6c Lecture 13: May 12, 2015

11 The completeness theorem

Our goal today is the following theorem:

Theorem 11.1 (Gödel). *If L is a language, S is a set of sentences in the language L, and ϕ is a sentence in the language L, then $S \models \phi$ iff $S \vdash \phi$.*

The fact that $S \vdash \phi \rightarrow S \models \phi$ (i.e. the soundness of our proof system) is an easy induction on proofs. We simply check that each of our logical axioms is logically valid, and that modus ponens preserves validity. It will be a homework exercise.

The direction $S \models \phi \rightarrow S \vdash \phi$ (i.e. the completeness of our proof system) is more difficult. Like our proof of the completeness of propositional logic, we will prove the contrapositive: $\neg S \vdash \phi \rightarrow \neg S \models \phi$.

Recall that a set S of formulas is *inconsistent* if there is some formula ψ such that both $S \vdash \psi$ and $S \vdash \neg\psi$. Otherwise, we say S is *consistent*. We claim since $\neg S \vdash \phi$, then $S \cup \{\neg\phi\}$ must be consistent. This is because if it were inconsistent, then $S \vdash \phi$ via proof by contradiction (which was a homework problem).

Now recall that $S \models \phi$ means that for every model M of S, we have $M \models \phi$. Hence, to show $\neg S \models \phi$ we must find some model M of $S \cup \{\neg\phi\}$ (i.e. $M \models S$ and $M \models \neg\phi$).

Let's recap. To prove the completeness theorem, it is enough to show that if a theory T is consistent, then there is a model M of T. (And then we can apply this to $T = S \cup \{\neg\phi\}$ above).

To make this model, we will begin by finding a complete consistent theory extending T and then we will construct the model M from this complete theory (much as we constructed a valuation from a complete theory in our proof of the completeness theorem for propositional logic).

Recall that a theory T is *complete* if for any sentence ϕ we have either $\phi \in T$ or $\neg\phi \in T$. Note that if T is consistent and complete, the for any sentence ϕ, exactly one of $\phi \in T$ or $\neg\phi \in T$.

We begin with a key definition:

Definition 11.2. A *Henkin theory* is a theory T so that for every every formula ϕ and variable x, there is a constant symbol c such that $(\exists x\phi \rightarrow \phi^x_c) \in T$, where ϕ^x_c is the formula ϕ replacing each free occurrence of the variable x by the constant c. The constant c is called a *Henkin witness* for $\exists x\phi$.

The rough idea here is that in a Henkin theory, we have an actual constant naming a witness to $\exists x\phi$ whenever it is true.

For an example, lets consider the theory T of the structure $\langle \mathbb{N}; +, \cdot, 0, 1\rangle$. For the formula $\phi = \forall y(x + y = y)$, we can use the constant 0 for x to witness $\exists x\phi$. That is, $\exists x\forall y(x + y = y) \rightarrow \forall y(0 + y = y)$ is true. However, for the formula $\phi = \forall y(y + y = x \cdot y)$, there is no constant c so that $\exists x\forall y(y + y = x \cdot y) \rightarrow \forall y(y + y = c\dot{y})$ is true, since we don't have a constant symbol for the number 2 in our language. (You might be tempted to use the term $1+1$ here as a witness, but remember that formally, this is a term (it comes from one of our functions applied to constants) and it isn't a constant itself). Hence, T is not a Henkin theory.

Now consider the structure $\langle \mathbb{N}; +, \cdot, 0, 1, 2, \ldots\rangle$ containing a constant for each natural number. Then it is easy to see that the theory of this structure has the Henkin property; for each formula ϕ, if $\exists x\phi$ is true, then there must be some constant n such that ϕ is true when we replace x by n, and hence $\exists x\phi \rightarrow \phi^x_n$. If $\exists x\phi$ is false, then clearly $\exists x\phi \rightarrow \phi^x_c$ is vacuously true for any constant c.

Now we will prove two important lemmas about Henkin theories which will together immediately prove the completeness theorem.

Lemma 11.3. *Suppose L is a language and T is a consistent theory in the language L. Then there is a first order language $L' \supseteq L$ obtained by adding some constant symbols to L and a theory $T'' \supseteq T$ such that T'' is consistent, complete, and has the Henkin property.*

Proof. We do this proof in two steps.

Step 1: We will make a first-order language $L' \supseteq L$ by adding constant symbols to L, and a consistent Henkin theory $T' \supseteq T$ in the language L'.

Step 2: We will then finish by constructing a complete consistent $T'' \supseteq T'$ in the language L'. Note that since T' has the Henkin property, so does T'', since they are in the same language and T'' contains all the sentences of T'. We will omit this second step of the proof and leave it for homework, since it is essentially the same technique we used in our proof the completeness theorem for propositional logic for extending any consistent theory to a complete consistent theory.

We do the first step. We will define a sequence $L_0 \subseteq L_1 \subseteq L_2 \subseteq \ldots$ of languages and a sequence of theories $T_0 \subseteq T_1 \subseteq T_2 \subseteq \ldots$ such that each L_{n+1} is obtained from L_n by adding new constant symbols, and each theory T_n is consistent. To begin, set $L_0 = L$, and $T_0 = T$.

Now given L_n, let L_{n+1} be the language obtained by adding a new constant $c_{\exists x\phi}$ for each sentence $\exists x\phi$ in the language L_n. Let

$$T_{n+1} = T_n \cup \{\exists x\phi \rightarrow \phi^x_{c_{\exists x\phi}} : \phi \text{ is a sentence in the language } L_n\}$$

be the theory obtained by adding to T_n the sentence $\exists x\phi \rightarrow \phi^x_c$ for each sentence ϕ in L_n.

Define $T' = \bigcup_n T_n$ and $L' = \bigcup_n L_n$. Clearly L' is obtained from L by adding constants. Next, we show T' is a Henkin theory. If ϕ is a sentence in L', then $\phi \in L_n$ for some n, and hence by definition the sentence $\exists x\phi \rightarrow \phi^x_{c_{\exists x\phi}}$ is in T_{n+1} and thus in T'.

To finish, we show that T' is consistent. Since proofs are finite, if $T' \vdash \phi$, then $T_n \vdash \phi$ for some large enough n. Thus, it is enough to show that each T_n is consistent. We can do this by induction. For our base case, note by assumption that T_0 is consistent. Inductively, suppose T_n is consistent. We must now show T_{n+1} is consistent.

So assume by way of contradiction that T_{n+1} is inconsistent. Then since T_{n+1} is obtained from T_n by adding formulas of the form $\exists x\phi \to \phi^x_{c_{\exists x\phi}}$, there must be some finite number of such formulas $\phi_1, \dots, \phi_k$ such that

$$T_n \cup \{\exists x\phi_i \to (\phi_i)^x_{c_{\exists x\phi_i}} : i \leq k\}$$

is inconsistent. Then using proof by contradiction,

$$T_n \cup \{\exists x\phi \to (\phi_i)^x_{c_{\exists x\phi_i}} : i \leq k-1\} \vdash \neg(\exists x\phi_k \to (\phi_k)^x_{c_{\exists x\phi_k}})$$

Let $Q = T_n \cup \{\exists x\phi \to (\phi_i)^x_{c_{\exists x\phi_i}} : i \leq k-1\}$. Then since $(\neg(p \to q)) \to p$ and $(\neg(p \to q)) \to \neg q$ are tautologies, using our first logical axiom,

$$Q \vdash (\exists x\phi_i)$$

and

$$Q \vdash \neg(\phi_i)^x_{c_{\exists x\phi_i}}.$$

But then by generalization on constants applied to this last fact, we see $Q \vdash \forall x(\neg\phi_i)$, and hence $Q \vdash \neg\exists x\phi_i$, so Q is inconsistent.

Proceeding this way we can eliminate all the sentences $\exists x\phi \to (\phi_i)^x_{c_{\exists x\phi_i}}$ for $i = k, \dots, 1$ and show that T_n itself is inconsistent, which is a contradiction. □

Lemma 11.4. *Suppose L' is a language and T' is a complete consistent Henkin theory. Then there is a model M of T'.*

Proof. Consider the set X of all terms in the language L' that do not contain variables (we call these *closed terms*). Now define an equivalence relation $\sim$ on X by $s \sim t$ iff the sentence $s = t$ is in T'.

First, we have to check that this is actually an equivalence relation. For all closed terms r, s, t:

- $t \sim t$ since for each term t, $t = t$ is a logical axiom hence $T' \vdash t = t$, so $t = t \in T'$ since T' is complete and consistent.
- $s \sim t$ implies $t \sim s$ since if $s = t \in T'$, then since $s = t \to t = s$ is a logical axiom, $T' \vdash t = s$, so $t = s \in T'$.
- $r \sim s$ and $s \sim t$ implies $r \sim t$ since if $r = s$ and $s = t \in T'$, then $T' \vdash r = s \wedge s = t$, and since $(r = s \wedge s = t) \to r = t$ is a logical axiom, we have $T' \vdash r = t$, so $r = t \in T'$.

For each closed term s, denote by $[s] = \{t : t \sim s\}$ the equivalence class of the closed term s, and let $A = \{[t] : t \text{ is a closed term}\}$ be the set of all equivalence classes of $\sim$. This will be the universe of our model M. We now need to define the interpretations of the relations and function symbols of L'.

For each n-ary function symbol f define its interpretation in M by

$$f^M([t_1], \ldots [t_n]) = [f(t_1, \ldots, t_n)].$$

(Where $f(t_1, \ldots, t_n)$ is a term in our language, and $[f(t_1, \ldots, t_n)$ is its equivalence class. We need to show this is well defined. That is, if $[s_1] = [t_1], \ldots, [s_n] = [t_n]$, then $f^M([s_1], \ldots [s_n]) = f^M([t_1], \ldots [t_n])$. This follows from the fact that $s_1 = t_1 \wedge \ldots \wedge s_n = t_n \rightarrow f(s_1, \ldots, s_n) = f(t_1, \ldots, t_n)$ is a logical axiom.

Note that in the special case of constant c (i.e. a 0-ary function), we have that the interpretation c^M of c is $[c]$. Indeed, for any closed term t we have that $t^M = [t]$, which can be proved by induction on the construction of t.

Similarly, for each n-ary relation, R, we define its interpretation in M by $R^M([t_1], \ldots [t_n])$ iff $R(t_1, \ldots, t_n) \in T$. This is similarly well-defined since $s_1 = t_1 \wedge \ldots \wedge s_n = t_n \rightarrow R(s_1, \ldots, s_n) \rightarrow R(t_1, \ldots, t_n)$ is a logical axiom.

We have now finished our definition of the structure M. All we need to do to finish is to show that $M \vDash T'$. We will do this by showing that for each sentence ψ of L', we have $M \vDash \psi$ iff $\psi \in T'$.

We prove this by induction on the number of connectives and quantifiers in sentences. For our base case, we consider atomic sentences. We first consider atomic sentences of the form $s = t$ where s and t are closed terms. Then we have that $M \vDash s = t$ iff $s^M = t^M$ iff $[s] = [t]$ iff $s \sim t$ iff $s = t \in T'$. Similarly, for atomic sentences of the form $R(t_1, \ldots, t_n)$ we have $M \vDash R(t_1, \ldots, t_n)$ iff $R^M(t_1^M, \ldots, t_n^M)$ iff $R^M([t_1], \ldots, [t_n])$ iff $R(t_1, \ldots, t_n) \in T'$

Now for our induction step, suppose we have sentences ϕ and ψ.

($\wedge$): Then $M \vDash \phi \wedge \psi$ iff $M \vDash \phi$ and $M \vDash \psi$ iff $\phi \in T'$ and $\psi \in T'$ by our induction hypothesis iff $\phi \wedge \psi \in T'$ since T' is consistent and complete.

($\neg$): $M \vDash \neg\phi$ iff it is not the case that $M \vDash \phi$ iff $\phi \notin T'$ by our induction hypothesis iff $\neg\phi \in T'$ since T' is consistent and complete.

(The other logical connectives $\rightarrow$ and $\vee$ are similar).

To finish, we want to show for sentences of the form $\forall x \phi$, we have $M \vDash \forall x \phi$ iff $\forall x \phi \in T'$.

Now since T' is a Henkin theory, and recalling that $\exists x$ is an abbreviation for $\neg \forall x \neg$, there is a constant c and a formula of the form

$$\neg \forall x \neg(\neg\phi) \rightarrow (\neg\phi)^x_c \in T'$$

and since T' is complete,

$$\neg \forall x \phi \rightarrow (\neg\phi)^x_c \in T'$$

and so

$$\phi^x_c \rightarrow \forall x \phi \in T'$$

If $M \models \forall x\phi$, then certainly $M \models \phi[x \mapsto [c]]$ and so $M \models \phi^x_c$ which implies $\phi^x_c \in T'$ by our induction hypothesis and since $\phi^x_c \to \forall x\phi \in T'$ and T' is complete we have $\forall x\phi \in T'$.

Conversely, assume $\forall x\phi \in T'$. Then since for every closed term t, $\forall x\phi \to \phi^x_t$ is a logical axiom and T' is complete, we have $\phi^x_t \in T'$ and hence by our induction hypothesis $M \models \phi^x_t$, so $M \models \phi[x \mapsto t^M]$ so $M \models \phi[x \mapsto [t]]$. But since every element of M is of the form $[t]$, the we see $M \models \forall x\phi$. □

To finish our proof of the completeness theorem, start with Lemma 11.3 and obtain a complete consistent Henkin theory $T' \supseteq T$ in the language $L' \supseteq L$. Then use Lemma 11.4 to obtain a model M of T'. Finally, by restricting the model M to just the language L, we have a model of T.