Notes on First Order Logic

Notes for PHIL370

Eric Pacuit*

November 30, 2012

1 The Language of First-Order Logic

The language of predicate logic is constructed from a number of different pieces of syntax: variables, constants, function symbols and predicate symbols. Both function and predicate symbols are associated with an *arity*: the number of arguments that are required by the function or predicate. We start by defining **terms**. Let $\mathcal{V}$ be a finite (or countable) set of **variables** and $\mathcal{C}$ a set of **constants**.

Definition 1.1 (Terms) Let $\mathcal{V}$ be a set of variable, $\mathcal{C}$ a set of constant symbols and $\mathcal{F}$ a set of function symbols. Each function symbol is associated with an **arity** (a positive integer specifying the number of arguments). Write $f^{(n)}$ if the arity of f is n. A term τ is constructed as follows:

- Any variable $x \in \mathcal{V}$ is a term.
- Any constant $c \in \mathcal{C}$ is a term.
- If $f^(n) \in \mathcal{F}$ is a function symbol (i.e., f accepts n arguments) and $\tau_1, \ldots, \tau_n$ are terms, then $f(\tau_1, \ldots, \tau_n)$ is a term.
- Nothing else is a term.

Let $\mathcal{T}$ be the set of terms. ◁

Examples. Let $\mathcal{F} = \{f^{(2)}, g^{(1)}\}$, $\mathcal{C} = \{c, d\}$ and $\mathcal{V} = \{x, y, z\}$. Examples of terms include x, c, $f(x, y)$, $g(c)$, $f(c, d)$, and $g(f(g(c), f(x, f(c, g(z)))))$.

Terms are used to construct atomic formulas:

Definition 1.2 (Atomic Formulas) Let $\mathcal{P}$ be a set of predicate symbols. Each predicate symbol is associated with an arity (the number of objects that are related by P). We write $P^{(n)}$ if the arity of P is n. Suppose that P is an atomic predicate symbol with arity n. If $\tau_1, \ldots, \tau_n$ are terms, then $P(\tau_1, \ldots, \tau_n)$ is an atomic formula. To simplify the notation, we may write $P\tau_1\tau_2 \cdots \tau_n$. ◁

* *Webpage*: ai.stanford.edu/∼epacuit, *Email*: epacuit@umd.edu

Typically, it is assumed that the language contains a special predicate symbol '$=$' whose intended interpretation is *equality*.

Example. Suppose that $\mathcal{P} = \{P^{(1)}, Q^{(3)}\}$ and the function symbols, constants and variables are given in the above example. Examples of atomic formulas include $P(x)$, $P(f(c,x))$, $Q(f(c,y), z, g(f(d,x)))$ and $P(g(f(g(c), f(x, f(c, g(z))))))$.

A **signature** is a description of the constant symbols, functions symbols, predicate symbols (and often the logical constants). We are now ready to define formulas:

Definition 1.3 (Formulas) Formulas are constructed as follows:

- Atomic formulas $P(\tau_1, \ldots, \tau_n)$ are formulas;
- If φ is a formula, then so is $\neg\varphi$;
- If φ and ψ are a formulas, then so is $\varphi \wedge \psi$;
- If φ is a formula, then so is $(\forall x)\varphi$, where x is a variable;
- Nothing else is a formula.

The other boolean connectives $(\vee, \rightarrow, \leftrightarrow)$ are defined as usual. In addition, $(\exists x)\varphi$ is defined as $\neg(\forall x)\neg\varphi$. ◁

Example. Given the signature defined above, examples of formulas include: $P(x) \wedge P(y)$, $(\forall x)P(x)$, $(\forall y)P(x)$, $(\forall x)P(x) \rightarrow P(c)$, and $(\forall x)Q(x,c,y) \rightarrow (P(f(c)) \wedge (\exists z)P(g(z,d,f(c))))$

Definition 1.4 (Free Variable) Suppose that x is a variable. Then, x **occurs free in** φ is defined as follows:

1. If φ is an atomic formula, then x occurs free in φ provided x occurs in φ (i.e., is a symbol in φ).
2. x occurs free in $\neg\psi$ iff x occurs free in ψ
3. x occurs free in $\psi_1 \wedge \psi_2$ iff x occurs free in ψ_1 or x occurs free in ψ_2
4. x occurs free in $(\forall y)\psi$ iff x occurs free in ψ and $x \neq y$
5. x occurs free in $(\exists y)\psi$ iff x occurs free in ψ and $x \neq y$ ◁

The set of free variables in φ, denoted $\mathsf{Fr}(\varphi)$, is defined by recursion as follows:

1. If φ is an atomic formula, then $\mathsf{Fr}(\varphi)$ is the set of all variables (if any) that occur in φ
2. If φ is $\neg\psi$, then $\mathsf{Fr}(\neg\varphi) = \mathsf{Fr}(\varphi)$
3. If φ is $\psi_1 \wedge \psi_2$, then $\mathsf{Fr}(\varphi) = \mathsf{Fr}(\psi_1) \cup \mathsf{Fr}(\psi_2)$
4. If φ is $(\forall x)\psi$, then $\mathsf{Fr}(\psi) = \mathsf{Fr}(\psi)$ after removing x, if present.

A variable x that is not free is said to be **bound**. Formulas that do not contain any free variables are called sentences:

Definition 1.5 (Sentence) If φ is a formula and $\mathsf{Fr}(\varphi) = \emptyset$ (i.e., there are no free variables), then φ is a **sentence**. ◁

1.1 Substitutions

If τ and τ' are terms, we write $\tau[x/\tau']$ for the terms where x is replaced by τ'. We can formally define this operation by recursion:

- $x[x/\tau'] = \tau'$
- $y[x/\tau'] = y$ for $x \neq y$
- $c[x/\tau'] = c$
- $F(\tau_1, \ldots, \tau_n)[x/\tau'] = F(\tau_1[x/\tau'], \ldots, \tau_n[x/\tau'])$

The same notation can be used for formulas $\varphi[x/\tau]$ which means replace all free occurrences of x with τ in a formula φ. This is defined as follows:

- $P(\tau_1, \ldots, \tau_n)[x/\tau] = P(\tau_1[x/\tau], \ldots, \tau_n[x/\tau])$
- $\neg\psi[x/\tau] = \neg(\varphi[x/\tau])$
- $(\psi_1 \wedge \psi_2)[x/\tau] = \psi_1[x/\tau] \wedge \psi_2[x/\tau]$
- $((\forall x)\varphi)[x/\tau] = (\forall x)\varphi$
- $((\forall y)\varphi)[x/\tau] = (\forall y)\varphi[x/\tau]$, where $y \neq x$

The following are key examples of this operation:

1. $(x = y)[y/x]$ is $x = x$ and $(x = y)[x/y]$ is $y = y$,
2. $(\forall x(x = y))[x/y]$ is $(\forall x)x = y$,
3. $(\forall x(x = y))[y/x]$ is $(\forall x)x = x$,
4. $(\forall x)P(x) \to P(x)[x/\tau]$ is $(\forall x)P(x) \to P(\tau)$,
5. $(\forall x)\neg(\forall y)(x = y) \to (\neg\forall y(x = y))[x/y]$ is $(\forall x)\neg(\forall y)(x = y) \to \neg\forall y(y = y)$.

Definition 1.6 (Substitutability) A term τ is **substitutable for x in φ** is defined as follows:

- For an atomic formula φ, τ is always substitutable for x in φ (there are no quantifiers, so t can always be substituted for x)
- τ is substitutable for x in $\neg\psi$ iff τ is is substitutable for x in ψ
- τ is substitutable for x in $\psi_1 \wedge \psi_1$ iff τ is is substitutable for x in ψ_1 and τ is is substitutable for x in ψ_2
- τ is substitutable for x in $(\forall y)\psi$ iff either
    1. x does not occur free in $(\forall y)\psi$
    2. y does not occur in τ and τ is substitutable for x in ψ. $\triangleleft$

2 Models

2.1 Interpreting Terms

Suppose that W is a set. An **interpretation** I (for W) associates with each functions symbol F a function on W of the appropriate arity, denoted F^I, and to each constant c an element of W, denoted c^I. If W is a set and I an interpretation, then for a function symbol F of arity n,

$$F^I : \underbrace{W \times \cdots \times W}_{n \text{ times}} \to W$$

For each constant symbol, c, we have

$$c^I \in W$$

Our goal is to show how to associate with each term and element of a set W. We first need the notion of a substitution:

Definition 2.1 (Substitution) Suppose that W is a nonempty set. A **substitution** is a function $\mathbf{s} : \mathcal{V} \to W$. ◁

Definition 2.2 (Interpretation of Terms) Suppose that I is an interpretation for W and $\mathbf{s} : \mathcal{V} \to W$ is a substitution. We define the function $(I, \mathbf{s}) : \mathcal{T} \to W$ by recursion as follows:

- $(I, \mathbf{s})(x) = \mathbf{s}(x)$
- $(I, \mathbf{s})(c) = c^I$
- $(I, \mathbf{s})(F(\tau_1, \ldots, \tau_n)) = F^I((I, \mathbf{s})(\tau_1), \ldots, (I, \mathbf{s})(\tau_n))$ ◁

Suppose that $\mathbf{s} : \mathcal{V} \to W$ is a substitution. If $a \in W$, we define a new substitution $\mathbf{s}[x/a]$ as follows:

$$\mathbf{s}[x/a](y) = \begin{cases} a & \text{if } y = x \\ \mathbf{s}(x) & \text{otherwise} \end{cases}$$

Suppose that $\mathbf{s} : \mathcal{V} \to W$ and $\mathbf{s}' : \mathcal{V} \to W$ are two substitutions. For each variable $x \in \mathcal{V}$, we define a relation on the set of substitutions as follows:

$$\mathbf{s} \sim_x \mathbf{s}' \text{ iff } \mathbf{s}(y) = \mathbf{s}'(y) \text{ for all } y \neq x$$

Hence, $\mathbf{s} \sim_x \mathbf{s}'$ provided there is some $a \in W$ such that $\mathbf{s}' = \mathbf{s}[x/a]$.

Lemma 2.3 *Suppose that I is an interpretation for W and $\mathbf{s} : \mathcal{V} \to W$ is a substitution. For all terms τ and σ and variables x,*

$$(I, \mathbf{s})(\tau[x/\sigma]) = (I, \mathbf{s}[x/(I, \mathbf{s})(\sigma)])(\tau)$$

Proof. Let W be a nonempty set, I an interpretation for W and $\mathbf{s} : \mathcal{V} \to W$ a substitution. We prove by induction on the structure of τ that for all terms σ, $(I, \mathbf{s})(\tau[x/\sigma]) = (I, \mathbf{s}[x/(I, \mathbf{s})(\sigma)])(\tau)$.

<u>Base Case</u> There are two cases:

1. τ is x. Then, we have

$$\begin{array}{rcl}(I,\mathbf{s})(x[x/\sigma]) & = & (I,\mathbf{s})(\sigma) \\ & = & \mathbf{s}[x/(I,\mathbf{s})(\sigma)](x) \\ & = & (I,\mathbf{s}[x/(I,\mathbf{s})(\sigma)])(x)\end{array}$$

2. τ is $y \neq x$. Then, we have

$$\begin{array}{rcl}(I,\mathbf{s})(y[x/\sigma]) & = & (I,\mathbf{s})(y) \\ & = & s(y) \\ & = & s[x/(I,\mathbf{s})(\sigma)](y) \\ & = & (I,\mathbf{s}[x/(I,\mathbf{s})(\sigma)])(y)\end{array}$$

3. τ is c. Then, we have

$$\begin{array}{rcl}(I,\mathbf{s})(c[x/\sigma]) & = & (I,\mathbf{s})(c) \\ & = & c^I \\ & = & (I,\mathbf{s}[x/(I,\mathbf{s})(\sigma)])(c)\end{array}$$

Induction Step Suppose that the statement holds for $\tau_1,\ldots,\tau_n$ and τ is $F(\tau_1,\ldots,\tau_n)$. Then, we have:

$$\begin{array}{rcl}(I,\mathbf{s})(F(\tau_1,\ldots,\tau_n)[x/\sigma]) & = & (I,\mathbf{s})(F(\tau_1[x/\sigma],\ldots,\tau_n[x/\sigma])) \\ & = & F^I((I,\mathbf{s})(\tau_1[x/\sigma]),\ldots,(I,\mathbf{s})(\tau_n[x/\sigma])) \\ & = & F^I((I,\mathbf{s}[x/(I,\mathbf{s})(\sigma)])(\tau_1),\ldots,(I,\mathbf{s}[x/(I,\mathbf{s})(\sigma)])(\tau_n)) \\ & = & (I,\mathbf{s}[x/(I,\mathbf{s})(\sigma)])(F(\tau_1,\ldots,\tau_n))\end{array}$$

QED

2.2 First Order Models

Definition 2.4 (Model) A model is a pair $\mathfrak{A} = \langle W, I\rangle$ where W is a nonempty set (called the domain) and I is a function (called the interpretation) assigning to each function symbol F, a function denoted F^I, to each constant symbol, an element of W denoted c^I and to each predicate symbol P, a relation on W of the appropriate arity. If P has arity n, then we have

$$P^I \subseteq \underbrace{W \times \cdots \times W}_{n \text{ times}}$$

If $\mathfrak{A}$ is a model, we write $|\mathfrak{A}|$ for the domain of $\mathfrak{A}$, and we write $F^{\mathfrak{A}}$, $c^{\mathfrak{A}}$ and $P^{\mathfrak{A}}$ to denote F^I, c^I and P^I, respectively. $\triangleleft$

We say $\mathbf{s}$ is a substitution for $\mathfrak{A}$ provided $\mathbf{s} : \mathcal{V} \to |\mathfrak{A}|$. Let $\mathfrak{A} = \langle W, I\rangle$ be a model. For each term τ, we write $\tau^{\mathfrak{A},\mathbf{s}}$ for $(I,\mathbf{s})(\tau)$.

Definition 2.5 (Truth) Suppose that $\mathfrak{A}$ is a model and $\mathbf{s}$ is a substitution for $\mathfrak{A}$. The formula φ is true in $\mathfrak{A}$ (given $\mathbf{s}$), denoted $\mathfrak{A},\mathbf{s}\varphi$, is defined by recursion as follows:

- $\mathfrak{A},\mathbf{s} \models P(\tau_1,\ldots,\tau_n)$ iff $(\tau_1^{\mathfrak{A},\mathbf{s}},\ldots,\tau_n^{\mathfrak{A},\mathbf{s}}) \in P^{\mathfrak{A}}$

- $\mathfrak{A}, \mathbf{s} \models \neg\psi$ iff $\mathfrak{A}, \mathbf{s} \not\models \psi$
- $\mathfrak{A}, \mathbf{s} \models \psi_1 \wedge \psi_2$ iff $\mathfrak{A}, \mathbf{s} \models \psi_1$ and $\mathfrak{A}, \mathbf{s} \models \psi_2$
- $\mathfrak{A}, \mathbf{s} \models (\forall x)\psi$ iff for all substitutions $\mathbf{s}'$ for $\mathfrak{A}$ if $\mathbf{s} \sim_x \mathbf{s}'$, then $\mathfrak{A}, \mathbf{s}' \models \psi$ ◁

Example 2.6 Suppose that $\mathfrak{A} = \langle D, I\rangle$ where $D = \{1,2,3,4,5\}$, $a^I = 1$, $b^I = 3$, $c^I = 5$, $P^I = \{1,3,5\}$, $Q^I = \{2,4\}$, and $R^I = \{\langle 1,2\rangle, \langle 5,3\rangle, \langle 3,5\rangle\}$. Also suppose $\mathbf{s}$ is the following variable assignment $\mathbf{s}(x) = 2$, $\mathbf{s}(y) = 4$ and $\mathbf{s}(z) = 5$.

- $\mathfrak{A}, \mathbf{s} \models R(z, b)$. This holds because $\langle z^{\mathfrak{A},\mathbf{s}}, b^{\mathfrak{A},\mathbf{s}}\rangle = \langle \mathbf{s}(z), b^I\rangle = \langle 5, 3\rangle \in R^I$. Note that $R(z, b)$ is not a sentence.
- $\mathfrak{A}, \mathbf{s} \models (\exists x)R(x, b)$. This holds because $\mathfrak{A}, \mathbf{s}[x/5] \models R(x, b)$, so there is a $\mathbf{s}'$ such that $\mathbf{s} \sim_x \mathbf{s}'$ and $\mathfrak{A}, \mathbf{s}' \models R(x, b)$. The latter holds because $\langle x^{\mathfrak{A},\mathbf{s}[x/5]}, b^{\mathfrak{A},\mathbf{s}}\rangle = \langle \mathbf{s}[x/5](x), b^I\rangle = \langle 5, 3\rangle \in R^I$.
- $\mathfrak{A}, \mathbf{s} \not\models (\forall x)Px$, since $\mathbf{s}[x/2](x) = 2 \notin P^I$.
- $\mathfrak{A}, \mathbf{s} \not\models (\forall x)Qx$, since $\mathbf{s}[x/1](x) = 1 \notin Q^I$.
- $\mathfrak{A}, \mathbf{s} \not\models (\forall x)Px \vee (\forall x)Qx$, since $\mathfrak{A}, \mathbf{s} \not\models (\forall x)Px$, and $\mathfrak{A}, \mathbf{s} \not\models (\forall x)Qx$.
- $\mathfrak{A}, \mathbf{s} \models (\forall x)(Px \vee Qx)$. This is true since for each element u of D (we need to check all 5), $\mathbf{s}[x/u](x) \in P^I$ or $\mathbf{s}[x/u](x) \in Q^I$.
- $\mathfrak{A}, \mathbf{s} \not\models (\forall x)(Px \vee Qx) \to ((\forall x)Px \vee (\forall x)Qx)$ since $\mathfrak{A}, \mathbf{s} \models (\forall x)(Px \vee Qx)$ but $\mathfrak{A}, \mathbf{s} \not\models ((\forall x)Px \vee (\forall x)Qx)$.
- $\mathfrak{A}, \mathbf{s} \models ((\forall x)Px \vee (\forall x)Qx) \to (\forall x)(Px \vee Qx)$ since $\mathfrak{A}, \mathbf{s} \models (\forall x)(Px \vee Qx)$.

Observation 2.7 *Recall that $(\exists x)\varphi$ is* defined *to be $\neg(\forall x)\neg\varphi$. We can now show that this definition makes sense: Suppose that $\mathfrak{A}$ is a structure and $\mathbf{s}$ a substitution. Then,*

$$\mathfrak{A}, \mathbf{s} \models (\exists x)\psi \textit{ iff there is a } \mathbf{s}' \textit{ for } \mathfrak{A} \textit{ such that } \mathbf{s} \sim_x \mathbf{s}' \textit{ and } \mathfrak{A}, \mathbf{s}' \models \psi$$

Proof. Suppose that $\mathfrak{A}$ is a structure and $\mathbf{s}$ a substitution.

$$\begin{array}{lll} \mathfrak{A}, \mathbf{s} \models \neg(\forall x)\neg\varphi & \text{iff} & \mathfrak{A}, \mathbf{s} \not\models (\forall x)\neg\varphi \\ & \text{iff} & \text{it is not the case that for all } \mathbf{s}', \text{ if } \mathbf{s} \sim_x \mathbf{s}', \text{ then } \mathfrak{A}, \mathbf{s}' \models \neg\varphi \\ & \text{iff} & \text{there is a } \mathbf{s}' \text{ such that } \mathbf{s} \sim_x \mathbf{s}' \text{ and } \mathfrak{A}, \mathbf{s}' \not\models \neg\varphi \\ & \text{iff} & \text{there is a } \mathbf{s}' \text{ such that } \mathbf{s} \sim_x \mathbf{s}' \text{ and } \mathfrak{A}, \mathbf{s}' \models \varphi \end{array}$$

QED

Exercise 2.8 Let $\mathbb{N}$ be the set of natural numbers (i.e., the integers greater than or equal to 0). Consider a first order language with the constant symbol c, two functions symbols f and s and a predicate symbol P. Let $\mathfrak{N}$ be a first order structure where $|\mathfrak{N}| = \mathbb{N}$ and:

$$\begin{array}{rcl} c^{\mathfrak{N}} & = & 0 \\ f^{\mathfrak{N}}(n, m) & = & n + m \\ s^{\mathfrak{N}}(n) & = & n + 1 \\ P^{\mathfrak{N}} & = & \{(n, m) \mid n \leq m\} \end{array}$$

Determine which of the following formulas are true in $\mathfrak{N}$.

- $(\forall x)P(c, x)$
- $(\forall x)P(x, s(x))$
- $(\forall x)(\forall y)f(x, s(y)) = s(f(x, y))$
- $(\forall x)(\exists y)f(y, y) = x$
- $(\forall x)(\forall y)(\exists z)f(y, z) = x$

Definition 2.9 (Semantic Consequence) Let Γ be a set of sentences and φ a formula. If $\mathfrak{A}$ is a model and $\mathbf{s}$ a substitution for $\mathfrak{A}$, then we write $\mathfrak{A}, \mathbf{s} \models \Gamma$ provided $\mathfrak{A}, \mathbf{s} \models \psi$ for each $\psi \in \Gamma$. We say φ is a semantic consequence of Γ, denoted $\Gamma \models \varphi$, provided for all models $\mathfrak{A}$ and substitutions $\mathbf{s}$ for $\mathfrak{A}$, if $\mathfrak{A}, \mathbf{s} \models \Gamma$, then $\mathfrak{A}, \mathbf{s} \models \varphi$. ◁

Proposition 2.10 *Suppose that* $\mathbf{s}$ *and* $\mathbf{s}'$ *agree on all free variables in* φ*. Then,*

$$\mathfrak{A}, \mathbf{s} \models \varphi \quad \textit{iff} \quad \mathfrak{A}, \mathbf{s}' \models \varphi$$

Proof. The proof is by induction on the structure of φ. Suppose that $\mathbf{s}$ and $\mathbf{s}'$ are two substitutions for $\mathfrak{A}$ that agree on all free variables in φ.

<u>Base Case</u> φ is $P(\tau_1, \dots, \tau_n)$. Since $\mathbf{s}$ and $\mathbf{s}'$ agree on all free variables $P(\tau_1, \dots, \tau_n)$, it is easy to see that for each $i = 1, \dots, n$, $\tau_i^{\mathfrak{A},\mathbf{s}} = \tau_i^{\mathfrak{A},\mathbf{s}'}$ (we leave it to the reader to give the full proof by induction).

$$\begin{aligned}
\mathfrak{A}, \mathbf{s} \models P(\tau_1, \dots, \tau_n) \quad & \text{iff} \quad (\tau^{\mathfrak{A},\mathbf{s}}, \dots, \tau_n^{\mathfrak{A},\mathbf{s}}) \in P^{\mathfrak{A}} \\
& \text{(Definition of truth)} \\
& \text{iff} \quad (\tau^{\mathfrak{A},\mathbf{s}'}, \dots, \tau_n^{\mathfrak{A},\mathbf{s}'}) \in P^{\mathfrak{A}} \\
& \text{(since for each } i = 1, \dots, n,\ \tau_i^{\mathfrak{A},\mathbf{s}} = \tau_i^{\mathfrak{A},\mathbf{s}'}) \\
& \text{iff} \quad \mathfrak{A}, \mathbf{s}' \models P(\tau_1, \dots, \tau_n) \\
& \text{(Definition of truth)}
\end{aligned}$$

<u>Induction Step</u> φ is $\neg\psi$.

$$\begin{aligned}
\mathfrak{A}, \mathbf{s} \models \neg\psi \quad & \text{iff} \quad \mathfrak{A}, \mathbf{s} \not\models \psi \\
& \text{(Definition of truth)} \\
& \text{iff} \quad \mathfrak{A}, \mathbf{s}' \not\models \psi \\
& \text{(Induction Hypothesis)} \\
& \text{iff} \quad \mathfrak{A}, \mathbf{s}' \models \neg\psi \\
& \text{(Definition of truth)}
\end{aligned}$$

<u>Induction Step</u> φ is $\psi_1 \wedge \psi_2$.

$$\begin{aligned}
\mathfrak{A}, \mathbf{s} \models \psi_1 \wedge \psi_2 \quad & \text{iff} \quad \mathfrak{A}, \mathbf{s} \models \psi_1 \text{ and } \mathfrak{A}, \mathbf{s} \models \psi_2 \\
& \text{(Definition of truth)} \\
& \text{iff} \quad \mathfrak{A}, \mathbf{s}' \models \psi_1 \text{ and } \mathfrak{A}, \mathbf{s}' \models \psi_2 \\
& \text{(Induction Hypothesis)} \\
& \text{iff} \quad \mathfrak{A}, \mathbf{s}' \models \psi_1 \wedge \psi_2 \\
& \text{(Definition of truth)}
\end{aligned}$$

Induction Step φ is $(\forall x)\psi$. Note that since $\mathbf{s}$ and $\mathbf{s}'$ agree on all free variables in φ, if $\mathbf{s} \sim_x \mathbf{u}$ and $\mathbf{s}' \sim_x \mathbf{u}'$, then $\mathbf{u}$ and $\mathbf{u}'$ must agree on all free variables in ψ.

$$\begin{array}{ll} \mathfrak{A}, \mathbf{s} \models (\forall x)\psi & \text{iff} \quad \text{for all } \mathbf{u}, \text{ if } \mathbf{s} \sim_x \mathbf{u}, \text{ then } \mathfrak{A}, \mathbf{u} \models \psi \\ & \text{(Definition of truth)} \\ & \text{iff} \quad \text{for all } \mathbf{u}', \text{ if } \mathbf{s}' \sim_x \mathbf{u}', \text{ then } \mathfrak{A}, \mathbf{u}' \models \psi \\ & \text{(Induction Hypothesis, given the above observation)} \\ & \text{iff} \quad \mathfrak{A}, \mathbf{s}' \models (\forall x)\psi \\ & \text{(Definition of truth)} \end{array}$$

QED

Lemma 2.11 (Substitution Lemma) *Let $\mathfrak{A}$ be a model, $\mathbf{s}$ a substitution for $\mathfrak{A}$ and φ any formula. If τ is substitutable for x in φ, then*

$$\mathfrak{A}, \mathbf{s} \models \varphi[x/\tau] \quad \textit{iff} \quad \mathfrak{A}, \mathbf{s}[x/\tau^{\mathfrak{A},\mathbf{s}}] \models \varphi$$

Proof. The proof is by induction on the structure of φ. Let $\mathfrak{A}$ be a model, $\mathbf{s}$ a substitution for $\mathfrak{A}$ and φ any formula. Suppose that τ is substitutable for x in φ.
Base Case φ is $P(\tau_1, \ldots, \tau_n)$. Then,

$$\begin{array}{ll} \mathfrak{A}, \mathbf{s} \models P(\tau_1, \ldots, \tau_n)[x/\tau] & \text{iff} \quad \mathfrak{A}, \mathbf{s} \models P(\tau_1[x/\tau], \ldots, \tau_n[x/\tau]) \\ & \text{(Definition of term substitution)} \\ & \text{iff} \quad ((\tau_1[x/\tau])^{\mathfrak{A},\mathbf{s}}, \ldots, (\tau_n[x/\tau])^{\mathfrak{A},\mathbf{s}}) \in P^{\mathfrak{A}} \\ & \text{(Definition of truth)} \\ & \text{iff} \quad \left(\tau_1^{\mathfrak{A},\mathbf{s}[x/\tau^{\mathfrak{A},\mathbf{s}}]}, \ldots, \tau_n^{\mathfrak{A},\mathbf{s}[x/\tau^{\mathfrak{A},\mathbf{s}}]}\right) \in P^{\mathfrak{A}} \\ & \text{(by Lemma 2.3)} \\ & \text{iff} \quad \mathfrak{A}, \mathbf{s}[x/\tau^{\mathfrak{A},\mathbf{s}}] \models P(\tau_1, \ldots, \tau_n) \\ & \text{(Definition of truth)} \end{array}$$

Induction Step φ is $\neg\psi$. Then,

$$\begin{array}{ll} \mathfrak{A}, \mathbf{s} \models \neg\psi[x/\tau] & \text{iff} \quad \mathfrak{A}, \mathbf{s} \models \neg(\psi[x/\tau]) \\ & \text{(Definition of term substitution)} \\ & \text{iff} \quad \mathfrak{A}, \mathbf{s} \not\models \psi[x/\tau] \\ & \text{(Definition of truth)} \\ & \text{iff} \quad \mathfrak{A}, \mathbf{s}[x/\tau^{\mathfrak{A},\mathbf{s}}] \not\models \psi \\ & \text{(Induction Hypothesis)} \\ & \text{iff} \quad \mathfrak{A}, \mathbf{s}[x/\tau^{\mathfrak{A},\mathbf{s}}] \models \neg\psi \\ & \text{(Definition of truth)} \end{array}$$

Induction Step φ is $\psi_1 \wedge \psi_2$. Then,

$$\begin{array}{ll} \mathfrak{A}, \mathbf{s} \models (\psi_1 \wedge \psi_2)[x/\tau] & \text{iff} \quad \mathfrak{A}, \mathbf{s} \models \psi_1[x/\tau] \wedge \psi_2[x/\tau] \\ & \text{(Definition of term substitution)} \\ & \text{iff} \quad \mathfrak{A}, \mathbf{s} \models \psi_1[x/\tau] \text{ and } \mathfrak{A}, \mathbf{s} \models \psi_2[x/\tau] \\ & \text{(Definition of truth)} \\ & \text{iff} \quad \mathfrak{A}, \mathbf{s}[x/\tau^{\mathfrak{A},\mathbf{s}}] \models \psi_1 \text{ and } \mathfrak{A}, \mathbf{s}[x/\tau^{\mathfrak{A},\mathbf{s}}] \models \psi_2 \\ & \text{(Induction Hypothesis)} \\ & \text{iff} \quad \mathfrak{A}, \mathbf{s}[x/\tau^{\mathfrak{A},\mathbf{s}}] \models \psi_1 \wedge \psi_2 \\ & \text{(Definition of truth)} \end{array}$$

Induction Step Suppose that φ is $(\forall y)\psi$. Since τ is substitutable for x in φ we have two cases:

1. x does not occur free in ψ. Then $((\forall y)\psi)[x/\tau]$ is the same as $(\forall y)\psi$. Furthermore $\mathbf{s}$ and $\mathbf{s}[x/\tau]$ agree on all free variables in $(\forall y)\psi$. By Theorem **??**, we have $\mathfrak{A}, \mathbf{s} \models (\forall y)\psi[x/\tau]$ iff $\mathfrak{A}, \mathbf{s} \models (\forall y)\psi$ iff $\mathfrak{A}, \mathbf{s}[x/\tau](\forall y)\psi$.

2. x does occur free in ψ. Since τ is substitutable for x in $(\forall y)\psi$, according to the definition of *substitutability*, y does not occur in τ and τ is substitutable for x in ψ. Since y does not occur in τ, for each $a \in |\mathfrak{A}|$, we have

$$\tau^{\mathfrak{A},\mathbf{s}} = \tau^{\mathfrak{A},\mathbf{s}[y/a]}$$

We also have the following:

Claim 2.12 *Suppose that $d \in |\mathfrak{A}|$. Then, the following two statements are equivalent*

(a) For all $\mathbf{s}'$, if $\mathbf{s} \sim_y \mathbf{s}'$, then $\mathfrak{A}, \mathbf{s}'[x/d] \models \psi$.

(b) For all $\mathbf{s}'$, if $\mathbf{s}[x/d] \sim_y \mathbf{s}'$, then $\mathfrak{A}, \mathbf{s}' \models \psi$.

Proof of the claim. Suppose that (a) holds: for all $\mathbf{s}'$, if $\mathbf{s} \sim_y \mathbf{s}'$, then $\mathfrak{A}, \mathbf{s}'[x/d] \models \psi$. To prove (b), suppose that $\mathbf{s}[x/d] \sim_y \mathbf{s}'$. Since $\mathbf{s}'(z) = \mathbf{s}[x/d](z)$ for all $z \neq y$ and $x \neq y$, we have $\mathbf{s}'(x) = \mathbf{s}[x/d](x) = d$. Hence, $\mathbf{s}' = \mathbf{s}'[x/d]$. Furthermore, we have $\mathbf{s} \sim_y \mathbf{s}'[x/\mathbf{s}(x)]$. Hence, $\mathfrak{A}, \mathbf{s}'[x/\mathbf{s}(x)][x/d] \models \psi$. Now, since $\mathbf{s}'[x/\mathbf{s}(x)][x/d] = \mathbf{s}'[x/d]$, we have $\mathfrak{A}, \mathbf{s}'[x/d] \models \psi$. Finally, since $\mathbf{s}'[x/d] = \mathbf{s}'$, we have $\mathfrak{A}, \mathbf{s}' \models \psi$. Suppose that (b) holds: for all $\mathbf{s}'$, if $\mathbf{s}[x/d] \sim_y \mathbf{s}'$, then $\mathfrak{A}, \mathbf{s}' \models \psi$. To prove (a), suppose that $\mathbf{s} \sim_y \mathbf{s}'$. Then, $\mathbf{s}[x/d] \sim_y \mathbf{s}'[x/d]$. Hence, by (b), $\mathfrak{A}, \mathbf{s}'[x/d] \models \psi$, as desired. The completes the proof of the claim.

Returning to the proof of the main claim, since $x \neq y$, $((\forall y)\psi)[x/\tau]$ is $(\forall y)\psi[x/\tau]$. Hence,

$\mathfrak{A}, \mathbf{s} \models ((\forall y)\psi)[x/\tau]$ iff $\mathfrak{A}, \mathbf{s} \models (\forall y)\psi[x/\tau]$
(Definition of term substitution)
iff for all $\mathbf{s}'$, if $\mathbf{s} \sim_y \mathbf{s}'$, then $\mathfrak{A}, \mathbf{s}' \models \psi[x/\tau]$
(Definition of truth)
iff for all $\mathbf{s}'$, if $\mathbf{s} \sim_y \mathbf{s}'$, then $\mathfrak{A}, \mathbf{s}'[x/\tau^{\mathfrak{A},\mathbf{s}'}] \models \psi$
(Induction Hypothesis)
iff for all $\mathbf{s}'$, if $\mathbf{s} \sim_y \mathbf{s}'$, then $\mathfrak{A}, \mathbf{s}'[x/\tau^{\mathfrak{A},\mathbf{s}}] \models \psi$
(Since $\mathbf{s}' = \mathbf{s}[y/a]$ for some $a \in |\mathfrak{A}|$, $\tau^{\mathfrak{A},\mathbf{s}} = \tau^{\mathfrak{A},\mathbf{s}'}$)
iff for all $\mathbf{s}'$, if $\mathbf{s}[x/\tau^{\mathfrak{A},\mathbf{s}}] \sim_y \mathbf{s}'$, then $\mathfrak{A}, \mathbf{s}' \models \psi$
(By Claim 2.12)
iff $\mathfrak{A}, \mathbf{s}[x/\tau^{\mathfrak{A},\mathbf{s}}] \models (\forall y)\psi$
(Definition of truth)

This completes the proof. QED

3 Deductions in First Order Logic

The axiom system for first-order logic consists of the following four axioms:

1. All tautologies
2. $(\forall x)\varphi \to \varphi[x/t]$, where τ is substitutable for x in φ
3. $(\forall x)(\varphi \to \psi) \to ((\forall x)\varphi \to (\forall x)\psi)$
4. $\varphi \to (\forall x)\varphi$, where x does not occur free in φ

Definition 3.1 (Generalization) Given a formula φ, a **generalization of** φ is a formula of the form $(\forall x_1)\cdots(\forall x_n)\varphi$. ◁

Definition 3.2 (Tautology) A tautology (in FOL) is any formula obtained by replacing each atomic proposition with a first-order formula. ◁

Definition 3.3 (Deduction) We write $\Gamma \vdash \varphi$ iff there is a finite sequence of formulas $\varphi_1, \ldots, \varphi_n$ such that $\varphi_n = \varphi$, each φ_i is either a generalization of one of the above axioms for follows from earlier formulas on the list by modus ponens. ◁

Example 1. $\vdash Px \to \exists y Py$. The following is a derivation of this formula:

1.	$(\forall y \neg Py \to \neg Px) \to (Px \to \neg\forall y \neg Py)$	(Instance of a tautology)
2.	$\forall y \neg Py \to \neg Px$	(Instance of Axiom 2)
3.	$Px \to \neg\forall y \neg Py$	(MP 1 & 2)
4.	$Px \to \exists y Py$	(Definition of $\exists$)

We can extend this derivation to show that $\vdash (\forall x)(Px \to (\exists y)Py)$. In fact, we have the following meta-theorem:

Theorem 3.4 (Generalization Theroem) *If* $\Gamma \vdash \varphi$ *and* x *does not occur free in any formula in* Γ*, then* $\Gamma \vdash (\forall x)\varphi$

Proof. The proof is by induction on the length of a derivation. Suppose that $\Gamma \vdash \varphi$ and x does not occur free in any formula in Γ.

Base Case φ is a generalization of an axiom. Then, $(\forall x)\varphi$ is also a generalization of an axiom.

Suppose that $\varphi \in \Gamma$. Then, since x does not occur free in φ, $\varphi \to (\forall x)\varphi$ is an axiom. Hence, the following is a derivation of $(\forall x)\varphi$:

1.	φ	element of Γ
2.	$\varphi \to (\forall x)\varphi$	Axiom 3
3.	$(\forall x)\varphi$	MP from 1 & 2

Induction Hypothesis Suppose that φ follows from earlier formulas by Modus Ponens:

1.	ψ	$\Gamma \vdash \psi$
2.	$(\forall x)\psi$	IH
3.	$\psi \to \varphi$	$\Gamma \vdash \psi \to \varphi$
4.	$(\forall x)(\psi \to \varphi)$	IH
5.	$(\forall x)(\psi \to \varphi) \to ((\forall x)\psi \to (\forall x)\varphi)$	Axiom 3
6.	$(\forall x)\psi \to (\forall x)\varphi$	MP from 4 & 5
7.	$(\forall x)\varphi$	MP from 2 & 6

QED

We also have the following meta-theorems (the reader is invited to verify each of these theorems):

- Tautology Rule 1: $\Gamma \vdash \varphi$ iff $\Gamma \cup \Lambda$ tautologically implies φ (Λ is the set of all generalizations of the above axioms).
- Tautology Rule 2: If $\Gamma \vdash \varphi_1$, ..., $\Gamma \vdash \varphi_n$ and $\{\varphi_1, \ldots, \varphi_n\}$ tautologically implies ψ, then $\Gamma \vdash \psi$.
- Deduction Theorem: $\Gamma \cup \{\alpha\} \vdash \beta$ iff $\Gamma \vdash \alpha \to \beta$
- Contraposition: $\Gamma \cup \{\alpha\} \vdash \neg\beta$ iff $\Gamma, \beta \vdash \neg\alpha$
- Reductio Ad Absurdum: If $\Gamma \cup \{\varphi\}$ is inconsistent, then $\Gamma \vdash \neg\varphi$

Example 2. $\vdash \exists x \forall y \varphi \to \forall y \exists x$. To prove this, it suffices to show that $\{\exists x \forall y \varphi\} \vdash \forall y \exists x \varphi$, by the deduction theorem. It suffices to show that $\{\exists x \forall y \varphi\} \vdash \exists x \varphi$, by the Generalization Theorem. This is equivalent to showing $\{\neg \forall x \neg \forall y \varphi\} \vdash \neg \forall x \neg \varphi$. It suffices to show that $\{\forall x \neg \varphi\} \vdash \forall x \neg \forall y \varphi$, by Contraposition. It suffices to show that $\{\forall x \neg \varphi\} \vdash \neg \forall y \varphi$, by the Generalization Theorem. It suffices to show that $\{\forall x \neg \varphi, \forall y \varphi\}$ is inconsistent, by Reductio Ad Absurdum. The following deduction shows that $\{\forall x \neg \varphi, \forall y \varphi\} \vdash \bot$:

1.	$\forall x \neg \varphi \to \neg \varphi[x/x]$	(Instance of Axiom 2)
2.	$\forall x \neg \varphi$	(Assumption)
3.	$\neg \varphi$	(MP 1 & 2 ($\varphi[x/x]$ is φ))
4.	$\forall y \varphi \to \varphi[y/y]$	(Instance of Axiom 2)
5.	$\forall y \varphi$	(Assumption)
6.	φ	(MP 4 & 5 ($\varphi[y/y]$ is φ))
7.	$(\varphi \wedge \neg \varphi) \to \bot$	(Tautology)
8.	$\bot$	(MP 6 & 7)

Example 3. Prove that $\vdash \exists x(\alpha \wedge \beta) \to \exists x \alpha \wedge \exists x \beta$.

Proof. To prove $\vdash \exists x(\alpha \wedge \beta) \to \exists x \alpha \wedge \exists x \beta$ it is enough to show $\{\exists x(\alpha \wedge \beta)\} \vdash \exists x \alpha \wedge \exists x \beta$ by the Deduction Theorem. For this it is enough to show, $\{\exists x(\alpha \wedge \beta)\} \vdash \exists x \alpha$ and $\{\exists x(\alpha \wedge \beta)\} \vdash \exists x \beta$ by Tautology Rule 2 (noting that $\{\exists x \alpha, \exists x \beta\}$ tautologically implies $\exists x \alpha \wedge \exists x \beta$). To show $\{\exists x(\alpha \wedge \beta)\} \vdash \exists x \alpha$, we must show $\{\neg \forall x \neg(\alpha \wedge \beta)\} \vdash \neg \forall x \neg \alpha$ by the definition of '$\exists$'. For this, it is enough to show $\{\forall x \neg \alpha\} \vdash \forall x \neg(\alpha \wedge \beta)$ by Contraposition. For this, it is enough to show $\{\forall x \neg \alpha\} \vdash \neg(\alpha \wedge \beta)$ by Generalization. For this, it is enough to show $\{\forall x \neg \alpha\} \vdash \neg \alpha$ by Tautology Rule 2 since $\{\neg \alpha\}$ tautologically implies $\neg(\alpha \wedge \beta)$. Using an instance of axiom 2 and Modus Ponens, we can directly show that $\{\forall x \neg \alpha\} \vdash \neg \alpha$. The proof that $\{\exists x(\alpha \wedge \beta)\} \vdash \exists x \beta$ is similar. QED

This can be turned into a complete derivation as follows:

1.	$\forall x(\neg\alpha \to \neg(\alpha \wedge \beta))$	Instance of Axiom 1
2.	$\forall x(\neg\alpha \to \neg(\alpha \wedge \beta)) \to (\forall x\neg\alpha \to \forall x\neg(\alpha \wedge \beta))$	Instance of Axiom 3
3.	$\forall x\neg\alpha \to \forall x\neg(\alpha \wedge \beta)$	MP 1,2
4.	$(\forall x\neg\alpha \to \forall x\neg(\alpha \wedge \beta)) \to (\neg\forall x\neg(\alpha \wedge \beta) \to \neg\forall x\neg\alpha)$	Instance of Axiom 1
5.	$\neg\forall x\neg(\alpha \wedge \beta) \to \neg\forall x\neg\alpha$	MP 3,4
6.	$\exists x(\alpha \wedge \beta) \to \exists x\alpha$	Definition of '$\exists$'
7.	$\forall x(\neg\beta \to \neg(\alpha \wedge \beta))$	Instance of Axiom 1
8.	$\forall x(\neg\beta \to \neg(\alpha \wedge \beta)) \to (\forall x\neg\beta \to \forall x\neg(\alpha \wedge \beta))$	Instance of Axiom 3
9.	$\forall x\neg\beta \to \forall x\neg(\alpha \wedge \beta)$	MP 7,8
10.	$(\forall x\neg\beta \to \forall x\neg(\alpha \wedge \beta)) \to (\neg\forall x\neg(\alpha \wedge \beta) \to \neg\forall x\neg\beta)$	Instance of Axiom 1
11.	$\neg\forall x\neg(\alpha \wedge \beta) \to \neg\forall x\neg\beta$	MP 9,10
12.	$\exists x(\alpha \wedge \beta) \to \exists x\beta$	Definition of '$\exists$'
13.	$(\exists x(\alpha \wedge \beta) \to \exists x\alpha) \to ((\exists x(\alpha \wedge \beta) \to \exists x\beta)$ $\to (\exists x(\alpha \wedge \beta) \to (\exists x\alpha \wedge \exists x\beta)))$	Instance of Axiom 1
14.	$(\exists x(\alpha \wedge \beta) \to \exists x\beta) \to (\exists x(\alpha \wedge \beta) \to (\exists x\alpha \wedge \exists x\beta))$	MP 6,13
15.	$\exists x(\alpha \wedge \beta) \to (\exists x\alpha \wedge \exists x\beta)$	MP 12, 14

Theorem 3.5 (Generalization on Constants) *Suppose that $\Gamma \vdash \varphi$ and that c is a constant symbol that does not occur in Γ. Then there is a variable y (which does not occur in φ such that $\Gamma \vdash \forall y\varphi[c/y]$ (where $\varphi[c/y]$ is φ with c replaced by y). Furthermore, there is a deduction of $\forall y\varphi[c/y]$ from Γ in which c does not occur.*

Theorem 3.6 (Existence of Alphabetic Variants) *Suppose that φ is a formula, x a variable and τ a term. Then, we can find a formula φ' (that differs from φ only in the choice of quantified variables) such that*

1. *$\varphi \vdash \varphi'$ and $\varphi' \vdash \varphi$.*
2. *τ is substitutable for x in φ'*

4 Soundness and Completeness

Theorem 4.1 (Soundness) *For all sets of formulas Γ, $\Gamma \vdash \varphi$ implies $\Gamma \models \varphi$.*

Proof. We show that for all derivations from Γ, if φ is the last formula of the derivation, then $\Gamma \models \varphi$. The proof is by induction on the length of a derivation.
Base Case The base case is a derivation of length 1. In this case, the last formula of the sequence φ must either be an element of Γ or a generalization of an axiom. If $\varphi \in \Gamma$, then it is obvious that $\Gamma \models \varphi$. For the second case, we first show that each of the axioms are true in any model and substitution.

Claim 4.2 *Suppose that τ is substitutable for x in φ. Then, $\forall x\varphi \to \varphi[x/\tau]$ is valid.*

Proof of Claim. Suppose that τ is substitutable for x in φ, and that $\mathfrak{A}$ is a model and $\mathbf{s}$ is a substitution for $\mathfrak{A}$. Then, we have:

$$
\begin{array}{rcll}
\mathfrak{A}, \mathbf{s} \models \forall x \varphi & \text{iff} & \text{for all substitutions } \mathbf{s}', \text{ if } \mathbf{s} \sim_x \mathbf{s}' \text{ then } \mathfrak{A}, \mathbf{s}' \models \varphi & \text{(Definition of Truth)} \\
& \text{implies} & \mathfrak{A}, \mathbf{s}[x/\tau^{\mathfrak{A}}] \models \varphi & (\text{Since } \mathbf{s} \sim_x \mathbf{s}[x/\tau^{\mathfrak{A},\mathbf{s}}]) \\
& \text{iff} & \mathfrak{A}, \mathbf{s} \models \varphi[x/\tau] & \text{(Substitution Lemma)}
\end{array}
$$

QED (of Claim)

Claim 4.3 *Suppose that x does not occur in φ. Then, $\varphi \to \forall x\varphi$ is valid.*

Proof of Claim. Suppose that x does not occur in φ, and that $\mathfrak{A}$ is a model and $\mathbf{s}$ is a substitution for $\mathfrak{A}$. Suppose that $\mathfrak{A}, \mathbf{s} \models \varphi$. We will show that $\mathfrak{A}, \mathbf{s} \models \forall x\varphi$. Let $\mathbf{s}'$ be a substitution for $\mathfrak{A}$ with $\mathbf{s} \sim_x \mathbf{s}'$. Hence, $\mathbf{s}'(y) = \mathbf{s}(y)$ for all $y \neq x$; and so, $\mathbf{s}$ and $\mathbf{s}'$ agree on all free variables that occur in φ. Therefore, by Theorem 2.10, $\mathfrak{A}, \mathbf{s} \models \varphi$ iff $\mathfrak{A}, \mathbf{s}' \models \varphi$. QED (of Claim)

Claim 4.4 $\forall x(\varphi \to \psi) \to (\forall x\varphi \to \forall x\psi)$ *is valid.*

Proof of Claim. Suppose that $\mathfrak{A}$ is a model and $\mathbf{s}$ is a substitution for $\mathfrak{A}$. Suppose that $\mathfrak{A}, \mathbf{s} \models \forall x(\varphi \to \psi)$ and $\mathfrak{A}, \mathbf{s} \models \forall x\varphi$. We will show that $\mathfrak{A}, \mathbf{s} \models \forall x\psi$. Let $\mathbf{s}'$ be a substitution for $\mathfrak{A}$ such that $\mathbf{s} \sim_x \mathbf{s}'$. Then, $\mathfrak{A}, \mathbf{s}' \models \varphi \to \psi$ and $\mathfrak{A}, \mathbf{s}' \models \varphi$. Hence, by propositional reasoning, $\mathfrak{A}, \mathbf{s}' \models \psi$. Therefore, $\mathfrak{A}, \mathbf{s} \models \forall x\psi$. QED (of Claim)

Since the substitutions were arbitrary in the above proofs, it is not hard to see that any generalization of the above axioms will also be true.

Induction Hypothesis Suppose that $\alpha_1, \ldots, \alpha_k$ is a derivation of length k where there are $i, j < k$ such that $\alpha_j = \alpha_i \to \alpha_k$. Since $\alpha_1, \ldots, \alpha_i$ and $\alpha_1, \ldots, \alpha_j$ are derivations of length less than k, by the induction hypothesis, we have $\Gamma \models \alpha_i$ and $\Gamma \models \alpha_i \to \alpha_k$. Let $\mathfrak{A}$ be any model and $\mathbf{s}$ a substitution for $\mathfrak{A}$ such that $\mathfrak{A}, \mathbf{s} \models \Gamma$. Then $\mathfrak{A}, \mathbf{s} \models \alpha_i$ and $\mathfrak{A}, \mathbf{s} \models \alpha_i \to \alpha_k$. It is straightforward to check that $\mathfrak{A}, \mathbf{s} \models \alpha_k$, as desired. QED

Definition 4.5 (Consistent Set) A set of formulas Γ is **inconsistent** provided $\Gamma \vdash \bot$ (where $\bot$ is a formula of the form $Px \wedge \neg Px$ with P a symbol in the language[1]). A set of formulas Γ is **consistent** if it is not inconsistent. ◁

Definition 4.6 (Maximally Consistent Set) A set of formulas Γ is **maximally consistent** provided it is consistent and if $\Gamma \subsetneq \Gamma'$, then Γ' is inconsistent. ◁

Recall the following properties of maximally consistent sets (the proofs are the same as in the propositional case):

- For all formulas φ, either $\varphi \in \Gamma$ or $\neg\varphi \in \Gamma$, but not both.
- $\neg\varphi \in \Gamma$ iff $\varphi \notin \Gamma$
- $\varphi \wedge \psi \in \Gamma$ iff $\varphi \in \Gamma$ and $\psi \in \Gamma$.
- If $\varphi \to \psi \in \Gamma$ and $\varphi \in \Gamma$, then $\psi \in \Gamma$

[1]Note that we assume the language contains at least one predicate symbol. Sometimes the only predicate symbol in the language is the equality symbol =.

- If $\Gamma \vdash \varphi$ then $\varphi \in \Gamma$ (hence, all generalizations of axioms are elements of Γ)

Lemma 4.7 (Lindenbaum's Lemma) *Suppose that Γ is a consistent set. There is a maximally consistent set Γ' such that $\Gamma \subseteq \Gamma'$.*

The proof is the same as in the propositional case.

We say that a set of formulas Γ is **satisfiable** provided there is a model $\mathfrak{A}$ and substitution $\mathbf{s}$ such that $\mathfrak{A}, \mathbf{s} \models \psi$ for each $\psi \in \Gamma$.

Lemma 4.8 *The following statements are equivalent:*

1. *If $\Gamma \models \varphi$ then $\Gamma \vdash \varphi$*
2. *Any consistent set of formulas is satisfiable*

Proof. (a) $\Rightarrow$ (b): Suppose that for all sets of formulas Γ and formulas φ, if $\Gamma \models \varphi$ then $\Gamma \vdash \varphi$. Suppose that Δ is a consistent set of formulas that is not satisfiable. Then, since there are no models of Δ, we have $\Delta \models \bot$. By (a), $\Delta \vdash \bot$. Since $\bot \to \neg(\forall x\varphi \to \varphi[x/x])$ is a tautology, we have $\Delta \vdash \bot \to \neg(\forall x\varphi \to \varphi[x/x])$. Hence, by modus ponens, $\Delta \vdash \neg(\forall x\varphi \to \varphi[x/x])$. But $\forall x\varphi \to \varphi[x/x]$ is an instance of an axiom, so $\Delta \vdash \forall x\varphi \to \varphi[x/x]$. Therefore, Δ is not consistent, a contradiction. Therefore, Δ is satisfiable.

(b) $\Rightarrow$ (a): Suppose that any consistent set is satisfiable. We will show $\Gamma \not\vdash \varphi$ implies $\Gamma \not\models \varphi$. Suppose that $\Gamma \not\vdash \varphi$. Then $\Gamma \cup \{\neg\varphi\}$ is consistent. Hence, by (b), $\Gamma \cup \{\neg\varphi\}$ is satisfiable. That is, there is a model $\mathfrak{A}$ and substitution $\mathbf{s}$ such that for all $\alpha \in \Gamma$, $\mathfrak{A}, \mathbf{s} \models \alpha$ and $\mathfrak{A}, \mathbf{s} \models \neg\varphi$ (which implies $\mathfrak{A}, \mathbf{s} \not\models \varphi$). Therefore, $\Gamma \not\models \varphi$. QED

We say that a first-order language is **countable** if it contains countably many constant, function and predicate symbols.

Theorem 4.9 (Gödel's Completeness Theorem) *Suppose that Γ is a set of formulas in a countable language. Then, for all formulas φ, if $\Gamma \models \varphi$, then $\Gamma \vdash \varphi$.*

Proof. Let Γ be a consistent set of formulas in a countable language $\mathcal{L}$. We will show that Γ is satisfiable. Suppose that $\mathcal{L}'$ extends $\mathcal{L}$ with countable many *new* constant symbols. I.e., $\mathcal{L}'$ contains the same function and predicate symbols as $\mathcal{L}$ and if $\mathcal{C}$ is the set of constant symbols in $\mathcal{L}$, then the constant symbols in $\mathcal{L}'$ is $\mathcal{C}' = \mathcal{C} \cup \{c_1, c_2, \ldots\}$ where for each $i = 1, \ldots$, $c_i \notin \mathcal{C}$.

Claim 4.10 *Γ is a consistent set of formulas in the new language $\mathcal{L}'$.*

Proof of Claim. Suppose not. Then $\Gamma \vdash \alpha \wedge \neg\alpha$ for some α in $\mathcal{L}'$. This deduction contains only finitely many new constant symbols. By Generalization on constants (Theorem 3.5), each constant can be replaced by a variable. This means we have a derivation of $\alpha' \wedge \neg\alpha'$ in the language $\mathcal{L}$, which contradicts the assumption that Γ is consistent. QED (of Claim)

For each formula φ in $\mathcal{L}'$ and each variable x, add

$$\neg\forall x\varphi \to \varphi[x/c] \in \Gamma'$$

where c is a new constant symbol.

Fix a listing of all pairs of formulas (in $\mathcal{L}'$) and variables x:

$$\langle\varphi_1, x_1\rangle, \langle\varphi_2, x_2\rangle, \ldots$$

Let δ_1 be the formula $\neg\forall x_1\varphi_1 \to \neg\varphi_1[x_1/c_1]$ where c_1 is a new constant symbol no occurring in φ_1. For each $n = 1, \ldots$, let δ_n be the formula $\neg\forall x_n\varphi_1 \to \neg\varphi_n[x_n/c_n]$ where c_n is the first of the new constant symbols not occurring in δ_k for any $k < n$ or φ_n. Let $\Delta = \{\delta_1, \delta_2, \ldots\}$.

Claim 4.11 *The set* $\Gamma \cup \Delta$ *is consistent.*

Proof of Claim. Suppose not. Then, since deductions are finite, we must have $\Gamma\cup\{\delta_1, \ldots, \delta_m, \delta_{m+1}\}$ is inconsistent for some $m \geq 0$. Let m^* be the least such m. Then, by Reductio Ad Absurdum,

$$\Gamma \cup \{\delta_1, \ldots, \delta_{m^*}\} \vdash \neg\delta_{m^*+1}$$

Now, δ_{m^*+1} is the formula $\neg\forall x\varphi \to \neg\varphi[x/c]$ for some x, c and φ. Hence, we have

$$\Gamma \cup \{\delta_1, \ldots, \delta_{m^*}\} \vdash \neg(\neg\forall x\varphi \to \neg\varphi[x/c])$$

And so (using the Tautology Rule 1),

$$\Gamma \cup \{\delta_1, \ldots, \delta_{m^*}\} \vdash \neg\forall x\varphi \qquad \text{and} \qquad \Gamma \cup \{\delta_1, \ldots, \delta_{m^*}\} \vdash \varphi[x/c]$$

Since c is a constant that does not appear in any formula in $\Gamma \cup \{\delta_1, \ldots, \delta_{m^*}\}$, by Generalization on Constants, we have

$$\Gamma \cup \{\delta_1, \ldots, \delta_{m^*}\} \vdash \forall x\varphi$$

Hence $\Gamma \cup \{\delta_1, \ldots, \delta_{m^*}\}$ is inconsistent. If $m^* > 0$, then this contradicts the assumption that m^* is the *least* m such that $\Gamma \cup \{\delta_1, \ldots, \delta_m, \delta_{m+1}\}$ is inconsistent. If $m^* = 0$, then this contradicts the assumption that Γ is consistent. QED (of Claim)

Since $\Gamma \cup \Delta$ is consistent, by Lindenbaum's Lemma, we can find a **maximally consistent set** Γ' such that $\Gamma \cup \Delta \subseteq \Gamma'$.

To conclude the proof, we construct a model for Γ'.

Definition 4.12 (Canonical Model) Let Γ be a maximally consistent set in the language $\mathcal{L}$. The **canonical model for** Γ is the structure $\mathfrak{C}^\Gamma = \langle W, I\rangle$ where

- W is the set of terms in $\mathcal{L}$, i.e., $W = \{\tau \mid \tau \text{ is a term in } \mathcal{L}\}$
- For each constant symbol c, $c^I = c$
- For each function symbols f, $f^I = f$
- For each predicate symbol P, we have $(\tau_1, \ldots, \tau_n) \in P^I$ iff $P(\tau_1, \ldots, \tau_n) \in \Gamma$.

The canonical substitution $\mathbf{s}^C : \mathcal{V} \to W$ is the identity map $\mathbf{s}^C(x) = x$. When Γ is clear from context, we write $\mathfrak{C}$ instead of $\mathfrak{C}^\Gamma$. $\triangleleft$

First, note the following straightforward fact:

Fact 4.13 *For all terms τ, $\tau^{\mathfrak{C},\mathbf{s}^{\mathbf{C}}} = \tau$.*

Lemma 4.14 (Truth Lemma) *Suppose that Γ is a maximally consistent set. Then, for all formulas φ, $\varphi \in \Gamma$ iff $\mathfrak{C}^\Gamma, \mathbf{s}^C \models \varphi$.*

Proof of Truth Lemma. Suppose that Γ is a maximally consistent set and let $\mathfrak{C}$ be the canonical model for Γ and $\mathbf{s}$ the canonical substitution. We prove by induction on the structure of φ that $\varphi \in \Gamma$ iff $\mathfrak{C}, \mathbf{s} \models \varphi$.

Base Case φ is $P(\tau_1, \ldots, \tau_n)$ where $\tau_1, \ldots, \tau_n$ are terms. We have,

$$\begin{array}{rlll} \mathfrak{C}, s \models P(\tau_1, \ldots, \tau_n) & \text{iff} & (\tau_1^s, \ldots, \tau_n^s) \in P^{\mathfrak{C}} & \text{(Definition of truth)} \\ & \text{iff} & (\tau_1, \ldots, \tau_n) \in P^{\mathfrak{C}} & \text{(Fact 4.13)} \\ & \text{iff} & P(\tau_1, \ldots, \tau_n) \in \Gamma & \text{(Def. of Canonical Model)} \end{array}$$

Induction Step There are three cases.

Case 1. φ is $\neg\psi$. We have,

$$\begin{array}{rlll} \mathfrak{C}, s \models \neg\psi & \text{iff} & \mathfrak{C}, s \not\models \psi & \text{(Definition of truth)} \\ & \text{iff} & \psi \notin \Gamma & \text{(IH)} \\ & \text{iff} & \neg\psi \in \Gamma & \text{(Properties of maximally consistent sets)} \end{array}$$

Case 2. φ is $\psi_1 \wedge \psi_2$. We have,

$$\begin{array}{rlll} \mathfrak{C}, s \models \psi_1 \wedge \psi_2 & \text{iff} & \mathfrak{C}, s \models \psi_1 \text{ and } \mathfrak{C}, s \models \psi_2 & \text{(Definition of truth)} \\ & \text{iff} & \psi_1 \in \Gamma \text{ and } \psi_2 \in \Gamma & \text{(IH)} \\ & \text{iff} & \psi_1 \wedge \psi_2 \in \Gamma & \text{(Properties of maximally consistent sets)} \end{array}$$

Case 3. φ is $\forall x\psi$. We show that if $\forall x\psi \notin \Gamma$, then $\mathfrak{C}, \mathbf{s} \not\models \forall x\psi$ and if $\mathfrak{C}, \mathbf{s} \not\models \forall x\psi$, then $\forall x\psi \notin \Gamma$. The two statements are handled differently:

$$\begin{array}{rlll} (\forall x)\psi \notin \Gamma' & \text{iff} & \neg\forall x\psi \in \Gamma' & \text{(Since } \Gamma' \text{ is a maximally consistent set)} \\ & \text{implies} & \neg\psi[x/c] \in \Gamma' & \text{(Since } \neg\forall x\psi \to \neg\psi[x/c] \in \Gamma'\text{)} \\ & \text{iff} & \psi[x/c] \notin \Gamma' & \text{(Since } \Gamma' \text{ is a maximally consistent set)} \\ & \text{iff} & \mathfrak{C}, \mathbf{s} \not\models \psi[x/c] & \text{(Induction Hypothesis)} \\ & \text{iff} & \mathfrak{C}, \mathbf{s}[x/c^{\mathfrak{A},\mathbf{s}}] \not\models \psi & \text{(Substitution Lemma)} \\ & \text{iff} & \mathfrak{C}, \mathbf{s} \not\models \forall x\psi & \text{(Since } \mathbf{s} \sim_x \mathbf{s}[x/c^{\mathfrak{A},\mathbf{s}}]\text{)} \end{array}$$

$$
\begin{array}{lll}
\mathfrak{C}, \mathbf{s} \not\models (\forall x)\psi & \text{iff} & \text{there is a } \mathbf{s}' \text{ such that } \mathbf{s} \sim_x \mathbf{s}' \text{ and } \mathfrak{C}, \mathfrak{s}' \not\models \psi \\
& & \text{(Definition of truth)} \\
& \text{iff} & \text{there is a term } \tau \text{ such that } \mathfrak{C}, \mathfrak{s}[x/\tau] \not\models \psi \\
& & \text{(Since } |\mathfrak{C}| \text{ is the set of terms \& the definition of } \sim_x) \\
& \text{iff} & \text{there is a term } \tau \text{ such that } \mathfrak{C}, \mathfrak{s}[x/\tau] \not\models \psi' \\
& & \text{(where } \psi' \text{ is an alphabetic variant in which } \tau \text{ is substitutable for } x) \\
& \text{iff} & \text{there is a term } \tau \text{ such that } \mathfrak{C}, \mathfrak{s} \not\models \psi'[x/\tau] \\
& & \text{(by the Substitution Theorem)} \\
& \text{iff} & \text{there is a term } \tau \text{ such that } \psi'[x/\tau] \notin \Gamma' \\
& & \text{(by the Induction Hypothesis)} \\
& \text{iff} & (\forall x)\psi' \notin \Gamma' \\
& & \text{(Since } \forall x \psi' \to \psi'[x/\tau] \in \Gamma') \\
& \text{iff} & (\forall x)\psi \notin \Gamma' \\
& & \text{(By the equivalence of alphabetic variants)}
\end{array}
$$

This completes the proof of the Truth Lemma. QED (of Lemma)

Returning to the main proof, since Γ' is a maximally consistent set, it is satisfiable in the canonical model $\mathfrak{C}^{\Gamma'}$ with respect to the substitution $\mathbf{s}^{\mathbf{C}}$. This is a model in the extended language $\mathcal{L}'$. To obtain a model in the original language $\mathcal{L}$, restrict $\mathfrak{C}^{\Gamma'}$ to the symbols in $\mathcal{L}$. This completes the proof. QED