MATH 406/PHIL 306: MATHEMATICAL LOGIC AND MODEL THEORY NOTES

STEPHEN LIU

Abstract. Basically, the aim of the course is to prove Gödel's First Incompleteness Theorem ($G1$) from first principles. We will use the book "The Incompleteness Phenomenon" by Martin Goldstern and Haim Judah.

Contents

1. Peano Arithmetic - Introduction 1
2. Inductive Structures 1
3. Sentential Logic 3
4. First Order Logic 5
5. Completeness 11
6. The Incompleteness Theorem 17

1. Peano Arithmetic - Introduction

The weak version of $G1$ that we will be studying is a statement about Peano Arithmetic (PA) , however, we first need to understand what PA is in orer to prove statements about it as a whole. PA is a set of axioms that really articulate what we mean when we talk about arithmetic on the natural numbers. The axioms, followed by a more detailed discussion, are listed below:

(1) $\forall x \neg(sx = 0)$
(2) $\forall x, y(sx = sy \to x = y)$
(3) $\forall x, (0 + x = x)$
(4) $\forall x, y[sy + x = s(y + x)]$
(5) $\forall x(0x = 0)$
(6) $\forall x, y[syx = (yx) + x]$
(7) $\forall x(x^0 = s0)$
(8) $\forall x, y(x^{sy} = x^y x)$
(9) Induction Scheme: For any first order formula $\phi(x)$, the following is an axiom: $[\phi(0) \wedge \forall x(\phi(x) \to \phi(sx))] \to \forall x(\phi(x))$.

Note that usually we omit the quantifiers and implicitly understand that the x, y in the axioms mean any arbitrary x, y. We are given the objects 0 and s which is the unary successor function, and with these two objects and the axioms above we "define" arithmetic. Axioms $P3$ and $P4$ give us addition, $P5$ and $P6$ give us multiplication, $P7$ and $P8$ give us exponentiation, but the last axiom, the induction scheme is the most powerful because it allows us to prove general statements about arithmetic.

2. Inductive Structures

Mathematical induction as we usually understand it is that if the property is true for 0, and the property being true for n implies the property is true for sn, then the property is preserved for all n. However, if we think about this this is really induction over the specific structure of the natural numbers $\mathbb{N}$. We can generalize this so we can perform induction over other structures with the notion of an inductive structure. An inductive structure is made

Date: Spring 2017.

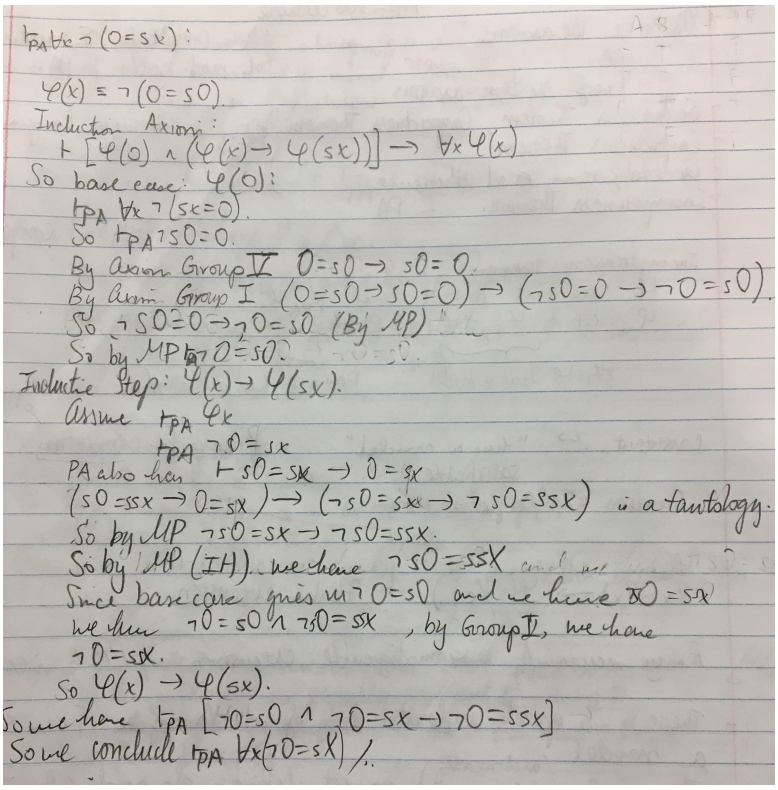

FIGURE 1.1. Proof $PA \vdash \forall x \neg(0 = sx)$

up of a set of blocks, typically denoted by B, and a set of operators, typically denoted by K. These blocks are the starting objects, they are given to us a priori. The set of operators are also given to us as initial methods, and by applying them to our blocks or applying them recursively to the objects we create by applying them to blocks, we generate the whole structure. This final collection of objects, or the whole structure, is denoted $C(B, K)$. For example, in the case of the natural numbers $\mathbb{N}$, $B = \{0\}$ and $K = \{s\}$ where s is the successor function. We generate the whole set $\mathbb{N}$ this way by having first 0, then taking $s0$, then $ss0$, then $sss0$, and so on. Below is a more formal definition of $C(B, K)$:

Definition. $C(B, K)$ is the set generated from B by K where

Case 1. every element of B is also in $C(B, K)$ and
Case 2. if F is an n- place operator in K and takes arguments $c_1, c_2, ..., c_k$ which are in $C(B, K)$, then $F(c_1, c_2, ..., c_k)$ is also in $C(B, K)$, and
Case 3. every element of $C(B, K)$ is obtained by Case 1 or Case 2 described above. If $C = C(B, K)$, then (B, K) is called an inductive structure on C.

So this basically formalizes the idea that a whole set can be generated with a sete of base objects and recursively applying operations on them. It is this property, that a structure is made up of smaller parts, that allows us to perform induction on it. However, before we get there, we still need to have one more definition:

Definition. Let $C = C(B, K)$ be an inductive structure, and let P be a property that elements of C may or may not have. Let F be an n- place operator in K, then we say F preserves P if whenever $a_1, a_2, ..., a_n$ satisfy P, then $F(a_1, a_2, ..., a_n)$ satisfy P.

With these two definitions together we can define the induction law.

Definition (The Induction Law)**.** Let $C = C(B, K)$ be an inductive structure with blocks B and operators K, such that the following is true: (a) every block satisfies the property P and (b) every operator satisfies the property P, then we say every element of C satisfies the property P.

3. Sentential Logic

In this section we will define the language of sentential logic as an inductive structure. Let $B = \{A_1, A_2, ...\}$ be a set of distinct symbols. We will call these the atomic sentences or sentential symbols. Let $F = \{\neg, \wedge, \vee, \rightarrow, \leftrightarrow, |\}$. We will call these elements in F connectives. Our alphabet for sentential logic will be the set $S = B \cup F \cup \{(,)\}$. S^+ is the collection of all finite sequences of S. However, this is just the set of symbols we will use, and this set doesn't say anything about correct syntax in constructing sentences in SL. We need to define some operations over S^+. Let the following operations be defined:

$$\begin{aligned} F_\neg(\alpha) &= (\neg\alpha) \\ F_\wedge(\alpha, \beta) &= (\alpha \wedge \beta) \\ F_\vee(\alpha, \beta) &= (\alpha \vee \beta) \\ F_\rightarrow(\alpha, \beta) &= (\alpha \rightarrow \beta) \\ F_\leftrightarrow(\alpha, \beta) &= (\alpha \leftrightarrow \beta) \\ F_|(\alpha, \beta) &= (\alpha|\beta) \end{aligned}$$

If we let K be the set of all these elements, then we define $\mathcal{L} = C(B, K)$ and we say $\mathcal{L}$ is the language of SL. Elements of $\mathcal{L}$ are called sentential formulas. Another way of defining $\mathcal{L}$ (due to Dedekind) is to say that the set $\mathcal{L}$ is the smallest set composed of B and K and where the elements of $\mathcal{L}$ is closed under the operators in K. Note here that though we have defined the language and structure of $\mathcal{L}$, we have said nothing about its semantics. We will eventually get to that, however, now that we have the structure of $\mathcal{L}$, we have to proceed to a discussion on what that means in terms of its syntax.

One key feature of the language that we want to ensure is that sentences or formulas in the language are uniquely readable, that is, if we write down a formula, there is no ambiguity as to which connectives operate on which symbols. We want to be able to prove that $\mathcal{L}$ as we have defined it satisfies this property, however, before we are able to get to it, we have to prove some lemmas along the way.

Lemma. *Let $\alpha \in \mathcal{L}$, then the number of left parentheses in α is equal to the number of right parentheses in α.*

Proof. We proceed with a proof by induction on $\mathcal{L}$. Let P be the property of having an equal number of left parentheses and right parentheses. Then:

□

Case 1. If $\alpha \in B$ then we are done because all elements of B have no parentheses, and thus satisfy P.
Case 2. Assume that α, β satisfy P, then $F_\neg(\alpha) = (\neg\alpha)$ satisfies P and $F_@(\alpha, \beta) = (\alpha@\beta)$ satisfy P, where $@ = \wedge, \vee, \rightarrow, \leftrightarrow, |$.

In general a lot of our proofs on $\mathcal{L}$ will follow this format of induction.

Definition. We say that $\alpha \in S^+$ is an initial segment of $\beta \in S^+$ if there exists a $\gamma \in S^+$ such that $\alpha\gamma = \beta$.

Lemma. *If $\alpha \in \mathcal{L}$ and α' is an initial segment of α, then α' has more left parentheses than right parentheses.*

Proof. We prove by induction. Base: If $\alpha \in B$ then α has no initial segments and we are done. Inductive:

□

Case 1. If $\alpha = (\neg\beta)$ then we have four cases to consider: (a): $\alpha' = ($, (b): $\alpha' = (\neg$, (c): $\alpha' = (\neg\beta'$, (d): $\alpha' = (\neg\beta$. In cases (a) and (b) it is obvious that they have more left parentheses than right parentheses. Case (c): By the induction hypothesis β' has more left parentheses than right parentheses, and since α' has the same number or right parentheses than β' and one more left parenthesis, α' has more left parentheses than right parentheses. Case (d): By the first lemma β has a balanced number of parentheses, so α', in adding one more left parenthesis, has more left parentheses than right parentheses.

Case 2. If $\alpha = (\beta@\gamma)$ then we have six cases to consider: (a): $\alpha' = ($, (b): $\alpha' = (\beta$, (c): $\alpha' = (\beta'$, (d): $\alpha' = (\beta@$, (e): $\alpha' = (\beta@\gamma'$, (f): $\alpha' = (\beta@\gamma$. In cases (a),(b), (d) and (f), β and/or γ have a balanced number of parentheses, so α' has one more left parentheses than right parentheses. In cases (c) and (e) by the inductive hypothesis β' and γ' have more left parentheses than right parentheses, and so α' will have more left than right parentheses, proving the theorem.

Corollary. *From the lemma above we have the corollary above that no initial segment a member of $\mathcal{L}$ is a member of $\mathcal{L}$.*

Proof. This is a rather straight forward proof because of the two lemmas above. Our first lemma gives us that every member of $\mathcal{L}$ has balanced parentheses, and the second lemma gives us that every initial segment of a member of $\mathcal{L}$ has unbalanced parentheses. Therefore, every initial segment of a member of $\mathcal{L}$ is not a member of $\mathcal{L}$.

□

We have this final lemma to prove before moving onto proving the unique readability theorem:

Lemma. *If $\alpha \in \mathcal{L}$, then α is a block or the first symbol of α is a left parenthesis.*

Proof. If α is a block, then we are done. If α is not a block, then α is of the form one of the elements of K, all of which start with a left parenthesis, so we are done.

□

Unique Readability Theorem of $\mathcal{L}$. *If $\alpha \in \mathcal{L}$, then α falls exactly into one of the following cases:*

Case 1. α is in B
Case 2. there exists a unique γ in $\mathcal{L}$ such that $\alpha = (\neg\gamma)$
Case 3. there exist unique β, γ in $\mathcal{L}$ such that $\alpha = (\beta@\gamma)$.

Proof. If α is in B then we are done. Otherwise, we proceed by a proof by contradiction, that is, we assume uniqueness does not hold. So we have the three following cases:

□

Case 1. $\alpha = (\neg\delta)$ and $\alpha = (\beta@\gamma)$. Therefore $(\neg\delta) = (\beta@\gamma)$, then $\neg\delta = \beta@\gamma$, which implies the first symbol of β is $\neg$, which contradicts our third lemma. Therefore $\beta \notin \mathcal{L}$.

Case 2. $\alpha = (\neg\delta)$ and $\alpha = (\neg\beta)$. Taking segments, we therefore conclude $\delta = \beta$.

Case 3. $\alpha = (\epsilon@\delta)$ and $\alpha = (\beta\triangle\gamma)$ where $\triangle$ is one of $\wedge, \vee, \rightarrow, \leftrightarrow, |$. We therefore have ϵ is either an initial segment of β or β is an initial segment of ϵ or $\epsilon = \beta$. But by the corollary, the first two cases cannot be so we have $\epsilon = \beta$. By the same argument, $\delta = \gamma$. Therefore we conclude $@ = \triangle$, and so the theorem is proven.

Definition. A complete truth assignment S is a function from the set of all atomic blocks $\{A_1, A_2, ...\}$ to the two element set $\{T, F\}$. When we write $S(A_1) = T$, we mean that the truth assignment S makes A_1 true.

A truth assignment is the same thing as a complete truth assignment except that S goes from a subset of the set of all atomics to $\{T, F\}$.

Definition. Let S be a complete truth assignment. We define the truth function $\bar{S} : \mathcal{L} \to \{T, F\}$ inductively as follows:

Case 1. If α is a block then $\bar{S}(\alpha) = S(\alpha)$.

Case 2. If $\alpha = (\neg\beta)$ for some $\beta \in \mathcal{L}$ then $\bar{S}(\alpha) = \begin{cases} T & \bar{S}(\beta) = F \\ F & \bar{S}(\beta) = T \end{cases}$

Case 3. If $\alpha = (\beta \wedge \gamma)$ then $\bar{S}(\alpha) = \begin{cases} T & \bar{S}(\beta) = T = \bar{S}(\gamma) \\ F & otherwise \end{cases}$

Case 4. If $\alpha = (\beta \vee \gamma)$ then $\bar{S}(\alpha) = \begin{cases} F & \bar{S}(\beta) = F = \bar{S}(\gamma) \\ T & otherwise \end{cases}$

Case 5. If $\alpha = (\beta \to \gamma)$ then $\bar{S}(\alpha) = \begin{cases} F & \text{if } \bar{S}(\beta)\text{=T and } \bar{S}(\gamma) = F \\ T & otherwise \end{cases}$

Case 6. If $\alpha = (\beta \leftrightarrow \gamma)$ then $\bar{S}(\alpha) = \begin{cases} T & \bar{S}(\alpha) = \bar{S}(\gamma) \\ F & otherwise \end{cases}$

Case 7. If $\alpha = (\beta|\gamma)$ then $\bar{S}(\alpha) = \begin{cases} F & \bar{S}(\beta) = T = \bar{S}(\gamma) \\ T & otherwise \end{cases}$

Definition. We say S satisfies α if $\bar{S}(\alpha) = T$. We say a set of sentential formulas ϕ is satisfiable if there is a truth assignment S which satisfies ϕ.

4. First Order Logic

Definition. The alphabet of a First Order Language (FOL) $\mathcal{L}$ contains a set of individual constants $\mathcal{C}$, a set of variables x_i where we index variables by the natural numbers, relation symbols R_j^n where j is the index and n is the arity of the symbol, function symbols F_j^n where j is the index and n is the arity of the symbol, the set of logical connectives $\{\neg, \wedge, \vee, \to, \leftrightarrow, |\}$ and the quantifier symbol $\forall$.

Note that because FOL is complete without the $\exists$ quantifier we do not formally include it in "alphabet" but we do use it as "syntactic sugar".

What we want to do is from this alphabet get to formulas in the language where we can make logical statements, but before we do that we first need to define what a term is, which we define as an inductive structure.

Definition. A term is:

(1) $B = \mathcal{C} \cup \{x_i\}$ where $\mathcal{C}$ is my set of constants and x_i are all my variables
(2) $K = F_j^n$ where F_j^n are all my functions.
(3) Then $C = C(B, K)$ is my set of terms. τ is a term if $\tau \in C$. We apply F_j^n on terms to get new terms in the set like so: if $\tau_1, \tau_2, .., \tau_n$ are terms then $F_j^n(\tau_1, \tau_2, ..., \tau_n)$ is a term.

Next we define formulas, which we also define as an inductive structure.

Definition. A formula is:

(1) $B =$ atomic formulas where the atomic formulas are:

Case 1. $\tau = \sigma$ where τ, σ are terms.
Case 2. $R^n_j \tau_1, \tau_2, ..., \tau_n$ where $\tau_1, \tau_2, ..., \tau_n$ are terms

(2) $K = \{\forall, \neg, \wedge, \vee, \rightarrow, \leftrightarrow, |\}$.

Definition. The language $\mathcal{L}$ is the set of all formulas from the alphabet.

So we see that we have made extensive use of inductive structures to define $\mathcal{L}$.

In the previous section we defined the inductive structure of formulas using the inductive structure of terms. In this section we define formulas inductively using prime formulas. Using this definition will allow us to define a truth table for a formula in $\mathcal{L}$ in a very nice way.

Definition. A prime formula is any formula that is atomic or begins with a quantifier. It is not a sentential compound of formulas.

We can now define the inductive structure of formulas.

Definition. A formula is:

(1) $B=$ the set of all prime formulas
(2) $K = \{\forall, \neg, \wedge, \vee, \rightarrow, \leftrightarrow, |\}$.

Now we can define the notion of a truth table in our first order language.

Definition. A truth table row for a formula $\alpha \in \mathcal{L}$ is a sequence $(\alpha_1, T/F), (\alpha_2, T/F), ...$ of formulas such that every α_i in the sequence is prime or a sentential compound of earlier ones and α is the last one and:

(1) If $\alpha_j = \neg\alpha_k$ and $k < j$ then α_j is assigned the opposite of α_k
(2) If $\alpha_j = (\alpha_k \wedge \alpha_m)$ and $k, m < j$ then α_j is assigned T if and only if both α_k and α_m are both assigned T. Otherwise, α_j is assigned F.
(3) If $\alpha_j = (\alpha_k \vee \alpha_m)$ and $k, m < j$ then α_j is assigned F if and only if both α_k and α_m are both assigned F. Otherwise, α_j is assigned T.
(4) If $\alpha_j = (\alpha_k|\alpha_m)$ and $k, m < j$ then α_j is assigned F if and only if both α_k and α_m are both assigned T. Otherwise, α_j is assigned T.
(5) If $\alpha_j = (\alpha_k \rightarrow \alpha_m)$ and $k, m < j$ then α_j is assigned F if and only if α_k is assigned T and α_m are is assigned F. Otherwise, α_j is assigned T.
(6) If $\alpha_j = (\alpha_k \leftrightarrow \alpha_m)$ and $k, m < j$ then α_j is assigned T if and only if both α_k and α_m are both assigned T or F. Otherwise, α_j is assigned F.

Definition. A formula $\alpha \in \mathcal{L}$ is called a tautology if it is true in every row in its truth table.

Now that we have defined the language, what we essentially have is the syntax, and now we need to talk about interpreting the language, which we do by defining the concept of a model of a language.

Definition. A model $\mathcal{M}$ of a language $\mathcal{L}$ is given by

(1) A set, called the domain or universe, denoted by $|\mathcal{M}|$ that is non-empty.
(2) For every constant $c_i \in \mathcal{L}$, an element $c_i^{\mathcal{M}} \in |\mathcal{M}|$.

(3) For every function symbol $F_j^n \in \mathcal{L}$, a function $F_j^{n\mathcal{M}} : |\mathcal{M}|^n \to |\mathcal{M}|$. An equivalent definition is $F_j^{n\mathcal{M}} \subseteq |\mathcal{M}|^{n+1}$ as we have for every $< m_1, m_2, ..., m_n > \in |\mathcal{M}|^n$, there is a unique $m_{n+1} \in |\mathcal{M}|$ such that $< m_1, m_2, ..., m_n, m_{n+1} > \in F_j^{n\mathcal{M}}$.
(4) For every relation symbol $R_j^n \in \mathcal{L}$ a subset $R_j^{n\mathcal{M}} \subseteq |\mathcal{M}|^n$.

This is cool, but before we can define anything like truth in a model we need to define the concepts of $\mathcal{M}$-terms and $\mathcal{M}$-formulas that correspond to terms and formulas in a language $\mathcal{L}$.

Definition. An $\mathcal{M}$-term is defined with an inductive structure where

(1) $B =$ variables x_i, constants of $\mathcal{L}$ c_i, and all elements of $|\mathcal{M}|$.
(2) $K =$ For each function symbol F_j^n of $\mathcal{L}$, and any $\mathcal{M}$-terms $\tau_1, \tau_2, ..., \tau_n$ then $F_j^n \tau_1, \tau_2, ..., \tau_n$ is a $\mathcal{M}$-term.

Definition. An $\mathcal{M}$-term is closed if it does not contain any free variables (more on this later).

Now we need to define what it means to interpret any closed $\mathcal{M}$-term:

Definition. To any closed $\mathcal{M}$-term τ we associate a value $\tau^{\mathcal{M}}$ which is an element of $|\mathcal{M}|$ and we call the interpretation of τin $\mathcal{M}$ by:

Case 1. If τ is the constant symbol c, then $\tau^{\mathcal{M}}$ is the element $c^{\mathcal{M}}$.
Case 2. If $\tau$is an element of $|\mathcal{M}|$, then $\tau^{\mathcal{M}} = \tau$.
Case 3. If τ is the term $F_j^n \tau_1, \tau_2, ..., \tau_n$, where $\tau_1, ..., \tau_n$ are $\mathcal{M}$-terms, then $\tau^{\mathcal{M}} = F_j^{n\mathcal{M}}(\tau_1^{\mathcal{M}}, \tau_2^{\mathcal{M}}, ..., \tau_n^{\mathcal{M}})$. Note that we put the brackets here because these things are no longer just symbols, they have values that, in principle, are computed.

It is important to know the difference between symbols and the interpretation of a symbol. For example, Let $|\mathcal{M}| = \mathbb{R}$, then π is a $\mathcal{M}$-term, but it is not an element of $\mathbb{R}$, rather $\pi^{\mathcal{M}} \in |\mathcal{M}|$ where $\pi^{\mathcal{M}}$ is the interpretation of the symbol πas the constant pi we know and love.

Now of course we need to define $\mathcal{M}$-formulas and the interpretation of $\mathcal{M}$-formulas.

Definition. An $\mathcal{M}$-formula is defined with an inductive structure where

(1) The blocks B are either:

Case 1. $\tau_1 = \tau_2$ where τ_1, τ_2 are $\mathcal{M}$-terms.
Case 2. $R_j^n \tau_1, \tau_2, ..., \tau_n$ for any $\mathcal{M}$-terms $\tau_1, \tau_2, ..., \tau_n$.

(2) For any $\mathcal{M}$-formuals α, β, K is either:

Case 1. $\alpha = (\neg\beta)$
Case 2. $(\alpha @ \beta)$ where $@ = \{\neg, \wedge, \vee, \to, \leftrightarrow, |\}$
Case 3. $\forall x_i \alpha$

We interpret $\mathcal{M}$-formuals with the notion of a closed formula being true in the model $\mathcal{M}$. Another way of saying this is if α is a closed formula that is "true" in $\mathcal{M}$, then we say $\mathcal{M}$ satisfies $\alpha$and we write $\mathcal{M} \models \alpha$.

But before we do this we need to first define substitution, so we can handle variables that might pop up in our formulas.

Definition. Let $\mathcal{M}$ be a model:

(1) For any $\mathcal{M}$-terms τ, σ and any variable x, we define $\sigma(x/\tau)$ inductively as follows:

(a) If σ is a constant symbol c or an element of $|\mathcal{M}|$, then $\sigma(x/\tau) = \sigma$.

(b) If σ is a variable y, then $\sigma(x/\tau) = \begin{cases} \sigma & x \neq y \\ \tau & x = y \end{cases}$
(c) If $\sigma = F\sigma_1, ..., \sigma_n$ then $\sigma(x/\tau) = F\sigma_1(x/\tau), ..., \sigma_n(x/\tau)$.

(2) For any $\mathcal{M}$-formula α, $\mathcal{M}$-term τ and variable x, we define $\alpha(x/\tau)$ inductively as follows:

(a) If α is an equation $\sigma_1 = \sigma_2$ then $\alpha(x/\tau)$ is the equation $\sigma_1(x/\tau) = \sigma_2(x/\tau)$.
(b) If $\alpha = R\sigma_1, ..., \sigma_n$ then $\alpha(x/\tau) = R\sigma_1(x/\tau), ..., \sigma_n(x/\tau)$.
(c) If $\alpha = (\neg\beta)$ then $\alpha(x/\tau) = (\neg\beta(x/\tau))$.
(d) If $\alpha = (\beta@\gamma)$ then $\alpha(x/\tau) = (\beta(x/\tau)@\gamma(x/\tau))$.
(e) If $\alpha = (\forall y\beta)$ where y is a variable different from x then $\alpha(x/\tau) = (\forall y\beta(x/\tau))$.

Now we can define what we mean when we say $\mathcal{M} \models \alpha$ ($\mathcal{M}$ satisfies α):

Definition. If α is a closed $\mathcal{M}$-formula, then we define $\mathcal{M} \models \alpha$ as follows:

(1) If α is the formula $\tau_1 = \tau_2$ where τ_1, τ_2 are closed $\mathcal{M}$-terms, then $\mathcal{M} \models \alpha$ is true if and only if $\tau_1^{\mathcal{M}}$ and $\tau_2^{\mathcal{M}}$ are the same element of $|\mathcal{M}|$.
(2) If α is $R(\tau_1, ..., \tau_n)$ where R is a relation symbol then $\mathcal{M} \models \alpha$ is true if and only if the n-tuple $(\tau_1^{\mathcal{M}}, ..., \tau_n^{\mathcal{M}}) \in R^{\mathcal{M}}$.
(3) If $\alpha = (\alpha_1 \wedge \alpha_2)$ then $\mathcal{M} \models \alpha$ is true if and only if both $\mathcal{M} \models \alpha_1$ and $\mathcal{M} \models \alpha_2$ are true.
(4) If $\alpha = (\alpha_1 \vee \alpha_2)$ then $\mathcal{M} \models \alpha$ is true if and only if $\mathcal{M} \models \alpha_1$ or $\mathcal{M} \models \alpha_2$ or both are true.
(5) If $\alpha = (\alpha_1 | \alpha_2)$ then $\mathcal{M} \models \alpha$ is true if and only if not both $\mathcal{M} \models \alpha_1$ and $\mathcal{M} \models \alpha_2$ are true.
(6) If $\alpha = (\alpha_1 \rightarrow \alpha_2)$ then $\mathcal{M} \models \alpha$ is true if and only if either $\mathcal{M} \models \alpha_1$ is false or $\mathcal{M} \models \alpha_2$ is true.
(7) If $\alpha = (\alpha_1 \leftrightarrow \alpha_2)$ then $\mathcal{M} \models \alpha$ is true if and only if either both or neither of the statements$\mathcal{M} \models \alpha_1$ is false or $\mathcal{M} \models \alpha_2$ are true.
(8) If $\alpha = (\neg\beta)$ then $\mathcal{M} \models \alpha$ is true if and only if $\mathcal{M} \models \beta$ is false.
(9) If $\alpha = (\forall x\beta)$ then $\mathcal{M} \models \alpha$ is true if and only if for all elements $m \in |\mathcal{M}|$, $\mathcal{M} \models \beta(x/m)$ is true.

So we have satisfaction for closed $\mathcal{M}$-formulas, but what about $\mathcal{M}$-formulas that have free variables in them? We first need to define something called the universal closure of a formula to handle this case.

Definition. If the free variables of α are $y_1, ..., y_n$ then we call the formula $\forall y_1 ... \forall y_n \alpha$ the universal closure of α.

Now we can define what satisfaction is for non-closed formulas.

Definition. If α is not closed, then $\mathcal{M} \models \alpha$ if and only if $\mathcal{M} \models \beta$ where β is the universal closure of α.

Finally we define the notion of a valid formula.

Definition. A formula α is valid if and only if for all models $\mathcal{M}$, $\mathcal{M} \models \alpha$. We write this as $\models \alpha$.

Theorem. *All tautologies of a first order language $\mathcal{L}$ are valid formulas.*

Proof. Let α be a tautology in $\mathcal{L}$ and let $\mathcal{M}$ be an arbitrary model of $\mathcal{L}$. Assume that the free variables of $\alpha$are $V = \{x_1, ..., x_n\}$ and let $m_1, ..., m_n$ be elements of $|\mathcal{M}|$. Let P be the set of prime formulas β with free variables among those in V. Let $\mathcal{L}_0$ be the inductive structure obtained from blocks in P and operators $\{\wedge, \neg\}$ only. Note that $\beta \in \mathcal{L}_0$. Now we define a sentential truth assigment s on P by $\forall \beta \in P, s(\beta) = \begin{cases} T & if \mathcal{M} \models \beta(x_1/m_1, ..., x_k/m_k) \\ F & if \mathcal{M} \not\models \beta(x_1/m_1, ..., x_k/m_k) \end{cases}$

Let $\bar{s}$ be the complete truth assigment defined by extending s. Then since α is a tautology, $\bar{s}(\beta) = T$. So $M \models \beta(x_1/m_1, ..., x_k/m_k)$ for every row. So $\mathcal{M} \models \alpha$ (or at least, the universal closure of α).

□

Definition. A theory Γ in language $\mathcal{L}$ is a set of closed formulas or sentences of $\mathcal{L}$. We take all the members of Γ to be true (ie, they are our "axioms").

So for example, PA is a theory of $\mathcal{L}$ that includes the successor, addition, multiplication, exponentiation symbols and the 0 symbol.

Definition. For any theory Γ and any formula α, we write $\Gamma \models \alpha$ if and only if for every model $\mathcal{M}$ of $\mathcal{L}$, for all $\beta \in \Gamma$, we have $\mathcal{M} \models \beta$ and $\mathcal{M} \models \alpha$. We say Γ entails α.

Obviously, for every $\gamma \in \Gamma$, we have $\Gamma \models \gamma$.

We can now start to build a proof system for FoL. We begin by defining the pure axioms of FoL:

Definition (The Pure Axioms)**.**

(1) Group I: Tautology Axioms:

   (a) Let α be a formula. If α is a tautology, then α is a pure axiom.

(2) Group II: Distributivity Axioms:

   (a) If α, β are formulas and x is a variable then $\forall x(\alpha \to \beta) \to (\forall x\alpha \to \forall x\beta)$ is a pure axiom.

(3) Group III: Substitution Axioms:

   (a) If α is a formula, x is a variable and $\tau$is a term, then $\forall x\alpha \to \alpha(x/\tau)$ is a pure axiom.

(4) Group IV: Generalization Axioms:

   (a) If α is a formula, x is a variable and x is not a free variable of α, then $\alpha \to \forall x\alpha$ is a pure axiom.

(5) Group V: Equality Axioms:

   (a) If x, y, z are variables, then:

      (i) $x = x$
      (ii) $x = y \to y = x$
      (iii) $x = y \wedge y = z \to x = z$ are pure axioms.

(6) Group VI: Equivalence Axioms for Relations:

   (a) Assume R is a k-ary relation symbol and $x_1, \ldots, x_k$ and $y_1, \ldots, y_k$ are variables, then $x_1 = y_1 \wedge \ldots \wedge x_k = y_k \to (R(x_1, \ldots, x_k) \leftrightarrow R(y_1, \ldots, y_k))$ is a pure axiom.

(7) Group VII: Equivalence Axioms for Functions:

   (a) Assume F is a k-ary relation symbol and $x_1, \ldots, x_k$ and $y_1, \ldots, y_k$ are variables, then $x_1 = y_1 \wedge \ldots \wedge x_k = y_k \to F(x_1, \ldots, x_k) = F(y_1, \ldots, y_k)$ is a pure axiom.

Definition. A logical axiom is a pure axiom preceeded by any number of $\forall x_i$ symbols where x_i is a variable.

Now that we have our axioms, we can define what a derivation or proof is in our derivation system.

Definition. Let $<\alpha_1,\ldots,\alpha_n>=\bar{\alpha}$ be a sequence of formulas. We say that $\bar{\alpha}$is a derivation (Written as $\vdash \bar{\alpha}$ to say $\bar{\alpha}$ is a logical theorem) if for each j with $1 \leq j \leq n$ at least one of the following conditions holds:

(1) α_j is a logical axiom.
(2) There exists $i,k<j$ such that α_j is derived from α_i and α_k by modus ponens. (This means α_i is the formula $\alpha_k \to \alpha_j$.)

Basically what we want in our derivation system is that a derivation is a sequence of formulas where each succeeding formula we write down is either an axiom or follows from two previous formulas in the sequence by modus ponens.

But this definition of a derivation doesn't allow for using the sentences in Γ as axioms. And so we have the following definition:

Definition. $\bar{\alpha}$ is a derivation from Γ if it is a logical theorem (see above), or it is a derivation with the conditions:

(1) α_j is a logical axiom.
(2) There exists $i,k<j$ such that α_j is derived from α_i and α_k by modus ponens. (This means α_i is the formula $\alpha_k \to \alpha_j$.)
(3) α_j is in Γ (In this case α_j is a nonlogical axiom).

Now that we have a basic framework for proving statements in FoL, we prove a few theorems that are immensely helpful:

The Deduction Theorem. *If* $\Gamma \cup \{\alpha\} \vdash \beta$ *if and only if* $\Gamma \vdash \alpha \to \beta$.

Proof. First we need to show that if $\Gamma \vdash \alpha \to \beta$, then $\Gamma \cup \{\alpha\} \vdash \beta$:

We have the derivation

$$\begin{array}{c} \vdots \\ \alpha \to \beta \end{array}$$

After that last line we assume α (ie, we are taking $\Gamma \cup \{\alpha\}$ instead of Γ as our axiom set, so the derivation becomes

$$\begin{array}{ccc} & \cdots & \\ \alpha & \to & \beta \\ \textit{Assume} & & \alpha \\ & & \beta \end{array}$$

Now we need to show that if $\Gamma \cup \{\alpha\} \vdash \beta$ then $\Gamma \vdash \alpha \to \beta$:

Let

$$\begin{array}{rcl} \Gamma \cup \{\alpha\} & \vdash & \beta_1 \\ & \vdots & \\ \Gamma \cup \{\alpha\} & \vdash & \beta_n \end{array}$$

be a derivation of $\Gamma \cup \{\alpha\} \vdash \beta$. We can produce a derviation of $\Gamma \vdash \alpha \to \beta$ by replacing each line $\Gamma \cup \{\alpha\} \vdash \beta_k$ by one or three lines of the form $\Gamma \vdash \ldots$ ending in $\Gamma \vdash \alpha \to \beta_k$. We can do this because, by induction, we have

If β_k is a logical axiom or a nonlogical axiom in Γ, then we replace $\Gamma \cup \{\alpha\} \vdash \beta_k$ by

$$\Gamma \vdash \beta_k \tag{4.1}$$

$$\Gamma \vdash \beta_k \to (\alpha \to \beta_k) \tag{4.2}$$

$$\Gamma \vdash \alpha \to \beta_k \tag{4.3}$$

The first line was because β_k is an axiom, the second line is a tautology, and the third line follows from the previous two by modus ponens.

If β_k is the nonlogical axiom α then $\alpha \to \beta_k$ is really just $\alpha \to \alpha$ which is a tautology, so we can replace $\Gamma \cup \{\alpha\} \vdash \beta_k$ by $\Gamma \vdash \alpha \to \beta_k$.

Otherwise, if β_k was obtained by modus ponens, say from $\Gamma \cup \{\alpha\} \vdash \beta_i \to \beta_k$ and $\Gamma \cup \{\alpha\} \vdash \beta_i$, then these two lines have already been replaced by lines that contain the lines $\Gamma \vdash \alpha \to (\beta_i \to \beta_k)$ and $\Gamma \vdash \alpha \to \beta_i$. So we replace $\Gamma \cup \{\alpha\} \vdash \beta_k$ by

$$\begin{aligned} \Gamma &\vdash (\alpha \to (\beta_i \to \beta_k)) \to ((\alpha \to \beta_i) \to (\alpha \to \beta_k)) \\ \Gamma &\vdash (\alpha \to \beta_i) \to (\alpha \to \beta_k) \\ \Gamma &\vdash \alpha \to \beta_k \end{aligned}$$

The first line is a tautology and the next two lines follow by modus ponens.

After replacing all n lines we have a derivation of $\Gamma \vdash \alpha \to \beta$.

□

That was the deduction theorem. What this theorem does is basically encapsulate the concept of direct proof. We prove conditions $p \to q$ by first assuming p in a subproof and then showing that q follows. This theorem tells us that we can do that. We now state and prove the generalization theorem.

Generalization Theorem. *Let Γ be a set of formulas and x be a variable such that x is not free for every formula in Γ. If $\Gamma \vdash \alpha$ then $\Gamma \vdash \forall x \alpha$.*

Proof. We prove by induction on the length of the proof.

If α is a logical axiom then $\forall x \alpha$ is a logical axiom.

If $\alpha \in \Gamma$ then x is not free in α, therefore, $\alpha \to \forall x \alpha$ is a logical axiom (Group IV Generalization Axiom). Therefore, $\Gamma \vdash \alpha \to \forall x \alpha$ and since $\Gamma \vdash \alpha$ by modus ponens we have $\Gamma \vdash \forall x \alpha$.

Otherwise, there is a formula β such that $\Gamma \vdash \beta$ and $\Gamma \vdash \beta \to \alpha$. By induction $\Gamma \vdash \forall x \beta$ and $\Gamma \vdash \forall x (\beta \to \alpha)$. Then by modus ponens and the distributivity axiom we have $\Gamma \vdash \forall x \beta \to \forall x \alpha$. And again by modus ponens we have $\Gamma \vdash \forall x \beta$.

□

The generalization theorem allows us to prove $\forall$ introduction, which encapsulates our idea of universal generalization.

Finally we have this one last fact:

Fact. *For any Γ, α if $\Gamma \vdash \alpha$ then there is a finite subset $\Gamma_0 \subseteq \Gamma$ such that $\Gamma_0 \vdash \alpha$.*

A sketch of a proof of this statement is that any initial segment of a sequence (in this case, our derivation), is finite, and so deriving any formula only requires a finite set of axioms.

5. Completeness

Definition. A theory Γ is inconsistent if there is a α such that $\Gamma \vdash \alpha$ and $\Gamma \vdash \neg\alpha$. Γ is consistent if it is not inconsistent.

Theorem. *Γ is inconsistent if and only if $\Gamma \vdash \beta$ for every β.*

Proof. Since $\neg\alpha \to (\alpha \to \beta)$ is a tautology. Hence, if $\Gamma \vdash \alpha$ and $\Gamma \vdash \neg\alpha$ (if Γ is inconsistent) then applying modus ponens we have $\Gamma \vdash \beta$.

□

Theorem. *$\Gamma \vdash \alpha$ if and only if $\Gamma \cup \{\neg\alpha\}$ is inconsistent.*

Proof. Assume $\Gamma \vdash \alpha$. Then $\Gamma \cup \{\neg\alpha\} \vdash \alpha$ since $\Gamma \vdash \alpha$ alone. But that means we also have $\Gamma \cup \{\neg\alpha\} \vdash \neg\alpha$ because we assumed $\neg\alpha$. Therefore, $\Gamma \cup \{\neg\alpha\}$ is inconsistent.

Now assume $\Gamma \cup \{\neg\alpha\}$ is inconsistent. That means Γ derives every formula, including α. So we have $\Gamma \cup \{\neg\alpha\} \vdash \alpha$. By the deduction theorem, we have $\Gamma \vdash \neg\alpha \to \alpha$. Since $(\neg\alpha \to \alpha) \to \alpha$ is a tautology, we have $\Gamma \vdash \alpha$.

□

The Soundness Theorem. *Let Γ be a set of sentences. Assume that $\mathcal{M}$ is an arbitrary model such that $\mathcal{M} \models \Gamma$. Let α be a formula. If $\Gamma \vdash \alpha$, then $\mathcal{M} \models \alpha$. In other words, if $\Gamma \vdash \alpha$, then $\Gamma \models \alpha$.*

Proof. We prove by induction:

If $\alpha \in \Gamma$ then by assumption $\mathcal{M} \models \alpha$.

If α is a logical axiom then $\mathcal{M} \models \alpha$ because $\alpha$is valid.

If α is derived from $\beta \to \alpha$ then by the inductive hypothesis, $\mathcal{M} \models \beta \to \alpha$ and $\mathcal{M} \models \beta$ so $\mathcal{M} \models \beta \to \alpha$.

□

The point of this proof is to capture the notion that everything we can proof in our proof system should be true in the model.

The converse of the soundness theorem is the completeness theorem, which encapsulates the idea that everything that is true should be derivable in our derivation system. This theorem is proved a few sections down.

We need to begin with enumerable sets:

Definition. A set A is enumerable if it is finite or if there is an onto function $f : \mathbb{N} \to A$. We call $\{f(0), f(1), ...\}$ an enumeration of A. Equivalently, A is enumerable if there is a one-to-one function $g : A \to \mathbb{N}$ or if there is a bijection $h : A \to B$ where B is some initial segment of $\mathbb{N}$.

From here we have a few quick and easy results:

(1) Every non-empty subset of an enumerable set is enumerable.
(2) The union of enumerably many sets is enumerable.
(3) The cartesian product of finitely many enumerable sets is enumerable.

We can prove 2 the same way we normally prove the rational numbers are countable.

Proof. Proof of 3:

Let X and Y be enumerable infinite sets. $X \times Y = \{(x_0, y) : y \in Y\} \cup \{(x_1, y) : y \in Y\} \cup \cdots$, where $\{x_0, x_1, \ldots\}$ is an enumeration of X. Each $\{(x_i, y) : y \in Y\}$ is enumerable (as there are as many elements in the set as there are elements in Y). Since $X \times Y$ is the union of all these sets, by 2 we conclude that $X \times Y$ is enumerable. We can use induction to prove the product of n enumerable sets is enumerable for $n \in \mathbb{N}$.

□

Here's where things get relevant for us:

(1) Let $\mathcal{L}$ be a FoL language. Then the set of terms of $\mathcal{L}$ is enumerable.
(2) The set of all sentences in $\mathcal{L}$ is enumerable.
(3) The set of all formulas $\alpha(x)$ with x as sole free variable is enumerable.

We can prove this all at a stroke by proving the more general theorem:

Theorem. *For any enumerable set S, the set W of all strings of elements of S (words of S) is enumerable.*

Proof. For any natural number $n \in \mathbb{N}$, the set W_n of strings of elements of S of length n is enumerable because it is the cartesian product of enumerable sets. And since $W = W_0 \cup W_1 \cup \cdots$, W is enumerable.

□

Compactness Theorem for Sentential Logic. *Let ϕ be an enumerable set of sentential formulas. Then ϕ is satisfiable if and only if every finite subset of ϕ is satisfieable.*

The analogous theorem for this in terms of FoL is given below:

Compactness Theorem for First Order Logic. *Let Γ be a theory, then Γ is consistent if and only if every finite subtheory of Γ is consistent.*

Definition. A consistent first order theory Γ is said to be complete if for every sentence α of $\mathcal{L}(\Gamma)$, either $\Gamma \vdash \alpha$ or $\Gamma \vdash \neg\alpha$.

Theorem. *For every consistent FoL theory Γ there is a complete, consistent extension Γ' ($\Gamma \subseteq \Gamma'$ where Γ' is in the same language of Γ).*

Proof. Enumerate all sentences of the language $\alpha_0, \alpha_1, \ldots$ Define $\Gamma_0 = \Gamma$. We define Γ_{n+1} recursively by having $\Gamma_{n+1} = \Gamma_n \cup \{\alpha_n\}$ if $\Gamma_n \cup \{\alpha_n\}$ is consistent and $\Gamma_{n+1} = \Gamma_n$ if $\Gamma_n \cup \{\alpha_n\}$ is inconsistent. We define $\Gamma' = \cup_{n\in\mathbb{N}}\Gamma_n$.

Proof by contradiction that Γ' is consistent:

Assume Γ' was inconsistent. Then some finite subtheory would be inconsistent (see Compactness Theorem for FoL). But since each finite subtheory is contained in some Γ_{n+1} and so some first Γ_{n+1} is the first to be inconsistent, which by construction is impossible. Since there is no first finite subtheory Γ_{n+1} that it inconsistent, no finite subtheory of Γ' is inconsistent. So we have a contradiction.

Proof that Γ' is complete:

Let αbe a sentence of Γ. We denote $\alpha_i = \alpha$. If $\alpha_i \in \Gamma'$ then $\Gamma' \vdash \alpha_i$. If $\alpha_i \notin \Gamma'$ then it must be the case that $\Gamma_i \cup \{\alpha_i\}$ is inconsistent. That means $\Gamma_i \vdash \neg\alpha_i$. So $\Gamma' \vdash \neg\alpha$. So either $\Gamma' \vdash \alpha$ or $\Gamma \vdash \neg\alpha$ and not both.

□

It must be noted that this process of producing an extension of Γ is not computationally feasible, as there is no efficient way of deciding whether $\Gamma \cup \{\alpha\}$ is consistent or not and we have to decide that for infinitely many sentences.

The Completeness Theorem. *A theory Γ is consistent if and only if it has a model.*

Proof. Proof that if Γ has a model, then Γ is consistent:

Suppose Γ has a model $\mathcal{M}$. If Γ is inconsistent, let α be a sentence such that $\Gamma \vdash \alpha \wedge \neg\alpha$. By the soundness theorem, this means we have $\mathcal{M} \models \alpha \wedge \neg\alpha$, which means $\mathcal{M} \models \alpha$ and $\mathcal{M} \models \neg\alpha$, which is impossible. So Γ is consistent.

□

Now that we have this proof, let us introduce an equivlant formulation of the completeness theorem:

Theorem. *If α is a formula, then $\Gamma \vdash \alpha$ if and only if $\Gamma \models \alpha$.*

This encapsulates the idea I mentioned earlier, that the completeness theorem means that everything that is true in our model should be deriveable in our derivation system.

To see that the completeness theorem implies the theorem above, note that $\Gamma \vdash \alpha \to \Gamma \models \alpha$ is the soundness theorem. Now, if $\Gamma \models \alpha$, then α is valid in every model of Γ. That means $\Gamma \cup \{\neg\alpha\}$ does not have a model. By the completeness theorem, that would mean $\Gamma \cup \{\neg\alpha\}$ is inconsistent, which we showed earlier means $\Gamma \vdash \alpha$.

To see that the theorem above implies the completeness theorem, note that Γ has no model if and only if $\Gamma \models \alpha \wedge \neg\alpha$ for any α. By the theorem above, $\Gamma \models \alpha \wedge \neg\alpha$ implies $\Gamma \vdash \alpha \wedge \neg\alpha$, so Γ is inconsistent.

To prove the completeness theorem we need to construct a model for Γ (assuming Γ is consistent).

We start by defining the idea of a Henkin Theory:

Definition. A theory Γ is Henkin if for every sentence of the form $\exists x\theta$ in $\mathcal{L}(\Gamma)$, there is a constant c in $\mathcal{L}(\Gamma)$ such that $\Gamma \vdash \exists x\theta \to \theta(x/c)$.

If we have a constant c such that $\Gamma \vdash \theta(x/c)$, then we say c is a witness for θ.

Lemma. *If Γ is a consistent theory in the language $\mathcal{L}$, then there is a language $\mathcal{L}' \supseteq \mathcal{L}$ and a consistent theory $\Gamma' \supseteq \Gamma$ such that, for all sentences of the form $\exists y\alpha$ in $\mathcal{L}$ there is a constant symbol $c \in \mathcal{L}'$ suc hthat the formula $\exists y\alpha \to \alpha(y/c)$ is in Γ'.*

Proof. Construction:

Let $\{\alpha_0, \alpha_1, \ldots\}$ be an enumeration of the sentences of $\mathcal{L}$ which are of the form $\exists x\theta$. So for each natural number i there is a formula θ_i and a variable y_i such that $\alpha_i = \exists y_i\theta_i$. For each i we will add a constant c_i to $\mathcal{L}$ to make $\mathcal{L}'$ and add the axiom $\exists y_i\theta_i \to \theta(y_i/c_i)$ to Γ to make Γ'

Proof Γ' is consistent:

Assume Γ' is inconsistent. Then some finite subtheory of Γ' must be inconsistent. So there is some k such that $\Gamma \cup \{\exists y_i\theta_i \to \theta_i(y_i/c_i) : i < k\}$ is consistent but $\Gamma \cup \{\exists y_i\theta_i \to \theta_i(y_i/c_i) : i < k\} \cup \{\exists y_k\theta_k \to \theta_k(y_k/c_k)\}$ is inconsistent. Let Γ^* be $\Gamma \cup \{\exists y_i\theta_i \to \theta_i(y_i/c_i) : i < k\}$. The $\Gamma^* \vdash \neg(\exists y_k\theta_k \to \theta_k(y_k/c_k)$. So expanding out the $\neg$, we have $\Gamma^* \vdash \exists y_k\theta_k \wedge \neg\theta_k(y_k/c_k)$. So $\Gamma^* \vdash \exists y_k\theta_k$. However, since $\Gamma^* \vdash \neg\theta_k(y_k/c_k)$, by the generalization theorem we also have $\Gamma^* \vdash \forall x\neg\theta_k(y_k/x)$ which is a contradiction. Therefore, Γ' is consistent.

□

So by this construction we have constructed a theory Γ' which is consistent and very α in $\mathcal{L}$ has a witness. We repeat the process so that every $\alpha \in \mathcal{L}'$ also has a witness to create a consistent Henkin extension of Γ:

Theorem. *Assume Γ is a consistent theory in the language $\mathcal{L}$, then there exists a $\mathcal{L}_H \supseteq \mathcal{L}$ and a consistent theory $\Gamma_H \supseteq \Gamma$ in the language $\mathcal{L}_\mathcal{H}$ such that Γ_H is Henkin.*

Proof. Let $\Gamma_0 = \Gamma$ and $\mathcal{L}_0 = \mathcal{L}$. For each n, we construct Γ_{n+1} and $\mathcal{L}_{n+1}$ from Γ_n and $\mathcal{L}_n$ as above. We then take $\Gamma_H = \cup\{\Gamma_n : n \in \mathbb{N}\}$. Γ_H is Henkin and $\mathcal{L}(\Gamma_H)$ is enumearable. We call Γ_H the Henkinization of Γ.

Proof Γ_H is consistent:

If Γ_H is inconsistent, then there must be a finite inconsistent subset, so for some n, Γ_n must be inconsistent. But by the lemma above, we see that Γ_n must be consistent. Since there is no first finite inconsistent subset, Γ_H must be consistent.

□

Our strategy moving on will be to show that the Henkin extension of Γ, Γ' has a model $\mathcal{M}'$. We then show that we can get a model $\mathcal{M}$ of Γ from $\mathcal{M}'$. We formalize this correspondence below:

Definition. Let $\mathcal{L}$ be a language and let $\mathcal{L}'$ be an extension of $\mathcal{L}$. Let $\mathcal{M}$ be a model for $\mathcal{L}$, and let $\mathcal{M}'$ be a model for $\mathcal{L}'$. We say that $\mathcal{M}'$ is an expansion of $\mathcal{M}$ if

(1) $|\mathcal{M}'| = |\mathcal{M}|$
(2) Every relation symbol R in $\mathcal{L}$ is also a relation symbol R in $\mathcal{L}'$ and $R^\mathcal{M} = R^{\mathcal{M}'}$.
(3) Every function symbol F in $\mathcal{L}$ is also a function symbol F in $\mathcal{L}$' and $F^\mathcal{M} = F^{\mathcal{M}'}$.
(4) Every constant c in $\mathcal{L}$ is also a constant $\mathcal{L}'$ and $c^\mathcal{M} = c^{\mathcal{M}'}$.

We call $\mathcal{M}$ a restriction of $\mathcal{M}'$ to $\mathcal{L}$, and we write $\mathcal{M} = \mathcal{M}'|\mathcal{L}$.

Now we establish this correspondence between $\mathcal{M}'$ and $\mathcal{M}$ with the following lemma:

Lemma (Restriction Lemma)**.** *Let Γ and Γ' be theories in language $\mathcal{L}$ and $\mathcal{L}'$ respectively, with Γ' an extension of Γ and $\mathcal{L}'$ an extension of $\mathcal{L}$. Let $\mathcal{M}'$ be a model of Γ', then $\mathcal{M}'|\mathcal{L}$ is a model of Γ.*

Proof. Let $\mathcal{M} = \mathcal{M}'|\mathcal{L}$. We claim:

□

(1) For all closed $\mathcal{M}$-terms τ, $\tau^{\mathcal{M}} = \tau^{\mathcal{M}'}$.
(2) For all closed $\mathcal{M}$-formulas α, $\mathcal{M} \models \alpha$ if and only if $\mathcal{M}' \models \alpha$.
(3) $\mathcal{M} \models \Gamma$.

Proof. We prove 1 by induction on τ, using our definition above.

We prove 2 by induction on α using our definition above.

Proof of 3: Every formula $\alpha \in \Gamma$ is also in Γ', hence $\mathcal{M}' \models \alpha$, so by 2, $\mathcal{M} \models \alpha$.

□

Now we are ready to construct a model from the constants. Let Γ be a consistent theory. Let Γ_H be the Henkinization of Γ. We know that Γ_H must be consistent. Hence, Γ_H has a complete, simple extension Γ^*. Since Γ^* is in the same language of Γ_H, Γ^* is also a Henkin theory. We will show that Γ^* has a model $\mathcal{M}^*$. By the restriction lemma above, this will be enough to show that $\mathcal{M}^*|\mathcal{L}(\Gamma)$ is a model of Γ.

We define a relation $\sim$on the set of closed terms of $\mathcal{L}(\Gamma^*)$ by $\tau \sim \mu$ if and only if $\Gamma^* \vdash \tau = \mu$. $\sim$ is an equivalence relation. From this we can start to construct our model and universe for Γ^*. For $\mathcal{M}^*$, we define

(1) $|\mathcal{M}^*|$ to be the set of equivalence classes under $\sim$.
(2) If c is a constant symbol in $\mathcal{L}(\Gamma^*)$, we let $c^{\mathcal{M}^*} = [c]$ where $[\tau]$ denotes the equivalence class of τ.
(3) Let F be an n-ary function symbol, we define $F^{\mathcal{M}^*}([\tau_1] \cdots [\tau_n]) = [F\tau_1 \cdots \tau_n]$.
(4) For any n-ary relation symbol R, we define $R^{\mathcal{M}^*} = \{([\tau_1], \cdots, [\tau_n]) \in \mathcal{M}^* | \Gamma^* \vdash R\tau_1 \cdots \tau_n\}$

Using this definition, we can move on to claim and prove $\mathcal{M}^* \models \Gamma^*$, but before we do so, we need to show that each closed term in $\mathcal{L}(\Gamma^*)$ has the required interpretation in $\mathcal{M}^*$:

Lemma. *If τ is a closed term in $\mathcal{L}(\Gamma^*)$, then $\tau^{\mathcal{M}^*} = [\tau]$.*

Proof. By induction on τ. Since τ is a closed term, it does not contain any variables and must be either a constant symbol or of the form $F\tau_1 \cdots \tau_n$.

If $\tau$is a constant symbol c, then by definition we have $\tau^{\mathcal{M}^*} = c^{\mathcal{M}^*} = [c] = [\tau]$.

If $\tau = F\tau_1 \cdots \tau_n$, then

$$\begin{aligned} \tau^{\mathcal{M}^*} &= (F\tau_1 \cdots \tau_n)^{\mathcal{M}^*} \\ &= F^{\mathcal{M}^*}((\tau_1)^{\mathcal{M}^*} \cdots (\tau_n)^{\mathcal{M}^*}) \\ &= F^{\mathcal{M}^*}([\tau_1] \cdots [\tau_n]) \\ &= [F\tau_1 \cdots \tau_n] \\ &= [\tau] \end{aligned}$$

□

We can now prove the following:

Theorem. *If α is a sentence of $\mathcal{L}(\Gamma^*)$, then $\Gamma^* \vdash \alpha$ if and only if $\mathcal{M}^* \models \alpha$.*

Proof. By induction on α. If $\alpha$is atomic then α is $R\tau_1 \cdots \tau_n$, or $\alpha$is $\tau_1 = \tau_2$.

If $\alpha = R\tau_1 \cdots \tau_n$, then $\Gamma^* \vdash \alpha$ if and only if

$$\begin{array}{rll} \Gamma^* & \vdash & R\tau_1 \cdots \tau_n \\ & iff & ([\tau_1] \cdots [\tau_n]) \in R^{\mathcal{M}^*} \\ & iff & ((\tau_1)^{\mathcal{M}^*} \cdots (\tau_n)^{\mathcal{M}^*}) \in R^{\mathcal{M}^*} \\ iff & \mathcal{M}^* \models & R(\tau_1 \cdots \tau_n) \\ & \models & \alpha \end{array}$$

If α is $\tau = \mu$, then $\Gamma^* \vdash \alpha$ if and only if

$$\begin{array}{rll} \Gamma^* & \vdash & \tau = \mu \\ & iff & \tau \sim \mu \\ & iff & [\tau] = [\mu] \\ & iff & \tau^{\mathcal{M}^*} = \mu^{\mathcal{M}^*} \\ & iff & \mathcal{M}^* \models \tau = \mu \\ & iff & \mathcal{M}^* \models \alpha \end{array}$$

If $\alpha = (\neg\beta)$, then $\Gamma^* \vdash \alpha$ iff

$$\begin{array}{rll} \Gamma^* & \vdash & \neg\beta \\ iff & \Gamma^* \nvdash & \beta \\ iff & \mathcal{M}^* \not\models & \beta \\ iff & \mathcal{M}^* \models & \neg\beta \\ iff & \mathcal{M}^* \models & \alpha \end{array}$$

If $\alpha = (\beta \wedge \gamma)$, then we have $\Gamma^* \vdash \alpha$ iff

$$\begin{array}{rll} \Gamma^* & \vdash & \beta \wedge \gamma \\ iff & \Gamma^* \vdash \beta & and \Gamma^* \vdash \gamma \\ iff & \mathcal{M}^* \models \beta & and \mathcal{M}^* \models \gamma \\ iff & \mathcal{M}^* \models & \beta \wedge \gamma \\ iff & \mathcal{M}^* \models & \alpha \end{array}$$

Finally, if $\alpha = \forall x\beta$. Since Γ^* is a Henkin theory, for every closed term $\tau \in \mathcal{L}(\Gamma^*)$ there is a constant c in $\mathcal{L}^*$ such that $\mathcal{L}^* \models c = \tau$. Assume $\Gamma^* \vdash \forall x\beta$. Then for any element $m \in \mathcal{M}^*$, we can find a term τ such that $m = [\tau]$ is the equivalence class of τ. So we have $\Gamma^* \vdash \beta(x/\tau)$. By the induction hypothesis, we have $\mathcal{M}^* \models \beta(x/\tau)$. As $\tau^{\mathcal{M}^*} = [\tau] = m$, we have $\mathcal{M}^* \models \beta(x/m)$. This can be done for any m, so $\mathcal{M}^* \models \forall x\beta$. Conversely, assume $\mathcal{M}^* \models \forall x\beta$. Let c be a constant such that the formula $\exists x\neg\beta \rightarrow \neg\beta(x/c)$ is in Γ^*. So $\Gamma^* \vdash \beta(x/c) \rightarrow \neg\exists x\neg\beta$, which is equivalent to $\Gamma^* \vdash \beta(x/c) \rightarrow \forall x\beta$. Since $\mathcal{M}^* \models \forall x\beta$, we have $\mathcal{M}^* \models \beta(x/c)$, so by the induction hypothesis, $\Gamma^* \vdash \beta(x/c)$, so by modus ponens, we have $\Gamma^* \vdash \forall x\beta$.

□

So we have the corollary $\mathcal{M}^*$ is a model of Γ^*, completing the completeness theorem.

Definition. Let $\mathbb{N} =< \mathbb{N}, +, \cdot, <, 0, s >$ be the natural numbers with the usual operations and relations. Let $Th(\mathbb{N})$ ("the theory of $\mathbb{N}$") be the set of all sentences that are valid in $\mathbb{N}$. Every model $\mathcal{M}$ of $Th(\mathbb{N})$ is called a model of arithmetic. The model $\mathbb{N}$ is called the standard or natural model of arithmetic.

In this section we will show that PA can't restrict us to only the standard model or arithmetic, and in a sense, leaves enough room for "junk".

Definition. For every term τ and for every natural number n there is a term $s^n\tau$ that is defined inductively as follows: $s^0\tau = \tau$, $s^{n+1}\tau = s(s^n\tau)$.

So basically, every $s^i\tau$ is a term with a finite number of s symbols in front of it.

Definition. Let $\mathcal{M}$ be a model of arithmetic. The finite or standard elements of $\mathcal{M}$ are the elements that are of the form $a = (s^n 0)^{\mathcal{M}}$. The other elements are called nonstandard or infinite. We define $M_{fin} = \{(s^n 0)^{\mathcal{M}} \in \mathcal{M} : n \in \mathbb{N}\}$. If $\mathcal{M}$ has no nonstandrad elements, it it called standard, otherwise, it is called nonstarndard.

Theorem. *There are nonstandard models of arithmetic.*

Proof. We prove this theorem by constructing one from some well chosen axioms. Let $\Gamma = Th(\mathbb{N}) \cup \{0 < c, s0 < c, ss0 < c, \ldots\}$ where c is a constant symbol. To show that Γ has a model, by the completeness theorem, it is enough to show that it is consistent.

Proof Γ is consistent:

By the compactness theorem of FoL, it is enough to show that evey finite subset of Γ is consistent. Let $\Gamma_0 \subseteq Th(\mathbb{N}) \cup \{0 < c, s0 < c, ss0 < c, \ldots\}$ be finite, then $\Gamma_0 \subseteq Th(\mathbb{N}) \cup \{0 < c, s0 < c, ss0 < c, \ldots, s^n 0 < c\}$ for some $n \in \mathbb{N}$. Let $\mathbb{N}' =< \mathbb{N}, +, \cdot, <, 0, s, n+1 >$. So $\mathbb{N} = \mathbb{N}'|\mathcal{L}$ and $c^{\mathbb{N}'} = n+1$. Then $\mathbb{N}' \models \Gamma_0$. By the completeness theorem, since Γ_0 has a model, it is a consistent. So Γ is consistent, which means it has a model.

Now we need to define this model and show that it contains nonstandard elements.

Let $\mathcal{M} =< M, +^{\mathcal{M}}, \cdot^{\mathcal{M}}, <^{\mathcal{M}}, 0^{\mathcal{M}}, s^{\mathcal{M}}, c^{\mathcal{M}} >$ be a model of Γ and let $\kappa = c^{\mathcal{M}}$. Then for every natural number n, $\mathcal{M} \models s^n 0 < c$. Since $\forall x(x < y \to x \neq y) \in Th(\mathbb{N})$ (We've defined the $<$ operator in $\mathbb{N}$), $\mathcal{M} \models s^n 0 < c$ implies $\mathcal{M} \models c \neq s^n 0$, and therefore $\mathcal{M} \models \kappa \neq s^n 0$. Let $\mathcal{M}' = \mathcal{M}|\mathcal{L}$. Then $\mathcal{M}'$ is a nonstandard model of arithmetic, because $\mathcal{M}' \models \kappa \neq s^n 0$ for all natural number n.

□

What this means is that κ is a sort of "infinite" number (we call it a nonstandard number). Since PA proves $\forall x(x = 0 \vee \exists y(x = sy))$, there are x's that are successors of κ. So we have infinitely many $\kappa, \kappa + 1, \kappa + 2, \cdots$. Similarly, there are $\left\lfloor \frac{\kappa}{2} \right\rfloor, \left\lfloor \frac{\kappa}{2} \right\rfloor + 1, \left\lfloor \frac{\kappa}{2} \right\rfloor + 2, \cdots$ as well. So there is actually an infinite amount of junk that lies beyond all the standard numbers that PA leaves room for. PA can't actually restrict us to just the standard numbers.

6. The Incompleteness Theorem

Definition. Whenever n is a natural number, $\underline{n}$ is the term $s \cdots s0$ with n s's. $\underline{n}$ is called a numeral and is a syntactical object, whereas n is a semantic object.

We the symbol $+$ represents addition because in general we have $PA \vdash \underline{n} + \underline{m} = \underline{n+m}$. And so on for the other arithmetic symbols. We can prove these things using the PA axioms and induction.

We can code finite sequences of integers by single integers.

Definition. Let $p_1 := 2, p_2 := 3, p_3 := 5, \ldots, p_k :=$ the k-th prime number. We will code a sequence $< a_1, \ldots, a_n >$ of positive natural numbers by the single natural number $\#(a_1, \ldots, a_n) := 2^{a_1} \cdot 3^{a_2} \cdot \cdots \cdot p_n^{a_n}$, which we will call the code of the sequence $< a_1, \ldots, a_n >$. For any number c and for any $k > 0$, we let $(c)_k = max\{l : p_k^l | c\}$ where $x|y \leftrightarrow \exists z(x \cdot z = y) \leftrightarrow x$ divides y .

We have unique readability in this system, and this relies on the fact that every number has a unique prime factorization. Note, however, that every sequence has a number code but not every number is a code of a sequence.

To represent these relations within the language of PA, we define the following formulas:

Definition. $prime(x) := (x > \underline{1}) \wedge \neg(\exists y \exists z x = y \cdot z \wedge y > \underline{1} \wedge z > \underline{1})$.

So the corresponding set $prime := \{p : \mathbb{N} \models prime(p)\}$ is the set of all prime numbers.

Definition. $nextprime(x, y) := prime(x) \wedge prime(y) \wedge x < y \wedge \forall z(prime(z) \to z \leq x \vee z \geq y)$

So $\mathbb{N} \models nextprime(p, q)$ if and only if p and q are adjacent primes. As mentioned before, not all numbers are codes of sequences, so let's define a set seq which gives the set of all codes of sequences by $seq := \{p_1^{e_1} \cdots p_n^{e_n} : n \in \mathbb{N}, e_1, \ldots, e_n > 0\}$. Let's encode this in PA:

Definition. $seq(c) \coloneqq \forall p, q : (prime(p) \wedge prime(q) \wedge p < q \wedge q|c \to p|c) \wedge c > \underline{1}$.

So $\mathbb{N} \models seq(n)$ if and only if $n \in seq$.

How do we obtain a formula $\alpha(x, y)$ that decides whether x is the y-th prime number, we first consider the following set of numbers:

Let $products \coloneqq \{\#(1), \#(1, 2), \#(1, 2, 3), \cdots\}$.

Encoding this in PA, we have:

Definition. $products(x) \coloneqq \underline{2}|x \wedge \neg\underline{4}|x \wedge seq(x) \wedge \forall p \forall q \forall e(nextprime(q, p) \wedge p|x \to (q^e|x \leftrightarrow p^{se}|x))$

So $products = \{n \in \mathbb{N} : \mathbb{N} \models products(n)\}$.

Using this, we can define the following:

Definition. $nthprime(p, k) \coloneqq prime(p) \wedge \exists w products(w) \wedge p^k|w \wedge \neg p^{k+1}|w$.

Then $\mathbb{N} \models nthprime(p, k)$ if and only if $p = p_k$, the k-th prime number.

Now recall that $e = (c)_k$ if and only if $p_k^e|c$ and $p_k^{e+1} \nmid c$. Let's define the set $entry \coloneqq \{(c, k, e) : c \in seq, e = (c)_k\}$. Informally, $(c, k, e) \in entry$ if and only if c codes a sequence whose k-th entry is e.

Encoding this in PA, we have:

Definition. $entry(c, k, e) \coloneqq seq(c) \wedge (k > 0) \wedge (e > 0) \wedge \forall p(prime(p) \wedge nthprime(p, k) \to p^e|c \wedge \neg p^{e+1}|c)$

So $\mathbb{N} \models entry(c, k, e)$ if and only if c codes a sequence, $(c)_k = e$ and $e > 0$. And whenever $< e_1, \cdots, e_n >$ is a finite sequence of positive integers, there is a code $c \in seq$ such that for all $k \leq n : \mathbb{N} \models entry(c, k, e_k)$.

Now we code terms and formulas of PA as natural numbers, using a system called Godel Numbering:

First we define $\ulcorner x \urcorner$ for all symbols x in our language according to the following table:

x	$\ulcorner x \urcorner$
0	2
s	4
$=$	6
$+$	8
$\cdot$	10
$\uparrow$	12
$\forall$	14
$\wedge$	16
$\neg$	18
x_n	$2n + 1$

Now we define $\ulcorner \tau \urcorner$ by induction on terms as follows:

(1) Blocks $\ulcorner 0 \urcorner$ or $\ulcorner x_n \urcorner$ are as given.

(2) Operators: $\ulcorner \tau_1 + \tau_2 \urcorner = 2^8 \cdot 3^{\ulcorner \tau_1 \urcorner} \cdot 5^{\ulcorner \tau_2 \urcorner}$, and so on.

Since the inductive system on terms in uniquely readable, this system is also uniquely readable.

With this we can define similar a similar formula $term(x)$ such that $\mathbb{N} \models term(n)$ if and only if $n = \ulcorner t \urcorner$ for some term t. We start by defining the atomics:

Definition. $var(x) \coloneqq (\exists y x = \underline{2} \cdot y + \underline{1})$. $zero(x) \coloneqq (x = \underline{\ulcorner 0 \urcorner})$.

We can express the fact that a term t is obtained by joining terms $t_1$and t_2 with an operation o by the following formula $yieldtm(t_1, t_2, o, t)$:

Definition. $yieldtm(t_1, t_2, o, t) \coloneqq [(o = \underline{\ulcorner + \urcorner} \vee o = \underline{\ulcorner \cdot \urcorner} o = \underline{\ulcorner \uparrow \urcorner}) \wedge (t = \underline{2}^o \cdot \underline{3}^{t_1} \cdot \underline{5}^{t_2})] \vee (o = \underline{\ulcorner s \urcorner} \wedge t = \underline{2}^o \cdot \underline{3}^{t_1})$

Informally, $yieldtm(m, n, o, p)$ says: "p codes the term obtained by joining the two terms coded by m and n, using the operation symbol coded by o).

If c is the code of a sequence, $obtained(c, k, l, o, n)$ will express the fact that the n-th element of the sequence was obtained from the k-th and the l-th element, which were joined by the symbol coded by o:

Definition. $obtained(c,k,l,o,n) \coloneqq \exists e,f,g[entry(c,k,e) \wedge entry(c,l,f) \wedge entry(c,n,g) \wedge yieldtm(e,f,o,g)]$

We define a few more formulas:

Definition. $atomicterm(c,n) \coloneqq \exists e entry(c,n,e) \wedge (zero(e) \vee var(e))$.

So $\mathbb{N} \models atomicterm(c,n)$ if and only if the n-th entry in c is the Godel number of an atomic term.

We say that a sequence $< t_1, t_2, \cdots, t_n >$ "builds" the term t if and only if t appears in the sequence, and for all $k \leq n$, t_k is either the constant term 0 or a varialbe x_n, or is of the form $t_i + t_j$, or $t_i \cdot t_j$, or $t_i^{t_j}$ for $i,j < k$, or $t_k = st_j$ for some $j < k$.

Definition. $buildTerm(c,t) \coloneqq seq(c) \wedge \exists n entry(c,n,t) \wedge \forall n \forall n : entry(c,n,e) \to atomicterm(c,n) \vee \exists k,l < n \exists o obtained(c,k,l,o,n)$

This formula expresses the fact that c is the code of a sequence $< \ulcorner t_1 \urcorner, \ldots, \ulcorner t_n \urcorner >$, where $< t_1, \ldots, t_n >$ builds t.

So now we cand define $term(t)$:

Definition. $term(t) \coloneqq \exists c buildTerm(c,t)$.

So $\mathbb{N} \models term(n)$ if and only if there is a term t such that $n = \ulcorner t \urcorner$.

Similarly, we can define formulas:

(1) $formula(f)$ where $\mathbb{N} \models formula(n)$ if and only if there is a formula α such that $n = \ulcorner \alpha \urcorner$.
(2) $theorem(f)$ where $\mathbb{N} \models theorem(\ulcorner \alpha \urcorner)$ if and only if $PA \vdash \alpha$.

The two formulas above require defining a lot of other formulas I'm not really interested in taking the time to define at the moment.

Now we can actually move on to proving the incompleteness theorem:

Definition. If α is a formula, we define $\mathcal{D}(\alpha)$ (the diagonalisation of α) by $\mathcal{D}(\alpha) \coloneqq \alpha(x_0/\underline{\ulcorner \alpha \urcorner})$.

Intuitively, we say given α which expresses a property of x_0 (eg. $\alpha(x_0) \coloneqq x_0$ is prime), then $\mathcal{D}(\alpha)$ says $\ulcorner \alpha \urcorner$ has that property.

We can come up with a formula $\alpha(x_0) \coloneqq$ there is no proof of the diagonalization of x_0, or x_0 is not a theorem. Then $\sigma = \mathcal{D}(\alpha)$ says "I am not a theorem of PA."

Let's examine σ:

Case 1. If $\sigma$is not a theorem of PA, $\sigma$is true, so that means it is not a theorem of PA.
Case 2. If $\sigma$is not a theorem of PA, since this is exactly what σclaims, $\sigma$is valid in $\mathbb{N}$.

So there is a closed formula of PA, which is true but is not provable in PA.