What is Model Theory?

Michael Lieberman
Kalamazoo College

Math Department Colloquium

October 16, 2013

What is model theory?

Model theory is an area of mathematical logic that seeks to use the tools of logic to solve concrete mathematical problems. Given a class of interesting objects (graphs, groups, vector spaces, etc.),

- we isolate the basic vocabulary needed to describe them, and
- identify the rules (expressed in this vocabulary) that characterize precisely the objects of interest.

Based on the size and complexity of this set of rules—and a little bit of first-order logic—we can often draw new and surprising conclusions...

The goals for today are fairly modest. We'll focus on

- getting a sense of how this looks in practice,
- think about the Compactness Theorem for first order logic,
- and use it to prove a couple of non-obvious facts.

In particular:

Theorem (De Bruijn-Erdős, '48)

If any finite part of an infinite graph G can be colored with k colors, then the entire graph can be colored with just k colors.

Proposition (A. Robinson, '60)

There is a version of the real numbers containing infinitesimals—numbers $\alpha > 0$ with the property that $\alpha < 1/n$ for any positive integer n.

Begin with the real numbers, $\mathbb{R}$. Lots of ways we might think about them, with corresponding basic vocabularies:

- Additive reals: $\langle \, \rangle$

Begin with the real numbers, $\mathbb{R}$. Lots of ways we might think about them, with corresponding basic vocabularies:

- Additive reals: $\langle \; \rangle$
  Constant symbol 0

Begin with the real numbers, $\mathbb{R}$. Lots of ways we might think about them, with corresponding basic vocabularies:

- Additive reals: $\langle 0, \rangle$

Begin with the real numbers, $\mathbb{R}$. Lots of ways we might think about them, with corresponding basic vocabularies:

- Additive reals: $\langle 0, \ \rangle$
  Binary function symbol $+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$
  $$+ : (x, y) \mapsto x + y$$

Begin with the real numbers, $\mathbb{R}$. Lots of ways we might think about them, with corresponding basic vocabularies:

- Additive reals: $\langle 0, +, \ \rangle$

Begin with the real numbers, $\mathbb{R}$. Lots of ways we might think about them, with corresponding basic vocabularies:

- Additive reals: $\langle 0, +, \rangle$
  Unary function symbol $- : \mathbb{R} \to \mathbb{R}$
  $$- : x \mapsto -x$$

Begin with the real numbers, $\mathbb{R}$. Lots of ways we might think about them, with corresponding basic vocabularies:

- Additive reals: $\langle 0, +, - \rangle$

Begin with the real numbers, $\mathbb{R}$. Lots of ways we might think about them, with corresponding basic vocabularies:

- Additive reals: $\langle 0, +, - \rangle$
- Multiplicative reals: $\langle \, \rangle$

Begin with the real numbers, $\mathbb{R}$. Lots of ways we might think about them, with corresponding basic vocabularies:

- Additive reals: $\langle 0, +, - \rangle$
- Multiplicative reals: $\langle \; \rangle$
  Constant symbol 1

Begin with the real numbers, $\mathbb{R}$. Lots of ways we might think about them, with corresponding basic vocabularies:

- Additive reals: $\langle 0, +, - \rangle$
- Multiplicative reals: $\langle 1, \ \rangle$

Begin with the real numbers, $\mathbb{R}$. Lots of ways we might think about them, with corresponding basic vocabularies:

- Additive reals: $\langle 0, +, - \rangle$
- Multiplicative reals: $\langle 1, \; \rangle$
  Binary function symbol $\times : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$
  $\times \; : \; (x, y) \mapsto x \cdot y$

Begin with the real numbers, $\mathbb{R}$. Lots of ways we might think about them, with corresponding basic vocabularies:

- Additive reals: $\langle 0, +, - \rangle$
- Multiplicative reals: $\langle 1, \times, \rangle$

Begin with the real numbers, $\mathbb{R}$. Lots of ways we might think about them, with corresponding basic vocabularies:

- Additive reals: $\langle 0, +, - \rangle$
- Multiplicative reals: $\langle 1, \times, \;\rangle$
  Unary function symbol $(\;)^{-1} : \mathbb{R} \to \mathbb{R}$
  $$(\;)^{-1} : x \mapsto 1/x$$

Begin with the real numbers, $\mathbb{R}$. Lots of ways we might think about them, with corresponding basic vocabularies:

- Additive reals: $\langle 0, +, - \rangle$
- Multiplicative reals: $\langle 1, \times, (\ )^{-1} \rangle$

Begin with the real numbers, $\mathbb{R}$. Lots of ways we might think about them, with corresponding basic vocabularies:

- Additive reals: $\langle 0, +, - \rangle$
- Multiplicative reals: $\langle 1, \times, (\ )^{-1} \rangle$
- Combined: $\langle 0, 1, +, \times, -, (\ )^{-1} \rangle$

Begin with the real numbers, $\mathbb{R}$. Lots of ways we might think about them, with corresponding basic vocabularies:

- Additive reals: $\langle 0, +, - \rangle$
- Multiplicative reals: $\langle 1, \times, (\ )^{-1} \rangle$
- Combined: $\langle 0, 1, +, \times, -, (\ )^{-1} \rangle$
- Ordered reals: $\langle 0, 1, +, \times, -, (\ )^{-1}, \ \rangle$

Begin with the real numbers, $\mathbb{R}$. Lots of ways we might think about them, with corresponding basic vocabularies:

- Additive reals: $\langle 0, +, - \rangle$
- Multiplicative reals: $\langle 1, \times, (\ )^{-1} \rangle$
- Combined: $\langle 0, 1, +, \times, -, (\ )^{-1} \rangle$
- Ordered reals: $\langle 0, 1, +, \times, -, (\ )^{-1}, \ \rangle$
  Binary relation $\leq$ on $\mathbb{R} \times \mathbb{R}$
  $\leq (x, y)$ if and only if $x \leq y$

Begin with the real numbers, $\mathbb{R}$. Lots of ways we might think about them, with corresponding basic vocabularies:

- Additive reals: $\langle 0, +, - \rangle$
- Multiplicative reals: $\langle 1, \times, (\ )^{-1} \rangle$
- Combined: $\langle 0, 1, +, \times, -, (\ )^{-1} \rangle$
- Ordered reals: $\langle 0, 1, +, \times, -, (\ )^{-1}, \leq \rangle$

Begin with the real numbers, $\mathbb{R}$. Lots of ways we might think about them, with corresponding basic vocabularies:

- Additive reals: $\langle 0, +, - \rangle$
- Multiplicative reals: $\langle 1, \times, (\ )^{-1} \rangle$
- Combined: $\langle 0, 1, +, \times, -, (\ )^{-1} \rangle$
- Ordered reals: $\langle 0, 1, +, \times, -, (\ )^{-1}, \leq \rangle$

For now, let's forget about $\times$ and $(\ )^{-1}$, and focus on the restricted vocabulary

$$\langle 0, 1, +, -, \leq \rangle$$

What other objects are we used to talking about using this same language? Among others: $\mathbb{Z}$, the integers. If we take $\mathbb{R}$ and $\mathbb{Z}$ with the standard interpretations of symbols in $\langle 0, 1, +, -, \leq \rangle$, though, there are serious differences between the two...

What other objects are we used to talking about using this same language? Among others: $\mathbb{Z}$, the integers. If we take $\mathbb{R}$ and $\mathbb{Z}$ with the standard interpretations of symbols in $\langle 0, 1, +, -, \leq \rangle$, though, there are serious differences between the two...

Example:
In $\mathbb{R}$, for any distinct x and y, say $x < y$, there is a real number z with $x < z < y$ (the ordering is *dense*).

What other objects are we used to talking about using this same language? Among others: $\mathbb{Z}$, the integers. If we take $\mathbb{R}$ and $\mathbb{Z}$ with the standard interpretations of symbols in $\langle 0, 1, +, -, \leq \rangle$, though, there are serious differences between the two...

Example:
In $\mathbb{R}$, for any distinct x and y, say $x < y$, there is a real number z with $x < z < y$ (the ordering is *dense*).

In $\mathbb{Z}$, distinct elements are separated by chasms: there is no integer z such that $0 < z < 1$.

What other objects are we used to talking about using this same language? Among others: $\mathbb{Z}$, the integers. If we take $\mathbb{R}$ and $\mathbb{Z}$ with the standard interpretations of symbols in $\langle 0, 1, +, -, \leq \rangle$, though, there are serious differences between the two...

Example:
In $\mathbb{R}$, for any distinct x and y, say $x < y$, there is a real number z with $x < z < y$ (the ordering is *dense*).

In $\mathbb{Z}$, distinct elements are separated by chasms: there is no integer z such that $0 < z < 1$.

To make this distinction clear, we need a precise and unambiguous language.

Given a vocabulary $\mathcal{V}$, we form a *first-order language* through finite combinations of symbols from $\mathcal{V}$, the equals sign, parentheses, variable symbols, and an array of logical symbols:

Given a vocabulary $\mathcal{V}$, we form a *first-order language* through finite combinations of symbols from $\mathcal{V}$, the equals sign, parentheses, variable symbols, and an array of logical symbols:

$\wedge$	"and"
$\vee$	"or"
$\neg$	"not"
$\rightarrow$	"implies"
$\forall x$	"for all x"
$\exists x$	"there exists an x"

Given a vocabulary $\mathcal{V}$, we form a *first-order language* through finite combinations of symbols from $\mathcal{V}$, the equals sign, parentheses, variable symbols, and an array of logical symbols:

$\wedge$	"and"
$\vee$	"or"
$\neg$	"not"
$\rightarrow$	"implies"
$\forall x$	"for all x"
$\exists x$	"there exists an x"

Examples:

$$\neg\exists z(0 < z \wedge z < 1) \qquad (\phi_1)$$

$$\forall x \forall y[(x < y) \rightarrow \exists z(x < z \wedge z < y)] \qquad (\phi_2)$$

To ensure that we consider only objects that behave like $\mathbb{R}$, then, we should restrict our attention to those that satisfy the density condition.

In fact, to eliminate as much bad behavior as possible, we consider

$$\mathrm{Th}(\langle \mathbb{R}, 0, 1, +, -, \leq \rangle),$$

the *theory of* $\mathbb{R}$, the set of **all** first-order sentences in this vocabulary that are true in $\mathbb{R}$.

If we restrict to objects with interpretations of 0, 1, $+$, $-$, and $\leq$ satisfying all of these sentences, we've gone a long way toward characterizing $\mathbb{R}$.

Problem: First order logic isn't very expressive.

Problem: First order logic isn't very expressive.

To avoid infinitesimal numbers, we might like to say "there is no number $x > 0$ such that $x < 1/n$ for all positive integers n."

Problem: First order logic isn't very expressive.

To avoid infinitesimal numbers, we might like to say "there is no number $x > 0$ such that $x < 1/n$ for all positive integers n."

$$\neg\exists\alpha[(0 < \alpha < 1) \wedge (0 < \alpha < 1/2) \wedge (0 < \alpha < 1/3) \wedge \dots]$$

But this requires an infinite "and," which isn't allowed.

Problem: First order logic isn't very expressive.

To avoid infinitesimal numbers, we might like to say "there is no number $x > 0$ such that $x < 1/n$ for all positive integers n."

$$\neg\exists\alpha[(0 < \alpha < 1) \wedge (0 < \alpha < 1/2) \wedge (0 < \alpha < 1/3) \wedge \dots]$$

But this requires an infinite "and," which isn't allowed.

So first order logic alone isn't enough to protect us from infinitesimals, among other things...

So why use it? Because, in short, it's compact:

Theorem (Compactness Theorem)

Version 1: Let Γ *be an infinite set of first order sentences. If* Γ *is inconsistent, then there is a finite set of sentences* $\Gamma' \subset \Gamma$ *that is itself inconsistent.*

Version 2: Let Γ *be an infinite set of first order sentences. If for any finite* $\Gamma' \subset \Gamma$ *there is an object* $X_{\Gamma'}$ *obeying all of the sentences in* Γ'*, then there is a single object that obeys the entire infinite list* Γ*.*

The word "graph" means many things to many people. Here, a graph consists of

- A set of vertices X.
- An edge relation E. For x and y in X, xEy means "there is an edge from x to y."

Less formally, a graph consists of a family of nodes, some of which are connected to others by edges...

We say a graph can be colored with k colors (or is k-colorable) if it is possible to color each of the nodes one of k colors in such a way that no two connected nodes are colored the same.

We say a graph can be colored with k colors (or is k-colorable) if it is possible to color each of the nodes one of k colors in such a way that no two connected nodes are colored the same.

[There's some geometric content here: if a graph G is 2-colorable, it cannot contain any triangles, among other things. If it's 3-colorable, it can't contain any tetrahedra, among other things.]

We say a graph can be colored with k colors (or is k-colorable) if it is possible to color each of the nodes one of k colors in such a way that no two connected nodes are colored the same.

[There's some geometric content here: if a graph G is 2-colorable, it cannot contain any triangles, among other things. If it's 3-colorable, it can't contain any tetrahedra, among other things.]

If we are given an infinite graph (i.e. infinitely many nodes), we can imagine verifying that any finite part is, say, 2-colorable, but how can we conclude this for the graph as a whole?

We say a graph can be colored with k colors (or is k-colorable) if it is possible to color each of the nodes one of k colors in such a way that no two connected nodes are colored the same.

[There's some geometric content here: if a graph G is 2-colorable, it cannot contain any triangles, among other things. If it's 3-colorable, it can't contain any tetrahedra, among other things.]

If we are given an infinite graph (i.e. infinitely many nodes), we can imagine verifying that any finite part is, say, 2-colorable, but how can we conclude this for the graph as a whole?

Magic...

Theorem (De Bruijn-Erdős)

If every finite subgraph of an infinite graph G is k-colorable, then G itself is k-colorable.

Proof: The language of G contains only the edge relation E. Add constant symbols v_x for each node $x \in X$, and $c_1, \ldots, c_k$ for each of the k colors, and let f be a function symbol capturing the coloring (the sentence $f(v_x) = c_i$ meaning that node x gets the color c_i). With the language thus enriched, we can form sentences:

- C_x: x is colored with one of $c_1, \ldots, c_k$
- $A_{x,y}$: If xEy, x and y receive different colors.

Let Γ be the set of all such sentences.

Take a finite subset $\Gamma' \subset \Gamma$. The sentences in Γ' only involve finitely many nodes from G, say $x_1, \ldots, x_n$. Let G' be the finite subgraph of G involving only these vertices.

The sentences in Γ specify that

- every node of G' should be colored with one of the c_i
- any connected nodes in G' should receive different colors

That is, the sentences in Γ' will be satisfied if G' is k-colorable. True by assumption.

Take a finite subset $\Gamma' \subset \Gamma$. The sentences in Γ' only involve finitely many nodes from G, say $x_1, \ldots, x_n$. Let G' be the finite subgraph of G involving only these vertices.

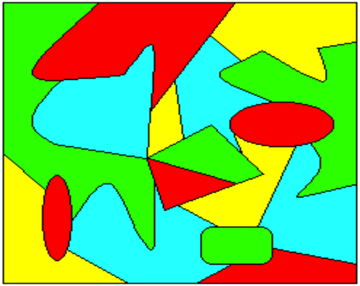
The sentences in Γ specify that

- every node of G' should be colored with one of the c_i
- any connected nodes in G' should receive different colors

That is, the sentences in Γ' will be satisfied if G' is k-colorable. True by assumption.

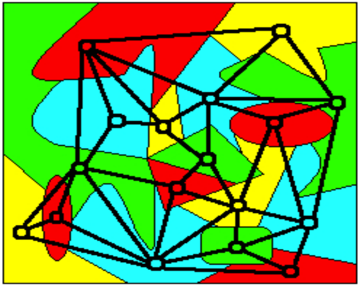
By the Compactness Theorem, all of the sentences in Γ can be satisfied simultaneously. That is, G is k-colorable.

The Four Color Theorem (Appel-Hakken, '76) states that any planar map can be colored using only four colors in such a way that no adjacent regions are given the same color.



We can turn this into a graph coloring question...

Draw a node in each region, and connect it to the nodes in each adjoining region.



From this perspective, our theorem guarantees that if every finite chunk of an infinite planar map is 4-colorable, so is the map itself.

The Four Color Theorem was the first major result proved by computer. The reduction to the finite was an essential part of the process: there were only so many (~ 1900) finite configurations to consider, and the task of testing each case was given to the computer.

We've all seen the definition of the derivative via limits:

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

We've all seen the definition of the derivative via limits:

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

This is a 19th century formalization of a much earlier idea of Leibniz (and Newton), who thought of the derivative as (roughly speaking) a quotient

$$\frac{df}{dx} = \frac{f(x+\alpha) - f(x)}{\alpha}$$

where α is not a quantity that goes to 0, but rather infinitesimal —closer to 0 than any nonzero real number.

We've all seen the definition of the derivative via limits:

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

This is a 19th century formalization of a much earlier idea of Leibniz (and Newton), who thought of the derivative as (roughly speaking) a quotient

$$\frac{df}{dx} = \frac{f(x+\alpha) - f(x)}{\alpha}$$

where α is not a quantity that goes to 0, but rather infinitesimal —closer to 0 than any nonzero real number.

A weird idea, seemingly paradoxical, and the subject of much contemporary criticism (cf. Berkeley, *The Analyst*).

We saw that first order logic isn't robust enough to prohibit a version of the real numbers containing infinitesimal elements. We now see that such a version actually exists.

Proposition (A. Robinson, '60)

There is a version of the real numbers containing an infinitesimal element: $\alpha > 0$ with $\alpha < 1/n$ for all positive integers n.

How do we proceed? Via the Compactness Theorem, naturally.

To the basic vocabulary of the ordered reals, $\langle \mathbb{R}, 0, 1, +, -, \leq \rangle$, we add a new constant symbol α. Let Γ be the set of sentences

$$T \cup \{\alpha < 1, \alpha < 1/2, \alpha < 1/3, \dots\}$$

where T is the theory of the ordered reals. Any finite subset Γ' of Γ will consist of sentences from T (rules concerning the behavior of the reals) and finitely many of the α-sentences, say $\alpha < 1, \dots, \alpha < 1/k$.

Is there an object that behaves like the reals and contains an element $\alpha < 1/k$? Of course: the reals! Take $\alpha = 1/(2k)$.

Since every finite subset of the rules contained in Γ can be satisfied, there is an object $\mathfrak{R}$ that obeys all of them simultaneously: both the rules guaranteeing $\mathbb{R}$-like behavior, and those ensuring the existence of an infinitesimal.

Since every finite subset of the rules contained in Γ can be satisfied, there is an object $\mathfrak{R}$ that obeys all of them simultaneously: both the rules guaranteeing $\mathbb{R}$-like behavior, and those ensuring the existence of an infinitesimal.

By construction, then, elements of this $\mathfrak{R}$ can be added, multiplied, and inverted just like real numbers, but there is a new, impossibly small element.

Since every finite subset of the rules contained in Γ can be satisfied, there is an object $\mathfrak{R}$ that obeys all of them simultaneously: both the rules guaranteeing $\mathbb{R}$-like behavior, and those ensuring the existence of an infinitesimal.

By construction, then, elements of this $\mathfrak{R}$ can be added, multiplied, and inverted just like real numbers, but there is a new, impossibly small element.

This represents a kind of vindication of Leibniz's vision of infinitesimals, although it was a very long time (almost 300 years) coming...