteq \mathscr{P}(X)$. This set could just be $\bigcup \mathscr{S}$. Any set that is shattered by $\mathscr{S}$ is a subset of $\bigcup \mathscr{S}$.

Examples.

1. Let $\mathscr{S}$ be the set of half-lines $\{x \in \mathbb{R} \colon ax + b > 0\}$. Then $\mathscr{S}$ shatters $\{0, 1\}$ (and every other pair of real numbers), but no set of size 3. Thus $\mathrm{VC}(\mathscr{S}) = 2$. One should note that the elements of $\mathscr{S}$ are defined by two parameters.

2. Let $\mathscr{S}$ be the set of half-planes $\{(x, y) \in \mathbb{R}^2 \colon ax + by + c > 0\}$ (defined by three parameters). Then $\mathscr{S}$ shatters $\{(0, 0), (0, 1), (1, 0)\}$ (and every other set of three non-collinear points of $\mathbb{R}^2$), but no set of four points, so $\mathrm{VC}(\mathscr{S}) = 3$.

3. The set of rectangles $[a, b] \times [c, d]$ shatters $\{(\pm 1, 0), (0, \pm 1)\}$, so it has dimension at least 4.

Definition. Given the set-system $\mathscr{S}$ and a set A, we define

$$A \cap \mathscr{S} = \{A \cap S \colon S \in \mathscr{S}\}$$

If also $n \in \omega$, we define

$$\pi_{\mathscr{S}}(n) = \max\{|A \cap \mathscr{S}| \colon |A| = n\}.$$

Now we have that

$$\begin{aligned}\mathrm{VC}(\mathscr{S}) &= \max\{|A|\colon |A\cap\mathscr{S}| = 2^{|A|}\}\\ &= \max\{n\in\omega\colon \pi_{\mathscr{S}}(n) = 2^n\}.\end{aligned}$$

The following is apparently due to Sauer and Shelah independently. The proof is adapted from van den Dries [6].

Theorem 1. *If* $\mathrm{VC}(\mathscr{S}) = d$ *and* $d \leqslant n$, *then*

$$\pi_{\mathscr{S}}(n) \leqslant \sum_{i=0}^{d}\binom{n}{i}.$$

Proof. We show that, if the last inequality fails, then A itself has a subset larger than d that is shattered by $\mathscr{S}$, so $\mathrm{VC}(\mathscr{S}) > d$. That is, under the hypotheses

$$|A| = n, \qquad |A\cap\mathscr{S}| > \sum_{i=0}^{d}\binom{n}{i}$$

(which themselves entail $d < n$), the set A has a subset B such that

$$|B| = d+1, \qquad B\cap\mathscr{S} = \mathscr{P}(B).$$

This is easily true if $d = 0$ or $d = n-1$. Indeed, in case $d = 0$, the set $A\cap\mathscr{S}$ has an element b, so we can let $B = \{b\}$. In case $d = n-1$, then

$$|A\cap\mathscr{S}| \geqslant \sum_{i=0}^{n-1}\binom{n}{i} + 1 = 2^n,$$

so A itself is the desired set B.

We continue by induction on n. We have just treated the case where $n = d+1$. Suppose the claim holds when $n = m$, but now

$$|A| = m+1, \qquad |A\cap\mathscr{S}| > \sum_{i=0}^{d}\binom{m+1}{i}.$$

We may assume $0 < d < m$ (since we have taken care of the other cases). Let $b \in A$. If $(A \smallsetminus \{b\}) \cap \mathscr{S}$ is so large that

$$|(A \smallsetminus \{b\}) \cap \mathscr{S}| > \sum_{i=0}^{d} \binom{m}{i},$$

then by inductive hypothesis we are done. So suppose $(A \smallsetminus \{b\}) \cap \mathscr{S}$ is not so large. We make the analysis

$$(A \smallsetminus \{b\}) \cap \mathscr{S} = \mathscr{A}_1 \cup \mathscr{A}_2,$$

a disjoint union, where

$$X \in \mathscr{A}_2 \iff X \in A \cap \mathscr{S} \;\;\&\;\; X \cup \{b\} \in A \cap \mathscr{S}.$$

That is, $\mathscr{A}_2$ consists of the elements of $(A \smallsetminus \{b\}) \cap \mathscr{S}$ that have two pre-images in $A \cap \mathscr{S}$ under the map $Y \mapsto Y \smallsetminus \{b\}$; but each element of $\mathscr{A}_1$ has one pre-image. Then we compute

$$\begin{aligned}
|A \cap \mathscr{S}| &= |\mathscr{A}_1| + 2 \cdot |\mathscr{A}_2| \\
&= |\mathscr{A}_1| + |\mathscr{A}_2| + |\mathscr{A}_2| \\
&= |(A \smallsetminus \{b\}) \cap \mathscr{S}| + |\mathscr{A}_2|,
\end{aligned}$$

so that

$$\begin{aligned}
|\mathscr{A}_2| &= |A \cap \mathscr{S}| - |(A \smallsetminus \{b\}) \cap \mathscr{S}| \\
&\geqslant \sum_{i=0}^{d} \left(\binom{m+1}{i} - \binom{m}{i} \right) \\
&= \sum_{i=1}^{d} \binom{m}{i-1} \\
&= \sum_{i=0}^{d-1} \binom{m}{i}.
\end{aligned}$$

Now, since

$$|(A \smallsetminus \{b\}) \cap \mathscr{S}| \geqslant |\mathscr{A}_2|,$$

we have by inductive hypothesis that $A \smallsetminus \{b\}$ must have a subset C such that

$$|C| = d, \qquad\qquad C \cap \mathscr{S} = \mathscr{P}(C).$$

But we need, and have, more than this. We may assume

$$C \cap \mathscr{A}_2 = \mathscr{P}(C).$$

But each element of this set has two pre-images in $(C \cup \{b\}) \cap \mathscr{S}$ under $Y \mapsto Y \smallsetminus \{b\}$. That is, if $X \in C \cap \mathscr{A}_2$, then $(C \cup \{b\}) \cap \mathscr{S}$ contains both X and $X \cup \{b\}$. This ensures $(C \cup \{b\}) \cap \mathscr{S} = \mathscr{P}(C \cup \{b\})$, so we are done. $\square$

We can compute

$$\sum_{i=0}^{d} \binom{n}{i} = \frac{1}{d!} \cdot n^d + \text{lower terms}.$$

So if $\mathrm{VC}(\mathscr{S})$ is finite, $\pi_{\mathscr{S}}(n)$ is eventually bounded by a polynomial in n of degree no greater than $\mathrm{VC}(\mathscr{S})$.

Definition. If $\mathrm{VC}(\mathscr{S})$ is finite, then the **Vapnik–Chervonenkis density** of $\mathscr{S}$ is

$$\limsup_{n\to\infty} \frac{\log \pi_{\mathscr{S}}(n)}{\log n}.$$

This can be denoted by

$$\mathrm{vc}(\mathscr{S}).$$

Then $\mathrm{vc}(\mathscr{S}) \leqslant \mathrm{VC}(\mathscr{S})$.

2. Logic

The examples above are of the form where

$$\mathscr{S} = \{\varphi^{\mathfrak{M}}(x, b) \colon b \in M_y\} \tag{$*$}$$

for some formula $\varphi(x,y)$ in the signature of a structure $\mathfrak{M}$. Here x and y are tuples of variables, and M_y is the set of tuples of elements of M indexed by the entries in y, so that[1]

$$\varphi^{\mathfrak{M}}(x,b) = \{a \in M_x : \mathfrak{M} \models \varphi(a,b)\}.$$

Definition. When $\mathscr{S}$ is as in $(*)$, we may write φ in place of $\mathscr{S}$ in compound symbols; so $\mathrm{VC}(\varphi)$ means $\mathrm{VC}(\mathscr{S})$, and so on.

The set-system $\mathscr{S}$ as in $(*)$ shatters a subset A of M_x if and only if, for all subsets C of A, there is an element b_C of M_y such that

$$C = A \cap \varphi^{\mathfrak{M}}(x, b_C),$$

that is, for all a in A,

$$a \in C \iff \mathfrak{M} \models \varphi(a, b_C).$$

If $\mathscr{S}$ shatters all finite subsets of M_x, this is equivalent to the model-theoretic property of φ called *NIP,* defined as follows.

Definition. Let T be a complete theory. We say that $\varphi(x,y)$ has the **Independence Property** in T if for all n in ω,

$$T \vdash \exists(x_i : i < n)\ \exists(y_W : W \in \mathscr{P}(n)) \bigwedge_{i<n} \bigwedge_{W\subseteq n} \varphi^{i\in W}(x_i, y_W),$$

where

$$\varphi^{i\in W} \text{ is } \begin{cases} \varphi, & \text{if } i \in W, \\ \neg\varphi, & \text{if } i \notin W. \end{cases}$$

If $\varphi(x,y)$ does *not* have the Independence Property, it has **NIP.** The theory T itself has the Independence Property, if some formula has the Independence Property in T. Otherwise T is NIP.

Some sources may reverse the roles of x and y in the definition. This does not matter, as the definition is symmetric, by the next theorem below.

[1] This notation readily allows $\mathfrak{M}$ to have several sorts. If $x = (x_i : i < m)$, then elements of M_x are of the form $(a_i : i < m)$, where a_i belongs to the sort that x_i ranges over. For the set $\varphi^{\mathfrak{M}}(x,b)$, the paper writes $\varphi^{\mathfrak{M}}(M_x,b)$, but this seems needless, unless one wants to be able to write $\varphi^{\mathfrak{M}}(A,b)$ for $A \cap \varphi^{\mathfrak{M}}(x,b)$.

Examples. These are from Poizat [4].

1. If T is the theory of $(\mathbb{N}, \mid)$, then $x \mid y$ has the independence property in T. For, if $(p_i : i \in \omega)$ is the sequence of primes, and for all finite subsets W of ω, if

$$b_W = \prod_{i \in W} p_i,$$

then for all i in ω, for all W in $\mathscr{P}_\omega(\omega)$ (the set of finite subsets of ω),

$$T \vdash p_i \mid b_W \iff i \in W.$$

2. If T is the theory of an infinite Boolean algebra, then $x \leqslant y$ has the independence property in T. For, in every model of T, there is a sequence $(a_i : i \in \omega)$ of elements such that

$$i \neq j \implies a_i \wedge a_j = 0.$$

If $W \in \mathscr{P}_\omega(\omega)$, let

$$b_W = \bigvee_{i \in W} a_i.$$

Then

$$T \vdash a_i \leqslant b_W \iff i \in W.$$

Theorem 2. *The definition of the Independence Property is symmetric in x and y.*

Proof. Suppose $\varphi(x, y)$ has the Independence Property in T. Then for all n in ω,

$$T \vdash \exists(x_W : W \subseteq n)\ \exists(y_{\mathcal{V}} : \mathcal{V} \subseteq \mathscr{P}(n)) \bigwedge_{W \subset n} \bigwedge_{\mathcal{V} \subseteq \mathscr{P}(n)} \varphi^{W \in \mathcal{V}}(x_W, y_{\mathcal{V}}).$$

Now we can take away all but n of the $\mathcal{V}$. In particular, if $i < n$, we let

$$\mathcal{V}(i) = \{W \in \mathscr{P}(n) : i \in W\}.$$

Now we take away all of the $\mathcal{V}$, except for the $\mathcal{V}(i)$; and we can use i in place of $\mathcal{V}(i)$ as an index. This leaves us with

$$T \vdash \exists(x_W : W \in \mathscr{P}(n))\ \exists(y_i : i < n) \bigwedge_{W \in \mathscr{P}(n)} \bigwedge_{i < n} \varphi^{i \in W}(x_W, y_i). \quad \square$$

The Independence Property is a special case of a more general property:

Definition. We say that $\varphi(x, y)$ has the **Order Property** in T if for all n in ω,

$$T \vdash \exists(x_i : i < n)\, \exists(y_j : j < n) \bigwedge_{i<n} \bigwedge_{j<n} \varphi^{i<j}(x_i, y_j).$$

In the definition, by re-indexing, we can replace the condition $i < j$ with $i \leqslant j$ or $i \geqslant j$. In particular, the definition is symmetric in x and y. Also we have the following.

Theorem 3. *If $\varphi(x, y)$ has the Independence Property in T, then it has the Order Property in T.*

Proof. In the definition of the Independence Property, throw out all W except those that are elements of n. □

Example. If T is the theory of $(\mathbb{Q}, <)$, then $x < y$ has the Order Property in T, but not the Independence Property. (Moreover, no formula has the Independence Property in T; but this would take work to prove.)

Another way to define the Independence Property is given by the following. We shall use it to prove the theorem after this one. (The development is based on Poizat [3, 4].)

Theorem 4. *The following are equivalent.*

1. *The formula $\varphi(x, y)$ has the Independence Property in T.*
2. *In some model $\mathfrak{N}$ of T, there is an indiscernible sequence $(a_n : n \in \omega)$ and an element b such that both $\{n \in \omega : \mathfrak{N} \models \varphi(a_n, b)\}$ and $\{n \in \omega : \mathfrak{N} \models \neg\varphi(a_n, b)\}$ are cofinal in ω.*
3. *In some model $\mathfrak{N}$ of T, there is an indiscernible sequence $(a_n : n \in \omega)$ and an element b such that, for all n in ω,*

$$\mathfrak{N} \models \varphi^{2|n}(a_n, b).$$

Proof. Condition (3) is a special case of (2), that is, (3)⇒(2); the converse follows by throwing out terms of the indiscernible sequence and re-indexing.

(3)⇒(1). For every n in ω, for every W in $\mathscr{P}(n)$, there is σ in ${}^n\omega$ such that

$$\sigma(0) < \cdots < \sigma(n-1), \qquad\qquad 2 \mid \sigma(i) \iff i \in W.$$

By (3) then, we have

$$\mathfrak{N} \models \exists y \bigwedge_{i<n} \varphi^{i\in W}(a_{\sigma(i)}, y),$$

so by indiscernibility

$$\mathfrak{N} \models \exists y \bigwedge_{i<n} \varphi^{i\in W}(a_i, y).$$

(1)⇒(3). Under the hypothesis, by Compactness, in some model $\mathfrak{M}$ of T, there are collections $(a'_n : n \in \omega)$ and $(b_W : W \in \mathscr{P}(\omega))$ such that

$$\mathfrak{M} \models \varphi(a'_n, b_W) \iff n \in W.$$

Let $\mathfrak{N}$ be an elementary extension of $\mathfrak{M}$, and let $\mathscr{U}$ be an ultrafilter on ω. For every subset A of N, the set of formulas $\psi(x)$ over $M \cup A$ such that

$$\{n \in \omega : \mathfrak{N} \models \psi(a'_n)\} \in \mathscr{U}$$

is a complete type; for,

- it is closed under conjunction (since $\mathscr{U}$ is closed under intersection),
- it contains $\psi(x)$ or $\neg\psi(x)$, for every $\psi(x)$ over $M \cup A$ (since $\mathscr{U}$ contains W or $\omega \smallsetminus W$, for every W in $\mathscr{P}(\omega)$),
- every element is satisfied in $\mathfrak{N}$.

Assume further that $\mathscr{U}$ is nonprincipal, so it contains all cofinite subsets of ω. Then the type just defined is not realized by any element of $M \cup A$. (If it were realized by c in this set, then it would contain the formula $x = c$, and then $c = a'_n$ for some n. This n would be unique, since all of the a'_m are distinct; but then $\{n\} \in \mathscr{U}$.)

Assume further that $\mathfrak{N}$ is $|M|$-saturated. We can now obtain a sequence $(a_n : n \in \omega)$ from $\mathfrak{N}$ such that for all n in ω, for all formulas $\psi(x)$ over $M \cup \{a_i : i < n\}$,

$$\mathfrak{N} \models \psi(a_n) \iff \{k \in \omega : \mathfrak{N} \models \psi(a'_k)\} \in \mathscr{U}.$$

Indeed, a_n just realizes the appropriate complete type over $M \cup \{a_i : i < n\}$. Also a_n does not belong to this set. Moreover, each a_{n+i} realizes this type. Therefore the sequence $(a_n : n \in \omega)$ is indiscernible over M.

Suppose for some ψ over M we have

$$\mathfrak{N} \models \neg\psi(a_0, \dots, a_n).$$

Then by definition of a_n we must have, for some k in ω,

$$\mathfrak{N} \models \neg\psi(a_0, \dots, a_{n-1}, a'_k).$$

Continuing in this matter, we obtain a sequence σ on $n + 1$ such that

$$\mathfrak{N} \models \neg\psi(a'_{\sigma(0)}, \dots, a'_{\sigma(n)}).$$

Now consider the contrapositive of this result. By hypothesis, for all m in ω, for all sequences σ on $2m$, we have

$$\mathfrak{N} \models \exists y \bigwedge_{i|2m} \varphi^{2|i}(a'_{\sigma(i)}, y).$$

Therefore

$$\mathfrak{N} \models \exists y \bigwedge_{i|2m} \varphi^{2|i}(a_i, y).$$

By saturation, b exists as desired. $\square$

Condition (2) in the theorem is that $\varphi(x, b)$ **splits** the indiscernible sequence $(a_n : n \in \omega)$.

Theorem 5. *If a complete theory T has the Independence Property, then a formula $\varphi(x, y)$, where $|x| = 1$, has the Independence Property in T.*

Proof. Let $|x| = 1$, and suppose y is minimal such that some formula $\varphi((x,y),z)$ has the Independence Property in T. By the symmetry of this property, and the last theorem, there is a model $\mathfrak{M}$ of T and an indiscernible sequence $(a_n\colon n \in \omega)$ and (b,c) from $\mathfrak{M}$ such that

$$\mathfrak{M} \models \varphi^{2|n}(b,c,a_n).$$

We shall show that $(a_n\colon n \in \omega)$ is indiscernible over c. In that case, by the last theorem, $\varphi(x,(y,z))$ has the Independence Property in T.

We use induction. Let T_m be the theory that entails:

- T itself;
- the indiscernibility of $(a_n\colon n \in \omega)$;
- the sentences $\varphi^{2|n}(b,c,a_n)$;
- the sentences

  $$\theta(c,a_{\sigma(0)},\dots,a_{\sigma(m-1)}) \leftrightarrow \theta(c,a_0,\dots,a_{m-1}),$$

  where $\sigma(0) < \dots < \sigma(m-1)$ (and θ has no parameters).

So $T_0 \subseteq T_1 \subseteq \dots$ We want to show that $\bigcup_{m\in\omega} T_m$ is consistent.

By hypothesis, T_0 is consistent.

Suppose T_m is consistent. Supposing $\sigma \in {}^m\omega$, and $\sigma(0) < \dots < \sigma(m-1)$, we let

$$a_\sigma = (a_{\sigma(0)},\dots,a_{\sigma(m-1)}).$$

Then $\bigl((a_\sigma,a_n)\colon \sigma(m-1) < n < \omega\bigr)$ is an indiscernible sequence. Therefore, by the minimality of y, no formula $\psi(c,a_\sigma,y)$ splits the sequence $(a_n\colon n \in \omega)$, but either all a_n for n sufficiently large satisfy the formula, or all such a_n satisfy its negation.

Therefore, by throwing out some indices (always in adjacent pairs, to maintain parity), we may assume that $\psi^{\mathfrak{M}}(c,a_\sigma,y)$ includes the whole set $\{a_n\colon \sigma(m-1) < n\}$, or else $\neg\psi^{\mathfrak{M}}(c,a_\sigma,y)$ does.

By Compactness then, there is a complete type p_σ in $(z_0,\dots,z_m)$ over c that is realized by each (a_σ,a_n) such that $\sigma(m-1) < n$.

Then again by Compactness, in some sufficiently saturated model of T_m, there must be some a_ω such that the sequence $(a_n\colon n < \omega+1)$ is indiscernible, and for each σ as above, (a_σ,a_ω) realizes p_σ.

Suppose σ_k are chosen in ${}^m\omega$ for each k in ω, so that

$$\sigma_0(0) < \cdots < \sigma_0(m-1) < \sigma_1(0) < \cdots$$

Then the sequence $\big((a_{\sigma_n}, a_\omega) \colon n \in \omega)\big)$ is indiscernible. Again by minimality of y, no formula $\theta(c, z_0, \ldots, z_m)$ can split the sequence. Therefore, by throwing out some indices again, we may assume that either $\theta(c, z_0, \ldots, z_m)$ or $\neg\theta(c, z_0, \ldots, z_m)$ is, for all σ, contained in p_σ.

Then all of the p_σ are the same type. Thus T_{m+1} is consistent. This completes the induction and the proof. □

We can now conclude every weakly o-minimal theory (that is, the theory of a linearly ordered structure in which the definable singulary relations are finite unions of convex sets) is NIP. In such a theory, we aim to show

$$\mathrm{vc}(\varphi(x, y)) \leqslant |y|.$$

If, instead, the theory is that of $\mathbb{Q}_p$, we aim to show

$$\mathrm{vc}(\varphi(x, y)) \leqslant 2 \cdot |y| - 1.$$

A. Compactness

The recent model-theory text of Tent and Ziegler [5] introduces the Compactness Theorem as follows:

> Its name is motivated by the results in Section 4.2 which associate to each theory a certain compact topological space.
>
> We call a theory T *finitely satisfiable* if every finite subset of T is consistent.
>
> **Theorem 2.2.1** (Compactness Theorem). *Finitely satisfiable theories are consistent.*[2]

The section referred to begins:

[2]The quotation is taken from what is called an early second edition, distributed to the Seminar by email as a `pdf` file.

We now endow the set of types of a given theory with a topology. The Compactness Theorem 2.2.1 then translates into the statement that this topology is compact, whence its name.

Fix a theory T. An n-type is a maximal set of formulas $p(x_1, \ldots, x_n)$ consistent with T. We denote by $\mathrm{S}_n(T)$ the set of all n-types of T. We also write $\mathrm{S}(T)$ for $\mathrm{S}_1(T)$.[3] [...]

Remark. The Stone duality theorem asserts that the map

$$X \mapsto \{C \mid C \text{ clopen subset of } X\}$$

yields an equivalence between the category of 0-dimensional compact spaces and the category of Boolean algebras. The inverse map assigns to every Boolean algebra $\mathcal{B}$ its Stone space $\mathrm{S}(\mathcal{B})$, the set of all ultrafilters (see Exercise 1.2.4) of $\mathcal{B}$. For more on Boolean algebras see [Givant and Halmos, *Introduction to Boolean algebras*].

Nothing is incorrect here. But one might be led to believe that the type spaces are *by definition* Stone spaces of Boolean algebras of logically equivalent formulas. Since Stone spaces are always compact, the type spaces are compact, and one might then conclude that the Compactness Theorem follows. But this would be a wrong conclusion, since this theorem fails in second-order logic, and yet Stone spaces of algebras of second-order formulas are still compact.

By the definition above, the type spaces are *dense subspaces* of certain Stone spaces. The Compactness Theorem is that these subspaces are compact. Since Stone spaces are Hausdorff, the type spaces must then be closed; therefore they are the whole Stone spaces. The point of this section is to spell out the details of these observations.

Fix some logic L that extends ordinary propositional logic: it could be a first-order logic, a second-order logic, or something else. There is a class **Mod** of structures interpreting L, and a set Sn of sentences of L. If $\sigma \in \mathrm{Sn}$, we can define

$$\mathbf{Mod}(\sigma) = \{\mathfrak{A} \in \mathbf{Mod} \colon \mathfrak{A} \models \sigma\}.$$

[3] $\mathrm{S}_0(T)$ can be considered as the set of all complete extensions of T, up to equivalence. [Footnote in source.]

If $\mathfrak{A} \in \mathbf{Mod}$, we can define

$$\mathrm{Th}(\mathfrak{A}) = \{\sigma \in \mathrm{Sn} \colon \mathfrak{A} \models \sigma\}.$$

If $\Gamma \subseteq \mathrm{Sn}$ and $\mathcal{K} \subseteq \mathbf{Mod}$, we define

$$\mathbf{Mod}(\Gamma) = \bigcap_{\sigma \in \Gamma} \mathbf{Mod}[\sigma], \qquad \mathrm{Th}(\mathcal{K}) = \bigcap_{\mathfrak{A} \in \mathcal{K}} \mathrm{Th}(\mathfrak{A}).$$

The classes $\mathbf{Mod}(\Gamma)$ are **elementary classes** (though usually this term assumes a first-order logic); the classes $\mathrm{Th}(\mathcal{K})$ are **theories.** The functions $\Gamma \mapsto \mathbf{Mod}(\Gamma)$ and $\mathcal{K} \mapsto \mathrm{Th}(\mathcal{K})$ determine a *Galois correspondence* between the theories and the elementary classes.[4]

Moreover, since

$$\mathbf{Mod}(\sigma \vee \tau) = \mathbf{Mod}(\sigma) \cup \mathbf{Mod}(\tau),$$

the elementary classes are the closed *classes* of a *topology* on **Mod** with basis consisting of the closed classes $\mathbf{Mod}(\sigma)$.

To say that the logic L has a **compactness theorem** is to say that if $\Gamma \subseteq \mathrm{Sn}$ and every finite subset of Γ has a model, then Γ itself has a model. But this just means that if $\{\mathbf{Mod}(\sigma) \colon \sigma \in \Gamma\}$ has the Finite Intersection Property, then $\bigcap_{\sigma \in \Gamma} \mathbf{Mod}(\sigma) \neq \varnothing$: that is, **Mod** is compact as a topological space.

A similar Galois correspondence arises in algebraic geometry. Suppose L/K is a field-extension, and X is an n-tuple of variables. If $f \in K[X]$, define

$$\mathrm{V}(f) = \{a \in L^n \colon f(a) = 0\}.$$

If $a \in L^n$, define

$$\mathrm{I}(a) = \{f \in K[X] \colon f(a) = 0\}.$$

[4]Evry relation R between sets or classes A and B induces a Galois correspondence between certain subsets of A and of B. In one direction this one-to-one, order-reversing correspondence is $X \mapsto \bigcap_{x \in X} \{y \in B \colon x \mathrel{R} y\}$. The original Galois correspondence is induced by the relation $\{(a, \sigma) \in K \times \mathrm{Aut}(K) \colon \sigma(a) = a\}$, where K is a field.

If $A \subseteq K[X]$ and $B \subseteq L^n$, define

$$\mathrm{V}(A) = \bigcap_{f \in A} \mathrm{V}(f), \qquad\qquad \mathrm{I}(B) = \bigcap_{a \in B} \mathrm{I}(a).$$

The sets $\mathrm{V}(A)$ are **algebraic sets** over K. The sets $\mathrm{I}(B)$ are radical ideals, but perhaps not every radical ideal of $K[X]$ is of this form, unless L is algebraically closed. In any case, there is a Galois correspondence between the algebraic sets and the radical ideals of the form $\mathrm{I}(B)$. Moreover, since

$$\mathrm{V}(fg) = \mathrm{V}(f) \cup \mathrm{V}(g),$$

the sets $\mathrm{V}(f)$ compose a basis of closed sets for a topology on L^n, the **Zariski topology,** in which the closed sets are just the algebraic sets. (Strictly, there is a Zariski topology for every subfield K of L.)

The radical ideals $\mathrm{I}(a)$ are prime ideals; but not necessarily every prime ideal is of this form, unless we have both that L is algebraically closed, *and* that the transcendence-degree of L/K is at least n.

Let us suppose this is so. If we identify points a and b of L^n if $\mathrm{I}(a) = \mathrm{I}(b)$, then the space becomes the **spectrum** of $K[X]$: the corresponding topological space whose underlying set consists of the prime ideals of $K[X]$. The spectrum need not be Hausdorff, since it is possible to have $\mathrm{I}(a) \subset \mathrm{I}(b)$, so that every closed set that contains $\mathrm{I}(a)$ contains $\mathrm{I}(b)$, but not conversely.[5] The spectrum is however compact, since L^n itself is compact. Indeed, suppose a collection $\{\mathrm{V}(f) \colon f \in A\}$ of basic closed subsets of L^n has the Finite Intersection Property. Since

$$\mathrm{V}(f) \cap \mathrm{V}(g) = \mathrm{V}(f, g),$$

the set A must generate a proper ideal of $K[X]$. This ideal then is included in a prime ideal $\mathrm{I}(a)$, so $a \in \bigcap_{f \in A} \mathrm{V}(f)$.

In the logical situation, we identify σ and τ if $\mathbf{Mod}(\sigma) = \mathbf{Mod}(\tau)$. Then Sn becomes a Boolean algebra in the usual way. Every subset $\mathrm{Th}(\mathcal{K})$ of Sn can now be understood as a *filter* of this algebra; and every subset $\mathrm{Th}(\mathfrak{A})$, as an *ultrafilter.* (Note that $\mathrm{Th}(\varnothing)$ is the improper filter Sn.)

[5] Of course the symbol $\subset$ here is to $\subseteq$ as $<$ is to $\leqslant$. Two errors of TEX are that `\subset` gives $\subset$ and not $\subseteq$, and `\leq` and `\le` give $\leq$ and not $\leqslant$.

However, not every ultrafilter of Sn need be the theory of some structure. For, such an ultrafilter is just a subset $\mathscr{U}$ with two properties:

1. Every finite subset of $\mathscr{U}$ has a model.
2. If $\sigma \notin \mathscr{U}$, then $\neg\sigma \in \mathscr{U}$.

In the second-order logic for $(\mathbb{N}, 1, x \mapsto x+1)$ with an additional constant-symbol c, the Peano axioms, together with the sentences $c \neq 1$, $c \neq 2$, and so on, are included in a proper filter, and therefore an ultrafilter; but they have no model.

In general, if two structures have the same theory, we may say the structures are **elementarily equivalent,** though again this term is usually reserved for first-order logic. We may denote the relation by $\equiv$. As in algebraic geometry, we may now consider $\mathbf{Mod}/\equiv$ instead of $\mathbf{Mod}$ itself. The points of $\mathbf{Mod}/\equiv$ can be considered as the theories of structures; that is, we assume

$$(\mathbf{Mod}/\equiv) = \{\mathrm{Th}(\mathfrak{A}) \colon \mathfrak{A} \in \mathbf{Mod}\}.$$

Then $\mathbf{Mod}/\equiv$ is a subspace of the Stone space $\mathrm{S}(\mathrm{Sn})$ of ultrafilters of Sn. We have seen that it may be a proper subspace.

However, it is a *dense* subspace. For, the basic closed subsets of $\mathrm{S}(\mathrm{Sn})$ are the subsets $[\sigma]$, where $\sigma \in \mathrm{Sn}$ and

$$[\sigma] = \{U \in \mathrm{S}(\mathrm{Sn}) \colon \sigma \in U\}.$$

(Here U stands for ultrafilter. Again, it is not necessarily the complete theory of some structure. Therefore $\sigma \in U$ should not be written as $U \vdash \sigma$.) Since

$$[\neg\sigma] = \mathrm{S}(\mathrm{Sn}) \smallsetminus [\sigma],$$

the basic closed sets are also basic open sets. If $U \in [\sigma]$, then $\sigma \neq \bot$, so σ has a model $\mathfrak{A}$, and then $\mathrm{Th}(\mathfrak{A}) \in [\sigma]$. Thus $\mathbf{Mod}/\equiv$ is dense in $\mathrm{S}(\mathrm{Sn})$.

The Stone space of a Boolean algebra can be identified with the spectrum of the corresponding Boolean ring. This is because prime ideals of a Boolean ring are maximal and are the duals of ultrafilters: If $\mathfrak{p}$ is a prime ideal, then $\{\neg x \colon x \in \mathfrak{p}\}$ is an ultrafilter. This ultrafilter is also the complement of $\mathfrak{p}$.

In particular, Stone spaces are compact. They are also Hausdorff, so that compact subspaces are closed. Therefore the following are equivalent:

1. L has a compactness theorem,
2. $\mathbf{Mod}/{\equiv}$ is compact,
3. $\mathbf{Mod}/{\equiv}$ is a closed subspace of $\mathrm{S}(\mathrm{Sn})$,
4. $\mathbf{Mod}/{\equiv}$ is all of $\mathrm{S}(\mathrm{Sn})$.

In case L is a first-order logic, we can give direct proofs of (2) and (3). (Poizat gives them both.) We use **ultraproducts** in each case, and **Łoś's Theorem.** Specifically, for every indexed family $(\mathfrak{A}_i \colon i \in \Omega)$ of structures in $\mathbf{Mod}$, for every ultrafilter $\mathscr{U}$ on Ω, there is a structure $\mathfrak{A}$ such that, for all σ in Sn,

$$\mathfrak{A} \models \sigma \iff \{i \in \Omega \colon \mathfrak{A}_i \models \sigma\} \in \mathscr{U}. \tag{†}$$

In fact, $\mathfrak{A}$ can be taken as the ultraproduct denoted by

$$\prod_{i \in \Omega} \mathfrak{A}_i / \mathscr{U};$$

and when one proves (†), one will allow σ to have constants $(a_i \colon i \in \Omega)$ from $\prod_{i \in \Omega} A_i$, interpreted in each $\mathfrak{A}_i$ as a_i. This does not really give a more general result, since we can now go back and assume those constants were part of the language from the beginning.

Proof of compactness. Write $[\sigma]$ for $\{T \in \mathbf{Mod}/{\equiv} \colon \sigma \in T\}$. Suppose the collection $\{[\sigma] \colon \sigma \in B\}$ has the Finite Intersection Property. Then it generates a proper filter of subsets of $\mathbf{Mod}/{\equiv}$, so it is included in an ultrafilter $\mathscr{U}$ on $\mathbf{Mod}/{\equiv}$. If $T \in \mathbf{Mod}/{\equiv}$, then T has a model $\mathfrak{A}_T$. Let $\mathfrak{A}$ be the ultraproduct

$$\prod_{T \in \mathbf{Mod}/\equiv} \mathfrak{A}_T / \mathscr{U}.$$

Suppose $\sigma \in B$, so that $[\sigma] \in \mathscr{U}$. We have

$$\{T \in \mathbf{Mod}/{\equiv} \colon \mathfrak{A}_T \models \sigma\} = \{T \in \mathbf{Mod}/{\equiv} \colon \sigma \in T\} = [\sigma].$$

By Łoś's Theorem, $\mathfrak{A} \models \sigma$, so $\mathrm{Th}(\mathfrak{A}) \in [\sigma]$. So $\mathbf{Mod}/{\equiv}$ is compact. □

We can streamline the proof by using that an arbitrary topological space is compact if and only if every ultrafilter on the underlying set includes the filter of neighborhoods of a point.[6] Then we can just start the proof with $\mathscr{U}$.

Moreover, using this criterion for compactness, we have a neat proof that the Stone space $\mathrm{S}(B)$ of an arbitrary Boolean algebra B is compact: For, if $\mathscr{U}$ is an ultrafilter on $\mathrm{S}(B)$, then it converges to the point

$$\{x \in B \colon [x] \in \mathscr{U}\},$$

where $[x] = \{F \in \mathrm{S}(B) \colon x \in F\}$. To see this, first note that the given 'point' is indeed a filter F on B, because the map $x \mapsto [x]$ from B to $\mathscr{P}(\mathrm{S}(B))$ is a Boolean algebra homomorphism; F is then an ultrafilter on B, because the homomorphism is injective. Finally, if $F \in [x]$, this just means $x \in F$, so $[x] \in \mathscr{U}$.

In this last proof, we can replace $\mathrm{S}(B)$ with an arbitrary subset Ω of it. We obtain that an ultrafilter $\mathscr{U}$ on Ω converges to the point

$$\{x \in B \colon [x] \cap \Omega \in \mathscr{U}\}$$

of $\mathrm{S}(B)$. The proof goes through as before, except that one needs to note also

$$[x] \cap \Omega \notin \mathscr{U} \iff [\neg x] \cap \Omega \in \mathscr{U}.$$

Now consider the case where B is Sn, and Ω is $\mathbf{Mod}/{\equiv}$. The limit of $\mathscr{U}$ is precisely the theory of the ultraproduct $\mathfrak{A}$ that we found above.

Indeed, Łoś's Theorem can be understood as being the statement that the limit of $\mathscr{U}$ is indeed the theory of this structure. In the original statement of the theorem, the indices are arbitrary; but we could treat the index of $\mathfrak{A}_i$ as $\mathrm{Th}(\mathfrak{A}_i)$ itself. We may have wanted $\mathfrak{A}_i$ and $\mathfrak{A}_j$ to be the same structure, or just to have the same theory, even though $i \neq j$; but we can deal with this by expanding the language.

In short, seen in the right light, the Compactness Theorem of first-order logic and Łoś's Theorem are the same, except that the latter theorem actually gives you the desired model.

As noted, we can also show directly that $\mathbf{Mod}/{\equiv}$ is closed:

[6] The topological reference that I happen to have on hand is Willard [7].

Proof of closedness. Let $U \in \mathrm{S}(\mathrm{Sn})$. Every element σ of U has a model $\mathfrak{A}_\sigma$. Also, the subsets $\{\tau\colon \tau \leqslant \sigma\}$ of U generate a proper filter, since

$$\{\tau\colon \tau \leqslant \sigma\} \cap \{\rho\colon \rho \leqslant \sigma\} = \{\tau\colon \tau \leqslant \sigma \wedge \rho\}.$$

(Remember that $\tau \leqslant \sigma$ means every model of τ is a model of σ; we can write this also as $\tau \vdash \sigma$.) Let $\mathscr{U}$ be an ultrafilter on U that includes this filter. Then the ultraproduct $\prod_{\sigma \in U} \mathfrak{A}_\sigma/\mathscr{U}$ is a model of U, since

$$\{\tau\colon \mathfrak{A}_\tau \models \sigma\} \supseteq \{\tau\colon \tau \leqslant \sigma\}. \qquad \square$$

B. Choice and Determinacy

In these notes, the Axiom of Choice has been tacitly assumed. The purpose of this section is to suggest that this Axiom is not 'obviously' or 'intuitively' correct, since it contradicts another set-theoretic axiom that might be considered 'obviously' or 'intuitively' correct. That axiom is the Axiom of Determinacy, according to which, in certain *games* of infinite length, one of the players always has a winning strategy.

We consider games with two players. Hodges [2] calls these players $\forall$ and $\exists$, after Abelard and Eloise; but I propose to call them simply 0 and 1, for notational purposes. A **game** that 0 and 1 can play is determined by a partition $A_0 \amalg A_1$ of the ${}^{\omega}2$ of binary sequences on ω. A particular **play** of the game can be analyzed as a sequence of **rounds,** indexed by ω. In round m, player 0 chooses an element a_{2m} of 2; this is the **move** of 0 in this round. Then player 1 moves by choosing an element a_{2m+1} of 2. The play itself is then the sequence $(a_n : n \in \omega)$ or a, which is an element of ${}^{\omega}2$. The play is **won** by that player e such that $a \in A_e$; and then player $1 - e$ has **lost.**

Each player e may use a **strategy,** namely a function f_e from $\bigcup_{m \in \omega} {}^{m+e}2$ to 2. (So f_e assigns to each finite binary sequence an element of 2; but f_1 need make no assignment to the empty sequence.) If both f_0 and f_1 are chosen, then a play is determined, namely the sequence $(a_n : n \in \omega)$ given by

$$a_{2m} = f_0(a_1, a_3, \dots, a_{2m-1}), \qquad a_{2m+1} = f_1(a_0, a_2, \dots, a_{2m}),$$

or simply by

$$a_{2m+e} = f_e(a_{1-e}, a_{3-e}, \dots, a_{2m-1+e}).$$

That is, f_e determines the move of player e from the previous moves by the *other* player. The player's own previous moves need not be formally considered, since they themselves were already determined by the player's strategy and the other player's previous moves.

Suppose player $1 - e$ has chosen strategy f_{1-e}. For every b in ${}^{\omega}2$, player e might choose a strategy f_e that is constant on each set ${}^{m+e}2$, having the value b_m there. The resulting play will be a, where

$$a_{2m+1-e} = f_{1-e}(b_0, b_1, \dots, b_{m-e}), \qquad\qquad a_{2m+e} = b_m.$$

This shows that, for every choice of f_{1-e}, there are continuum-many plays that can result if player $1 - e$ uses this strategy.

If, using a strategy f_e, player e wins all plays of a game, then f_e is a **winning** strategy for that game. The game is **determined** if one of the players has a winning strategy. The **Axiom of Determinacy** is that in every game, one of the players has a winning strategy: in other words, for every choice of the A_e, one of the following sentences of infinitary logic is true:

$$\exists x_0\ \forall x_1\ \exists x_2\ \cdots\ (x_0, x_1, x_2, \dots) \in A_0,$$
$$\forall x_0\ \exists x_1\ \forall x_2\ \cdots\ (x_0, x_1, x_2, \dots) \in A_1.$$

However, this Axiom is false under the assumption of the Axiom of Choice, or more precisely under the assumption that the Continuum can be well-ordered, so that there is a least ordinal, 2^{ω}, whose cardinality is that of ${}^{\omega}2$.

Indeed, every ordinal is $\alpha + n$ for some unique limit ordinal α and finite ordinal n. Then the original ordinal is even or odd, according as n is even or odd. Now we can list all possible strategies as $(f^{\alpha} \colon \alpha < 2^{\omega})$, where f^{α} will be a strategy for e if and only if $\alpha + e$ is even.

We can then define a list $(a^{\alpha} \colon \alpha < 2^{\omega})$ of possible plays (that is, elements of ${}^{\omega}2$) so that,

- for all α, if $\alpha + e$ is even, then e can use strategy f^α for the play a^α; that is,

$$a^\alpha_{2m+e} = f^\alpha(a^\alpha_{1-e}, a^\alpha_{3-e}, \dots, a^\alpha_{2m-1+e});$$

- for all distinct α and β, $a^\alpha \neq a^\beta$, at least if $\alpha + \beta$ is odd.

Indeed, we can proceed recursively. If $(a^\beta : \beta < \alpha)$ has been defined, and $\alpha < 2^\omega$, then since there are continuum-many plays in which the strategy f^α is used, one of them, to be called a^α, is not among those a^β such that $\beta < \alpha$ and $\beta + \alpha$ is odd.

Since, if $\alpha + e$ is even, player e can use strategy f^α for the play a^α, this means player $1 - e$ has *some* strategy that, with f^α, determines a^α. Therefore, if we now choose the partition of ${}^\omega 2$ so that

$$\{a^\alpha : \alpha + e \text{ even}\} \subseteq A_{1-e},$$

then neither player has a winning strategy for the game: the game is not determined.

References

[1] M. Aschenbrenner, A. Dolich, D. Haskell, D. Macpherson, and S. Starchenko, *Vapnik–Chervonenkis density in some theories without the independence property, I*, ArXiv e-prints (2011).

[2] Wilfrid Hodges, *Building models by games*, Dover Publications, Mineola, New York, 2006, original publication, 1985. MR MR812274 (87h:03045)

[3] Bruno Poizat, *Cours de théorie des modèles*, Bruno Poizat, Lyon, 1985, Une introduction à la logique mathématique contemporaine. [An introduction to contemporary mathematical logic]. MR 817208 (87f:03084)

[4] ———, *A course in model theory*, Universitext, Springer-Verlag, New York, 2000, An introduction to contemporary mathematical logic, Translated from the French by Moses Klein and revised by the author. MR 1757487 (2001a:03072)

[5] Katrin Tent and Martin Ziegler, *A course in model theory*, Lecture Notes in Logic, vol. 40, Association for Symbolic Logic, La Jolla, CA, 2012. MR 2908005

[6] Lou van den Dries, *Tame topology and o-minimal structures*, London Mathematical Society Lecture Note Series, vol. 248, Cambridge University Press, Cambridge, 1998. MR 1633348 (99j:03001)

[7] Stephen Willard, *General topology*, Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1970. MR MR0264581 (41 #9173)