Logic for Linguists: Lecture 3

Gregory Wilsenach

University of Cambridge

30th October 2019

First-Order Logic: A Review (1)

Last week we introduced the syntax and semantics of first-order logic.

We start with a vocabulary and then build up first-order formulas by first defining terms, from which we define atomic formulas, and then, by taking Boolean combinations and applying quantifiers, we define formulas.

For example, if we take $\tau := \{P, S\}$, where P and S are binary relation symbols, meant to denote the relations "biological parent of" and "sibling of", then the following is a first-order formula over τ:

$$\phi := \forall x \exists y \exists z [\neg(x = y) \wedge \neg(y = z) \wedge \neg(x = z) \\ \wedge (S(x, y) \wedge S(y, x) \wedge P(z, x) \wedge P(z, y))].$$

First-Order Logic: A Review (2)

In order to make sense of such a formula we need the notion of a model. We can think of a model as a concrete world or set of circumstances to which such a sentence might refer. To that end a model must specify a universe of possible objects that a formula might quantify over and must instantiate the relevant relation, constant, and function symbols.

We might consider the τ-model $\mathcal{M}$ with universe $\{a, b, c, d\}$ such that

- $P^{\mathcal{M}} = \{(a, b), (a, c), (a, d)\}$, and
- $S^{\mathcal{M}} = \{(a, b), (b, c), (c, a)\}$.

Is $\mathcal{M}$ a model of ϕ? (i.e. Do we have $\mathcal{M} \models \phi$?)

If not, what could be added to $\mathcal{M}$ to make it a model of ϕ?

Proofs and Their Importance

In this lecture we will discuss the theory of proofs. In the case of propositional logic we could check if a statement was valid by writing out the truth table, i.e. by checking the value of the formula for every valuation. The truth table is finite, but there are 2^v where v is the number of variables in the formula, so this is always possible to do.

In first-order logic in order to check if a statement is valid in the same sense we would need to check that it holds in every model. But this is a big problem! The set of all models is infinite in general!

In order to solve this problem we introduce a finite syntactic notion that allows us to verify that some new statement is true by starting with other statements, which we call axioms, and reasoning from them. We call these syntactic objects proofs.

In this lecture we will introduce the natural proof system.

The Natural Proof System

A proof from some set of axioms is a sequence of steps, where each step is either an axiom, an assumption, or a statement that follows from previous steps via some rule of inference. We will spend much of the rest of the lecture specifying these rules of inference. We write a rule of inference as follows

$$\frac{\phi_1 \qquad \dots \qquad \phi_n}{\psi}$$

This notation is intended to denote that from $\phi_1, \dots, \phi_n$ we can immediately derive ψ. We have for each connective ($\wedge$, $\vee$, $\neg$, $\rightarrow$) and each quantifier ($\forall$, $\exists$) rules of inference that introduce the given symbol and rules that eliminate the given symbol. We chain together these rules in order to build-up the proof in a treelike structure, which will look something like this:

$$\frac{\dfrac{\dfrac{A}{B} \qquad C}{D} \qquad \dfrac{\dfrac{E \qquad F}{G}}{H}}{J}$$

More on Proof Notation

When we are proving a result we write

$$\begin{array}{c}[\phi] \\ \psi\end{array}$$

to denote that by assuming ϕ we can prove ψ. Once the assumption has been used somewhere we say that it has been discharged and it is no longer considered an assumption for our proof. All assumptions that are not members of our starting set of axioms *must* be discharged.

Let τ be a vocabulary. Let Γ be a theory (i.e. a set of sentences) over τ and let ϕ be a formula over τ. We write

$$\Gamma \vdash \phi$$

to denote that there is a proof from the set of axioms Γ that derives ϕ. The previous example witnesses that

$$\{A, C, E, F\} \vdash J.$$

Conjunction Rules

The first rules we will consider are for conjunctions of propositions. We can derive from two propositions their conjunction and from their conjunction we can derive each proposition. We write this in symbols as follows:

$$\frac{\phi \qquad \psi}{\phi \wedge \psi} \qquad \frac{\phi \wedge \psi}{\phi} \qquad \frac{\phi \wedge \psi}{\psi}$$

Examples:

$$\frac{\exists x L(x) \qquad \forall x Y(x)}{(\exists x L(x)) \wedge (\forall x Y(x))} \qquad \frac{(\exists x L(x)) \wedge (\forall x Y(x))}{\exists x L(x)} \qquad \frac{(\exists x L(x)) \wedge (\forall x Y(x))}{\forall x Y(x)}$$

We could really put any propositions in here.

Implication Rules

We now consider rules for implication.

$$\frac{\begin{matrix}[\phi] \\ \psi\end{matrix}}{\phi \to \psi} \qquad\qquad \frac{\phi \qquad \phi \to \psi}{\psi}$$

Let I be the unary relation symbol meant to denote "from Iowa" and A be the unary relation symbol meant to denote "from America". We have the following examples:

$$\frac{\begin{matrix}[I(t)] \\ A(t)\end{matrix}}{I(t) \to A(t)} \qquad\qquad \frac{I(t) \qquad I(t) \to A(t)}{A(t)}$$

Disjunction Rules

We next consider rules for disjunction.

$$\frac{\phi}{\phi \vee \psi} \qquad \frac{\phi}{\psi \vee \phi} \qquad \frac{\phi \vee \psi \quad \begin{matrix}[\phi] \\ \theta\end{matrix} \quad \begin{matrix}[\psi] \\ \theta\end{matrix}}{\theta}$$

Let NY be the unary relation symbol meant to denote "from New York". We have the following examples:

$$\frac{I(t)}{I(t) \vee \psi} \qquad \frac{I(t) \vee NY(t) \quad \begin{matrix}[I(t)] \\ A(t)\end{matrix} \quad \begin{matrix}[NY(t)] \\ A(t)\end{matrix}}{A(t)}$$

Biconditional Rules

We next consider rules for Biconditional statements (i.e. if and only if).

$$\frac{\begin{matrix} [\phi] & [\psi] \\ \psi & \phi \end{matrix}}{\phi \leftrightarrow \psi} \qquad \frac{\phi \quad \phi \leftrightarrow \psi}{\psi} \qquad \frac{\psi \quad \phi \leftrightarrow \psi}{\phi}$$

We let BA denote the unary relation symbol meant to denote "from the big apple". We have the following examples:

$$\frac{\begin{matrix} [NY(t)] & [BA(t)] \\ BA(t) & NY(t) \end{matrix}}{NY(t) \leftrightarrow BA(t)} \qquad \frac{NY(t) \quad NY(t) \leftrightarrow BA(t)}{BA(t)}$$

Negation

We next consider rules for negation.

$$\frac{\begin{matrix}[\phi] & & [\phi] \\ \psi & & \neg\psi\end{matrix}}{\neg\phi}$$

$$\frac{\begin{matrix}[\neg\phi] & & [\neg\phi] \\ \psi & & \neg\psi\end{matrix}}{\phi}$$

From Propositional Logic to First-Order Logic, an Intermission

We should note that so far the proof rules we have considered involve conjunctions, disjunctions, negations, and implications. These are the usual propositional connectives.

The proof system we have thus far built up can thus be applied perfectly well to propositional logic.

Before we go on to discuss those inference rules particular to first-order logic we first discuss an example.

Example (1)

We should like to prove:

$$\vdash (A \to B) \to ((A \wedge C) \to (B \wedge C))$$

The theorem we want to prove is an implication, i.e. it is of the form $\psi \to \phi$ where $\psi = A \to B$ and $\phi = (A \wedge C) \to (B \wedge C)$. When constructing a proof it is often useful to work backwards. We need to introduce the implication symbol, so we know our proof must be of the form

$$\dfrac{\begin{array}{c}[A \to B] \\ (A \wedge C) \to (B \wedge C)\end{array}}{(A \to B) \to ((A \wedge C) \to (B \wedge C))}$$

Note 1: This is not a proof! We have not shown how to derive that statement just above the line yet.

Note 2: We can use the assumption $A \to B$ as many times as we like, we know for sure it will be discharged in the final step.

Example (2)

We next notice that $(A \wedge C) \to (B \wedge C)$ is also an implication, and so we can similarly prove it by assuming $A \wedge C$ and proving $B \wedge C$, and so our proof now looks something like:

$$\dfrac{\dfrac{\begin{array}{c}[A \to B],\ [A \wedge C] \\ B \wedge C\end{array}}{(A \wedge C) \to (B \wedge C)}}{(A \to B) \to ((A \wedge C) \to (B \wedge C))}$$

Example (3)

So now how do we prove $B \wedge C$? We could prove B and C separately and then use the $\wedge$-introduction rule. Our proof is now looks something like:

$$\dfrac{\dfrac{\dfrac{\begin{matrix}[A \to B], [A \wedge C] \\ B\end{matrix} \qquad \begin{matrix}[A \to B], [A \wedge C] \\ C\end{matrix}}{B \wedge C}}{(A \wedge C) \to (B \wedge C)}}{(A \to B) \to ((A \wedge C) \to (B \wedge C))}$$

Example (4)

Now, how shall we prove both B and C? Well we can prove C immediately from $A \wedge C$ using the $\wedge$-elimination rule. Our proof now looks something like:

$$\dfrac{\dfrac{\dfrac{\dfrac{[A \to B],\ [A \wedge C]}{B} \qquad \dfrac{[A \wedge C]}{C}}{B \wedge C}}{(A \wedge C) \to (B \wedge C)}}{(A \to B) \to ((A \wedge C) \to (B \wedge C))}$$

Example (5)

To prove B we notice that we can deduce A from $A \wedge C$ using the $\wedge$-elimination and then deduce B from $A \to B$ using $\to$-elimination. We arrive finally at:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[A \wedge C]}{A} \qquad [A \to B]}{B} \qquad \dfrac{[A \wedge C]}{C}}{B \wedge C}}{(A \wedge C) \to (B \wedge C)}}{(A \to B) \to ((A \wedge C) \to (B \wedge C))}$$

Q.E.D. Now back to inference rules for first-order logic.

Universal Quantification

We first need to discuss the notion of a substitution. For a term t, a variable x, and a formula ϕ we write $\phi[t/x]$ to denote the formula defined by replacing each instance of the variable x with the term t. We now consider the rules for universal quantification.

$$\frac{\phi[t/v]}{\forall v \phi} \qquad \frac{\forall v \phi}{\phi[t/v]}$$

We should note that both of the above rules assume that t does not appear in ϕ or any undischarged assumption in the proof of $\phi[t/v]$. Let F be the binary relation meant to denote "is friends with" and let Bob be a constant symbol. We have the following examples:

$$\frac{\exists x F(v,x)[t/v]}{\forall v \exists x F(v,x)} \qquad \frac{\forall v \exists x F(v,x)}{\exists x F(\text{Bob},x)}$$

Existential Quantification

The rules for universal quantification are as follows:

$$\frac{\phi[t/v]}{\exists v\phi}$$

$$\frac{\exists v\phi \qquad \begin{matrix}[\phi[t/v]] \\ \psi\end{matrix}}{\psi}$$

We should note that both of the above rules assume that t does not appear in $\exists v\phi$ or any undischarged assumption other than $\phi[t/v]$ in the proof of ψ. We have the following examples:

$$\frac{\exists x R(v,x)[\text{Bob}/v]}{\exists v \exists x R(v,x)}$$

$$\frac{\exists v M(\text{Mary}, v) \qquad \begin{matrix}[M(\text{Mary}, v)[v/t]] \\ \psi\end{matrix}}{\psi}$$

Equality

The final proof rules are those for equality.

$$\frac{}{t = t} \qquad \frac{\phi[s/v] \qquad s = t}{\phi[t/v]} \qquad \frac{\phi[s/v] \qquad t = s}{\phi[t/v]}$$

We have the following examples:

$$\frac{}{\text{Bob} = \text{Bob}} \qquad \frac{\exists x F(x, \text{Bob}) \qquad t = \text{Bob}}{\exists x F(x, t)}$$

Try and Prove:

$$\exists x \forall y R(x, y) \vdash \forall y \exists x R(x, y)$$

An instance of this problem: “if there is a person who is friends with everybody then for every person there exists at least one person who is their friend.”

What about the converse of this statement: “if every person has at least one friend then there is one person who is friends with everybody.” Is that true? What goes wrong if you try to prove

$$\forall x \exists y R(x, y) \vdash \exists y \forall x R(x, y)?$$

Soundness

Let τ be a vocabulary. Let Γ be a theory and ϕ a formula over τ. We recall that Γ entails ϕ if for every τ-model $\mathcal{M}$ if $\mathcal{M} \models \Gamma$ then $\mathcal{M} \models \phi$.

We say that a proof system is sound if for every vocabulary τ and every theory Γ and formula ϕ over γ if $\Gamma \vdash \phi$ then Γ entails ϕ. In other words, if ϕ is provable from Γ then if all of the sentences in Γ are true then ϕ is true. This might informally say that "all provable implications are valid."

It would be wonderful to have the converse, that "all valid implications are provable"...

Completeness

We say that a proof system is complete if for every theory Γ and every formula ϕ if Γ entails ϕ then $\Gamma \vdash \phi$.

Astoundingly, this proof system is complete for first-order logic.

Theorem (The Completeness Theorem)

The natural deduction system for first-order logic is both sound and complete.

I mentioned earlier that there are many other proof systems for first-order logic, including: Herlbert-style deduction, the sequent calculus, the tableaux method, resolution, etc. We can prove completeness for appropriate systems in each case.

Summary

In this lecture we

- introduced the notion of a formal proof and discussed the natural proof system,
- defined all of the rules for natural proofs for propositional logic and gave an example,
- extended this proof system in order to define the natural proof system for first-order logic, and
- defined the notions of soundness and completeness and stated the completeness theorem.

Reading Material

Here are some good notes on natural deduction:

- https://cs.anu.edu.au/courses/comp2600/lectures/PropND.pdf
- https://cs.anu.edu.au/courses/comp2600/lectures/FirstOrderND.pdf
- https://leanprover.github.io/logic_and_proof/natural_deduction_for_first_order_logic.html

Here is a very short paper critically discussing natural deduction in the context of natural language:

- http://www.cs.toronto.edu/~trebla/eq-wiltink.pdf

Here is a paper using natural deduction to automate inferring knowledge from a provided text:

- http://www.cs.vsb.cz/duzi/Paper143.pdf