2.5 Soundness and Completeness of the Predicate Logic

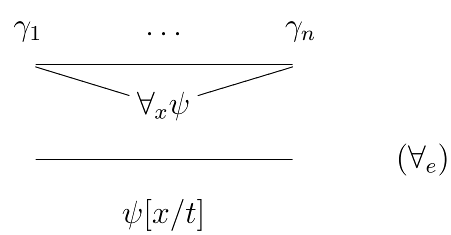
2.5.1 Soundness of the Predicate Logic

The soundness of predicate logic can be proved following the same idea used for the propositional logic. Therefore, we need to prove the following theorem:

Theorem 6 (Soundness of the predicate logic)**.** *Let* Γ *be a set of predicate formulas, if* $\Gamma \vdash \varphi$ *then* $\Gamma \models \varphi$*. In other words, if* φ *is provable from* Γ *then* φ *is a logical consequence of* Γ*.*

Proof. The proof is by induction on the derivation of $\Gamma \vdash \varphi$ similarly to the propositional case, and hence we focus just on the new rules: $(\forall_e)$, $(\forall_i)$, $(\exists_e)$, $(\exists_i)$.

If the last rule applied in the proof $\Gamma \vdash \varphi$ is $(\boldsymbol{\forall_e})$, then $\varphi = \psi[x/t]$ and the premise of the last rule is $\forall_x\psi$ as depicted in the following figure, where $\{\gamma_1, \ldots, \gamma_n\}$ is the subset of formulas in Γ used in the derivation.



The subtree rooted by the formula $\forall_x\psi$ and with open leaves labeled by formulas in Γ, corresponds to a derivation for the sequent $\Gamma \vdash \forall_x\psi$, that by induction hypothesis implies $\Gamma \models \forall_x\psi$. Therefore, for all interpretations that make the formulas in Γ true, also $\forall_x\psi$ would be true: $I \models \Gamma$ implies $I \models \forall_x\psi$. The last implies that for all $a \in \mathtt{D}$, where $\mathtt{D}$ is the domain of I, $I\frac{x}{a} \models \psi$, and in particular, $I\frac{x}{t^I} \models \psi$. Consequently, $I \models \psi[x/t]$. Therefore, one has that for any interpretation I, such that $I \models \Gamma$, $I \models \psi[x/t]$, which implies $\Gamma \models \psi[x/t]$.

If the last rule applied in the proof of $\Gamma \vdash \varphi$ is $(\boldsymbol{\forall_i})$, then $\varphi = \forall_x\psi$ and the premise of the last rule is $\psi[x/x_0]$ as depicted in the following figure:

$$\dfrac{\begin{array}{c}\gamma_1 \quad \ldots \quad \gamma_n \\ \hline \psi[x/x_0]\end{array}}{\forall_x \psi} \quad (\forall_i)$$

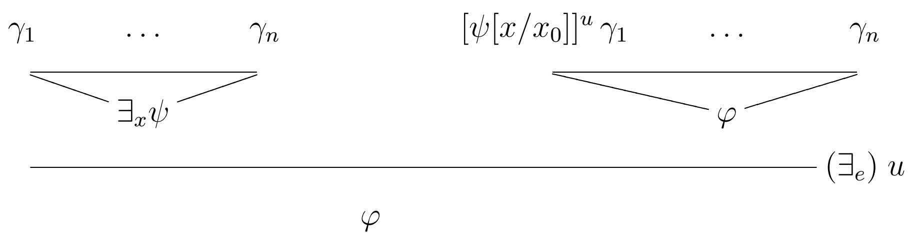
The subtree rooted by the formula $\psi[x/x_0]$ and with open leaves labeled by formulas in $\{\gamma_1, \ldots, \gamma_n\} \subset \Gamma$, corresponds to a derivation for the sequent $\Gamma \vdash \psi[x/x_0]$, in which no open assumption contains the variable x_0. This variable can be selected in such a manner that it does not appear free in any formula of Γ. By induction hypothesis, we have that $\Gamma \models \psi[x/x_0]$. This implies that all interpretations that make the formulas in Γ true, also make $\psi[x/x_0]$ true: $I \models \Gamma$ implies $I \models \psi[x/x_0]$. Since x_0 does not occurs in Γ, for all $a \in \mathtt{D}$, where $\mathtt{D}$ is the domain of I, $I\frac{x}{a} \models \Gamma$ and also $I\frac{x_0}{a} \models \psi[x/x_0]$ or, equivalently, $I\frac{x}{a} \models \psi$. Hence $\Gamma \models \forall_x \psi$.

If the last rule applied in the proof of $\Gamma \vdash \varphi$ is $(\exists_i)$, then $\varphi = \exists_x \psi$ and the premise of the last rule is $\psi[x/t]$ as depicted in the following figure, where again $\{\gamma_1, \ldots, \gamma_n\}$ is the subset of formulas of Γ used in the derivation:

$$\dfrac{\begin{array}{c}\gamma_1 \quad \ldots \quad \gamma_n \\ \hline \psi[x/t]\end{array}}{\exists_x \psi} \quad (\exists_i)$$

The subtree rooted by the formula $\psi[x/t]$ and with open leaves labeled by formulas of Γ, corresponds to a derivation of the sequent $\Gamma \vdash \psi[x/t]$, that by induction hypothesis implies $\Gamma \models \psi[x/t]$. Therefore, any interpretation I that makes the formulas in Γ true, also makes $\psi[x/t]$ true. Thus, since $I \models \psi[x/t]$ implies $I\frac{x}{t^I} \models \psi$, one has that $I \models \exists_x \psi$. Therefore, $\Gamma \models \exists_x \psi$.

Finally, for a derivation of the sequent $\Gamma \vdash \varphi$ that finishes with an application of the rule $(\exists_e)$, one has as premises the formulas $\exists_x \psi$ and φ. The former labels a root of a subtree with open leaves labeled by assumptions in $\{\gamma_1, \ldots, \gamma_n\} \subset \Gamma$ that corresponds to a derivation for the sequent $\Gamma \vdash \exists_x \psi$; the later labels a subtree with open leaves in $\{\gamma_1, \ldots, \gamma_n\} \cup \{\psi[x/x_0]\}$ and corresponds to a derivation for the sequent $\Gamma, \psi[x/x_0] \vdash \varphi$, where x_0 is a variable that does not occur free in $\Gamma \cup \{\varphi\}$, as depicted in the figure below:



By induction hypothesis, one has $\Gamma \models \exists_x \psi$ and $\Gamma, \psi[x/x_0] \models \varphi$. The first means that for any interpretation I such that $I \models \Gamma$, $I \models \exists_x \psi$. Thus, there exists some $a \in \mathtt{D}$, the domain of I, such that $I\frac{x}{a} \models \psi$. Notice also that since x_0 does not occur in Γ, one has that $I\frac{x_0}{a} \models \Gamma$. From the second, since $I\frac{x_0}{a} \models \Gamma, \psi[x/x_0]$, one has that $I\frac{x_0}{a} \models \varphi$. But, since x_0 does not occurs in φ, one concludes that $I \models \varphi$. $\square$

Exercise 30. *Complete all other cases of the proof of the Theorem 6 of soundness of predicate logic.*

2.5.2 Completeness of the Predicate Logic

The completeness proof for the predicate logic is not a direct extension of the completeness proof for the propositional logic. The completeness theorem was first proved by Kurt Gödel, and here we present the general idea of a proof due to Leon Albert Henkin (for nice complete presentations see references mentioned in the chapter on suggested readings).

The kernel of the proof is based on the fact that *every consistent set of formulas is satisfiable*, where consistency of the set Γ means that the absurd is not derivable from Γ:

Definition 28. *A set* Γ *of predicate formulas is consistent if not* $\Gamma \vdash \bot$.

Note that if we assume that **every consistent set is satisfiable** then the completeness can be easily obtained as follows:

Theorem 7 (Completeness)**.** *Let* Γ *be a set of predicate formulas. If* $\Gamma \models \varphi$ *then* $\Gamma \vdash \varphi$.

Proof. We prove that not $\Gamma \vdash \varphi$ implies not $\Gamma \models \varphi$. From not $\Gamma \vdash \varphi$ one has that $\Gamma \cup \{\neg\varphi\}$ is consistent because if $\Gamma \cup \{\neg\varphi\}$ were inconsistent then $\Gamma \cup \{\neg\varphi\} \vdash \bot$ by definition, and one could prove φ as follows:

$$\dfrac{\begin{array}{c}\Gamma, [\neg\varphi]^a \\ \vdots \\ \bot\end{array}}{\varphi}\ (\text{PBC})\ a$$

Therefore, $\Gamma \vdash \varphi$, which contradicts the supposition that not $\Gamma \vdash \varphi$. Now, since $\Gamma \cup \{\neg\varphi\}$ is consistent, by the assumption that consistent sets are satisfiable, we have that $\Gamma \cup \{\neg\varphi\}$ is satisfiable. Therefore, we conclude that not $\Gamma \models \varphi$. □

Our goal from now on is to prove that **every consistent set of formulas is satisfiable**. The idea is, given a consistent set of predicate formulas Γ, to build a model I for Γ, and since the sole available information is its consistency, this must be done by purely syntactical means, that is by using the language to build the desired model.

The key concepts in Henkin's proof are the notion of *witnesses* of existential formulas and extension of consistent sets of formulas to *maximally consistent* sets.

Definition 29 (Witnesses and maximally consistency)**.** *Let* Γ *be a set of formulas*

Γ *contains witnesses if and only if for every formula of the form* $\exists_x\varphi$ *in* Γ*, there exists a term* t *such that* $\Gamma \vdash \exists_x\varphi \to \varphi[x/t]$.

Γ *is maximally consistent if and only if for each formula* φ*,* $\Gamma \vdash \varphi$ *or* $\Gamma \vdash \neg\varphi$.

Notice that from the definition, for any possible extension of a maximally consistent set Γ, say Γ' such that $\Gamma \subseteq \Gamma'$, $\Gamma' = \Gamma$. Maximally consistent sets are also said to be *closed for negation.*

The proof is done in two steps, and uses the fact that every subset of a satisfiable set is also satisfiable:

1. every consistent set can be extended to a maximally consistent set containing witnesses;

2. every maximally consistent set containing witnesses has a model.

If Γ does not contain witnesses, these formulas can not be built in a straightforward manner, since one can not choose any arbitrary term t to be witness of the existential formula without changing the semantics. Nevertheless, any consistent set can be extended to another consistent set containing witnesses. The simplest case, is when the language is countable and the set Γ uses only a finite set of free variables, that is $\texttt{fv}(\Gamma)$ is finite. Since the set of existential formulas is also countable and there are infinite unused variable (those that do not appear free in Γ). Then these variables can be used as witnesses without any conflict. The other cases are more elaborated and are left as research exercises to the reader (Exercises 32 and 33): the case in which the language is countable, but Γ uses infinitely many free variables and the case in which the language is not countable.

In the sequel we will treat the simplest case in which the set of constant, function and predicate symbols occurring in Γ is at most countable and there are only finitely many variables occurring in Γ. The next two lemmas complete the first part of the proof: a consistent set might be extended to a maximally consistent set with witnesses. This is done proving first how variables might be used to include witnesses and then how a consistent set with witnesses can be extended to a maximally consistent set.

Lemma 4 (Construction of witnesses)**.** *Let Γ be a consistent set over a countable language such that $\texttt{fv}(\Gamma)$ is finite. There exists an extension $\Gamma' \supseteq \Gamma$ over the same language, such that Γ' is consistent and contains witnesses.*

Proof. Let $\exists_{x_1}\varphi_1, \exists_{x_2}\varphi_2, \ldots$ be an enumeration of all the existential formulas built over the language. Let $y_1, y_2, \ldots$ be an enumeration of the variables not occurring free in Γ, and consider the formulas below, for $i > 0$:

$$(\exists_{x_i}\varphi_i) \rightarrow \varphi_i[x_i/y_i]$$

Let Γ_0 be defined as Γ, and Γ_n, for $n > 0$ be defined as below:

$$\Gamma_n = \Gamma_{n-1} \cup \{(\exists_{x_n}\varphi_n) \rightarrow \varphi_n[x_n/y_n]\}$$

We will prove the consistence of Γ' defined as $\Gamma' = \bigcup_{n\in\mathbb{N}} \Gamma_n$ by induction on n. The base case is trivial since Γ is consistent by hypothesis. For $k > 0$, suppose Γ_{k-1} is consistent, but Γ_k is not, i.e.

$$\Gamma_k = \Gamma_{k-1} \cup \{(\exists_{x_k}\varphi_k) \to \varphi_k[x_k/y_k]\} \vdash \bot \tag{2.1}$$

Now consider the following derivation:

$$\dfrac{(\text{LEM})\ (\exists_{x_k}\varphi_k)\lor\neg(\exists_{x_k}\varphi_k) \qquad \begin{matrix}\Gamma_{k-1}\ [\exists_{x_k}\varphi_k]^a\\ \nabla_1\\ \bot\end{matrix} \qquad \begin{matrix}\Gamma_{k-1}\ [\neg\exists_{x_k}\varphi_k]^b\\ \nabla_2\\ \bot\end{matrix}}{\bot}\ (\lor e)\, a\, b$$

where

$$\nabla_1 : \quad \dfrac{[\exists_{x_k}\varphi_k]^a \qquad \dfrac{\Gamma_{k-1} \qquad \dfrac{[\varphi_k[x_k/y_k]]^u}{\exists_{x_k}\varphi_k \to \varphi_k[x_k/y_k]}\ (\to i)\emptyset}{\bot}\ (2.1)}{\bot}\ (\exists e)u$$

and

$$\nabla_2 : \quad \dfrac{\Gamma_{k-1} \qquad \dfrac{\dfrac{[\neg\exists_{x_k}\varphi_k]^b}{\neg\varphi_k[x_k/y_k] \to \neg\exists_{x_k}\varphi_k}\ (\to i)\emptyset}{\exists_{x_k}\varphi_k \to \varphi_k[x_k/y_k]}\ (\text{CP})}{\bot}\ (2.1)$$

But this is a proof of $\Gamma_{k-1} \vdash \bot$ which contradicts the assumption that Γ_{k-1} is consistent.

Therefore, Γ_k is consistent. □

In the previous proof, note that if $\Gamma_{i-1} \vdash \exists_{x_i}\varphi_i$ then it must be the case that $\Gamma_i \vdash \varphi_i[x_i/y_i]$ in order to preserve the consistency. Therefore, $\varphi_i[x_i/y_i]$ might be added to the set of formulas, but not its negation, as will be seen in the further construction of maximally consistent sets.

Now we prove that every maximally consistent set containing witnesses has a model.

Lemma 5 (Lindenbaum)**.** *Each consistent set of formulas* Γ *over a countable language is contained in a maximally consistent set* Γ^* *over the same language.*

Proof. Let $\delta_1, \delta_2, \ldots$ be an enumeration of the formulas built over the language. In order to build a consistent expansion of Γ we recursively define the family of indexed sets of formulas Γ_i as follows:

- $\Gamma_0 = \Gamma$
- $\Gamma_i = \begin{cases} \Gamma_{i-1} \cup \{\delta_i\}, & \text{if } \Gamma_{i-1} \cup \{\delta_i\} \text{ is consistent;} \\ \Gamma_{i-1}, & \text{otherwise.} \end{cases}$

Now let $\Gamma^* = \bigcup_{i \in \mathbb{N}} \Gamma_i$. We claim that Γ^* is maximally consistent. In fact, if Γ^* is not maximally consistent then there exists a formula $\gamma \notin \Gamma^*$ such that $\Gamma^* \cup \{\gamma\}$ is consistent. But by the above enumeration, there exists $k \geq 1$ such that $\gamma = \delta_k$, and since $\Gamma_{k-1} \cup \{\gamma\}$ should be consistent, $\delta_k \in \Gamma_{k+1}$. Hence $\delta_k = \gamma \in \Gamma^*$. □

From the previous lemmas (4 and 5), one has that every consistent set of formulas built over a countable set of symbols and with finitely many free variables can be extended to a maximally consistent set which contains witnesses. In this manner we complete the first step of the prove.

Now, we will complete the second step of the proof, that is that any maximally consistent set that contain witnesses is satisfiable. We start with two auxiliary definitional observations.

Lemma 6. *Let* Γ *be a maximally consistent set of formulas. Then for any formula* φ *either* $\varphi \in \Gamma$ *or* $\neg\varphi \in \Gamma$*.*

Lemma 7. *Let Γ be a maximally consistent set. For any formula φ, $\Gamma \vdash \varphi$ if, and only if $\varphi \in \Gamma$.*

Proof. Suppose $\Gamma \vdash \varphi$. From Lemma 6, either $\varphi \in \Gamma$ or $\neg\varphi \in \Gamma$. If $\neg\varphi \in \Gamma$ then Γ would be inconsistent:

$$\dfrac{\begin{array}{c}\Gamma\\ \nabla\\ \varphi\end{array} \qquad \begin{array}{c}\Gamma\\ \nabla\\ \neg\varphi\end{array}}{\bot}\ (\neg_e)$$

Therefore, $\varphi \in \Gamma$. □

We now define a model, that is called the *algebra* or *structure of terms* for the set Γ which is assumed to be maximally consistent and containing witnesses. The model, denoted as I_Γ, is built from Γ by taking as domain, the set D of all terms built over the countable language of Γ as given in the definition of terms 13. The designation d for each variable is the same variable and the interpretation of each non variable term is itself too: $t^{I_\Gamma} = t$. Notice that since our predicate language does not deal with equality symbol, different terms are interpreted as different elements of D. The map m of I_Γ maps each n-ary function symbol in the language, f, in the function f^{I_Γ} such that for all terms $t_1, \ldots, t_n$, $(f(t_1, \ldots, t_n))^{I_\Gamma} = f^{I_\Gamma}(t_1^{I_\Gamma}, \ldots, t_n^{I_\Gamma}) = f(t_1, \ldots, t_n)$, and for each n-ary predicate symbol p, p^{I_Γ} is the relation defined as

$$(p(t_1, \ldots, t_n))^{I_\Gamma} = p^{I_\Gamma}(t_1^{I_\Gamma}, \ldots, t_n^{I_\Gamma}) \text{ if and only if } p(t_1, \ldots, t_n) \in \Gamma$$

With these definitions we have that for any atomic formula φ, $\varphi \in \Gamma$ if and only if $I_\Gamma \models \varphi$. In addition, according to the interpretation of quantifiers, for any atomic formula $\forall_{x_1} \ldots \forall_{x_n}\varphi \in \Gamma$ if and only if $I_\Gamma \models \forall_{x_1} \ldots \forall_{x_n}\varphi$ and $\exists_{x_1} \ldots \exists_{x_n}\varphi \in \Gamma$ if and only if $I_\Gamma \models \exists_{x_1} \ldots \exists_{x_n}\varphi$.

Using the assumptions that Γ has witnesses and is maximally consistent, formulas can be correctly interpreted in I_Γ as below.

1. $\bot^{I_\Gamma} = F$ and $\top^{I_\Gamma} = T$
2. $\varphi^{I_\Gamma} = T$, iff $\varphi \in \Gamma$, for any atomic formula φ
3. $(\neg\varphi)^{I_\Gamma} = T$, iff $\varphi^{I_\Gamma} = F$
4. $(\varphi \wedge \psi)^{I_\Gamma} = T$, iff $\varphi^{I_\Gamma} = T$ and $\psi^{I_\Gamma} = T$
5. $(\varphi \vee \psi)^{I_\Gamma} = T$, iff $\varphi^{I_\Gamma} = T$ or $\psi^{I_\Gamma} = T$
6. $(\varphi \to \psi)^{I_\Gamma} = T$, iff $\varphi^{I_\Gamma} = F$ or $\psi^{I_\Gamma} = T$
7. $(\exists_x\varphi)^{I_\Gamma} = T$, iff $(\varphi[x/t])^{I_\Gamma} = T$, for some term $t \in \mathtt{D}$
8. $(\forall_x\varphi)^{I_\Gamma} = T$, iff $(\varphi[x/t])^{I_\Gamma} = T$, for all $t \in \mathtt{D}$.

Indeed, this interpretation is well-defined only under the assumption that Γ has witnesses and is maximally consistent. For instance, the item 3 is well-defined since $\neg\varphi \in \Gamma$ if and only if not $\varphi \in \Gamma$. For the item 5, if $(\varphi \vee \psi) \in \Gamma$ and not $\varphi \in \Gamma$, by maximally consistency one has that $\neg\varphi \in \Gamma$; thus, from $(\varphi \vee \psi)$ and $\neg\varphi$, it is possible to derive ψ (by simple application of rules $(\vee_e)$ and $(\neg_e)$ and $(\bot_e)$). Similarly, if we assume $(\varphi \vee \psi) \in \Gamma$ and not $\psi \in \Gamma$, we can derive φ. For the item 6, suppose $(\varphi \to \psi) \in \Gamma$ and $\varphi \in \Gamma$, then one can derive ψ (by application of $(\to_e)$); otherwise, if $(\varphi \to \psi) \in \Gamma$ and not $\psi \in \Gamma$, by maximally consistency, $\neg\psi \in \Gamma$, from which one can infer $\neg\varphi$ (by application of contraposition). For the item 7, if we assume $\exists_x\varphi \in \Gamma$, by the existence of witnesses, there is a term t such that $\exists_x\varphi \to \varphi[x/t] \in \Gamma$, and from these two formulas we can derive $\varphi[x/t]$ (by a simple application of rule $(\to_e)$).

Exercise 31. *Complete the analysis well-definedness for all the items in the interpretation of formulas I_Γ, for a set Γ that contains witnesses and is maximally complete.*

Theorem 8 (Henkin)**.** *Let Γ be a maximally consistent set containing witnesses. Then for all φ,*

$$I_\Gamma \models \varphi, \text{ if, and only if } \Gamma \vdash \varphi.$$

Proof. The proof is done by induction on the structure of φ. If φ is an atomic formula then $\varphi \in \Gamma$ iff $(\varphi)^{I_\Gamma} = T$, by definition.

If $\varphi = \neg\varphi_1$ then:

$$\begin{array}{lcl} \neg\varphi_1 \in \Gamma & \iff & \text{(because } \Gamma \text{ is maximally consistent)} \\ \varphi_1 \notin \Gamma & \iff & \text{(by induction hypothesis)} \\ \text{not } I_\Gamma \models \varphi_1 & \iff & \text{(by definition)} \\ I_\Gamma \models \neg\varphi_1. & & \end{array}$$

If $\varphi = \varphi_1 \wedge \varphi_2$ then:

$$\begin{array}{lcl} \varphi_1 \wedge \varphi_2 \in \Gamma & \iff & \text{(by definition)} \\ \varphi_1 \in \Gamma \text{ and } \varphi_2 \in \Gamma & \iff & \text{(by induction hypothesis for both } \varphi_1 \text{ and } \varphi_2\text{)} \\ I_\Gamma \models \varphi_1 \text{ and } I_\Gamma \models \varphi_2 & \iff & \text{(by definition)} \\ I_\Gamma \models \varphi_1 \wedge \varphi_2. & & \end{array}$$

If $\varphi = \varphi_1 \vee \varphi_2$ then:

$$\begin{array}{lcl} \varphi_1 \vee \varphi_2 \in \Gamma & \iff & \text{(by definition)} \\ \varphi_1 \in \Gamma \text{ or } \varphi_2 \in \Gamma & \iff & \text{(by induction hypothesis for both } \varphi_1 \text{ and } \varphi_2\text{)} \\ I_\Gamma \models \varphi_1 \text{ or } I_\Gamma \models \varphi_2 & \iff & \text{(by definition, no matter the condition holds for } \varphi_1 \text{ or } \varphi_2\text{)} \\ I_\Gamma \models \varphi_1 \vee \varphi_2. & & \end{array}$$

If $\varphi = \varphi_1 \to \varphi_2$ then we split the proof into two parts. Firstly, we show that $\varphi_1 \to \varphi_2 \in \Gamma$ implies $I_\Gamma \models \varphi_1 \to \varphi_2$. We have two subcases:

1. $\varphi_1 \in \Gamma$: In this case, $\varphi_2 \in \Gamma$. In fact, if $\varphi_2 \notin \Gamma$ then $\neg\varphi_2 \in \Gamma$ by the maximality of Γ, and Γ becomes contradictorily inconsistent:

$$\dfrac{\dfrac{\varphi_1 \to \varphi_2 \qquad \varphi_1}{\varphi_2}\,(\to_e) \qquad \neg\varphi_2}{\bot}\,(\neg_e)$$

Thus, by induction hypothesis one has:

$$\begin{array}{lll} \varphi_1 \in \Gamma \text{ and } \varphi_2 \in \Gamma & \Longleftrightarrow & \text{(by induction hypothesis for both } \varphi_1 \text{ and } \varphi_2) \\ I_\Gamma \models \varphi_1 \text{ and } I_\Gamma \models \varphi_2 & \Longrightarrow & \text{(by definition)} \\ I_\Gamma \models \varphi_1 \to \varphi_2. & & \end{array}$$

2. $\varphi_1 \notin \Gamma$: In this case, $\neg\varphi_1 \in \Gamma$ by the maximality of Γ. Therefore,

$$\begin{array}{lll} \neg\varphi_1 \in \Gamma & \Longleftrightarrow & \text{(by induction hypothesis)} \\ I_\Gamma \models \neg\varphi_1 & \Longleftrightarrow & \text{(by definition)} \\ \text{not } I_\Gamma \models \varphi_1 & \Longrightarrow & \text{(by definition)} \\ I_\Gamma \models \varphi_1 \to \varphi_2. & & \end{array}$$

Now we prove that $I_\Gamma \models \varphi_1 \to \varphi_2$ implies $\varphi_1 \to \varphi_2 \in \Gamma$. By definition of the semantics of implication, there are two cases:

1. $\varphi_1^{I_\Gamma} = F$: In this case, we have that $(\neg\varphi_1)^{I_\Gamma} = T$, and hence $\neg\varphi_1 \in \Gamma$, by induction hypothesis. We can now derive $\varphi_1 \to \varphi_2$ as follows, and conclude by Lemma 7:

$$\dfrac{\dfrac{\dfrac{\neg\varphi_1 \qquad [\varphi_1]^a}{\bot}\,(\neg_e)}{\varphi_2}\,(\bot_e)}{\varphi_1 \to \varphi_2}\,(\to_i)\ a$$

2. $\varphi_2^{I_\Gamma} = T$: By induction hypothesis $\varphi_2 \in \Gamma$, and we derive $\varphi_1 \to \varphi_2$ as follows, and conclude by Lemma 7:

$$\dfrac{\varphi_2}{\varphi_1 \to \varphi_2}\,(\to_i)\ \emptyset$$

If $\varphi = \exists_x \varphi_1$ then:

$$\begin{array}{lll} \exists_x \varphi_1 \in \Gamma & \iff & \text{(for some } t \in \mathtt{D}\text{, since } \Gamma \text{ contains witnesses)} \\ \varphi_1[x/t] \in \Gamma & \iff & \text{(by induction hypothesis)} \\ I_\Gamma \models \varphi_1[x/t] & \iff & \text{(by definition)} \\ I_\Gamma \models \exists_x \varphi_1. & & \end{array}$$

If $\varphi = \forall_x \varphi_1$ then:

$$\begin{array}{lll} \forall_x \varphi_1 \in \Gamma & \iff & \text{(otherwise } \Gamma \text{ becomes inconsistent as shown below)} \\ \varphi_1[x/t] \in \Gamma, \text{ for all } t \in \mathtt{D} & \iff & \text{(by induction hypothesis)} \\ I_\Gamma \models \varphi_1[x/t], \text{ for all } t \in \mathtt{D} & \iff & \text{(by definition)} \\ I_\Gamma \models \forall_x \varphi_1. & & \end{array}$$

For the first equivalence, note that if $\forall_x \varphi_1 \in \Gamma$ then $\varphi_1[x/t] \in \Gamma$, for all term $t \in \mathtt{D}$, otherwise Γ becomes contradictorily inconsistent:

$$\dfrac{\neg\varphi_1[x/t] \qquad \dfrac{\forall_x \varphi_1}{\varphi_1[x/t]}\,(\forall_e)}{\bot}\,(\bot_e)$$

□

Using as a model I_Γ, it is possible to conclude, in this case, that consistent sets are satisfiable.

Corollary 2 (Consistency implies satisfiability)**.** *If* Γ *is a consistent set of formulas over a countable language with a finite set of free variables then* Γ *is satisfiable.*

Proof. Initially, Γ is consistently enlarged obtaining the set Γ' including witnesses according to the construction in Lemma 4; afterwards, Γ' is closed maximally obtaining the set $(\Gamma')^*$according to the construction in Lindenbaum's Lemma (5). This set contains witnesses and is maximally consistent; then, by Henkin's Theorem (8), I_Γ is a model of $(\Gamma')^*$, hence a model of Γ too. □

Exercise 32. **(*)** *Research in the suggested related references how a consistent set built over a countable set of symbols, but that uses infinite free variables can be extended to a maximal consistent set with witnesses. The problem, is that in this case there are no new variables that can be used as witnesses. Thus, one needs to extend the language with new constant symbols that will act as witnesses, but each time a new constant symbol is added to the language the set of existential formulas change.*

Exercise 33. **(*)** *Research the general case in which the language is not restricted, that is the case in which Γ is built over a non countable set of symbols.*

2.5.3 Compactness Theorem and Löwenheim-Skolem Theorem

The connections between $\models$ and $\vdash$ as well as between consistence and satisfiability provided in this section, give rise to other additional important consequences that relate semantic and syntactic elements of the predicate logic. Here we present two important theorems that are related with the scope and limits of the expressiveness of predicate logic.

Theorem 9 (Compactness)**.** *Given a set Γ of predicate formulas and a formula φ, the following holds*

i. $\Gamma \models \varphi$ *if and only if there is a finite set* $\Gamma_0 \subseteq \Gamma$ *such that* $\Gamma_0 \models \varphi$

ii. Γ *is satisfiable if and only if for all finite set* $\Gamma_0 \subseteq \Gamma$*,* Γ_0 *is satisfiable.*

Proof. i. For necessity, if $\Gamma \models \varphi$, by completeness one has that there exists a derivation ∇ for $\Gamma \vdash \varphi$. The derivation ∇ uses only a finite subset of assumptions, say $\Gamma_0 \subseteq \Gamma$. Thus, $\Gamma_0 \vdash \varphi$ and, by correctness, one concludes that $\Gamma_0 \models \varphi$. For sufficiency, suppose that $\Gamma_0 \models \varphi$, for a finite set $\Gamma_0 \subseteq \Gamma$. By completeness there exists a derivation ∇ for $\Gamma_0 \vdash \varphi$. But ∇ is also a derivation for $\Gamma \vdash \varphi$; hence, by correctness one concludes that $\Gamma \models \varphi$.

ii. Necessity is proved by contraposition: if Γ_0 were unsatisfiable for some finite set $\Gamma_0 \subseteq \Gamma$, then Γ_0 would be inconsistent, since consistency implies satisfiability (Corollary 2); thus, $\Gamma_0 \vdash \bot$, which implies also that $\Gamma \vdash \bot$ and by correctness that $\Gamma \models \bot$. Hence, Γ would be unsatisfiable. Sufficiency is proved also by contraposition: if we assume that Γ is unsatisfiable, then since there exists no model for Γ, $\Gamma \models \bot$ holds. By completeness also, $\Gamma \vdash \bot$ and hence, there exists a finite set $\Gamma_0 \subseteq \Gamma$, such that $\Gamma_0 \vdash \bot$, which by correctness implies that $\Gamma_0 \models \bot$. Thus, we conclude that Γ_0 is unsatisfiable.

□

The compactness theorem has several applications that are useful for restricting the analysis of consistency and satisfiability of arbitrary sets of predicate formulas to only finite subsets. This also has important implications in the possible cardinality of models of sets of predicate formulas such as the one given in the following theorem.

Theorem 10 (Löwenheim-Skolem). *Let Γ be a set of formulas such that for any natural $n \in \mathbb{N}$, there exists a model of Γ with a domain of cardinality at least n. Then Γ has also infinite models.*

Proof. Consider an additional binary predicate symbol E and the formulas φ_n for $n > 0$, defined as

$$\forall_x \; E(x,x) \wedge \exists_{x_1,\ldots,x_n} \bigwedge_{i \neq j; i,j=1}^{n} \neg E(x_i, x_j)$$

For instance, the formulas φ_1 and φ_3 are given respectively as $\forall_x E(x,x)$ and $\forall_x \; E(x,x) \wedge \exists_{x_1} \exists_{x_2} \exists_{x_3} \; (\neg E(x_1,x_2) \wedge \neg E(x_1,x_3) \wedge \neg E(x_2,x_3))$.

Notice that φ_n has models of cardinality at least n. It is enough to interpret E just as a the reflexive relation among the elements of the domain of the interpretation. Thus, pairs of different elements of the domain do not belong to the interpretation of E.

Let Φ be the set of formulas $\{\varphi_n \mid n \in \mathbb{N}\}$. We will prove that all finite subsets of the set of formulas $\Gamma \cup \Phi$ are satisfiable and then by the compactness theorem conclude that $\Gamma \cup \Phi$ is satisfiable too. An interpretation $I \models \Gamma \cup \Phi$ should have an infinite model, since also $I \models \Phi$

and all formulas in Φ are true in I only if there are infinitely many elements in the domain of I.

To prove that any finite set $\Gamma_0 \subset \Gamma \cup \Phi$ is satisfiable, let k be the maximum k such that $\varphi_k \in \Gamma_0$. Since Γ has models of arbitrary finite cardinality, let I' be a model of Γ with at least k elements in its domain $\mathtt{D}$. I' can be extended in such a manner that the binary predicate symbol E is interpreted just as the reflexive relation over $\mathtt{D}$. Let I be the extended interpretation. It is clear that $I \models \Gamma$ since E is a new symbol and also $I \models \Gamma_0 \cap \Phi$ since the domain has at least k different elements. Also, since $I \models \Gamma$, we have that $I \models \Gamma \cap \Gamma_0$. Hence, $I \models \Gamma_0$ and so we conclude that Γ_0 is satisfiable. □

Exercise 34. *Prove that there is no predicate formula φ that holds exclusively for all finite interpretations.*

Exercise 35. *Let E be a binary predicate symbol, e a constant and $\cdot$ and -1 be binary and unary function symbols, respectively. The theory of groups is given by the models of the set of formulas Γ_G:*

$$\forall_x \; E(x, x)$$
$$\forall_{x,y} \; (E(x, y) \rightarrow E(y, x))$$
$$\forall_{x,y,z} \; (E(x, y) \wedge E(y, z) \rightarrow E(x, z))$$
$$\forall_x \; E(x \cdot e, x)$$
$$\forall_x \; E(x \cdot x^{-1}, e)$$
$$\forall_{x,y,z} \; E((x \cdot y) \cdot z, x \cdot (y \cdot z))$$

Notice that according to the three first axioms the symbol E should be interpreted as an equivalence relation such as the equality. Indeed, the three other axioms are those related with group theory itself: the fourth one states the existence of an identity element, the fifth one the inverse function and the sixth one the associativity of the binary operation.

Prove the existence of infinite models by proving that for any $n \in \mathbb{N}$, the structure of arithmetic modulo n is a group of cardinality n. The elements of this structure are all integers modulo n (i.e. the set $\{0, 1, \ldots, n\text{-}1\}$), with addition and identity element 0.

Exercise 36. *A graph is a structure of the form* $G = \langle V, E \rangle$*, where* V *is a finite set of vertices and* $E \subset V \times V$ *a set of edges between the vertices. The problem of reachability in graphs is the question whether there exists a finite path of consecutive edges, say* $(u, u_1), (u_1, u_2), \ldots, (u_{n-1}, v)$*, between two given nodes* $u, v \in V$*.*

Prove that there is no predicate formula that expresses reachability in graphs.

Hint: the key observation to conclude is that the problem of reachability between two nodes might be answered positively whenever there exists a path of arbitrary length.

2.6 Undecidability of the Predicate Logic

The gain of expressiveness obtained in predicate logic w.r.t. to the propositional logic comes at a price. Initially, remember that for a given propositional formula φ, one can always answer whether φ is valid or not by analyzing its truth table. This means that there is an algorithm that receives an arbitrary propositional formula as input and **always** answers after a finite amount of time `yes`, if the given formula is valid; or `no`, otherwise. The algorithm works as follows: build the truth table for φ and check whether it is true for all interpretations. Note that this algorithm is not efficient because the (finite) number of possible interpretations grows exponentially w.r.t. the number of propositional variables occurring in φ.

In general, a computational question with a `yes` or `no` answer depending on the parameters is known as a *decision problem*. A decision problem is said to be *decidable* whenever there exists an algorithm that correctly answers `yes` or `no` for each instance of the problem, and when such algorithm does not exist the decision problem is said to be *undecidable* . Therefore, we conclude that that *validity* is decidable in propositional logic.

The natural question that arises at this point is whether validity is decidable or not in predicate logic. Note that the truth table approach is no longer possible because the number of different interpretations for a given predicate formula φ is not finite. In fact, as stated in the previous paragraph the gain of expressiveness of the predicate logic comes at a price: validity is undecidable in predicate logic. This fact is usually known as the *undecidability of predicate logic*, and has several important consequences. In fact, it is straightforward from the completeness of predicate logic that provability is also undecidable, i.e. there is no algorithm that receives a predicate formula φ as input and returns `yes` if $\vdash \varphi$, or `no` if not $\vdash \varphi$.

The standard technique for proving the undecidability of the predicate logic consists in reducing a known undecidable problem to the validity of the predicate logic in such a way that decidability of validity of the predicate logic entails the decidability of the other problem leading to a contradiction. In what follows, we consider the word problem for a specific monoid introduced by G. S. Tseitin, and that is well-known to be undecidable.

A semigroup is an algebraic structure with a binary associative operator $\cdot$ over a given set A. When in addition the structure has an identity element id it is called a monoid. By associativity, one understands that for all x, y, z in A, $x \cdot (y \cdot z) = (x \cdot y) \cdot z$, and for all $x \in A$ the identity satisfies the properties $id \cdot x = x$ and $x \cdot id = x$. In general, the word problem in a given semigroup with a given set of equations E (between pairs of elements of A), is the problem of answering whether two words are equal *applying* these equations.

By an application of an equation, say $u = v$ in E, one understand an equational transformation of the form below, where x y are any elements of A.

$$x \cdot (u \cdot y) = x \cdot (v \cdot y)$$

Hence, the word problem consists in answering for any pair of elements $x, y \in A$ if there exists a finite chain, possibly of length zero, of applications of equations that transform x in y:

$$x \equiv x_0 \stackrel{u_1=v_1}{=} x_1 \stackrel{u_2=v_2}{=} x_2 \stackrel{u_3=v_3}{=} \dots \stackrel{u_n=v_n}{=} x_n \equiv y \tag{2.2}$$

In the chain above, the notation $\equiv$ is used for syntactic equality and $\stackrel{u_i=v_i}{=}$ for highlighting that the equation applied in the application step is $u_i = v_i$.

Tseitin's monoid is given by the set Σ^* of words freely generated by the quinary alphabet $\Sigma = \{a, b, c, d, e\}$. In this structure the binary associative operator is the concatenation of words and the empty word plays the role of the identity. The set of equations is given below. For simplicity, we will omit parentheses and the concatenation operator.

$$
\begin{aligned}
ac &= ca \\
ad &= da \\
bc &= cb \\
bd &= db \\
ce &= eca \\
de &= edb \\
cdca &= cdcae
\end{aligned} \tag{2.3}
$$

As previously mentioned Tseitin introduced this specific monoid with the congruence generated by this set of equations and proved that the word problem in this structure is undecidable.

In order to reduce the above problem to validity of the predicate logic, we choose a logical language with a constant symbol $\square$, five unary function symbols f_a, f_b, f_c, f_d and f_e, and a binary predicate P. The constant $\square$ will be interpreted as the empty word, and each function symbol, say $f_\star$ for $\star \in \Sigma$, as the concatenation of the symbol $\star$ to the left of the term given as argument of $f_\star$. For example, the word *baaecde* will be encoded as $f_b(f_a(f_a(f_e(f_c(f_d(f_e(\square)))))))$, which for brevity will be written simply as $f_{baaecde}(\square)$. The binary predicate P will play the role of equality, i.e. $P(x, y)$ is interpreted as x is equal to y (modulo the congruence induced by the set of equations above, which would be assumed as axioms).

Our goal is, given an instance of the word problem $x, y \in \Sigma^*$ specified above, to build a formula $\varphi_{x,y}$ such that x equals y in this structure if and only if $\models \varphi_{x,y}$. The formula $\varphi_{x,y}$ is of the form

$$\varphi' \to P(f_x(\square), f_y(\square)) \tag{2.4}$$

where φ' is the following formula:

$$
\begin{aligned}
&\forall_x (P(x,x)) \wedge \\
&\forall_x \forall_y (P(x,y) \to P(y,x)) \wedge \\
&\forall_x \forall_y, \forall_z (P(x,y) \wedge P(y,z) \to P(x,z)) \wedge \\
&\forall_x \forall_y (P(x,y) \to P(f_{ac}(x), f_{ca}(y))) \wedge \\
&\forall_x \forall_y (P(x,y) \to P(f_{ad}(x), f_{da}(y))) \wedge \\
&\forall_x \forall_y (P(x,y) \to P(f_{bc}(x), f_{cb}(y))) \wedge \\
&\forall_x \forall_y (P(x,y) \to P(f_{bd}(x), f_{db}(y))) \wedge \\
&\forall_x \forall_y (P(x,y) \to P(f_{ce}(x), f_{eca}(y))) \wedge \\
&\forall_x \forall_y (P(x,y) \to P(f_{de}(x), f_{edb}(y))) \wedge \\
&\forall_x \forall_y (P(x,y) \to P(f_{cdca}(x), f_{cdcae}(y))) \wedge \\
&\forall_x \forall_y (P(x,y) \to P(f_{a}(x), f_{a}(y))) \wedge \\
&\forall_x \forall_y (P(x,y) \to P(f_{b}(x), f_{b}(y))) \wedge \\
&\forall_x \forall_y (P(x,y) \to P(f_{c}(x), f_{c}(y))) \wedge \\
&\forall_x \forall_y (P(x,y) \to P(f_{d}(x), f_{d}(y))) \wedge \\
&\forall_x \forall_y (P(x,y) \to P(f_{e}(x), f_{e}(y)))
\end{aligned}
\tag{2.5}
$$

Suppose $\models \varphi_{x,y}$. Our goal is to find a model for $\varphi_{x,y}$ which tells us if there is a solution to the instance $x, y \in \Sigma^*$. Consider the interpretation I with domain Σ^* and such that:

- the constant $\Box$ is interpreted as the empty word;
- each unary function symbol $f_\star$, for $\star \in \Sigma$, is interpreted as the function $f_\star^I : \Sigma^* \to \Sigma^*$ that appends the symbol $\star$ to the word $x \in \Sigma^*$ given as argument, i.e. $f_\star^I(x) = \star x$;
- and the binary predicate P is interpreted as follows:
  $P(x,y)^I$ if and only if there exists a chain, possibly of length zero, of applications of the equations (2.3) that transform x into the word y.

We claim that $I \models \varphi'$. Let us consider each case:

- $I \models \forall x(P(x,x))$: take the empty chain.

- $I \models \forall_x \forall_y (P(x,y) \to P(y,x))$: for any x, y such that $I \models P(x,y)$, take the chain given for $P(x,y)$ in reverse order.

- $I \models \forall_x \forall_y \forall_z (P(x,y) \wedge P(y,z) \to P(x,z))$: for any x, y, z such that $I \models P(x,y)$ and $I \models P(y,z)$, append the chains given for $P(x,y)$ and $P(y,z)$.

- $I \models \forall_x \forall_y (P(x,y) \to P(f_{ac}(x), f_{ca}(y)))$: for any x, y such that $I \models P(x,y)$, take the chain given for $P(x,y)$ and use this for the chain of equations for $acx = acy$; then add an application of the equation $ac = ca$ to obtain cay. A similar justification is given for all other cases related with equations (2.3), but the last.

- $I \models \forall_x \forall_y (P(x,y) \to P(f_\star(x), f_\star(y)))$ where $\star \in \Sigma$: for any x, y such that $I \models P(x,y)$, take the chain given for $P(x,y)$ and use it for the chain for the equation $\star x = \star y$.

Since $I \models \varphi_{x,y}$ and $I \models \varphi'$, we conclude that $I \models P(f_x(\Box), f_y(\Box))$. Therefore, the instance x, y of the word problem has a solution.

Conversely, suppose the instance x, y of the word problem has a solution in Tseitin's monoid; i.e., there is a chain of applications of the equations (2.3) from x resulting in the word y as given in the chain (2.2). We will suppose that this chain is of length n.

We need to show that $\varphi_{x,y}$ is valid; i.e., that $\models \varphi_{x,y}$. Let us consider an arbitrary interpretation I' over a domain $\mathtt{D}$ with an element $\Box^{I'}$, five unary functions $f_a^{I'}$, $f_b^{I'}$, $f_c^{I'}$, $f_d^{I'}$, $f_e^{I'}$ and a binary relation $P^{I'}$. Since $\varphi_{x,y}$ is equal to $\varphi' \to P(f_u(\Box), f_v(\Box))$, we have to show that if $I' \models \varphi'$ then $I' \models P(f_u(\Box), f_v(\Box))$.

We proceed by induction in n, the length of the chain of applications of equations (2.3) for transforming x in y.

IB: case $n = 0$, we have that $x \equiv y$ and if $I' \models \varphi'$, $I' \models \forall_x P(x,x)$ which also implies that $I' \models P(x,x)$.

IS: case $n > 0$, the chain of applications of equations to transform x in y is of the form

$$x \equiv x_0 \overset{u_1 = v_1}{=} x_1 \overset{u_2 = v_2}{=} x_2 \overset{u_3 = v_3}{=} \ldots x_{n-1} \overset{u_n = v_n}{=} x_n \equiv y$$

By induction hypothesis we have that $I' \models P(x, x_{n-1})$. If we prove that $I' \models P(x_{n-1}, y)$, we can conclude that $I' \models P(x, y)$, since $I' \models \forall_x \forall_y \forall_z P(x, y) \wedge P(y, z) \rightarrow P(x, z)$ because we are assuming that $I' \models \varphi'$.

Thus, the proof resumes to prove that equalities obtained by one step of application of equations in (2.3) hold in I': in particular if we suppose that $u_n = v_n$ is the equation $u = v$ in (2.3), $x_{n-1} \equiv wuz$ and $y = wvz$, we need to prove that $I' \models P(f_{wuz}(\Box), f_{wvz}(\Box))$, which is done by the following three steps:

1. First, one has that $I' \models P(f_z(\Box), f_z(\Box))$, since $I' \models \forall_x P(x, x)$.
2. Second, since $u = v$ in (2.3), $I' \models \forall_x \forall_y (P(x, y) \rightarrow P(f_u(x), f_v(y)))$. Thus, by the previous item one has that $I' \models P(f_{uz}(\Box), f_{vz}(\Box))$;
3. Third, $I' \models P(f_{wuz}(\Box), f_{wvz}(\Box))$ is obtained from the last item, inductively on the length of w, since $I' \models \forall_x \forall_y (P(x, y) \rightarrow P(f_\star(x), f_\star(y)))$, for all $\star \in \Sigma$.

To conclude the undecidability of validity of the predicate logic, if we suppose the contrary, we will be able to answer for any $x, y \in \Sigma^*$ if $\models P(f_x(\Box), f_y(\Box))$ answering consequently if x equals y in Tseitin's monoid, which is impossible since the word problem in this structure is undecidable.

Theorem 11 (Undecidability of the Predicate Logic)**.** *Validity in the predicate logic, that is answering whether for a given formula φ, $\models \varphi$ is undecidable.*

Notice, that by Gödel completeness theorem undecidability of validity immediately implies undecidability of derivability in the predicate logic. Indeed, in the above reasoning one can use the completeness theorem to alternate between validity and derivability.

Exercise 37. *Accordingly to the three steps above to prove $I' \models P(f_{wuz}(\Box), f_{wvz}(\Box))$, build a derivation for the sequent $\vdash P(f_{wuz}(\Box), f_{wvz}(\Box))$. Concretely, prove that:*

a. $\varphi' \vdash P(f_z(\Box), f_z(\Box))$, *for* $z \in \Sigma^*$;

b. $\varphi', P(f_z(\Box), f_z(\Box)) \vdash P(f_{uz}(\Box), f_{vz}(\Box))$, *for* $u = v$ *in the set of equations (2.3);*

c. $\varphi', P(f_{uz}(\Box), f_{vz}(\Box)) \vdash P(f_{wuz}(\Box), f_{wvz}(\Box))$, *for* $w \in \Sigma^*$;

d. $\varphi' \vdash P(f_{wuz}(\Box), f_{wvz}(\Box))$.