CHAPTER 4

Sounds like things are complete

This is the part of the chapter where we follow a pithy little narrative about the night some poor traveller (referred to mostly in second person) is spending in Hilbert's Hotel. We are now four chapters in, and our fellow denizen is now finally asleep.

I'm not one to impose dreams on them, but if I were to imagine what they are dreaming, I'd imagine that they're dreaming of *you*,[1] our fellow reader, suffering through the proof of the Substitution Lemma...

What a nightmare! Thankfully, there are better times ahead.

1. Tautologies are still true!

1.1. Universally valid formulas...

Definition 1.1.1. An $\mathcal{L}$-formula ϕ is called *universally valid* if for every $\mathcal{L}$-structure $\mathcal{M}$ and every assignment $\alpha : \mathsf{Var} \to \mathcal{M}$, we have that $\mathcal{M} \models \phi[\alpha]$. A *tautology* is a universally valid $\mathcal{L}$-sentence. We write $\models \phi$ to indicate that ϕ is universally valid.

Example 1.1.2. Let $\mathcal{L}$ be a language with a single binary relation symbol $\underline{R}$. Then:

(1) $(x = y) \to (\underline{R}(x, y) \leftrightarrow \underline{R}(y, x))$ is universally valid.

(2) $((\exists x)(\forall y)\underline{R}(x, y)) \to (\forall y)(\exists x)\underline{R}(x, y)$ is a tautology.

(3) The sentence $(\exists x)(x = x)$ is a tautology.

(4) The formula $(\exists x)x \neq y$ is not universally valid.

Exercise 1.1.3. Prove (1)-(4) in the previous example carefully.

The following exercise connects tautologies with universally valid formulas. Since $\forall$ is sometimes referred to as **universal quantification**, it also justifies our terminology.

[1] Different use of the pronoun?

Exercise 1.1.4. Show that the formula $\phi(x_1, \ldots, x_n)$ is universally valid if and only if the sentence $(\forall x_1) \cdots (\forall x_n)\phi$ is a tautology. The sentence $(\forall x_1) \cdots (\forall x_n)\phi$ is called the **universal closure** of $\phi(x_1, \ldots, x_n)$.

An important fact about universally valid formulas is that their universal validity is preserved under **expansions** (i.e. under the operation of adding more symbols to our language $\mathcal{L}$ to obtain a new language $\mathcal{L}' \supseteq \mathcal{L}$ and interpreting, in an $\mathcal{L}$-structure all symbols of $\mathcal{L}$ as we did previously).

Lemma 1.1.5. *Let ϕ be an $\mathcal{L}$ formula and consider a language $\mathcal{L}' \supseteq \mathcal{L}$. Then, ϕ is universally valid as an $\mathcal{L}$-formula if and only if it is universally valid as an $\mathcal{L}'$-formula.*

PROOF. It's enough to note that any $\mathcal{L}$-structure has an expansion to an $\mathcal{L}'$-structure. $\square$

Exercise 1.1.6. Why, in the proof of the previous lemma, is it enough to note that any $\mathcal{L}$-structure has an expansion to an $\mathcal{L}'$-structure?

1.2. ...and the Propositional Calculus (again). In all honesty if when we dropped the propositional logic stuff and started all of this first-order logic business you felt like all of that hard work was for nothing, well, you were somewhat right, but as a matter of fact, not *totally* right. Propositional logic gives us a way of generating universally valid formulas!

First, a little lemma connecting propositional logic with first-order logic in a totally expected way.

Lemma 1.2.1. *Let ϕ be a propositional formula, with $\mathsf{Var}(\phi) = \{A_1, \ldots, A_n\}$. Let $\psi_1, \ldots, \psi_n$ be $\mathcal{L}$-formulas. Let $\mathcal{M}$ be an $\mathcal{L}$-structure and $\alpha : \mathsf{Var} \to \mathcal{M}$ an assignment. Let $\mathcal{A}_\alpha : \mathsf{Var} \to \{T, F\}$ be the propositional assignment defined by:*

$$\mathcal{A}_\alpha : \mathsf{Var} \to \{T, F\}$$
$$A_i \mapsto \begin{cases} T & \textit{if } \mathcal{M} \vDash \psi_i[\alpha] \\ F & \textit{otherwise} \end{cases}$$

Then, the following are equivalent:

(1) $\mathcal{M} \vDash (\phi[\psi_1/A_1, \ldots, \psi_n/A_n])[\alpha]$.[2]

[2]Here, of course, I'm a notation abuser who is writing $\phi[\psi_1/A_1, \ldots, \psi_n/A_n]$ to mean the first-order formula obtained by replacing every instance of A_i in ϕ by ψ_i.

(2) $\phi[\mathcal{A}_\alpha] = T$.

Proof. We argue by induction on the structure of ϕ. Indeed, if ϕ is a propositional variable, then this is immediate. Now, if ϕ is of the form $\phi_1 \wedge \phi_2$, then we have:

$$\begin{aligned}(\phi_1 \wedge \phi_2)[\mathcal{A}_\alpha] = T &\text{ iff } \phi_1[\mathcal{A}_\alpha] \text{ and } \phi_2[\mathcal{A}_\alpha] = T\\ &\text{ iff } \mathcal{M} \vDash (\phi_1[\psi_1/A_1, \ldots, \psi_n/A_n])[\alpha] \text{ and } \mathcal{M} \vDash (\phi_2[\psi_1/A_1, \ldots, \psi_n/A_n])[\alpha]\\ &\text{ iff } \mathcal{M} \vDash (\phi_1 \wedge \phi_2[\psi_1/A_1, \ldots, \psi_n/A_n])[\alpha].\end{aligned}$$

The other cases follow similarly (see next exercise). CRUCIALLY we are doing induction on the structure of a PROPOSITIONAL formula, so we never need to consider quantifiers! ☐

Exercise 1.2.2. Finish the proof of the previous lemma.

In a sense, what this tells us is that if a first-order formula is built up in a propositional way (e.g. if it is quantifier-free), then to evaluate satisfaction, we just need to do a truth table!

Corollary 1.2.3. *In the notation of the previous lemma, if ϕ is a propositional tautology, then $\phi[\psi_1/A_1, \ldots, \psi_n/A_n]$ is a universally valid formula.*

Proof. For any $\mathcal{L}$-structure and any assignment α, we have that $\phi[\mathcal{A}_\alpha] = T$ (as in the previous lemma) is true, and thus we're done. ☐

Exercise 1.2.4. Show that the converse of this corollary is not true. More precisely, write down a first-order formula χ which is of the form $\phi[\psi_1/A_1, \ldots, \psi_n/A_n]$, for some propositional formula ϕ and first-order formulas $\psi_1, \ldots, \psi_n$, and such that χ is universally valid, but ϕ is not.

The upshot of this is that we have everything we discussed in the previous chapter concerning tautologies. So even in the first-order world, we may abbreviate $\wedge$ and $\vee$ and $\neg$ and all other logical connectives using only $\to$. We have already shown that $(\exists x)\phi$ semantically is just an abbreviation for $\neg(\forall x)\neg\phi$ and thus, it follows that, up to logical equivalence, all first-order formulas can be written using the logical symbols $\to$, $\neg$ and $\forall$. Syntactically, this is not quite that obvious (for starters, our syntax does not quite know that structures are supposed to be non-empty). Thus, for the syntactic business below, we will also keep track of $\exists$ (this can be handled in various ways, but that's what we're going with here).

2. New day, new axiome

First-order logic is a much more complicated beast than propositional logic. The old axioms will not suffice. Throughout this chapter, we will fix a first-order language $\mathcal{L}$, and everything we will be doing will be happening in $\mathcal{L}$ (or some expansion of it).

2.1. Axioms for Equality. We need to make sure that our proof system understands that the symbol $\doteq$ behaves like equality. The best (and only) way to do this is to hardcode it:

(E1) Reflexivity: $(\forall x)(x \doteq x)$.

(E2) Symmetry: $(\forall x)(\forall y)(x \doteq y \to y \doteq x)$.

(E3) Transitivity: $(\forall x)(\forall y)(\forall z)(x \doteq y \wedge y \doteq z \to x \doteq z)$.

(E4) For each n-ary relation symbol $\underline{R} \in \mathsf{Rel}(\mathcal{L})$ a "congruence" axiom:

$$(\forall x_1)\ldots(\forall x_n)(\forall y_1)\ldots(\forall y_n)\left(\bigwedge x_i = y_i \to (\underline{R}(x_1,\ldots,x_n) \to \underline{R}(y_1,\ldots,y_n))\right).$$

(E5) For each n-ary function symbol $\underline{f} \in \mathsf{Fun}(\mathcal{L})$ a "congruence" axiom:

$$(\forall x_1)\ldots(\forall x_n)(\forall y_1)\ldots(\forall y_n)\left(\bigwedge x_i = y_i \to (\underline{f}(x_1,\ldots,x_n) \doteq \underline{f}(y_1,\ldots,y_n))\right).$$

Exercise 2.1.1. Show that (E1)-(E5) are tautologies.

2.2. Quantifier Axioms. We also need some axioms that tell our proof system what the deal with quantifier is. We'll be somewhat economical here:

(Q1) For every $\mathcal{L}$-formula ϕ such that $x \notin \mathsf{Free}(\phi)$ and every $\mathcal{L}$-formula ψ, the axiom:

$$(\forall x)(\phi \to \psi) \to (\phi \to (\forall x)\psi).$$

(Q2) For every $\mathcal{L}$-formula ϕ and every $\mathcal{L}$-term t the axiom:

$$(\forall x)\phi \to \phi[t/x].$$[3]

(Q3) For every $\mathcal{L}$-formula ϕ and every $\mathcal{L}$-term t the axiom:

$$\phi[t/x] \to (\exists x)\phi$$

[3]This is where the Substitution Lemma comes in really handy. If we had defined substitutions in a more naive way, then we would need to restrict instances of this axiom to terms t such that there is no free occurrence of x in ϕ which lies within the scope of a quantification that binds a variable in t. We don't need to worry about that though!

(Q4) For every $\mathcal{L}$-formula ϕ, the axiom:

$$\neg(\forall x)\neg\phi \leftrightarrow \exists\phi.$$

Exercise 2.2.1. Show that (Q1)-(Q4) are tautologies.

[*Hint.* If you've proved that (Q2) and (Q3) are tautologies without having used the Substitution Lemma, then something's gone wrong.]

2.3. Propositional Axioms. Well we can't get away from propositional tautologies, can we. Let's recall our three axioms from back in the day, but now where we allow ϕ, ψ and χ to be arbitrary $\mathcal{L}$-formulas:

(A1) $(\phi \to (\psi \to \phi))$

(A2) $((\phi \to (\psi \to \chi)) \to ((\phi \to \psi) \to (\phi \to \chi)))$

(A3) $((\neg\phi \to \neg\psi) \to ((\neg\phi \to \psi) \to \phi))$

Since these are substitutions into propositional tautologies, we have the following, right away:

Remark 2.3.1. All instances of (A1)-(A3) are universally valid.

3. So what's a proof, again?

A proof system is the same thing it was last time we discussed proof systems. It'd be funny if the concept had changed. That being said, this time around for first-order logic, our proof system will have all the axioms we listed above. We'll need to throw in a new scary looking rule, too. Let's summarise:

- Axioms: (E1)-(E5), (Q1)-(Q4), and (A1)-(A3)
- Deduction Rules:

(MP) GIVEN: $\phi \to \psi$ and ϕ
DEDUCE: ψ

(Gen) GIVEN: ϕ
DEDUCE: $(\forall x)\phi$

A formal proof of formula ϕ is still a finite sequence:

$$(\phi_1, \ldots, \phi_n)$$

of formulas such that $\phi_n = \phi$ and for each $i \leq n$, one of the following holds:

- Either ϕ_i is an instance of an axiom;

- Or ϕ_i can be deduced from an instance of (MP) or (Gen) for some $j, k < i$.

If there is a formal proof of ϕ, we again write $\vdash_{\mathcal{L}} \phi$ (Indeed, if we were working in some $\mathcal{L}' \supseteq \mathcal{L}$ then the notion of $\vdash_{\mathcal{L}'}$ could be different, it'll turn out that it's not but okay we are trying to be formal right now). In this case, as before we shall call ϕ a **theorem** (of $\mathcal{L}$),

More generally (and just like before) let T be an $\mathcal{L}$-theory. We say that ϕ is **deducible** (in $\mathcal{L}$) from T if there is a finite sequence:

$$(\phi_1, \ldots, \phi_n)$$

of formulas such that $\phi_n = \phi$ and for each $i \leq n$, one of the following holds:

- Either ϕ_i is an instance of an axiom;
- Or $\phi_i \in T$
- Or ϕ_i can be deduced from an instance of (MP) or (Gen) for some $j, k < i$.

In this case, we write $T \vdash_{\mathcal{L}} \phi$.

3.1. Simplifying derivations. We showed in the previous chapter that from (A1)-(A3) and (MP), all propositional tautologies can be proved. Thus, we may assume that all first-order instances of propositional tautologies are in our axiom list.

Given the following remark, and me feeling a little splurgy, how about, this time around, since we already have all propositional tautologies that we forgo some formalities and keep all binary connectives in our language (algorithmically, we can substitute them out before a formal proof and then substitute them back in in every step of the proof). This will make our lives rather easy:

Example 3.1.1. Here are some easy yet important deductions:

(1) $\{\phi, \psi\} \vdash_{\mathcal{L}} \phi \wedge \psi$

$$\begin{aligned} \delta_1 :& \phi \to (\psi \to (\phi \wedge \psi)) && \text{(Taut)} \\ \delta_2 :& \psi \to (\phi \wedge \psi) && \text{(MP)} \\ \delta_3 :& \phi \wedge \psi && \text{(MP)} \end{aligned}$$

This would have been possible, but a real pain if we didn't shortcut the derivation of instances of propositional tautologies from (A1)-(A3). Thankfully, all the work we did when showing completeness of propositional logic allows us to do that.

(2) Suppose that $y \notin \mathsf{Var}(\phi)$. Then, $\vdash_{\mathcal{L}} (\forall y)\phi[y/x] \to \forall x\phi$. This is immediate, by how we've set up our axioms, since:

$$\begin{aligned}
&\delta_1 : (\forall y)\phi[y/x] \to (\phi[y/x])[x/y] && \text{(Q2)}\\
&\delta_2 : (\forall x)\,((\forall y)\phi[y/x] \to (\phi[y/x])[x/y]) && \text{(Gen)}\\
&\delta_3 : (\forall x)\,((\forall y)\phi[y/x] \to (\phi[y/x])[x/y]) \to ((\forall y)\phi[y/x] \to (\forall x)((\phi[y/x])[x/y])) && \text{(Q1)}\\
&\delta_4 : (\forall y)\phi[y/x] \to (\forall x)(\phi[y/x][x/y]) && \text{(MP)}
\end{aligned}$$

But, by Lemma 3.2.9, we have that $\phi[y/x][x/y] = \phi$. Hopefully, this starts to justify why we needed all those annoying syntactic lemmas in the previous chapter!

(3) $\vdash_{\mathcal{L}} (\forall x)\phi \to \phi$.

$$\delta_1 : (\forall x)\phi \to \phi[x/x] \quad \text{(Q2)}$$

But $\phi[x/x] = \phi$.

The following lemma gives us an important derivation:

Lemma 3.1.2. *If $\phi \to \psi$ is an instance of a propositional tautology, then*

$$\vdash (\forall x)\phi \to (\forall x)\psi.$$

PROOF. This is a simple derivation:

$$\begin{aligned}
&\delta_1 : && (\forall x)\phi \to \phi && \text{Example 3.1.1(3)}\\
&\delta_2 : && \phi \to \psi && \text{(Ass/Taut)}\\
&\delta_3 : && ((\forall x)\phi \to \phi) \to ((\phi \to \psi) \to ((\forall x)\phi \to \psi)) && \text{(Taut)}\\
&\delta_4 : && ((\phi \to \psi) \to ((\forall x)\phi \to \psi)) && \text{(MP)}\\
&\delta_5 : && (\forall x)(\phi \to \psi) && \text{(MP)}\\
&\delta_6 : && (\forall x)[(\forall x)\phi \to \psi] && \text{(Gen)}\\
&\delta_7 : && (\forall x)[(\forall x)\phi \to \psi] \to [(\forall x)\phi \to (\forall x)\psi] && \text{(Q1)}\\
&\delta_8 : && (\forall x)\phi \to (\forall x)\psi && \text{(MP).}
\end{aligned}$$

The tautology we used for δ_3 is of course:

$$(A \to B) \to ((B \to C) \to (A \to C)).$$

□

Exercise 3.1.3. Show that (Q3) is derivable from the rest of the axioms and deduction rules. More precisely, let ϕ be any formula and t any term. Show that there exists a sequence $(\phi_1, \dots, \phi_n)$, with ϕ_n being $\phi[t/x] \to (\exists x)\phi$, where for each $i \leq n$, one of the following holds:

- Either ϕ_i is an instance of an axiom (E1)-(E5), (Q1),(Q2),(Q4), (A1)-(A3);[4]
- Or ϕ_i can be deduced from an instance of (MP) or (Gen) for some $j, k < i$.

3.2. Sounds sound. Once again, the minimal requirement from $\vdash_{\mathcal{L}}$ is that it only produces deductions of universally valid formulas. More generally:

THEOREM 3.2.1 (Soundness). *Let T be an $\mathcal{L}$ theory and ϕ an $\mathcal{L}$-sentence. If $T \vdash_{\mathcal{L}} \phi$ then $T \vDash \phi$.*

PROOF. Combine Exercises 2.1.1 and 2.2.1 and Remark 2.3.1. Then prove that:

- If $T \vDash \phi$ and $T \vDash \phi \to \psi$ then $T \vDash \psi$.
- If $T \vDash \phi$ then $T \vDash (\forall x)\phi$.

Conclude by induction on the length of derivations. In fact, since it is possible that *formulas* (rather than sentences) show up in the derivation (because of the propositional axioms), we would need to argue a bit more generally, to show that if $T \vdash_{\mathcal{L}} \phi$ for a *formula* ϕ then, in every model of T the universal closure of ϕ is true. Let's not worry too much about that though. $\square$

That sure was fast. Well don't worry about it, that's what HW is for!

[4] By the previous discussion, you can use any propositional tautology here.

Homework 5

4. Building blocks

4.1. Something old. We'll start with things that we've seen before.

THEOREM 4.1.1 (The Deduction Lemma, Redux). *Let T be an $\mathcal{L}$-theory, ϕ an $\mathcal{L}$-sentence and ψ an $\mathcal{L}$-formula. Then, the following are equivalent:*

(1) $T \cup \{\phi\} \vdash_{\mathcal{L}} \psi$.

(2) $T \vdash_{\mathcal{L}} \phi \to \psi$.

PROOF. The structure of the proof is the same as that of the Deduction Lemma back in propositional logic, but now we also have to check what happens when in the induction δ_i is obtained from δ_j $(j \leq i)$ using (Gen). Suppose that this is the case. Then, we must have that δ_i is precisely $(\forall x)\delta_j$. By induction, we have a derivation $T \vdash_{\mathcal{L}} (\phi \to \delta_j)$. Then, by (Gen) we have $T \vdash_{\mathcal{L}} (\forall x)(\phi \to \delta_j)$, and by (Q1), since ϕ is a sentence, we have $T \vdash_{\mathcal{L}} (\forall x)(\phi \to \delta_j) \to (\phi \to (\forall x)\delta_j)$, and thus by (MP) we can conclude that $T \vdash_{\mathcal{L}} (\phi \to (\forall x)\delta_j)$ which is precisely $T \vdash_{\mathcal{L}} (\phi \to \delta_i)$. □

Exercise 4.1.2. Write out the rest of the proof of the Deduction Lemma.

Definition 4.1.3. Let T be an $\mathcal{L}$-theory. We say that T is *inconsistent* if for some sentence ϕ we have that $T \vdash_{\mathcal{L}} \phi$ and $T \vdash_{\mathcal{L}} \neg\phi$. If T is not inconsistent, then we say that it is *consistent*.

Example 4.1.4. Let $\mathcal{M}$ be an $\mathcal{L}$-structure. We define $\mathsf{Th}(\mathcal{M})$ to be:

$$\mathsf{Th}(\mathcal{M}) := \{\phi \text{ an } \mathcal{L}\text{-sentence} : \mathcal{M} \vDash \phi\}.$$

This is a consistent theory (by soundness).

Definition 4.1.5. An $\mathcal{L}$-theory T is *complete* if it is consistent and for all sentences ϕ we have that either $T \vdash_{\mathcal{L}} \phi$ or $T \vdash_{\mathcal{L}} \neg\phi$.

Remark 4.1.6. For any $\mathcal{L}$-structure $\mathcal{M}$, the theory $\mathsf{Th}(\mathcal{M})$ is complete.

Exercise 4.1.7. Let T be an $\mathcal{L}$-theory. Prove that:

(1) T is inconsistent if and only if for all $\mathcal{L}$-formulas ϕ we have that $T \vdash_{\mathcal{L}} \phi$. [*Hint.* You only need to worry about the left-to-right direction. Use the propositional tautology $(\neg\chi \wedge \chi) \to \phi$ for any formula ϕ.]

(2) If $T \vdash_{\mathcal{L}} \phi$ then there is some finite $T_0 \subseteq T$ such that $T_0 \vdash_{\mathcal{L}} \phi$. (You already saw a sketch of a proof for this in the proof of propositional compactness, I hope.)

(3) Deduce that if all finite subsets of T are consistent, then so is T (again, this is just to get you to revise compactness).

(4) Finally, deduce that if $(T_i)_{i\in I}$ is a family of consistent theories such that $T_i \subseteq T_j$ whenever $i, j \in I$, then $\bigcup_{i\in I} T_i$ is consistent.

Corollary 4.1.8. *Let T be an $\mathcal{L}$-theory and ϕ an $\mathcal{L}$-sentence. Then $T \vdash_{\mathcal{L}} \phi$ if and only if $T \cup \{\neg\phi\}$ is inconsistent.*

Proof. One just needs to note that if $T\cup\{\neg\phi\}$ is inconsistent, then $T\cup\{\neg\phi\} \vdash_{\mathcal{L}} \phi$ (by previous exercise). By the Deduction Lemma, we have that $T \vdash_{\mathcal{L}} (\neg\phi \to \phi)$ and since $(\neg\phi \to \phi) \to \phi$ is a propositional tautology we can conclude by (MP). [5] □

Example 4.1.9. Let T be a theory in a language $\mathcal{L}$ and let $(\forall x)\phi$ be an $\mathcal{L}$-sentence. Then $\vdash (\forall x)\neg\phi \to \neg(\forall x)\phi$. By the Deduction theorem, it suffices to show that $(\forall x)\neg\phi \vdash \neg(\forall x)\phi$. We show that $(T \cup \{(\forall x)\neg\phi\}) \cup \{\forall x\phi\}$ is inconsistent. Indeed, let t be any term. Then, $(T \cup \{(\forall x)\neg\phi\}) \cup \{\forall x\phi\} \vdash \phi[t/x]$. Thus, by (Q3) we have that $(T \cup \{(\forall x)\neg\phi\}) \cup \{\forall x\phi\} \vdash (\exists x)\phi$ and by (Q4) and (MP) we have that $(T \cup \{(\forall x)\neg\phi\}) \cup \{\forall x\phi\} \vdash \neg\forall x\neg\phi$, which is a sentence, and thus T is inconsistent.

4.2. Something new.

Now for new-er and exciting-er things.

Lemma 4.2.1 (Simulation of constants by variables)**.** *Let ψ be an $\mathcal{L}$-formula, T an $\mathcal{L}$-theory and C a set of constant symbols such that $C \cap \mathsf{Const}(\mathcal{L}) = \emptyset$. For any variable $x \in \mathsf{Var}$ and any constant symbol $\underline{c} \in C$, the following are equivalent:*

(1) $T \vdash_{\mathcal{L}} (\forall x)\psi$.

(2) $T \vdash_{\mathcal{L}\cup\{\underline{c}\}} \psi[\underline{c}/x]$.

(3) $T \vdash_{\mathcal{L}\cup\{\underline{c}\}} \psi$

Proof. To make our lives easier, we may as well assume that $T = \emptyset$ [Why "may we as well"?]

- (1) $\implies$ (3): Immediate by (Q2).

[5] This proof is too wordy. Can you come up with a better one?

- (3) $\implies$ (2) If $T \vdash_{\mathcal{L}\cup\{\underline{c}\}} \psi$ then $T \vdash_{\mathcal{L}\cup\{\underline{c}\}} (\forall x)\psi$ by (Gen). By (Q2) we have that $T \vdash_{\mathcal{L}\cup\{\underline{c}\}} ((\forall x)\psi)) \rightarrow \psi[\underline{c}/x]$ and thus, we are done by (MP).
- (2) $\implies$ (1): This is the heart of the proof. We assume that $T \vdash_{\mathcal{L}\cup\{\underline{c}\}} \psi[\underline{c}/x]$, i.e. that there is a sequence $\delta_1, \ldots, \delta_n$ of $\mathcal{L}\cup\{\underline{c}\}$-formulas which is an $\mathcal{L}\cup\{\underline{c}\}$-derivation of $\psi[\underline{c}/x]$. We want to transform this into an $\mathcal{L}$-derivation of ψ. We do this, of course, by induction on n. For an $\mathcal{L} \cup \{\underline{c}\}$-formula χ and a variable $y \notin \mathsf{Var}(\chi)$ we write $\tilde{\chi}$ for the formula $\chi[y/\underline{c}]$.[6] Here is what we have to show about this "tilde" operation, for any $\mathcal{L} \cup \{\underline{c}\}$-formulas ϕ, χ and ψ:

  (1) $\tilde{\phi}$ is an $\mathcal{L}$-formula.

  (2) $\tilde{\phi}[\underline{c}/y]$ is equal to ϕ.

  (3) If ϕ is an instance of (A1)-(A3), (E1)-(E5), (Q1)-(Q4) then so is $\tilde{\phi}$.

  (4) If ϕ is obtained using (MP) from ψ and χ then $\tilde{\phi}$ is obtained using (MP) from $\tilde{\psi}$ and $\tilde{\chi}$.

  (5) If ϕ is obtained using (Gen) from ψ then $\tilde{\phi}$ is obtained using (Gen) from $\tilde{\psi}$.

  Having proved (1)-(5) above, let $y \notin \bigcup_{i\leq n} \mathsf{Var}(\delta_i)$. Then it should be clear that for any $\mathcal{L} \cup \{\underline{c}\}$-formula ϕ, if $\delta_1, \ldots, \delta_n$ is an $\mathcal{L} \cup \{\underline{c}\}$-derivation of ϕ, then $\tilde{\delta}_1, \ldots, \tilde{\delta}_n$ is an $\mathcal{L}$-derivation of $\widetilde{\phi[c/x]}$, which is enough to conclude this implication.

  So, once we prove (1)-(5), above, we will be done. The hardest proof is the case of (Q3) in (3). Recall that an instance of (Q3) is a formula χ of the form:
$$\phi[t/x] \rightarrow (\exists x)\phi.$$
  By definition, we have that:
$$\tilde{\chi} = (\widetilde{\phi[t/x]}) \rightarrow (\exists x)\tilde{\phi}.$$
  and it is enough to show that this is an instance of (Q3). To do this, we need to show that:
$$(\widetilde{\phi[t/x]}) = \tilde{\phi}[\tilde{t}/x],$$
  where $\tilde{t}$ is naturally the term $t[y/\underline{c}]$. This is a straightforward induction (just like most things in life).

[6] This is the formula obtained by substituting every occurrence of $\underline{c}$ in χ by y. One defines this by induction.

$\square$

Exercise 4.2.2. Prove (1),(2), the remaining cases of (3), (4) and (5), from above.

The next corollary will seem rather innocuous, but it's extremely crucial. Essentially, it tells us that if a theory cannot prove a contradiction, then it cannot prove a contradiction even after we add new constants to it. The point here is that when we add new constants to a theory, the things we can prove amount to universal sentences in the original theory (by the previous lemma).

Corollary 4.2.3. *Let $\mathcal{L}$ be a first-order language, T an $\mathcal{L}$-theory and C a set of constant symbols such that $C \cap \mathsf{Const}(\mathcal{L}) = \emptyset$. Then, the following are equivalent:*

(1) T is consistent as an $\mathcal{L}$-theory (i.e. there is no $\mathcal{L}$-sentence ψ such that $T \vdash_{\mathcal{L}} \psi$ and $T \vdash_{\mathcal{L}} \neg\psi$).

(2) T is consistent as an $\mathcal{L} \cup C$-theory (i.e. there is no $\mathcal{L} \cup C$-sentence ψ' such that $T \vdash_{\mathcal{L}\cup C} \psi$ and $T \vdash_{\mathcal{L}\cup C} \neg\psi$).

Proof. (1) $\implies$ (2) is trivial, thus we need only prove (2) $\implies$ (1). It suffices to do this when $C = \{\underline{c}\}$ and then argue by induction on $|C|$, noting that any deduction only uses a finite number of constants. Suppose that T is inconsistent as an $\mathcal{L} \cup C$-theory and let ψ' be an $\mathcal{L} \cup \{\underline{c}\}$-sentence witnessing this. By the previous lemma we have that $T \vdash_{\mathcal{L}} (\forall x)\psi[x/c]$ and $T \vdash_{\mathcal{L}} (\forall x)\neg\psi[x/c]$. By Example 4.1.9 and (MP) we have that: $T \vdash_{\mathcal{L}} \neg(\forall x)\psi[x/c]$. It follows that T is inconsistent as an $\mathcal{L}$-theory. $\square$

5. Building models

Our next, and main for this chapter, goal is to prove the converse of the Soundness Theorem. This is important enough to have the name of an important logician:

Theorem 5.0.1 (Gödel's Completeness Theorem). *Let $\mathcal{L}$ be a fist order language and T a first-order theory. If $T \vDash \phi$ then $T \vdash_{\mathcal{L}} \phi$. In particular, (the universal closure of) every universally valid formula has a proof in $\vdash_{\mathcal{L}}$.*

Gödel proved (a version of) this theorem in 1929, in his PhD thesis. The proof I will present here is a modern adaptation of the proof given 20 years later by Henkin, also in his PhD thesis. The core idea of the proof will be somewhat similar to the proof of the completeness theorem for propositional logic. Indeed, we will prove a version of the Adequacy Lemma:

THEOREM 5.0.2 (The Adequacy Lemma, Redux). *If T is consistent, then T is satisfiable.*

and again, this is a funky statement: “T is consistent” (a SYMBOLIC notion) implies “T has a model” (a SEMANTIC notion). How do we do this, well, as the title suggests, we will have to build a model. Just so you know what's lying ahead, we will do this in two steps.

- Step 1: Any complete theory in a language $\mathcal{L}$ with enough constant symbols to witness every existential formula is satisfiable.
- Step 2. Any consistent theory can be extended to complete theory with enough constant symbols to witness every existential formula.

These steps, and especially the second step, may seem mysterious at the moment, and indeed, Henkin's construction is kind of magical, but once we've seen all the gore they will feel rather easy, hopefully.

5.1. Henkin Models. Let's turn the condition in Step 1 into a definition:

Definition 5.1.1. Let T be an $\mathcal{L}$-theory. We say that T has *Henkin witnesses* if for every $\mathcal{L}$-formula $\phi(x)$ there is a constant symbol $\underline{c} \in \mathsf{Const}(\mathcal{L})$ such that the formula $(\exists x)\phi(x) \to \phi(\underline{c})$ belongs to T.

Remark 5.1.2. If $\mathcal{L}$ is a language in which a theory T has Henkin witnesses, then $\mathsf{Const}(\mathcal{L}) \neq \emptyset$. Indeed, for the formula $x \doteq x$ (or any formula really), there is some $\underline{c} \in \mathsf{Const}(\mathcal{L})$ such that $(\exists x)(x \doteq x) \to (\underline{c} \doteq \underline{c}) \in T$.

Proposition 5.1.3. *If T is a complete $\mathcal{L}$-theory that has Henkin witnesses then T has a model.*

We'll have to pull a model out of a hat for this one. Well all we have is T and the language $\mathcal{L}$ and all those pesky Henkin witnesses, so that's what we ought to use.

PROOF. Let $\mathcal{T}_0$ be the set of all closed $\mathcal{L}$-terms.[7] By our previous remark, $\mathcal{T}_0 \neq \emptyset$. We will build an $\mathcal{L}$-structure $\mathcal{M}$ whose universe **is** (pretty much) the set $\mathcal{T}_0$, and we'll do this in the only way that we can. We'll let the interpretations of things be the things. This will be a bit of a braintwister, so bare with me.

[7]Recall an $\mathcal{L}$-term t is closed if $\mathsf{Var}(t) = \emptyset$.

We have to be a little bit careful now. We need to observe that in $\mathcal{T}_0$ there may be distinct closed terms that T proves are equal, which is rather annoying, because we want our interpretation of equality ($\doteq$) to be *true* equality.

Define a binary relation $\sim$ on $\mathcal{T}_0$ by setting:

$$t \sim t' \text{ if and only if } T \vdash t \doteq t'.$$

Claim 1. $\sim$ is an equivalence relation.

Proof of Claim 1. This follows quite literally from the equality axioms (E1)–(E3) and the quantifier axiom (Q2). ◀

Let us now set $\mathcal{T} = \mathcal{T}_0/_\sim$. We will define an $\mathcal{L}$-structure $\mathcal{M}$ on $\mathcal{T}$ in the more-or-less only way we can:

(1) For every $\underline{c} \in \mathsf{Const}(\mathcal{L})$ we set $c^{\mathcal{M}}$ to be $[\underline{c}]_\sim$ (recall that $\underline{c}$ is a closed term, so its equivalence class modulo $\sim$ is an element of the domain).

(2) If $\underline{R} \in \mathsf{Rel}(\mathcal{L})$ is an n-ary relation symbol, then, for all $[t_1]_\sim, \ldots, [t_n]_\sim \in \mathcal{T}$ we set:

$$R^{\mathcal{M}}([t_1]_\sim, \ldots, [t_n]_\sim) \text{ if and only if } T \vdash \underline{R}(t_1, \ldots, t_n).$$

(3) If $\underline{f} \in \mathsf{Fun}(\mathcal{L})$ is an n-ary function symbol, then, for all $[t_1]_\sim, \ldots, [t_n]_\sim \in \mathcal{T}$ we set:

$$f^{\mathcal{M}}([t_1]_\sim, \ldots, [t_n]_\sim) = [\underline{f}(t_1, \ldots, t_n)]_\sim,$$

where, again, the latter is indeed a closed term.

If you're trusting people you may have read the definition above and been perfectly happy with it, but unfortunately there is an issue! We have to check that things are *well-defined*. What does that mean? Well in (2) and (3), we made a choice on a bunch of equivalence classes which seemingly depended on their representatives! We need to prove that:

- If $[t_1]_\sim = [t_1']_\sim, \ldots, [t_n]_\sim = [t_n']_\sim$ then:

$$T \vdash \underline{R}(t_1, \ldots, t_n) \text{ iff } T \vdash \underline{R}(t_1', \ldots, t_n'),$$

for each n-ary relation symbol $\underline{R}$ and similarly:

- If $[t_1]_\sim = [t_1']_\sim, \ldots, [t_n]_\sim = [t_n']_\sim$ then:

$$T \vdash \underline{f}(t_1, \ldots, t_n) \doteq \underline{f}(t_1', \ldots, t_n'),$$

for each n-ary function symbol $\underline{f}$.

CLAIM 2. The interpretations in $\mathcal{M}$ are well defined.

PROOF OF CLAIM 2. This follows pretty much immediately by the equality axioms (E4), (E5) and (Q2). ◀

Now that we have our structure $\mathcal{M}$, we wish to show that $\mathcal{M} \models T$. First of all, observe that if ϕ is any atomic $\mathcal{L}$-sentence (i.e. an atomic formula with no free variables),[8] then we have that:

$$\mathcal{M} \models \phi \text{ if and only if } T \vdash \phi.$$

To see this, formally should consider cases:

- Case 1. ϕ is of the form $t_1 \doteq t_2$ for some $\mathcal{L}$-terms t_1 and t_2, where each of t_1 and t_2 are closed terms.
- Case 2. ϕ is of the form $\underline{R}(t_1, \ldots, t_n)$, where $\underline{R} \in \mathsf{Rel}(\mathcal{L})$ is an n-ary relation symbol and $t_1, \ldots, t_n$ are closed terms.

But in either case, the claim is immediate, by how we built $\mathcal{M}$ (it was for Case 1 here that we had to go through all this quotienting business).

Now, we will show that actually, for every $\mathcal{L}$-sentence ϕ we have that $\mathcal{M} \models \phi$ if and only if $T \vdash \phi$. This will suffice to show that $\mathcal{M} \models T$.

The argument is, of course, by induction on the structure of ϕ. We here are assuming that all formulas are built using only $\rightarrow, \neg, \forall$ and $\exists$. So we have to check the following cases:

- Case 1. ϕ is of the form $\neg\psi$. We know, by induction that $\mathcal{M} \models \psi$ if and only if $T \vdash \psi$. Since T is complete we have that $T \vdash \neg\psi$ if and only if it is not the case that $T \vdash \psi$, but by inductive hypothesis this is the case if and only if it is not the case that $\mathcal{M} \models \psi$. All in all:
$$\begin{aligned} T \vdash \neg\psi &\text{ iff } T \nvdash \psi && (T \text{ is complete}) \\ &\text{ iff } \mathcal{M} \not\models \psi && (\text{IH}) \\ &\text{ iff } \mathcal{M} \models \neg\psi && (\text{Truth}) \end{aligned}$$
- Case 2. ϕ is of the form $\psi \rightarrow \chi$. The cases here are very similar to the proof of the completeness theorem for propositional logic (see next exercise).

[8] Question to the audience, how do atomic $\mathcal{L}$-sentences look like? Spoilers: See next case distinction.

- Case 3. ϕ is of the form $(\forall x)\psi$. First, we show that if $T \vdash (\forall x)\psi$ then $\mathcal{M} \vDash (\forall x)\psi$. This is where the construction of $\mathcal{M}$ comes in handy:

$$\begin{aligned} T \vdash (\forall x)\psi \text{ implies that for all closed } t \text{ we have } T \vdash \psi[t/x] \quad & \text{(Q2)} \\ \text{implies that for all } t \in \mathcal{M} \text{ we have } T \vdash \psi[t/x] \quad & \text{(construction of } \mathcal{M}) \\ \text{implies that for all } t \in \mathcal{M} \text{ we have } \mathcal{M} \vDash \psi[t/x] \quad & \text{(IH)} \\ \text{implies that } \mathcal{M} \vDash (\forall x)\psi \quad & \text{(Truth).} \end{aligned}$$

  Next, we show that if $T \nvdash (\forall x)\psi$ then $\mathcal{M} \nvDash (\forall x)\psi$. This is where the completeness of T comes in handy.

  - If $T \nvdash (\forall x)\psi$, then $T \nvdash (\forall x)\neg\neg\psi$. To see this, suppose toward contradiction that $T \vdash (\forall x)\neg\neg\psi$. Since $\neg\neg\psi \to \psi$ is an instance of a propositional tautology, by Lemma 3.1.2 we have that $T \vdash (\forall x)\neg\neg\psi \to (\forall x)\psi$. And thus, $T \vdash (\forall x)\psi$, contradicting our original assumption.
  - Since T is complete we have that $T \vdash \neg(\forall x)\neg\neg\psi$. By (Q4) we have that $T \vdash \exists x \neg\psi$.

  So far so good, but you must have been asking yourselves (I hope), why all this Henkin witness business. Well, if T has Henkin witnesses, then $(\exists x)\neg\psi(x) \to \neg\psi(\underline{c}) \in T$, for some constant symbol $\underline{c}$.[9] But then:

  - $T \vdash \neg\psi(\underline{c})$, by (MP).
  - So, by Case 1, we have that $\mathcal{M} \vDash \neg\psi(c)$, and hence (by SEMANTICS!) $\mathcal{M} \vDash \neg(\forall x)\psi$.

  This concludes this case.

- Case 4. ϕ is of the form $(\exists x)\psi$. A simpler version of the argument above. One could also conclude by combining Case 1 and Case 3 (see next exercise).

□

Exercise 5.1.4. Write out the details of the missing cases in the proof above.

[9] A good question at this point is why ψ has at most one free variable x. The answer is, of course, because $(\forall x)\psi$ was assumed to be a sentence.