



UNIVERSITY OF
CAMBRIDGE

Introductory Logic

Lecture 5: Metatheorems, models, theories

Alan Mycroft

Computer Laboratory, University of Cambridge, UK
<http://www.cl.cam.ac.uk/~am21>

MPhil in ACS – 2011/12

This lecture is a draft and has typos

- A sound and complete proof system for FOL
- Consistency
- Models of a theory
- Isomorphism, Elementary Equivalence
- Theories and their properties
- Skolem-Löwenheim theorem

Induction principle

I meant to say this last lecture:

For a given set of sentences Γ , we call the set $\{\phi \mid \Gamma \vdash_R \phi\}$ the *theorems* of Γ . We identify Γ as a *theory*.

We have an induction principle: Suppose S is a set of wffs such that:

- S includes Γ
- S is closed under every rule in R (i.e. using every possible proof rule assuming every member of S as antecedents only results in existing members of S).

Then S contains every theorem of Γ .

Such sets S are called ‘deductively closed’, because of their property $S \vdash_R \phi$ iff $\phi \in S$.

Enderton's FOL proof system

For concreteness we present Enderton's set of inference rules for FOL.

This is a Hilbert system, with modus ponens as the single non-trivial inference rule, but with countably many *logical axioms* – which I have presented as rules with zero antecedents, but which Enderton treats as a set of axioms Λ .

Thus when I've said $\Gamma \vdash_R \phi$, Enderton might write $\Gamma \cup \Lambda \vdash \phi$. The differences are inessential.

The idea is that Λ contains axioms which are true in every logic, unlike axioms like $a + b = b + a$ which might hold in a particular FOL theory Γ

In general when we write $\Gamma \vdash \phi$ we mean $\Gamma \vdash_R \phi$ relative to the presumed set of rules R (for Enderton just (MP)).

Capture-avoiding substitution

Define *substitution* of a *term* u for a *variable* x in a given term t recursively by:

$$x[u/x] = u$$

$$y[u/x] = y \text{ if } y \text{ and } x \text{ are different variables}$$

$$F(t_1, \dots, t_k)[u/x] = F(t_1[u/x], \dots, t_k[u/x])$$

This simply extends to wffs, provided we are careful:

- We stop substituting for x within $\forall x \sigma$.
- We must not substitute a term t (containing y as a free variable) for x within a $\forall y \sigma$.

These concepts are familiar from programming languages and are easily computable. Sometimes it is necessary to ‘ α -rename’ or use an ‘alphabetic variant’ to enable substitution to be applicable.

Enderton's FOL proof system (2)

Letting x, y be variables ϕ, ψ be wffs, the logical axioms are:

- 1 Tautologies (obtained from Propositional Logic tautologies by replacing propositional variables consistently with wffs)
- 2 $\forall x\phi \rightarrow \phi[t/x]$ for some term t ; the notation $\phi[t/x]$ means capture-avoiding substitution.
- 3 $\forall x(\phi \rightarrow \psi) \rightarrow (\forall x\phi \rightarrow \forall x\psi)$
- 4 $\phi \rightarrow \forall x\phi$ provided x is not free in ϕ
- 5 $x = x$
- 6 $x = y \rightarrow \phi \rightarrow \phi'$ where ϕ' a version of ϕ obtained by replacing zero or more (but not necessarily all) of its occurrences of x with y .

Additionally, we allow any of these forms to be generalised.

Generalisation adds, for each wff τ above and for each n variables x_1, \dots, x_n the additional axiom: $\forall x_1 \dots \forall x_n \tau$.

Some authors treat generalisation as an inference rule (GEN) $\frac{\vdash \tau}{\vdash \forall x \tau}$.

We can talk about this, but Enderton remarks:

“But first we should admit that the above list of logical axioms may not appear very natural.”

Two metatheorems

Theorem (Generalisation Theorem)

If $\Gamma \vdash \phi$ and x does not occur free in any wff in Γ then $\Gamma \vdash \forall x\phi$.

Proof.

Unreasonably assuming(!) soundness and completeness enables a simple proof:

$$\begin{aligned}\Gamma \vdash \phi &\Rightarrow \Gamma \models_{\mathcal{I},v} \phi \text{ for all interpretations} \\ &\Rightarrow \Gamma \models_{\mathcal{I},v} \forall x\phi \text{ for all interpretations} \\ &\Rightarrow \Gamma \vdash \forall x\phi.\end{aligned}$$



But we may need to prove it by induction on syntax if we want to use it as part of a proof of soundness and completeness.

Two metatheorems (2)

Theorem (Deduction Theorem)

If $\Gamma; \gamma \vdash \phi$ then $\Gamma \vdash \gamma \rightarrow \phi$

(notation: $\Gamma; \gamma$ means $\Gamma \cup \{\gamma\}$)

This connects the first-order connective \rightarrow with the meta-level symbols \vdash and \models . (The reverse direction of ‘if-then’ is just modus ponens.)

Proof.

Again, assuming soundness and completeness makes things easy:

$$\begin{aligned} \Gamma; \gamma \vdash \phi &\Rightarrow \Gamma; \gamma \models \phi && \text{soundness} \\ &\Rightarrow \Gamma \models \gamma \rightarrow \phi && \text{same interpretations satisfy both} \\ &\Rightarrow \Gamma \vdash \gamma \rightarrow \phi && \text{completeness} \end{aligned}$$

□

But we may need to prove it by induction on syntax if we want to use it as part of a proof of soundness and completeness.

Proof of soundness

To prove soundness of an inference system we generally perform so-called *rule induction*. For each rule $\frac{\vdash \phi_1 \quad \dots \quad \vdash \phi_n}{\vdash \phi_0}$ we prove that whenever $\models \phi_i$ for $1 \leq i \leq n$ we also have $\models \phi_0$.

For the axioms $\frac{}{\vdash \phi_0}$ this reduces to showing that ϕ_0 is valid.

For the Hilbert-style system Enderton gives this is just a matter of showing each instance of the six axiom schemes is valid and then proving modus ponens preserves validity:

whenever $\models \phi$ and $\models \phi \rightarrow \psi$ we have $\models \psi$

This is straightforward from the definition of \models .

One issue to newcomers to this topic is spotting when proofs are trivial, like this one, or require much more work.

Simple proof of completeness

This is generally significantly harder than soundness (which is effectively just induction). It can be seen as a sort of ‘reachability’ problem: can we reach *all* valid sentences starting from the axioms by repeatedly applying the rules.

The proof for the whole of FOL is rather involved (Enderton pp. 135–145), so here we just consider the propositional logic subset.

However, if we continue to take the whole set of propositional tautologies as axioms, then completeness trivially holds.

So completeness questions in propositional logic tend to concern whether a given subset of tautologies are complete when combined with modus ponens . . .

Simple proof of completeness (2)

We show (following Paulson's Part 1B "Logic and Proof" lectures) that the following Propositional wffs as axioms (together with modus ponens) are complete for propositional truth (the axioms are sound because they are tautologies):

Axiom K: $\phi \rightarrow (\psi \rightarrow \phi)$

Axiom S: $(\phi \rightarrow \psi \rightarrow \theta) \rightarrow ((\phi \rightarrow \psi) \rightarrow (\phi \rightarrow \theta))$

Axiom DN: $(\neg\phi \rightarrow \neg\psi) \rightarrow ((\neg\phi \rightarrow \psi) \rightarrow \phi)$

(Paulson defines $\neg\phi$ as abbreviating $\phi \rightarrow \perp$ where \perp is syntax for the propositional constant falsity, so he can use the following simpler, and more recognisable version of DN (double negation), but we are sticking with Enderton's formalisation having \neg and \rightarrow being primitive.)

Axiom DN': $\neg\neg\phi \rightarrow \phi$

Simple Proof of completeness (3)

Basically we want to show $\Gamma \models \tau$ implies $\Gamma \vdash \tau$. Instead we show the *contrapositive*: $\Gamma \not\models \tau$ implies $\Gamma \not\vdash \tau$.

We repeatedly enlarge set Γ to obtain a maximal set Γ^* which also does not prove τ . First enumerate all possible wffs as ϕ_i ($i \in \mathbb{N}$) and define:

- $\Gamma_0 = \Gamma$
- $\Gamma_{i+1} = \Gamma_i$ if $\Gamma_i \cup \{\phi_i\} \vdash \tau$
= $\Gamma_i \cup \{\phi_i\}$ otherwise
- $\Gamma^* = \bigcup_{i \in \mathbb{N}} \Gamma_i$

It is important here that the inference rules never allow both ψ and $\neg\psi$ to become members of Γ^* – this would mean the inference rules were incomplete.

Simple Proof of completeness (4)

Now Γ^* has various special properties:

- $\Gamma^* \supseteq \Gamma$ and $\Gamma^* \not\vdash \tau$ and it is a maximal such set.
- It's 'truth-like': $\phi \in \Gamma^*$ iff $\neg\phi \notin \Gamma^*$. Also if Γ^* contains ϕ and $\phi \rightarrow \psi$ then it contains ψ etc.
- This means we can define a valuation v on propositional variables A which makes every sentence in Γ^* true and everything not in Γ^* (including τ) false; v is defined by:

$$\begin{aligned}v(A) &= \text{true if } A \in \Gamma^* \\ &= \text{false otherwise}\end{aligned}$$

- hence by construction $\Gamma^* \not\vdash \tau$.
- and hence $\Gamma \not\vdash \tau$ (since this valuation makes every wff in $\Gamma^* \supseteq \Gamma$ true).

This needs spelling out in more detail – sorry.

The Wikipedia article

http://en.wikipedia.org/wiki/Propositional_calculus

gives simple soundness and completeness proofs for the Propositional Calculus.

Satisfiability and consistency

We have already defined a theory Γ (a set of sentences) to be satisfiable if there is an interpretation which makes every member true.

Just as deducibility (\vdash) is a syntactic version of semantic entailment (\models) one might look for a syntactic version of satisfiability. Such a notion exists and is called *consistency*.

A theory Γ is inconsistent if there is a wff ϕ such that $\Gamma \vdash \phi$ and $\Gamma \vdash \neg\phi$ both hold.

Satisfiability and consistency are the semantic and syntactic sides of the same coin – and indeed coincide for FOL – but we need to define them separately so we can (meta-)prove them to be equivalent.

Inconsistency is contagious

Note that if a theory Γ is inconsistent then *every* wff τ is deducible, i.e. $\Gamma \vdash \tau$. This even includes the wff '0=1'. Why:

- Let ϕ be the exemplifying inconsistency in Γ , i.e. $\Gamma \vdash \phi$ and $\Gamma \vdash \neg\phi$ are derivable judgements.
- $A \rightarrow (\neg A \rightarrow B)$ is a tautology in Propositional Calculus, and hence its instances are valid (and deducible) in FOL. Take the instance $\phi \rightarrow (\neg\phi \rightarrow \tau)$
- Now use modus ponens twice to obtain $\Gamma \vdash \tau$.

So, inconsistency is very bad – indeed much of early 20th-century mathematics was involved in avoiding it (e.g. Russell's paradox). It turned eventually futile in that it was later shown that one can never prove the consistency of a theory capable of arithmetic (Gödel) – if indeed arithmetic is consistent!

Multiple models of a theory

Often when we axiomatise (give a theory Γ for) a mathematical idea (e.g. arithmetic or set theory) we would like to specify just one model for it.

Sometimes however, we want to axiomatise (say) group theory which has many models – at least one for each group.

We now want to discuss the idea of the *class of structures which model a theory Γ* – i.e. make every sentence in Γ true.

Note the word *class*. Just like the collection of all sets cannot be a set (Russell's paradox) the collection of all possible structures is also too large to be a set—however, provided one is careful with the uses made of classes then they do not cause trouble.

Isomorphism

Two forms of structures can never be distinguished by FOL – if one is a model of a theory then so is the other. These are *isomorphic structures* which differ only by renaming. We write $\mathcal{A} \cong \mathcal{B}$.

Recall a structure is a triple: a universe of discourse, interpretations for function symbols and interpretations for predicate symbols.

Say that two structures $\mathcal{A} = (D; F_i, P_i)$ and $\mathcal{B} = (D'; F'_i, P'_i)$ are *isomorphic* if there is a bijection $\theta : D \rightarrow D'$ having the properties

- $\theta(F_i(x_1, \dots, x_{\text{arity}(F_i)})) = F'_i(\theta(x_1), \dots, \theta(x_{\text{arity}(F_i)}))$
- $P_i(x_1, \dots, x_{\text{arity}(F_i)})$ iff $P'_i(\theta(x_1), \dots, \theta(x_{\text{arity}(F_i)}))$

Example: $(\mathbb{R}; 0, +; =)$ and $(\mathbb{R}^+; 1, \times; =)$ and the bijection $x \mapsto e^x$.

I've included '=' in the signature for pedantry, many people would omit it or assume it as part of the logic syntax rather than a relation symbol.

Elementary equivalence

As a way of determining the expressiveness of FOL at classifying structures, define two structures \mathcal{A}, \mathcal{B} over the same logic to be *elementarily equivalent*, written $\mathcal{A} \equiv \mathcal{B}$ if

for all wffs ϕ we have $\models_{\mathcal{A}} \phi$ iff $\models_{\mathcal{B}} \phi$.

Since ϕ cannot detect changes only to the names of the elements we have

$$\mathcal{A} \cong \mathcal{B} \text{ implies } \mathcal{A} \equiv \mathcal{B}$$

An example (see Löwenheim-Skolem later) of when the reverse implication fails is given by the structures $(\mathbb{Q}; ; <)$ and $(\mathbb{R}; ; <)$. These are not isomorphic (as there is no bijection between \mathbb{Q} and \mathbb{R} for cardinality reasons), but they do model exactly the same first-order properties. **Pedantry:** $(\mathbb{Q}; ; <)$ would normally be written $(\mathbb{Q}, <)$.

Elementary equivalence (2)

For finite structures (and a logic having equality) the reverse implication holds too:

$$\mathcal{A} \cong \mathcal{B} \text{ iff } \mathcal{A} \equiv \mathcal{B}$$

Wffs such as

$$\exists x \exists y \exists z (x \neq y \wedge y \neq z \wedge x \neq z \wedge \forall t (t = x \vee t = y \vee t = z))$$

fix the size of the structure (here 3), and further conjuncts can constrain the elements' behaviour to reach isomorphism.

Multiple models of a theory (2)

Given a set of first-order sentences Γ we write $Mod \Gamma$ to be the set of models of Γ , i.e. the set of interpretations which make every element of Γ true.

This captures the mathematical idea of (e.g.) $Mod \textit{Group}$ being the class of all groups when \textit{Group} is the set of axioms for a group. **Note again the semantic notion “class of all groups” and the syntactic one “set of axioms for a group”.**

Now we can classify classes \mathcal{K} of structures into those which are $Mod \Gamma$ for some Γ and those which aren't. We say \mathcal{K} is an *elementary class (EC)* if $\mathcal{K} = Mod \{\tau\}$ for some τ . Subtlety: we say \mathcal{K} is EC_{Δ} if it is $Mod \Gamma$ for some Γ .

We call a theory *categorical* if all its models are isomorphic. (“There's only the standard model”).

Multiple models of a theory (3)

Remarks:

- EC_{Δ} is somewhat more expressive than EC as it allows an infinite number of sentences in Γ , but note that any finite number can be treated as a singleton τ by simply conjoining (\wedge) them.
- Example: the class of all groups is EC. The class of all infinite groups is not EC, however it *is* EC_{Δ} (it takes an infinite number of sentences to ban all finite groups).
- The word ‘elementary’ here is a historical accident, it really means ‘first-order’.
- I believe the word ‘categorical’ here means ‘definite’ as in “a categorical ‘yes’ ” rather than as in “category theory”.

Theories

I've used the word 'theory' so far to refer to any set of sentences Γ . However, some authors, including Enderton, use it to mean 'deductively-closed' or 'entailment-closed' (these are of course equivalent concepts in FOL). See discussion on:

<http://planetmath.org/encyclopedia/FinitelyAxiomatizableTheory.html>

Such a closure is easy to express inductively: it's the smallest set of sentences Γ^* which includes Γ and whenever a sentence is deducible from (or entailed by) those in Γ^* then it is also in Γ^*

So now define (Enderton) a *theory* to be a set of sentences which is closed under semantic entailment. I.e. T is a theory if $T \models \tau$ implies $\tau \in T$. I'll try not to rely on the exact definition by using \vdash not \in .

There are various interesting properties which theories may or may not have ...

Properties of Theories

- A theory Γ is *consistent* iff for some sentence ϕ we have $\Gamma \not\vdash \phi$.
(This is an equivalent, but shorter, version than that given earlier – literature differs!)
- A theory Γ is *complete* iff for each sentence ϕ we have $\Gamma \vdash \phi$ or $\Gamma \vdash \neg\phi$. (Some authors require Γ to be consistent too.) Note that a *theory* being complete is a different (but related) concept to a *set of proof rules* being complete.
- A theory Γ is *axiomatisable* iff Γ includes a decidable subset Δ such that for every sentence ϕ we have $\Delta \vdash \phi$ iff $\Gamma \vdash \phi$. It is *finitely axiomatisable* if there is a finite such Δ . The typical theories one encounters, e.g. group theory, are axiomatisable since we give their axioms just by listing them—though note that axiom schemes, e.g. induction, represent a countable (and indeed effectively enumerable) set of instances.

Properties of Theories (2)

Note that while we personally might prefer theories to be consistent and axiomatisable, the issue of completeness is more complex. For example, group theory is incomplete because we want it *not* to specify whether or not $\exists x(x \cdot x \neq 1)$. On the other hand, we also want to know which mathematical structures can be specified precisely by a theory, remembering that FOL does not in general have an algorithm (Entscheidungsproblem) which determines whether or not a sentence ϕ holds in a theory Γ . An interesting result is that:

Theorem

Every complete axiomatisable theory Γ is decidable. I.e. there is an algorithm that given sentence ϕ as input yields 1 if $\Gamma \vdash \phi$ and 0 otherwise.

Proof.

See next slide.

Properties of Theories (3)

Proof.

Enumerate the sentences deducible from Γ . Wait until either ϕ or $\neg\phi$ appears (completeness guarantees one of them will). \square

There is also:

Theorem

Theorem. (Tarski) Every consistent theory Γ is included in a complete theory.

Proof.

Enumerate the sentences ϕ of the logic; whenever neither $\Gamma \vdash \phi$ nor $\Gamma \vdash \neg\phi$ holds then add either to Γ and continue with the new Γ . \square

Note that there's no reason for the resulting theory to be axiomatisable – its members may not be recursively enumerable.

Simple complete theories

As an example of Tarski's meta-theorem, note that group theory is not complete but if we add $\exists g (g \neq 1 \wedge \forall x (x = 1 \vee x = g))$ then the only model is the 2-element group.

An additional classical theory which is complete is that of Dense Linear Orders without endpoints. These satisfy, writing $x < y$ as shorthand for $x \leq y \wedge x \neq y$:

- $\forall x (x \leq x)$ (reflexivity)
- $\forall x \forall y (x \leq y \wedge y \leq x \rightarrow x = y)$ (anti-symmetry)
- $\forall x \forall y \forall z (x \leq y \wedge y \leq z \rightarrow x \leq z)$ (transitivity)
- $\forall x \forall y (x \leq y \vee y \leq x)$ (linear or total ordering)
- $\forall x \forall y (x < y \rightarrow \exists z (x < z \wedge z < y))$ (dense ordering)
- $\forall x \exists y \exists z (y < x \wedge x < z)$ (no endpoints)

Note that this theory only has infinite models (needs 'no endpoints').

Peano arithmetic

This provides arithmetic axioms for the signature $(\mathbb{N}; 0, S, +, \times; =)$

- 1 $\forall x \neg(S(x) = 0)$
- 2 $\forall x \forall y (S(x) = S(y) \rightarrow x = y)$
- 3 for every formula ϕ not involving x and having free variable y :
 $(\phi[0/y] \wedge (\forall x (\phi[x/y] \rightarrow \phi[S(x)/y]))) \rightarrow \forall x \phi[y/x]$
- 4 $\forall y (0 + y = y)$
- 5 $\forall x \forall y (S(x) + y = S(x + y))$
- 6 $\forall y (0 \times y = 0)$
- 7 $\forall x \forall y (S(x) \times y = y + (x \times y))$

Some authors give the signature $(\mathbb{N}; 0, S, +, \times; <, =)$ but $x < y$ can be treated as shorthand for $\exists t(x + t = y)$. The intended model is $(\mathbb{N}; 0, S, +, \times; =)$ (here we abuse notation by using (e.g.) $+$ both for the syntax (a function symbol) and its interpretation $+ : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$).

Peano arithmetic (2)

The Peano arithmetic axioms ('PA') have $(\mathbb{N}; \dots; \dots)$ as a model, but an important question is whether they sufficiently tie down the model to be categorical (only one model up to isomorphism). In other words, is PA a complete theory?

Seen from the perspective of group theory, we might hope the answer is 'yes', but it turns out the answer is 'no' (Gödel's incompleteness theorem (1931)).

Given a mathematical structure, for example $\mathcal{N} = (\mathbb{N}; \dots; \dots)$ arithmetic above, we write $Th \mathcal{N}$ for the set of sentences which hold of \mathcal{N} (this is trivially complete). The question is then whether we can find a set $\Gamma \subseteq Th \mathcal{N}$ which is recursively enumerable and whose deductive closure $Con \Gamma$ is equal to $Th \mathcal{N}$.

Peano arithmetic (3)

In particular not only PA is incomplete for \mathcal{N} but so is every first-order theory.

The proof again works by diagonalisation.

We assume we have an axiomatisation Γ (i.e. recursively enumerable set of axioms) with $Con \Gamma = Th \mathcal{N}$. We associate a *Gödel number* with each wff, proof derivation etc. which uniquely encodes it. (For example $2^i 3^j$ uniquely encodes pairs (i, j) of integers as a single integer.) We use these codes to create, from the putative Γ a sentence which is true in \mathcal{N} , but has no proof in Γ .

This is very similar to the construction of a Turing machine from the assumption that there is a Turing machine which detects whether (the description of) another Turing machine halts.

Compactness, finite and infinite models

We now return to the question of which structures can be characterised precisely by FOL sentences.

We now note that FOL obeys a similar compactness theorem as did propositional logic. However we express the intuitive version used there: “whenever $\Sigma \models \tau$ then there is a finite $\Sigma_0 \subseteq \Sigma$ such that $\Sigma_0 \models \tau$ ” with a more standard version (obtained by substituting τ with an unsatisfiable formulae such as $\phi \wedge \neg\phi$) and taking the contrapositive:

Theorem (Compactness)

If every finite subset of Σ has a model then so does Σ .

Proof.

This can be written (soundness and completeness) “if every finite subset of Σ is consistent then so is Σ ”. But this is trivially true since every inconsistency in Σ only uses a finite number of wffs in Σ to exhibit it. □

Theorem

Every finite structure \mathcal{I} (technically we need the logic to include '=') can be finitely axiomatised as Γ so that every model of Γ is isomorphic to \mathcal{I} .

Take as the first axiom the requirement that the number of elements is exactly $|\mathcal{I}|$. Each of the finite number of function and relation symbols in the logic only require a finite number of equations to exactly tabulate their behaviour.

Finite and infinite models (3)

We have another theorem about the limits of expressiveness of FOL.

Theorem

If a theory Γ has arbitrarily large finite models then it has an infinite model.

Remark: the converse is not true, see for example the Peano axioms.

Proof.

Write λ_k , ($k \in \mathbb{N}$) as shorthand for “there are k distinct values”, e.g. $\lambda_3 = \exists x \exists y \exists z (x \neq y \wedge x \neq z \wedge y \neq z)$ and consider whether $\Gamma' = \Gamma \cup \{\lambda_2, \lambda_3, \dots\}$ has a model. By hypothesis every finite subset of Γ' has a model, hence by compactness Γ' has a model (which, by construction, has more distinct elements than any given integer). \square

Löwenheim-Skolem Theorem

We now return to the question of which structures can be characterised precisely by FOL sentences, and the famous Löwenheim-Skolem theorem which implies a negative result:

If a countable first-order theory Γ has an infinite model, then for every infinite cardinal number κ it has a model of size κ .

Since we have not really discussed cardinal numbers in this course, we can specialise this to (exercise: why fix on \mathbb{N} rather than \mathbb{Q} , say?):

If a theory has an infinite model, then it has a model where the domain of discourse is \mathbb{N} and one where the domain of discourse is \mathbb{R} .

The result implies that first-order theories are unable to control the cardinality of their infinite models, and that no first-order theory with an infinite model can have exactly one model up to isomorphism.

Finite and infinite models (4)

I conclude this part of the lecture by observing that the above meta-theorems mean that only theories Γ which bound the size of a model have any possibility of specifying a unique model up to isomorphism.