Basic Logic 6

Frank Veltman
ILLC
University of Amsterdam

2009

3 The completeness theorem

Our goal in this section is to prove that the next statement is true:

Theorem 2 (Completeness Theorem) *Let φ be a formula and Δ a set of formulas of some language of first order predicate logic (with identity).*
If $\Delta \models \varphi$, then $\Delta \vdash \varphi$

Like in the case of propositional logic we will obtain this theorem as a corollary of the following *Henkin* proposition.

Proposition 5 *Let Δ be a consistent theory of some language $\mathcal{L}$ of first order predicate logic (with identity). There exists a model $\mathcal{M}$ such that $\mathcal{M} \models \varphi$ for all $\varphi \in \Delta$.*

Notice that the proposition is formulated for sets of *sentences*, whereas the completeness theorem is about *formulas* (possibly containing free variables). That's why we have to make small detour in the proof that shows that the Henkin proposition implies the completeness theorem.

Proof. Let $\Delta \cup \{\varphi\}$ be a set of formulas and assume that $\Delta \not\vdash \varphi$. Let $y_1, \ldots, y_n$ be the variables that occur freely somewhere in the formulas of $\Delta \cup \{\varphi\}$. Consider a sequence of individual constants $b_1, \ldots, b_n$ that occur nowhere in the formulas of $\Delta \cup \{\varphi\}$ and replace every variable y_i by b_i each time that y_i occurs freely in the formulas of $\Delta \cup \varphi$. This yields a new set of premises $\Delta' = \{\psi' \mid \psi' = \ldots [b_2/y_2][b_1/y_1]\psi$ for some $\psi \in \Delta\}$ and a new conclusion $\ldots [b_2/y_2][b_1/y_1]\varphi$. We can be certain that $\Delta' \not\vdash \varphi'$. (It is left as an exercise to check this).

This means that $\Delta' \cup \{\neg\varphi'\}$ is consistent. Given the Henkin proposition it follows that $\Delta' \cup \{\neg\varphi'\}$ has a model. So there exists some $\mathcal{M} = \langle D, I\rangle$ such that for every assignment v

$\mathcal{M} \models \ldots [b_2/y_2][b_1/y_1]\psi$ $[v]$ for every $\psi \in \delta$, whereas $\mathcal{M} \not\models \ldots [b_2/y_2][b_1/y_1]\varphi$ $[v]$.

Now consider some assignment assigning to each y_i the value $v'(y_i) = I(b_i)$. Repeated application of proposition 3 gives that

$\mathcal{M} \models \psi$ $[v']$ for every $\psi \in \Delta$, whereas $\mathcal{M} \not\models \varphi$ $[v']$. Therefore $\Delta \not\models \varphi$

Let us now concentrate on the Henkin proposition. Our task is this: given a consistent theory Δ, specify some domain D and interpretation I such that $\langle D, I \rangle \models \varphi$ for all $\varphi \in \Delta$. One way or the other we have to read off D and I from Δ, but how? In order not to complicate matters too much we will first look at languages without function symbols and identity.

Observation: If $Pa_1 \ldots a_n \in \Delta$ there must be objects $d_1, \ldots, d_n \in D$ such that $I(a_i) = d_i$ and such that $\langle d_1, \ldots, d_n \rangle \in I(P)$. *Idea:* Why not take for D the set of individual constants, and stipulate that the interpretation of any individual constant a is the constant itself: $I(a) = a$? For the predicates we could then in any case state that $\langle a_1, \ldots, a_n \rangle \in I(P)$ if $Pa_1 \ldots a_n \in \Delta$.

Problem: This gives only part of the model. We do not know yet what to do when $Pa_1 \ldots a_n \notin \Delta$. We cannot just state that in that case $\langle a_1, \ldots, a_n \rangle \notin I(P)$. After all it is very well possible that even though $Pa_1 \ldots a_n \notin \Delta$, $\mathcal{M} \models Pa_1 \ldots a_n$ for every model $\mathcal{M}$ of Δ.

Observation This problem does not occur if Δ happens to be a *maximal* consistent set of sentences. Because for such Δ it holds that for every atomic sentence if $Pa_1 \ldots a_n \notin \Delta$, make $\neg Pa_1 \ldots a_n \in \Delta$. Which means that every model $\mathcal{M}$ of Δ should render $Pa_1 \ldots a_n \in \Delta$ false.

Problem. But now suppose that Δ is maximal consistent and that for some one-place predicate all sentences of the form $\neg Pa \in \Delta$, whereas also $\exists x Px \in \Delta$. The model we are looking for should render this sentence true. But the model we described above does not do this.

Observation: This problem does not occur if Δ is not only maximal consistent but on top of that has the property that for every sentence of the form $\exists x\varphi$ the following holds: if $\exists x\varphi \in \Delta$, then $[a/x]\varphi \in \Delta$ for some individual constant a.

Let's call a set of sentences with this property a *Henkin theory.* Can we be certain that every maximal consistent Henkin theory has a model? Before we prove this, we will first state a lemma.

Lemma 1 *Let Δ be a maximal consistent Henkin theory. The following holds:*

(i) $\psi \in \Delta$ *iff* $\Delta \vdash \psi$

(ii) $\neg\psi \in \Delta$ *iff* $\psi \notin \Delta$

(iii) $\psi \wedge \chi \in \Delta$ *iff* $\psi \in \Delta$ *and* $\chi \in \Delta$

(iv) $\psi \vee \chi \in \Delta$ *iff* $\psi \in \Delta$ *or* $\chi \in \Delta$

(v) $\psi \to \chi \in \Delta$ *iff* $\psi \in \Delta$ *or* $\chi \in \Delta$

(vi) $\forall x\psi \in \Delta$ *iff* $[a/x]\psi \in \Delta$ *for all individual constants* a

(vii) $\exists x\psi \in \Delta$ *iff* $[a/x]\psi \in \Delta$ *for some individual constant* a*.*

Proof. We already proved (i) - (v). The proof of (vi) runs as follows.
$\Rightarrow$: Suppose that $\forall x\psi \in \Delta$. Then $\Delta \vdash [a/x]\psi$ for all a because of the fact that $E_\forall$ is a rule in our natural deduction system. Given (i) it follows that $[a/x]\psi \in \Delta$ for all individual constants a.
$\Leftarrow$: Suppose that $\forall x\psi \notin \Delta$. Then by (ii) $\neg\forall x\psi \in \Delta$.

From this it follows that $\Delta \vdash \exists x\neg\psi$ (exercise), which means that $\exists x\neg\psi \in \Delta$. Since Δ is a Henkin theory, this means that $[a/x]\neg\psi = \neg[a/x]\psi \in \Delta$ for some individual constant a. Which means that $[a/x]\psi \notin \Delta$.

The proof of (vii) is left to the reader.

Proposition 6 *Let* Δ *be a maximal consistent Henkin theory in a language* $\mathcal{L}$ *without function symbols and identity. There exists a model* $\mathcal{M} = \langle D, I\rangle$ *such that for all* $\varphi \in \Delta$ $\mathcal{M} \models \varphi$*.*

Proof. Let D be the set of individual constants of $\mathcal{L}$, and define I by
(i) $I(a) = a$ for all individual constats a; and
(ii) $I(P) = \{\langle a_1, \ldots, a_n\rangle \mid Pa_1 \ldots a_n \in \Delta\}$
We show by induction on the number of logical constatnts in φ that for every sentence φ

$\langle D, I\rangle \models \varphi$ iff $\varphi \in \Delta$.

The following cases are to be distinguished.

- $\varphi = Pa_1 \ldots a_n$. This is trivial because we defined our model such that $I(P) = \{\langle a_1, \ldots, a_n\rangle \mid Pa_1 \ldots a_n \in \Delta\}$

- $\varphi = \neg\psi$, $\varphi = \psi \wedge \chi$, $\varphi = \psi \vee \chi$, $\varphi = \psi \rightarrow \chi$

  For these cases, see the the proof of the completeness theorem of propositional logic.

- $\varphi = \forall x\psi$.

  $\langle D, I\rangle \not\models \forall x\psi$ iff there is some v such that $\langle D, I\rangle \not\models \psi$ $[v(x/a)]$ for some a.

  Now, proposition 3 teaches that

  $\langle D, I\rangle \not\models \psi$ $[v(x/a)]$ iff $\langle D, I\rangle \not\models [a/x]\psi$ $[v]$

  The formula $[a/x]\psi$ is a sentence, which means that

  $\langle D, I\rangle \not\models [a/x]\psi$ $[v]$ iff $\langle D, I\rangle \not\models [a/x]\psi$

  Now, with an appeal to the induction hypothesis we can state

  $\langle D, I\rangle \not\models [a/x]\psi$ iff $[a/x]\psi \notin \Delta$

The lemma above then gives

$[a/x]\psi \notin \Delta$ iff $\forall x\psi \notin \Delta$.

And so we have $\langle D, I\rangle \not\models \forall x\psi$ iff $\forall x\psi \notin \Delta$.

- $\varphi = \forall x\psi$. Similarly.

We know now that every maximal consistent Henkin theory has a model. We have to show that every consistent theory has a model. Wouldn't it be nice if we could prove that that every consistent theory is extendible to a maximal consistent Henkin theory.

At a first view this may seem impossible, if inly because it is very well possible that the language in which the theory is formulated is a language without any individual constants, and in such a case there is f course no way to find a witness $[a/x]\varphi$ for every sentence of the form $\exists x\varphi$. The next definition and proposition will make clear what our way out will be.

Definition 11 (Expansion)
Let $\mathcal{L}$ and $\mathcal{L}'$ be languages of first order predicate logic. $\mathcal{L}'$ is an expansion *of $\mathcal{L}$ iff all non-logical symbols of $\mathcal{L}$ are non-logical symbols of $\mathcal{L}'$. $\mathcal{L}'$ is a* simple expansion *of $\mathcal{L}$ iff the only difference between $\mathcal{L}'$ and $\mathcal{L}$ is that the first may have more individual constants than the latter.*

If $\mathcal{L}'$ is an expansion of $\mathcal{L}$, and $\mathcal{M}' = \langle D', I'\rangle$ is a model for $\mathcal{L}'$, then the reduct *of $\mathcal{M}'$ to $\mathcal{L}$ is the model $\mathcal{M} = \langle D, I\rangle$ for $\mathcal{L}$ defined by*

(i) $D' = D$;

(ii) for all individual a of $\mathcal{L}$, $I(a) = I'(a)$;

(iii) for all predicates P of $\mathcal{L}$, $I(P) = I'(P)$;

(iv) for all function symbols f of $\mathcal{L}$, $I(f) = I'(f)$.

In this case $\mathcal{M}'$ is also called an expansion *of $\mathcal{M}$.*

Proposition 7 *Let $\mathcal{L}'$ be an expansion of $\mathcal{L}$, $\mathcal{M}' = \langle D', I'\rangle$ is a model for $\mathcal{L}'$, and $\mathcal{M}$ the reduct $\mathcal{M}'$ to $\mathcal{L}$. Then for all formulas φ of $\mathcal{L}$*

$\mathcal{M} \models \varphi\ [v]$ *iff* $\mathcal{M}' \models \varphi\ [v]$

The proof is trivial.

Now, let Δ be a consistent theory in a language $\mathcal{L}$. Maybe we cannot extend Δ within $\mathcal{L}$ to a maximal consistent Henki theory, but why not take recourse to a simple expansion of $\mathcal{L}$ within which we can?

Lemma 2 *Now, let Δ be a consistent theory in a language $\mathcal{L}$. There exists a simple expansion $\mathcal{L}'$ of $\mathcal{L}$, and a theory Δ' in $\mathcal{L}'$ such that (i) $\Delta \subseteq \Delta'$, (ii) Δ' is consistent, and (iii) Δ' is a Henkin theory.*

Proof. We construct $\mathcal{L}'$ inductively, as follows. Let $\mathcal{L}_0 = \mathcal{L}$, and take $\mathcal{L}_{n+1}$ to be the simple expansion of $\mathcal{L}_n$ that one gets by adding one new individual constant, c_φ, to $\mathcal{L}_n$ for every sentence of the form $\exists x\varphi$ belonging to $\mathcal{L}_n$ but not to $\mathcal{L}_{n-1}$. Let $C_{\mathcal{L}_n}$ be the set of individual constants of $\mathcal{L}_n$. $\mathcal{L}'$ is the simple expansion of $\mathcal{L}$ that has as its individual constants all the elements of the set

$$\bigcup_{n\in \mathbb{N}} C_{\mathcal{L}_n}$$

Now let Δ_n be the set of all sentences ψ of $\mathcal{L}_n$ such that $\psi = \exists x\varphi \rightarrow\ [c_\varphi/x]\varphi$ for some sentence $\exists x\varphi$ belonging to $\mathcal{L}_n$ but not to $\mathcal{L}_{n-1}$. Define Δ' as the set of all sentences of $\mathcal{L}'$ that are derivable from

$$\Delta \cup \bigcup_{n\in \mathbb{N}} \Delta_n$$

Let's check (i)- (iii).

(i) Obviously $\Delta \subseteq \Delta'$.

(ii) Suppose Δ' is inconsistent. Then there is a finite subset Δ'' of Δ' such that $\Delta'' \vdash \bot$. Let k be the smallest number such that for some finite subset $\Gamma \subseteq \Delta$ and k sentences $\chi_1, \ldots, \chi_k \in \bigcup_{n\in \mathbb{N}} \Delta_n$

$$\Gamma, \chi_1, \ldots, \chi_k \vdash \bot \text{ (*)}$$

For every χ_i there is a smallest number m such that χ_i is a sentence of $\mathcal{L}_m$. Let n be the maximum of these minima. Without loss of generality we may assume that χ_k belongs to $\mathcal{L}_n$. Given (*) we have

$$\Gamma, \chi_1, \ldots, \chi_{k-1} \vdash \neg\chi_k$$

Given the choice of k, $\Gamma \cup \{\chi_1, \ldots, \chi_{k-1}\}$ is consistent. The sentence χ_k has the form $\exists x\varphi \rightarrow\ [c_\varphi/x]\varphi$, which means that

(a) $\Gamma, \chi_1, \ldots, \chi_{k-1} \vdash \exists x\varphi$, and

(b) $\Gamma, \chi_1, \ldots, \chi_{k-1} \vdash \neg[c_\varphi/x]\varphi$

By the choice of k and n we may assume that c_φ occurs nowhere in the sentences of $\Gamma \cup \{\chi_1, \ldots, \chi_{k-1}\}$. Given the proposition proved in Exercise 6 of section 1, this means that

$$\Gamma, \chi_1, \ldots, \chi_{k-1} \vdash \forall x\neg\varphi$$

But this contradicts (a) and the fact that $\Gamma \cup \{\chi_1, \ldots, \chi_{k-1}\}$ is consistent.

(iii) Δ' is a Henkin theory. Suppose $\exists x\varphi \in \Delta'$. There is a smallest n such that $\exists x\varphi$ is a sentence of $\mathcal{L}_n$. Then $\exists x\varphi \to [c_\varphi/x]\varphi \in \Delta_{n+1}$, which means that $\exists x\varphi \to [c_\varphi/x]\varphi \in \Delta'$. Given the definition of Δ' it follows that $[c_\varphi/x]\varphi \in \Delta'$.

Question. Explain why $\mathcal{L}'$ and $\cdot'$ cannot be difined in one step. Why do we need the inductive construction.

We do not have to prove anymore that every consistent set of sentences can be extended to a maximal consistent set. The proof would be the same as in the case of propositional logic.

We are now ready to prove the Henkin proposition

Proof of Proposition 5. Consider a simple expansion $\mathcal{L}'$ of $\mathcal{L}$ and an extension Δ' of Δ defined as in the proof of Lemma 2. Δ' is consistent and Henkin. Consider a maximal consistent extension Δ'' of Δ'. Notice that Δ'' remains a Henkin theory. By Proposition 6 there exists a model $\mathcal{M}'$ of Δ''. The reductreductreductreduct $\mathcal{M}$ of $\mathcal{M}'$ to $\mathcal{L}$ is a model of Δ.

It remains to extend the above to languages with function symbols and identity. It is not difficult to see that to this end the only thing we have to do is to extend Proposition 6 to

Proposition 8 *Let Δ be a maximal consistent Henkin theory in a language $\mathcal{L}$ of first order predicate logic with function symbols and identity. There exists a model $\mathcal{M} = \langle D, I\rangle$ such that for all $\varphi \in \Delta$ $\mathcal{M} \models \varphi$.*

Proof: The main reason why we cannot proceed as in the proof of Proposition 6 is because it is very well possible that there are two distinct individual constants a and b such that the sentence $a = b \in \Delta$. In that case a and b should refer to the very same object in the model that we want to construct, but this will not be the case if we define the domain D and interpretation function I the way we did in the proof of proposition 6. A second difference with proposition 6 is that we now also have complex terms in our languages, and these have to refer to something, too.

We solve these problems as follows. Let Θ be the set of closed terms of $\mathcal{L}$[1] Consider the following relation on Θ

$$t \backsim t' \text{ iff } t = t' \in \Delta$$

It is not difficult to prove that $\backsim$ is an equivalence relation. (Use exercise 7).

Now, define for every $t \in \Theta$ the set $[\,t\,]^{\backsim}$ as follows

$$[\,t\,]^{\backsim} = \{t' \in \Theta \mid t' \backsim t\}$$

[1] A term t is *closed* iff no variables occur freely in t.

In other words, $[t]^{\backsim}$ is the equivalence class generated by t under the relation $\backsim$.

The domain D of the model we are constructing will have these equivalence classes as its elements.

$$D = \{[t]^{\backsim} \mid t \in \Theta\}$$

And we define the interpretation of the individual constants as follows

$$\text{For every individual constant } a, I(a) = [a]^{\backsim}$$

Notice that this solves the problem mentioned above. If $a = b \in \Delta$, then $a \backsim b$, and therefore $[a]^{\backsim} = [b]^{\backsim}$, which means that $I(a) = I(b)$. The two constants a and b refer to the same elemet of D.

It remains to define the interpretation of predicate and function symbols. Thanks to the elimination rule for identity, we have

If $t_1 \backsim t'_1, \ldots, t_n \backsim t'_n$, then $Pt_1 \ldots t_n \in \Delta$ iff $Pt'_1 \ldots t'_n \in \Delta$.

Therefore we can rest assured that the following definition is okay:

$$I(P) = \{\langle [t_1]^{\backsim}, \ldots, [t_n]^{\backsim} \rangle \mid Pt_1 \ldots t_n \in \Delta\}$$

For the same reason, we also have: If $t_1 \backsim t'_1, \ldots, t_n \backsim t'_n, t_{n+1} \backsim t'_{n+1}$, then $f(t_1, \ldots, t_n) = t_{n+1} \in \Delta$ iff $f(t'_1, \ldots, t'_n) = t'_{n+1} \in \Delta$. That's why the following definition is fine, too.

$$I(f) = \{\langle [t_1]^{\backsim}, \ldots, [t_n]^{\backsim}, [t_{n+1}]^{\backsim} \rangle \mid f(t_1, \ldots, t_n) = t_{n+1} \in \Delta\}$$

By this the model $\mathcal{M} = \langle D, I \rangle$ has been fixed. Before we can check whether all $\varphi \in \Delta$ hold in $\mathcal{M}$, we first want to make sure that the interpretation $[t]_{\mathcal{M}}$ of each term t is the equivalence class $[t]^{\backsim}$. For individual constants this is of course the case as we choose $[a]_{\mathcal{M}} = I(a) = [a]^{\backsim}$. Now assume that this also holds for $t_1, \ldots, t_n$ (induction hypothesis). Then for the more complex term $f(t_1, \ldots, t_n)$ the following holds $[f(t_1, \ldots, t_n)]_{\mathcal{M}} = I(f)([t_1]_{\mathcal{M}}, \ldots, [t_n]_{\mathcal{M}}) = I(f)([t_1]^{\backsim}, \ldots, [t_n]^{\backsim)})$. Thanks to the introduction rule for identity, $f(t_1, \ldots, t_n) = f(t_1, \ldots, t_n) \in \Delta$, which together with the definition of $I(f)$ above leads to the conclusion that $I(f)([t_1]^{\backsim}, \ldots, [t_n]^{\backsim)}) = [f(t_1, \ldots, t_n)]^{\backsim}$. By induction on the complexity of the sentence φ, we now prove that

$$\text{for all } \varphi, \langle D, I \rangle \models \varphi \text{ iff } \varphi \in \Delta$$

For the *base case*, there are two possibilties. If $\varphi = Pt_1 \ldots t_n$, we proceed as follows.

$\langle D, I \rangle \models Pt_1 \ldots t_n$ iff $\langle [t_1]_{\mathcal{M}}, \ldots, [t_n]_{\mathcal{M}} \rangle \in I(P)$ (by semantics)

$\langle [t_1]_{\mathcal{M}}, \ldots, [t_n]_{\mathcal{M}} \rangle \in I(P)$ iff $\langle [t_1]^{\backsim}, \ldots, [t_n]^{\backsim} \rangle \in I(P)$ (see above)

$\langle [t_1]^{\backsim}, \ldots, [t_n]^{\backsim} \rangle \in I(P)$ iff $Pt_1 \ldots t_n \in \Delta$ (definition I).

The other possibility in the base case is that φ is an identity statement $t = t'$.

$\langle D, I \rangle \models t = t'$ iff $[t]_{\mathcal{M}} = [t']_{\mathcal{M}}$ (by semantics)

$[t]_{\mathcal{M}} = [t']_{\mathcal{M}}$ iff $[t]^{\backsim} = [t']^{\backsim}$ (see above)
$[t]^{\backsim} = [t']^{\backsim}$ iff $t \backsim t'$ (property of equivalence classes I).
$t \backsim t'$ iff $t = t' \in \Delta$ (definition of $\backsim$).

As for the *inductive step* we only check the case that $\varphi = \exists x\psi$.
Assume $\langle D, I\rangle \models \exists x\psi$. Then, by semantics,
$\langle D, I\rangle \models \psi\ [v(x/[t]^{\backsim})]$ for some v and t.
Given our observations above this means that
$\langle D, I\rangle \models \psi\ [v(x/[t]_{\mathcal{M}})]$.
By proposition 4 we can derive from this that
$\langle D, I\rangle \models [t/x]\psi\ [v]$.
By the induction hypothesis this implies that
$[t/x]\psi \in \Delta$.
And then, thanks to the rule $I_{\exists}$, it follows that $\exists x\psi \in \Delta$.
Conversely, assume that $\exists x\psi \in \Delta$. Since Δ is a Henkintheory, it follows that $[a/x]\psi \in \Delta$ for some constant a. By the induction hypothesis $\langle D, I\rangle \models [a/x]\psi$. This being so, $\langle D, I\rangle \models \psi$ [v(x/I(a))], and so $\langle D, I\rangle \models \exists x\psi$.

4 Beginning Model Theory

The completeness theorem has two important model theoretic corollaries.

Theorem 3 *(Compactness)*
Let Δ be some theory stated in a language of first order logic. Δ has a model iff every finite subset of Δ has a model.

Proof. Suppose Δ has no model. Then $\Delta \models \bot$. By the completeness theorem this means that $\Delta \vdash \bot$. In any derivation of Δ from $\bot$ only finitely many premisies can be used. Therefore there is some finite subset Δ' of Δ such that $\Delta' \vdash \bot$. By applying the correctness theorem, we get $\Delta' \models \bot$, which means that there is some finite subset of Δ that has no model.

Theorem 4 *(Skolem)*
Every consistent theory Δ has a finite or countable model.

Proof. In the previous chapter we gave a recipe how to build for any consistent theory a model. Given the way it is constructed the resulting model will always have at most countably many elements. (The number of elements in the domain is smaller than or equal to the number of terms of the language. The language will have at most countably many terms, because the number of individual constants will be countable, even though in expanding the language we may have added countably constants a countably number of times).

To those of you who are acquainted with for example Zermelo-Fraenkel set theory (ZF), Skolem's theorem may come a surprise. ZF is an axiomatization of set theory stated in a first -order language. In ZF one can prove that there are uncountably many sets. One would think that this implies that the domain of any model of ZF will have uncountably many elements. But no, according to Skolem's theorem some models of ZF will have countably many elements. This consideration is known as *Skolem's Paradox.* One might be tempted to conclude that ZF, and any other first order theory in which one can prove that there are uncountably many objects is bound to be inconsistent. Fortunately the paradox is only *apparant.* The only conclusion one can draw from it is that no first order theory can express the notion 'uncountable'. What is called "uncountable" in the *object* language, can from the outside be so interpreted that something countable — "countable" in the *metalanguage* – stands for it.

The languages of first order logic lack expressive power. This appears from the above, and it also appears from the next application of the Compactness theorem.

Proposition 9 *There is no first order theory Δ for which the following holds:*
$\mathcal{M}$ is a model of Δ iff the domain of $\mathcal{M}$ is finite.

Of course there are theories that enforce for any particular natural number n, that all models of the theory will have at least (and/or at most) n elements, but there is no way to say that any model with just an unspecified 'finite number' of elements will do.

Proof. Assume there is such a theory Δ and consider the theory Δ' consisting of the following infinite sequence $\varphi_2, \varphi_3, \ldots$ of sentences:

$\varphi_2 = \exists x \exists y (x \neq y)$
$\varphi_3 = \exists x \exists y \exists u (x \neq y \wedge x \neq u \wedge y \neq u)$
$\varphi_4 = \exists x \exists y \exists u \exists v (x \neq y \wedge x \neq u \wedge x \neq v \wedge y \neq u \wedge y \neq v \wedge u \neq v)$
...
...

Notice that φ_n states that there are at least n objects. Furthermore we have

(i) every finite subset of $\Delta \cup \Delta'$ has a model. Proof: consider any finite subset $\Delta'' \subseteq \Delta \cup \Delta'$. There are only finitely many sentences of Δ' in Δ''. Let i be the largest number such that $\varphi_i \in \Delta''$. There a exists a model of Δ with at least i elements. That model is also a model of Δ''.

(ii) From (i) it follows by the compactness theorem that $\Delta \cup \Delta'$ has a model.

(iii) But that's impossible. Because this model should on the one hand, as a model of Δ, have a finite domain, and on the other hand, as a model of Δ', have an infinite domain (Why?). Contradiction.

Definition 12 *Let $\mathcal{L}$ be a language of first order predicate logic, and $\mathcal{M} = \langle D.I \rangle$ and $\mathcal{M}' = \langle D', I' \rangle$ be two models for $\mathcal{L}$.*

(i) $\mathcal{M}$ *is* elementary equivalent *to* $\mathcal{M}'$, *iff for every sentence* φ *of* $\mathcal{L}$, $\mathcal{M} \models \varphi$ *iff* $\mathcal{M}' \models \varphi$.

(ii) $\mathcal{M}$ *is* isomorphic *to* $\mathcal{M}'$ *iff there exists a bijection* h *form* D *onto* D' *such that*

  (a) for all predicates P, $\langle d_1, \ldots, d_n \rangle \in I(P)$ *iff* $\langle h(d_1), \ldots, h(d_n) \rangle \in I'(P)$

  (b) for all function symbols f, $I(f)(d_1, \ldots, d_n) = d_{n+1}$ *iff* $I'(f)(h(d)_1, \ldots, h(d_n)) = h(d_{n+1})$.

  (c) for all individual constants a, $I'(a) = h(I(a))$.

  h *is called an* isomorphism *between* $\mathcal{M}$ *and* $\mathcal{M}'$.

When $\mathcal{M}$ is isomorphic to $\mathcal{M}'$ we will write $\mathcal{M} \cong \mathcal{M}'$. The intuitive idea behind the definition above is this. Sometimes two models a so similar that the one can be considered a *copy* of the other, which is obtained by replacing each object of the domain of one of the models by another object, thereby keeping the the same structure. Therefore it comes as no surprise that the following proposition holds.

Proposition 10 *Let* h *be an isomorphism between* $\mathcal{M}$ *and* $\mathcal{M}'$. *Let* v *be an assignment for the variables into* D, *and choose the assignment* v' *for the variables into* D' *such that for all vaiables* x, $v'(x) = h(v(x))$. *Then*

(i) for every term t, $h([t]^v_{\mathcal{M}}) = [t]^{v'}_{\mathcal{M}'}$

(ii) for every formula φ, $\mathcal{M} \models \varphi\ [v]$ *iff* $\mathcal{M}' \models \varphi\ [v']$

When $\mathcal{M}$ is elementary equivalent to $\mathcal{M}'$, we will write $\mathcal{M} \equiv \mathcal{M}'$. In elementary equivalent models the same sentences are true, which means that the models are highly similar, or at least that the differences cannot be expressed in the object language $\mathcal{L}$.

Corollary 1 *If* $\mathcal{M} \cong \mathcal{M}'$, *then* $\mathcal{M} \equiv \mathcal{M}'$.

It is hardly an exaggeration to say that model theory owes its existence to the fact that the converse of the above statement does not hold. Elementary equivalent models need not be isomorphic.

Recall the notion of a (syntactically) *complete* theory, and note that the following two statements are equivalent:

- The theory Δ is complete.

- All models of Δ are elementary equivalent.

Recall the theory of dense linear order without endpoints. We will prove that this theory is complete. The following theorem from set theory is part of the proof.

Theorem 5 (Cantor) *If $\langle T, R\rangle$ and $\langle T', R'\rangle$ are structures with the following properties:*

(i) $\mathcal{T}$ en $\mathcal{T}'$ are both countably infinite.

(ii) $\mathcal{R}$ en $\mathcal{R}'$ are both dense linear orderings without endpoints.

then there exists a bijection f from $\mathcal{T}$ onto$\mathcal{T}'$ such that for all $t, t' \in \mathcal{T}$, $\langle t, t'\rangle \in \mathcal{R}$ iff $\langle f(t), f(t')\rangle \in \mathcal{R}'$.

Proof: $\mathcal{T}$ is countable; let $t_0, t_1, t_2, \ldots$ be an enumeration of the elements of $\mathcal{T}$. Similarly, let $t'_0, t'_1, t'_2, \ldots$ be an enumeration of the elements of $\mathcal{T}'$.

Using the so-called zig-zag method, we will now first define a sequence of functions $f_0, f_1, f_2, f_3, \ldots$ with the following properties:

(i) $f_n \subseteq f_{n+1}$ for every $n \in \mathbb{N}$.

(ii) For every $n \in \mathbb{N}$ it holds that f_n is an isomorphism from $n+1$ elements of $\mathcal{T}$ onto $n+1$ elements of $\mathcal{T}'$.

(iii) $f = \bigcup_{n \in\ \mathbb{N}} f_n$ is the function we are looking for.

Here we go. f_0 is determined as follows: the domain of f_0 is $\{t_0\}$ and $f_0(t_0) = \{t'_0\}$. Assume that $f_0, f_1, f_2, f_3, \ldots, f_k$ are all defined and that the properties i) and (ii) hold for all $n \leq k$. Then f_{k+1} is is determined as follows:

1. *If* $k+1$ *is an even number*
   Let t_i be the first element in the enumeration $t_0, t_1, \ldots$ of $\mathcal{T}$ such that $t_i \notin dom(f_k)$. The following possibilities occur:

   (a) $\langle t_i, t_n\rangle \in \mathcal{R}$ for every t_n in the domain $dom(f_k)$ of f_k.
   Consider the structure $\langle dom(f_k), \mathcal{R}\restriction dom(f_k)\rangle$. This structure is a finite linear ordering; therefore it will have precisely one minimal element — say t_m. Assume that $f_k(t_m) = t'_j$. Notice that t'_j is the minimal element of the structure $\langle ran(f_k), \mathcal{R}'\restriction ran(f_k)\rangle$, where $ran(f_k)$ is the range of $f(k)$. Let t'_l be the first element in the enumeration $t'_0, t'_1, \ldots$ such that $\langle t'_l, t'_j\rangle \in \mathcal{R}'$. There must be such an element because $\mathcal{R}'$ has no endpoints. Choose $f_{k+1} = f_k \cup \{\langle t_l, t'_i\rangle\}$. Notice that the properties (i) and (ii) hold for every $n \leq k+1$.

(b) $\langle t_n, t_i \rangle \in \mathcal{R}$ for every $t_n \in dom(f_k)$.
The proof is analogous. Consider the maximum t_m of the structure $\langle dom(f_k), \mathcal{R} \restriction dom(f_k) \rangle$. Assume $f_k(t_m) = t'_j$. Let t'_l be the first element in the enumeration of $\mathcal{T}'$ such that $\langle t'_j, t'_l \rangle \in \mathcal{R}$. Choose $f_{k+1} = f_k \cup \{\langle t_i, t'_l \rangle\}$.

(c) $\langle t_n, t_i \rangle \in \mathcal{R}$ for some $t_n \in dom(f_k)$ and $\langle t_i, t_n \rangle \in \mathcal{R}$ for some $t_n \in dom(f_k)$.
Let A be the set $\{t_n \in dom(f_k) \mid \langle t_n, t_i \rangle \in \mathcal{R}\}$ and B the set $\{t_n \in dom(f_k) \mid \langle t_i, t_n \rangle \in \mathcal{R}\}$. The structure $\langle A, \mathcal{R} \restriction A \rangle$ has a maximal element, say t_m. The structure $\langle B, \mathcal{R} \restriction B \rangle$ has a minimal element, say t_l. Assume that $f_k(t_m) = t'_j$ and $f_k(t_l) = t'_h$. Since $\langle t_m, t_i \rangle \in \mathcal{R}$ and $\langle t_i, t_l \rangle \in \mathcal{R}$, also $\langle t_m, t_l \rangle \in \mathcal{R}$, . Moreover, there is no $t_n \in dom(f_k)$ such that $\langle t_m, t_n \rangle \in \mathcal{R}$ and $\langle t_n, t_l \rangle \in \mathcal{R}$. This implies that $\langle t'_j, t'_h \rangle \in \mathcal{R}'$. and that there is no $t'_n \in ran(f_k)$ such that $\langle t'_j, t'_n \rangle \in \mathcal{R}'$ and $\langle t'_n, t'_h \rangle \in \mathcal{R}'$.
Now, let t'_l be the first element in the enumeration of $\mathcal{T}'$ such that $\langle t'_j, t'_l \rangle \in \mathcal{R}'$ and $\langle t'_l, t'_h \rangle \in \mathcal{R}'$. There must be such an element because $\mathcal{R}'$ is dense. Choose $f_{k+1} = f_k \cup \{\langle t_i, t'_l \rangle\}$. Notice that the properties (i) and (ii) now hold for every $n \leq k+1$.

(d) There are no other possibilties

2. *If* $k+1$ *is an odd number*
Let t'_i be the first element in the enumeration of $\mathcal{T}'$ such that $t'_i \notin ran(f_k)$. The following possibilities obtain:

(a) $\langle t'_i, t'_n \rangle \in \mathcal{R}'$ for every $t'_n \in ran(f_k)$.
Consider the minimal element — say t'_m — of the structure $\langle ran(f_k), \mathcal{R}' \restriction ran(f_k) \rangle$. Suppose that $\breve{f}_k(t'_m) = t_j$. Let t_l be the first element in the enumeration of $\mathcal{T}$ such that $\langle t_i, t_j \rangle \in \mathcal{R}$. Choose $f_{k+1} = f_k \cup \{\langle t_l, t'_i \rangle\}$.

(b) $\langle t'_n, t'_i \rangle \in \mathcal{R}'$ for every $t'_n \in ran(f_k)$.
Analogously.

(c) $\langle t'_n, t'_i \rangle \in \mathcal{R}'$ for some $t'_n \in ran(f_k)$ and $\langle t'_i, t'_n \rangle \in \mathcal{R}'$ for some $t'_n \in ran(f_k)$. Left to the reader.

(d) There are no other possibilities.

Now, consider $f = \bigcup_{n \in \mathbb{N}} f_n$. We will show that $dom(f) = \mathcal{T}$ and $ran(f) = \mathcal{T}'$. Suppose for contradiction that $dom(f) \neq \mathcal{T}$. Let t_i be the first element in the enumeration of $\mathcal{T}$ such that $t_i \notin dom(f)$. And let k be the smallest number such that $t_j \in dom(f_k)$ for all $j < i$. Then $t_i \in dom(f_{k+2})$. That $ran(f) = \mathcal{T}'$ is shown in the same way. It is also clear that f is an injection (Proof left to the reader).

Finally, for all $t_i, t_j \in \mathcal{T}$ it holds that $\langle t_i, t_j \rangle \in \mathcal{R}$ iff $\langle f(t_i), f(t_j) \rangle \in \mathcal{R}'$. Suppose that $\langle t_i, t_j \rangle \in \mathcal{R}$; let k be the smallest number such that $t_i, t_j \in dom(f_k)$.

Clearly, $\langle f_k(t_i), f_k(t_j)\rangle \in \mathcal{R}'$; since $f_k \subseteq f$, also $\langle f(t_i), f(t_j)\rangle \in \mathcal{R}'$. The converse can be proved in the same manner.

Theorem 6 Θ *is complete.*

Proof: Assume Θ is not complete. Then there is a sentence φ such that neither φ nor $\neg\varphi$ is derivable from Θ. In that case both the theory $\Theta \cup \{\neg\varphi\}$ and the theory $\Theta \cup \{\varphi\}$ have a model.

The theory Θ has only infinite models. Therefore the theory $\Theta \cup \{\varphi\}$ has only infinite models. Given Skolem's theorem there exists a countable model $\langle \mathcal{T}, \mathcal{R}\rangle$ of $\Theta \cup \{\varphi\}$. For the same reasons there exists a countable model $\langle \mathcal{T}', \mathcal{R}'\rangle$ of $\Theta \cup \{\neg\varphi\}$.
$\langle \mathcal{T}, \mathcal{R}\rangle$ and $\langle \mathcal{T}', \mathcal{R}'\rangle$ are both countable structures with a dense linear ordering without endpoints. From the previous theorem it follows that $\langle \mathcal{T}, \mathcal{R}\rangle$ and $\langle \mathcal{T}', \mathcal{R}'\rangle$ are isomorphic. This being so, it holds for all sentences ψ of our first order language that if ψ is true in $\langle \mathcal{T}, \mathcal{R}\rangle$ desda ψ is also true in $\langle \mathcal{T}', \mathcal{R}'\rangle$, contradicting our earlier statement that φ is true in $\langle \mathcal{T}, \mathcal{R}\rangle$ and $\neg\varphi$ is true in $\langle \mathcal{T}', \mathcal{R}'\rangle$.

Now consider the following two models of the theory of dense linear orderings without endpoints:

1. The model $\langle D, I\rangle$ with $D = \mathbb{Q}$, the set of rational numbers, and $I(R)$ the 'smaller than' relation between rational numbers.

2. The model $\langle D', I'\rangle$ with $D' = \mathbb{R}$, the set of real numbers, and $I'(R)$ the 'smaller than' relation between real numbers.

Theorem 6 says that $\langle D, I\rangle \equiv \langle D', I'\rangle$. On the other hand it is not the case that $\langle D, I\rangle \cong \langle D', I'\rangle$, because there is no bijenction between the (countable) set of rational numbers $\mathbb{Q}$ and (uncountable) set of reals $\mathbb{R}$.

The above shows that two models can be elementary equivalent without being isomorphic. But the models figuring in the example have domains with different cardinality. The next example shows that there are also elemenrary equivalent models of the same cardinality that are not isomorphic.

Consider the so-called *standard model* of arithmetic, which is the model $\mathcal{N} = \langle D, I\rangle$ for the language of arithmetic with $D = \mathbb{N}$ and with the successor function as interpretation $I(S)$ of the S-operator, addition as interpretation $I(+)$ of the operator $+$, multiplication for $I(\times)$, and the number 0 as $I(0)$. Let $\Delta = \{\varphi \mid \mathcal{N} \models \varphi\}$. Notice that Δ is complete — a maximal consistent extension of Peano Arithmetic to be precise. Therefore all models of Δ are elementary equivalent. We will now construct a model for Δ that is not isomorphic with $\mathcal{N}$. Add one new individual constant c to the language of arithmetic, and extend Δ with the next infinite sequence of sentences:

$$c \neq 0, c \neq S(0), c \neq S(S(0)), c \neq S(S(S(0))), \ldots$$

Every finite subset Δ'' of Δ' has a model. After all, given such a finite subset Δ'' there is bound to be a largest number k such that

$$c \neq \underbrace{S(S(\ldots}_{k\times}(0).) \in \Delta''$$

By interpreting c as the number $k+1$ (or larger) we expand the standard model $\mathcal{N}$ to a model that renders all sentences in Γ true.

By compactness, Δ' has a model, and by Skolem's theorem it must have a countable model. This model cannot be isomorphic with $\mathcal{N}$ because $I(c)$ must be different from 0 and all successors of 0, and in $\mathcal{N}$ there are no such objects.

5 The theory of definitions

Let Σ be a set of sentences of a first-order language $\mathcal{L}$ with equality. We want to add a new n-place predicate symbol P to $\mathcal{L}$, and we do so by adding a sentence ϵ of the form

$$\forall x_1 \ldots \forall x_n (P x_1 \ldots x_n \leftrightarrow \varphi)$$

to Σ. Here φ can be any formula in which P does not occur, and in which at most the variables $x_1, \ldots, x_n$ occur free.

A sentence ϵ with these properties is called an *explicit definition* of P.

The next proposition explains why such sentences are rightly called definitions.

Proposition 11 *(i) Let ϵ be defined as above. The sentence ϵ allows elimination of P. That is, for every formula χ in which P occurs there exists a formula χ' in which P does not occur such that*

$$\Sigma \cup \{\epsilon\} \models \chi \leftrightarrow \chi'$$

(ii) The sentence ϵ is not creative. That is: for every χ in which P does not occur it holds that

$$\textit{If } \Sigma \cup \{\epsilon\} \models \chi, \textit{ then } \Sigma \models \chi$$

Proof.

(i) This is shown by induction on the complexity of χ. The only non trivial case is the base case, when $\chi = P t_1 \ldots t_n$. We cannot choose χ' to be simply $[t_1/x_1] \ldots [t_n/x_n]\varphi$ because some of the variables occuring in $t_1, \ldots, t_n$ might get bound by quantifiers occurring in φ. Therefore we proceed more carefully as follows. Let $y_1, \ldots, y_n$ be variables not occurring in φ or $t_1, \ldots, t_n$. Choose χ' to be the formula

$$\exists y_1 \ldots \exists y_n ([y_1/x_1] \ldots [y_n/x_n]\varphi \wedge t_1 = y_1 \wedge \ldots t_n = y_n)$$

Clearly

$$\Sigma \cup \{\epsilon\} \models Pt_1 \dots t_n \leftrightarrow \chi'$$

(ii) Assume P does not occur in χ, and $\Sigma \cup \{\epsilon\} \models \chi$. Suppose $\Sigma \not\models \chi$. Then there exists a model $\mathcal{M} = \langle D, I \rangle$ for the language in which Σ is stated such that $\mathcal{M} \models \psi$ for all $\psi \in \Sigma$ and $\mathcal{M} \not\models \chi$. Now, expand $\mathcal{M}$ to a model $\mathcal{M}' = \langle D, I' \rangle$ for the full language as follows:

$\langle d_1, \dots, d_n \rangle \in I'(P)$ iff $\mathcal{M} \models \varphi\ [v(x_1/d_1)\dots(x_n/d_n)]$ Clearly, $\mathcal{M}' \models \psi$ for all $\psi \in \Sigma$ and $\mathcal{M} \models \epsilon$. Hence we should have $\mathcal{M}' \models \chi$. Contradiction.

The next proposition says that all sentences one might want to call definitions can be written as explicit definitions.

Proposition 12 *Consider any sentence δ in which the predicate P occurs, and which allows elimination of P and is not creative. There exists an explicit definition ϵ of P such that*

$$\Sigma \models \delta \leftrightarrow \epsilon$$

Proof. Since δ allows elimination there must exist a formula φ such that

$$\Sigma \cup \{\delta\} \models Px_1 \dots x_n \leftrightarrow \varphi\ (^*)$$

Let ϵ be the sentence $\forall x_1 \dots \forall x_n (Px_1 \dots x_n \leftrightarrow \varphi)$. We will show that $\Sigma \models \delta \leftrightarrow \epsilon$.

That $\Sigma \models \delta \rightarrow \epsilon$ follows almost immediately from (*). To prove that $\Sigma \models \epsilon \rightarrow \delta$, note that ϵ is an explicit definition and therefore allows elimination. That is to say that one can find for δ a sentence δ' in which P does not occur such that

$$\Sigma \cup \{\epsilon\} \models \delta \leftrightarrow \delta'\ (^{**})$$

From this it follows that $\Sigma \models \epsilon \rightarrow (\delta \rightarrow \delta')$. We already know that $\Sigma \models \delta \rightarrow \epsilon$, and by combining these two things we get $\Sigma \models \delta \rightarrow (\delta \rightarrow \delta')$, and thus that $\Sigma \cup \{\delta\} \models \delta'$. P does not occur in δ', δ is not crative, hence $\Sigma \models \delta'$.

(**) also implies that $\Sigma \models \delta' \rightarrow (\epsilon \rightarrow \delta)$, and combining this with the last statement we get $\Sigma \models \epsilon \rightarrow \delta$.

In the next exercise we look at the problem of definitions from the opposite perspective. Suppose we have a theory Σ stated in a language $\mathcal{L}$. Is the vocabulary as economical as possible? Or are there any predicates we can eliminate from the theory?

Exercise 15 *Let $\mathcal{L}$ be a first-order language without function symbols, and P an n-place predicate. An* explicit definition *of P is a sentence of the form*

$$\forall x_1 \dots \forall x_n (P(x_1, \dots, x_n) \leftrightarrow \varphi)$$

Here, φ can be any formula in which P does not occur, and in which at most the variables $x_1, \ldots, x_n$ occur free.

Now, let Σ be a theory of $\mathcal{L}$. By definition, a predicate P is eliminable *from Σ iff there exists an explicit definition ϵ of P such that $\Sigma \models \epsilon$.*

Assume that P is eliminable from Σ by the explicit definition ϵ.

Consider Σ' given by

$$\Sigma' = \{\varphi \mid P \textit{ does not occur in } \varphi, \textit{ and } \Sigma \models \varphi\} \cup \{\epsilon\}$$

Show that for all ψ, $\Sigma \models \psi$ iff $\Sigma' \models \psi$.

In 1901, the italian logician Alessandro Padoa published a method by which one can show that a predicate P is *not* eliminable from a theory Σ. To this end it is sufficient to define two models $\mathcal{M} = \langle D, I \rangle$ and $\mathcal{M}' = \langle D', I' \rangle$ such that both $\mathcal{M}$ and $\mathcal{M}'$ are models of Σ, and (a) $D = D'$; (b) for all symbols $\sigma \neq P$, $I(\sigma) = I'(\sigma)$; but (c) $I(P) \neq I'(P)$. Obviously, if there are two such models, it will be impossible for there to be a an explicit definition $\epsilon = \forall x_1 \ldots \forall x_n (P x_1 \ldots x_n \leftrightarrow \varphi)$ such that $\Sigma \models \epsilon$.

Here is an example.

Exercise 16
Let Σ consist of the following sentences:

$\forall x \forall y \forall z ((Rxy \wedge Ryz) \rightarrow Rxz)$
$\forall x \forall y \forall z ((Exy \wedge Eyz) \rightarrow Exz)$
$\forall x \neg Rxx$
$\forall x \forall y ((Rxy \vee Ryx) \rightarrow \neg Exy)$
$\forall x \forall y ((Rxy \vee Ryx) \vee Exy)$

(Read 'Rxy' as 'x is prefered to y', and 'Exy' as 'x and y are equally prefered')

(i) Show that E is eliminable from Σ.

(ii) Show that R is not eliminable from Σ.

The reminder of this chapter is devoted to a proof of Beth's theorem which says that if a predicate is not eliminable, one can *always* show this using Padoa's method.

It will take some time before we are ready for the proof of this theorem. It will follow from another important theorem.

Definition 13 *Let Δ_1 be a theory stated in the language $\mathcal{L}_1$ and Δ_2 be a theory stated in the language $\mathcal{L}_2$. Let $\mathcal{L}_0$ be the language given by the vocabulary common to $\mathcal{L}_1$ and $\mathcal{L}_2$. A sentence θ of $\mathcal{L}_0$* separates *Δ_1 and Δ_2 iff $\Delta_1 \models \theta$ and $\Delta_2 \models \neg\theta$. Δ_1 and Δ_2 are* inseperable *iff no sentence separates them.*

If the union of two theories is consistent, then they are inseparable, of course. The converse is also true, and follows from the next theorem.

Theorem 7 *(Craig Interpolation Theorem) Let $\mathcal{L}$ be a first order language with identity. Let φ and ψ be sentences of $\mathcal{L}$ such that $\varphi \models \psi$. Then there exists a sentence θ such that*

(i) $\varphi \models \theta$ and $\theta \models \psi$

(ii) Every predicate, function symbol, or individual constant (excluding identity) occurring in θ occurs in both φ and ψ.

The sentence θ is called and *interpolant* for φ and ψ. The identity sign is allowed to occur in θ even in cases in which it does not occur in both φ and ψ. Here are some examples showing why this is necessary.

(a) $\varphi = Pa \wedge \neg Pa$, $\psi = \exists x Qx$

(b) $\varphi = \forall x \forall y(x = y)$, $\psi = \forall x \forall y(Px \to Px)$

Let's now turn to the proof of the theorem.

Proof. Suppose there is no interpolant for φ and ψ. We will construct a model of $\varphi \wedge \neg\psi$, thus showing that $\varphi \not\models \psi$.

Notice first that we may assume that $\{\varphi\}$ and $\{\neg\psi\}$ are consistent. If $\{\varphi\}$ were inconsistent $\forall x(x \neq x)$ would be an interpolant for φ and ψ, and $\{\neg\psi\}$ were inconsistent $\forall x(x = x)$ would be an interpolant, and we are assuming there is none.

We wil construct a model for $\varphi \wedge \neg\psi$ by intertwining the Henkin models for $\{\varphi\}$ and $\{\neg\psi\}$.

Let $\mathcal{L}_1$ be the language of the symbols occurring in φ, $\mathcal{L}_2$ be the the language of the symbols occurring in ψ, and $\mathcal{L}_0$ be the language of the symbols occurring in both. Add the same countable set of individual constants to the symbols of each of the languages $\mathcal{L}$, $\mathcal{L}_0$, $\mathcal{L}_1$, and $\mathcal{L}_2$ to obtain the languages $\mathcal{L}'$, $\mathcal{L}'_0$, $\mathcal{L}'_1$, and $\mathcal{L}'_2$.

The theories $\{\varphi\}$ and $\{\neg\psi\}$ are theories of $\mathcal{L}'_1$, and $\mathcal{L}'_2$. $\mathcal{L}'_0$ is the language given by the vocabulary common to $\mathcal{L}'_1$ and $\mathcal{L}'_2$.

The first thing we will show is that no sentence θ of $\mathcal{L}'_0$ separates $\{\varphi\}$ and $\{\neg\psi\}$. Suppose, for contradiction, that θ does. Then $\varphi \models \theta$, and $\neg\psi \models \neg\theta$, or $\theta \models \psi$. Let $c_1, \ldots, c_n$ be the 'new' individual constants occurring in θ. Let $y_1, \ldots, y_n$ be variables not occurring in φ or ψ. Given that $c_1, \ldots, c_n$ do not occur in φ, the above gives that $\varphi \models \forall y_1 \ldots \forall y_n [y_1/c_1] \ldots [y_n/c_n]\theta$, and $\forall y_1 \ldots \forall y_n [y_1/c_1] \ldots [y_n/c_n]\theta \models \psi$. But that makes $\forall y_1 \ldots \forall y_n [y_1/c_1] \ldots [y_n/c_n]\theta$ an interpolant for φ and ψ. Contradiction.

Now let

$\varphi_0, \varphi_1, \varphi_2, \ldots$
and
$\psi_0, \psi_1, \psi_2, \ldots$
be enumerations of the sentences of $\mathcal{L}'_1$ and $\mathcal{L}'_2$ respectively. Using these enumerations one can define two increasing sequences of theories,
$\{\varphi\} = \Phi_0 \subseteq \Phi_1 \subseteq \Phi_2 \subseteq \ldots$
and
$\{\neg\psi\} = \Psi_0 \subseteq \Psi_1 \subseteq \Psi_2 \subseteq \ldots$
where all Φ_i are stated in $\mathcal{L}'_1$ and all Ψ_i are stated in $\mathcal{L}'_2$ such that the following holds

(i) For all n, Φ_n and Ψ_n are inseparable.

(ii) For all n, if $\Phi_n \cup \{\varphi_n\}$ and Ψ_n are inseparable, then $\varphi_n \in \Phi_{n+1}$.
For all n, if $\Psi_n \cup \{\psi_n\}$ and Φ_{n+1} are inseparable, then $\psi_n \in \Psi_{n+1}$.

(iii) For all n, if $\varphi_n = \exists x\chi$ and $\varphi_n \in \Phi_{n+1}$, then $[c/x]\chi \in \Phi_{n+1}$ for some 'new' c.
For all n, if $\psi_n = \exists x\chi$ and $\psi_n \in \Psi_{n+1}$, then $[c/x]\chi \in \Psi_{n+1}$ for some 'new' c.

Now set

$$\Phi = \bigcup_{n \in \mathbb{N}} \Phi_n \text{ and } \Psi = \bigcup_{n \in \mathbb{N}} \Psi_n$$

Then Φ and Ψ are inseparable. It follows that Φ is consistent and Ψ is consistent. Our goal is to show that $\Phi \cup \Psi$ is consistent.

To this end, note first that Φ is a maximal consistent theory in $\mathcal{L}'_1$. Suppose, for contradiction, that neither $\varphi_n \in \Phi$ nor $\neg\varphi_n \in \Phi$. Then $\Phi_n \cup \{\varphi_n\}$ is separable from Ψ_n, which means that there exists a θ in $\mathcal{L}'_0$ such that

$$\Phi_n, \varphi \models \theta \text{ whereas } \Psi_n \models \neg\theta$$

As a consequence

$$\Phi \models \varphi \to \theta \text{ whereas } \Psi \models \neg\theta$$

By the same argument we find there must be a θ' in $\mathcal{L}'_0$ such that

$$\Phi \models \neg\varphi \to \theta' \text{ whereas } \Psi \models \neg\theta'$$

But then we have

$$\Phi \models \theta \vee \theta' \text{ whereas } \Psi \models \neg(\theta \vee \theta')$$

which contradicts the inseparabilty of Φ and Ψ.

So, Φ is maximal consistent. A similar argument shows that Ψ is maximal consistent.

Next we show that $\Phi \cap \Psi$ is a maximal consistent theory in $\mathcal{L}'_0$. To prove this let χ be a sentence of $\mathcal{L}'_0$. By the above either $\chi \in \Phi$ or $\neg\chi \in \Phi$, and either $\chi \in \Psi$ or $\neg\chi \in \Psi$. By inseparability we cannot have $\chi \in \Phi$ and $\neg\chi \in \Psi$, or $\neg\chi \in \Phi$ and $\chi \in \Psi$. Therefore either $\chi \in \Phi \cap \Psi$ or $\neg\chi \in \Phi \cap \Psi$.

With the proof of the completeness theorem in mind we know that we can construct models $\mathcal{M}_1$ and $\mathcal{M}_2$ for Φ and Ψ respectively, built on the set of 'new' constants. Given the fact that $\Phi \cap \Psi$ is a maximal consistent theory in $\mathcal{L}'_0$, the $\mathcal{L}'_0$-reducts of these models coincide. In fact we have here one $\mathcal{L}'$ model $\mathcal{M}$ for $\Phi \cup \Psi$ with a $\mathcal{L}'_1$ restriction $\mathcal{M}_1$ which is a model for Φ, and a $\mathcal{L}'_2$ restriction which is a model for Ψ. In this model φ is true and ψ is false.

Theorem 8 *(Beth's definability theorem) Let Σ be a theory. Suppose there are no two models $\mathcal{M} = \langle D, I\rangle$ and $\mathcal{M}' = \langle D', I'\rangle$ such that both $\mathcal{M}$ and $\mathcal{M}'$ are models of Σ, and (a) $D = D'$; (b) for all symbols $\sigma \neq P$, $I(\sigma) = I'(\sigma)$; but (c) $I(P) \neq I'(P)$.*

Then there exists an explicit definition $\epsilon = \forall x_1 \dots \forall x_n (Px_1 \dots x_n \leftrightarrow \varphi)$ such that $\Sigma \models \epsilon$.

Proof.

Let Σ' be just like Σ except that the predicate P is replaced by P'. Note that it follows from the premises that

$$\Sigma \cup \Sigma' \models \forall x_1 \dots \forall x_n (Px_1 \dots x_n \leftrightarrow P'x_1 \dots x_n)$$

Add new constants $c_1, \dots, c_n$ to the language. Then

$$\Sigma \cup \Sigma' \models Pc_1 \dots c_n \to P'c_1 \dots c_n$$

By the compactness theorem, there exist finite subsets $\Delta \subseteq \Sigma$ and $\Delta' \subseteq \Sigma'$ such that

$$\Delta \cup \Delta' \models Pc_1 \dots c_n \to P'c_1 \dots c_n$$

Let Δ'' be just like Δ' except that the predicate P' is everywhere replaced by P. Take the conjunction $\varphi(P)$ of all formulas in $\Delta \cup \Delta''$ and also the formula $\varphi(P')$ that results when P' is substituted for P everywhere in $\varphi(P)$. Notice that

$$\varphi(P) \land \varphi(P') \models Pc_1 \dots c_n \to P'c_1 \dots c_n$$

We can rearrange this in such a way that all predicates P appear on the one side and all predicates P' on the other

$$\varphi(P) \land Pc_1 \dots c_n \models \varphi(P') \to P'c_1 \dots c_n$$

Now apply Craig's interpolation theorem to find a sentence θ in which neither P nor P' occur such that

$$\varphi(P) \wedge Pc_1 \ldots c_n \models \theta \ (^*)$$

and

$$\theta \models \varphi(P') \rightarrow P'c_1 \ldots c_n$$

Given that P' does not occur in θ we will also have

$$\theta \models \varphi(P) \rightarrow Pc_1 \ldots c_n \ (^{**})$$

Combining $(^*)$ and $(^{**})$, we get

$$\varphi(P) \models Pc_1 \ldots c_n \leftrightarrow \theta$$

Recall that $\varphi(P)$ is built from Σ, so $c_1, \ldots, c_n$ do not occur in $\varphi(P)$. Therefore we will have, for variables $x_1, \ldots, x_n$ not occurring in θ,

$$\varphi(P) \models \forall x_1 \ldots \forall x_n (Px_1 \ldots x_n \leftrightarrow [x_1/c_1] \ldots [x_n/c_n]\theta)$$

Hence

$$\Sigma \models \forall x_1 \ldots \forall x_n (Px_1 \ldots x_n \leftrightarrow [x_1/c_1] \ldots [x_n/c_n]\theta)$$

QED.